

Supplementary Information: Exploring Chemistry and Catalysis by Biasing Skewed Distributions via Deep Learning

Zhikun Zhang¹ and GiovanniMaria Piccini^{2,*}

¹*Institute of Technical Thermodynamics, RWTH Aachen University, 52062 Aachen, Germany*

²*Dipartimento di Scienze Chimiche e Geologiche, Università degli Studi di Modena e Reggio Emilia, 41125 Modena, Italy*

CONTENTS

S1. Method Details	2
A. Warm-start Training	2
B. Parameters and Optimization	3
1. Wall Construction Parameters	3
2. Neural Network Architecture	3
3. Loss Function Weights	3
4. Batch Size and Dataset Handling	5
5. Training Strategy	5
C. Metastable States Identification	5
D. General Algorithm	5
S2. Case Studies – Computational Details	7
A. Model Potentials	7
1. Comparison of Different $\mathcal{L}_{\text{skew}}$	8
2. Multiple reaction pathways and selectivity with loxodynamics	9
3. Warmstarting efficiency	11
B. Gas Phase Chemical reactions	13
1. S _N 2 Reaction	13
2. Diels-Alder Reaction	13
3. Assessment of Skewencoder CVs in Conventional Enhanced Sampling Methods	15
C. Ethanol Dehydration in Acidic Chabazite under Operando Conditions	17
1. Simulation Setup	17
2. Restart Strategy for Exploring Multiple Potential Wells	18
3. Validation of the Reaction Mechanism Identified by Loxodynamics Simulations	18
D. 1-Butanol Dehydration in Acidic Chabazite	18
S3. Perceptually Uniform Visualizations	21
References	24

*) giovannimaria.piccini@unimore.it

S1. METHOD DETAILS

A. Warm-start Training

As previously discussed, the training data are fed incrementally into our model. Our goal is to ensure that each CV, derived from a sequence of latent spaces, remains effective and leverages all available data up to that point. To achieve this, we employ a warm-start training strategy, which has been widely successful in convex optimization¹.

Rather than training the model from scratch at each iteration, we initialize it with the parameters learned in the previous iteration and train it until convergence. Given the similarity between latent space manifolds in consecutive iterations—particularly in the early stages—this approach allows the model to retain previously learned information, significantly reducing training time, which is a critical overhead in our method.

However, a major challenge in this approach is catastrophic forgetting, which can lead to a loss of generality over multiple iterations. To mitigate this, we introduce a state-detection-informed restart-training strategy. Instead of reinitializing the model in every iteration through weight shrinking or fine-tuning only in later layers²⁻⁴, our approach dynamically determines whether to apply a warm-start or a cold-start based on the detected state of the system.

The Algorithm S1 provides a detailed breakdown of our training strategy, outlining the criteria for switching between warm-start and cold-start training.

As described in the training strategy, it is necessary to determine whether to apply warm-start training based on state detection. The criteria for this decision are system-dependent, and numerous state-of-the-art trajectory analysis tools are already available to identify stable or meta-stable states based on changes in geometry or bond orders. The outputs of these tools can be seamlessly integrated into our method. Therefore, without loss of generality, we maintain a **sequential list** of newly detected stable states.

Algorithm S1 The state-detection-informed restart training strategy of Skewencoder

Require: State list $\mathbf{L}_{\text{state}}$, Current state S_c , Previous state S_p

```

if  $S_c$  not in  $\mathbf{L}_{\text{state}}$  then
  APPEND( $\mathbf{L}_{\text{state}}$ ,  $S_c$ )
  WARM_START_TRAINING  $\leftarrow$  FALSE
else
  if  $S_c$  not stable then
    WARM_START_TRAINING  $\leftarrow$  TRUE
  else if  $S_p$  not stable then
    WARM_START_TRAINING  $\leftarrow$  FALSE
  else
    WARM_START_TRAINING  $\leftarrow$  TRUE
  end if
end if
 $S_p \leftarrow S_c$ 
if WARM_START_TRAINING is TRUE then
  Training from the checkpoint of last iteration.
else
  Training from scratch.
end if
return  $S_p$ ,  $\mathbf{L}_{\text{state}}$ 

```

After sampling with an unbiased MD simulation in the original stable state, the state list $\mathbf{L}_{\text{state}}$ is initialized as $[0, 1]$, where:

- **State 0** indicates that the system is confined by the wall, preventing access to any possible stable states.
- **State 1** represents the original stable state.

Any newly detected stable state is appended to the list cumulatively. **In general, the model will start training from scratch under two conditions:**

1. **The system reaches a new stable state** that has not been stored in the state list.
2. **The system returns to a previously known state** from another state (including state 0).

When the system reaches a new stable state, the previously stored state information **no longer reflects the properties of the new region**, necessitating a reinitialization of the NN parameters. Conversely, if the system

returns from a non-stable state on the PES to a known stable state, it implies that the search encountered a **potential barrier** that is difficult to overcome.

These types of barriers arise due to the complexity of the local PES, and sampling near them provides valuable information about the landscape surrounding known minima. In this case, we restart training from scratch to **redirect the system’s evolution** while simultaneously incorporating the supplementary information gathered near the barrier.

B. Parameters and Optimization

1. Wall Construction Parameters

For the construction of the wall, we set the default value of $\epsilon = 1.0$. The projection of samples to the latent space is always normalized to the interval $[-1, 1]$, and setting $\epsilon = 1.0$ shifts the harmonic wall by half of this range. This adjustment expands the relevant sampling region within the latent space, effectively increasing the speed and scope of PES exploration as more data accumulate in the global dataset under the multitask scheme.

2. Neural Network Architecture

The Autoencoder used in this work consists of several fully connected layers: an input layer, multiple hidden layers, a bottleneck layer with a single neuron, additional hidden layers, and an output layer. In each fully connected layer, the activation of a neuron is computed as a weighted sum of the activations from all neurons in the previous layer. This operation can be expressed as matrix multiplication, as shown in Equation S1:

$$\mathbf{y}_l = f_l(\mathbf{W}_l \mathbf{x}_{l-1} + \mathbf{b}_l) \quad (\text{S1})$$

where $\mathbf{y}_l$ and $\mathbf{x}_{l-1}$ represent the activations of layer l and the preceding layer $l - 1$, respectively. The function f_l is the activation function, $\mathbf{W}_l$ is the weight matrix defining the connections between layers, and $\mathbf{b}_l$ is the bias vector for layer l . The choice of activation function is crucial for introducing nonlinearity into the network, allowing it to effectively capture the nonlinear features of the one-dimensional manifold. The weights and biases serve as trainable parameters optimized during training to minimize the loss function.

The Autoencoder architecture should be tailored to the dimensionality of the input properties. However, we recommend adhering to the following guidelines:

- The encoder should have at least three layers to ensure sufficient depth for representing and approximating continuous multivariate functions⁵.
- The first hidden layer should have a node count greater than the input dimensionality to effectively capture relationships among features.
- Optionally, the final hidden layer before the latent space can be expanded to enhance nonlinearity in the transformation (see right scheme in Figure S2).

Both architectural patterns are illustrated in Figure S2. For activation functions, we use the shifted Softplus function, which offers improved smoothness and nonlinearity over the Exponential Linear Unit (ELU), enhancing both differentiation of input variables and NN convergence⁶.

3. Loss Function Weights

To balance the contributions of skewness loss and Autoencoder reconstruction, we set $\alpha = 0.1$, ensuring sensitivity to changes in skewness while preserving essential conformational structures during state transitions. Additionally, the weight of the L_2 regularization term is set to $\beta = 1 \times 10^{-5}$.

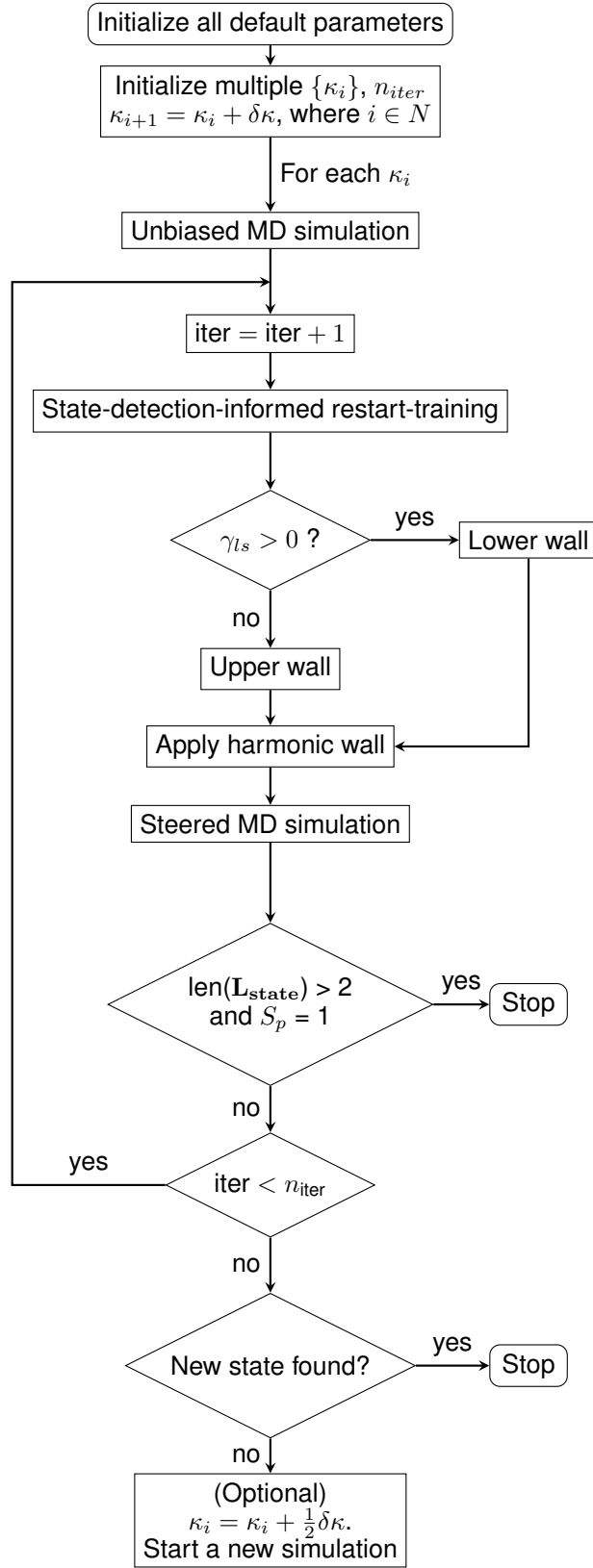


FIG. S1. **Workflow of the general biased sampling algorithm based on Skewencoder for searching the PES space.** In this flowchart, $\{\kappa_i\}$ represents a sequence of wall spring constants arranged in an arithmetic progression with a common difference of $\delta\kappa$. The parameter n_{iter} denotes the maximum number of iterations permitted for the Loxodynamics simulation. The quantity γ_{ls} refers to the projected sample skewness in latent space. The variable L_{state} indicates the list of identified metastable states, while S_p serves as the identifier for the current state.

4. Batch Size and Dataset Handling

For the local dataset, we set a default batch size of 100, as studies⁷ indicate that when sample sizes exceed 80, moment-based skewness estimation remains accurate for a mixture of two Gaussian distributions. For the global dataset, we define the batch size as $100n$, where n represents the current number of iterations, ensuring alignment between batch counts in the local and global datasets.

5. Training Strategy

Model parameters are optimized using gradient descent with the ADAM optimizer, employing a learning rate of 10^{-3} . The dataset is split into 80% training and 20% validation, and early stopping is implemented to mitigate overfitting and fluctuations caused by the skewness term. If the dataset size is not an exact multiple of the batch size, the last incomplete batch is discarded for simplicity.

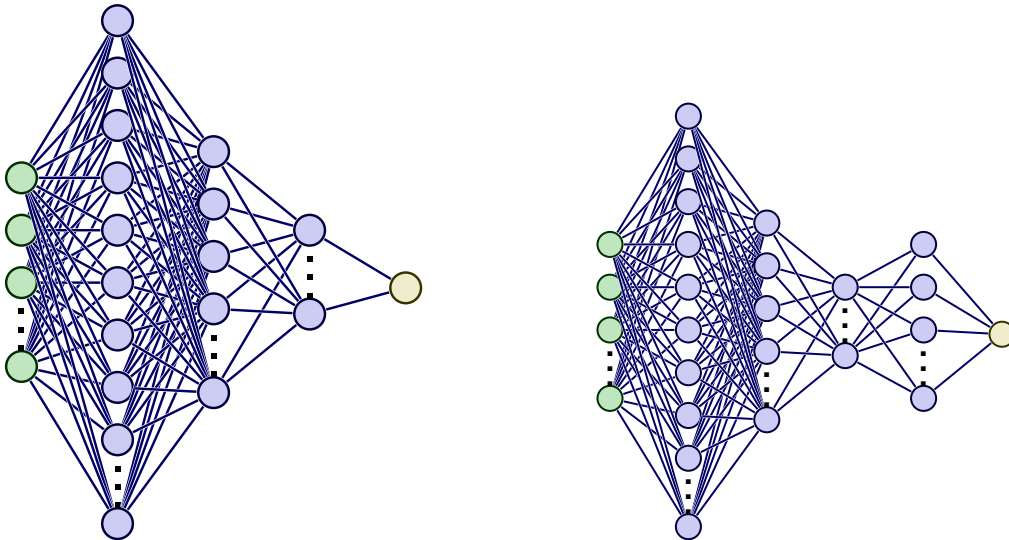


FIG. S2. Design pattern of the encoder part of Skewencoder

C. Metastable States Identification

In these model systems, metastable states are identified by the location of the average sample coordinates at each iteration. If the average value of sample coordinates falls within the yellow circles in Figure S3, the system is considered to be in a metastable state. Otherwise, the warm start training strategy described in Algorithm S1 is applied.

Metastable states in gas-phase and catalytic systems are identified by monitoring atomic contacts via input descriptors, as only relevant distances are affected by the loxodynamic bias. At each iteration, we compute the average of each descriptor over the last quarter of samples and determine bond status using Equation 8. If all bonds are either fully formed or broken, the system is classified as metastable. Conversely, if any average distance suggests an unstable state according to Equation 8, the warm start training strategy described in Algorithm S1 is applied.

It should be noted that our method is applicable to any criterion capable of distinguishing between metastable and non-stable states in the configurational space of a chemical system. The criterion described above is a simple example used to enhance the algorithm's efficiency.

D. General Algorithm

Figure S1 illustrates the iterative application of our biased sampling method for exhaustively exploring and identifying new states on the PES. This process integrates the training of the Skewencoder and the generation of the harmonic wall with respect to the CV in the latent space. The procedure is initialized with a sequence $\{\kappa_i\}$, where each value is

distributed at uniform intervals such that $\kappa_{i+1} = \kappa_i + \delta\kappa$, with $\delta\kappa$ denoting a constant increment. The parameter n_{iter} represents the number of iterations for which the simulation runs while keeping κ constant. To improve efficiency, the algorithm can be parallelized by running multiple replicas with different values of κ , allowing fine-tuning through simultaneous exploration.

In contrast to the simplified representation of a harmonic wall in latent space provided in Equation 7 of the main text, we differentiate between two types of biasing walls: the lower wall and the upper wall. These walls are distinguished by the regions in which they exert their influence. Specifically, during Loxodynamics simulations, a restraining harmonic wall functions as an upper wall when it is applied to configurations where the Skewencoder CV exceeds the sum of the projected mean $\mu(\hat{s}_{\text{LS}})$, sample variance $\sigma(\hat{s}_{\text{LS}})$, and sample skewness $\gamma(\hat{s}_{\text{LS}})$. Conversely, it serves as a lower wall when acting on configurations where the Skewencoder CV falls below this combined threshold. The designations "lower wall" and "upper wall" carry enhanced physical relevance, facilitating interpretation of how skewness relates to local anharmonicity and elucidating the mechanism by which Loxodynamics applies bias.

Beyond the stopping criterion based on n_{iter} , we implement an additional condition that allows early termination. The values $\mathbf{L}_{\text{state}}$ and S_p , returned from the training session as described in Algorithm S1, play a key role in this decision. Specifically, we define a reversible reaction pathway as one in which the system returns to the original state after detecting new intermediate states. When such a pathway is identified, the simulation can be terminated early to improve workflow efficiency.

The NN structure, along with its training and validation procedures, is implemented using PyTorch and the Lightning module, following the framework proposed in Bonati et al.⁸, with some modifications.

S2. CASE STUDIES – COMPUTATIONAL DETAILS

A. Model Potentials

As a proof of concept for our new method, we tested it on three simple 2D model potentials, illustrated in Figure S3.

Here we list the closed-form function of each model system illustrated in Figure S3. The parameters for Langevin dynamics simulation, mentioned in the main text, ensure that the system evolves under stochastic dynamics, where the chosen time step balances computational efficiency and integration accuracy, the temperature regulates thermal fluctuations, and the relatively high friction value provides strong coupling to the heat bath.

NN inputs are the particle's x, y positions. The encoder has layers with $[2, 60, 30, 1]$ nodes, and the total iterations, n_{iter} , is 20.

$$\begin{aligned}
 \text{Figure S3-(1): } & \frac{1}{1 + e^{\frac{-(y-x)}{4}}} \times (2.0x^2 + 0.25(y - 20.0)^2) + (1 - \frac{1}{1 + e^{\frac{-(y-x)}{4}}}) \times (0.25(x - 20.0)^2 + 2.5y^2) \\
 \text{Figure S3-(2): } & 1.34549x^4 + 1.90211x^3y + 3.92705x^2y^2 - 6.44246x^2 - 1.90211xy^3 + 5.58721xy + \\
 & 1.33481x + 1.34549y^4 - 5.55754y^2 + 0.904586y + 19 \\
 \text{Figure S3-(3): } & 0.15 \times (146.7 - 200 \times e^{-(x-1)^2-10y^2} - 100 \times e^{-x^2-10(y-0.5)^2} \\
 & - 170 \times e^{-6.5(x+0.5)^2+11(x+0.5)(y-1.5)-6.5(y-1.5)^2} + 15 \times e^{0.7(x+1)^2+0.6(x+1)(y-1)+0.7(y-1)^2})
 \end{aligned} \tag{S2}$$

Figure S4 illustrates the progression of exploration on the model PESs over selected iterations, showing how the system approaches new states. The blue points represent samples from the incremental dataset, which aggregates data from all previous and current iterations to form the global training dataset for the Autoencoder's reconstruction loss. In contrast, the red points correspond to samples from the current iteration, constituting the local training dataset for the skewness loss. Sampling results from specific consecutive iterations are presented based on the aforementioned criteria.

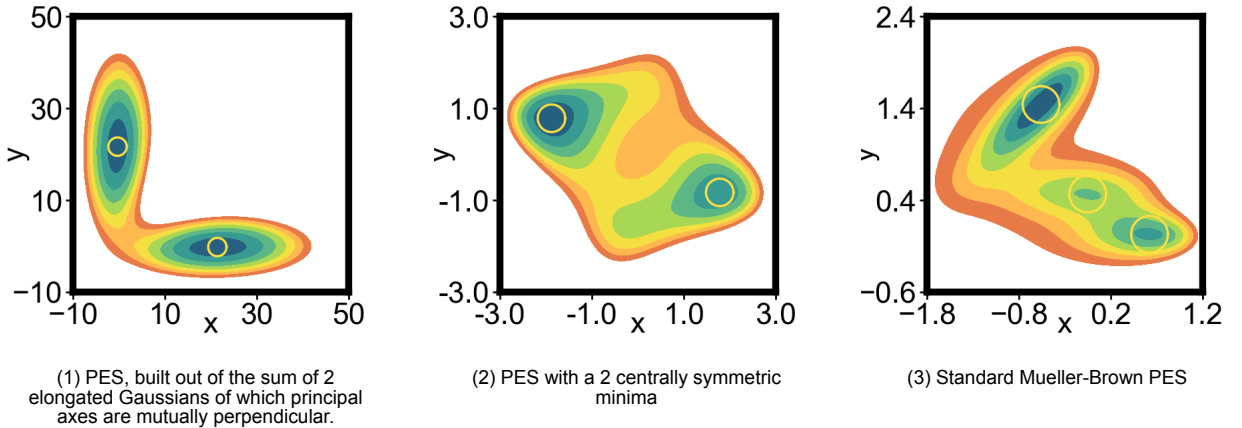


FIG. S3. **Model PESs for Langevin dynamics tests**^{9,10}. The darker region signifies an area with lower potential values, whereas the lighter region denotes an area with higher potential values. The isoline spacing is 15 for the first panel and 3 for the second and third panels. The yellow circles indicate the regions corresponding to the definition of metastable states in three different PESs. A copy of this figure using a perceptually uniform color palette is available in Section S3 of the Supplementary Information.

By biasing with our method, the sampling area progressively expands across iterations, ultimately reaching a new minimum on the PES.

In Figure S4, the evolution of the latent space during the simulation is illustrated. Initially, when sampling is confined to the starting metastable state, the latent CV captures only local features of the PES. As loxodynamic biases are applied and guide sampling toward saddle points on the PES, the latent space increasingly captures local anharmonicity. Significant changes in the latent CV occur when sampling reaches new metastable states. Upon termination of the algorithm, we infer that sufficient samples have been collected, allowing the latent space to approximate a meaningful CV and construct a qualitatively accurate representation of the reaction pathway.

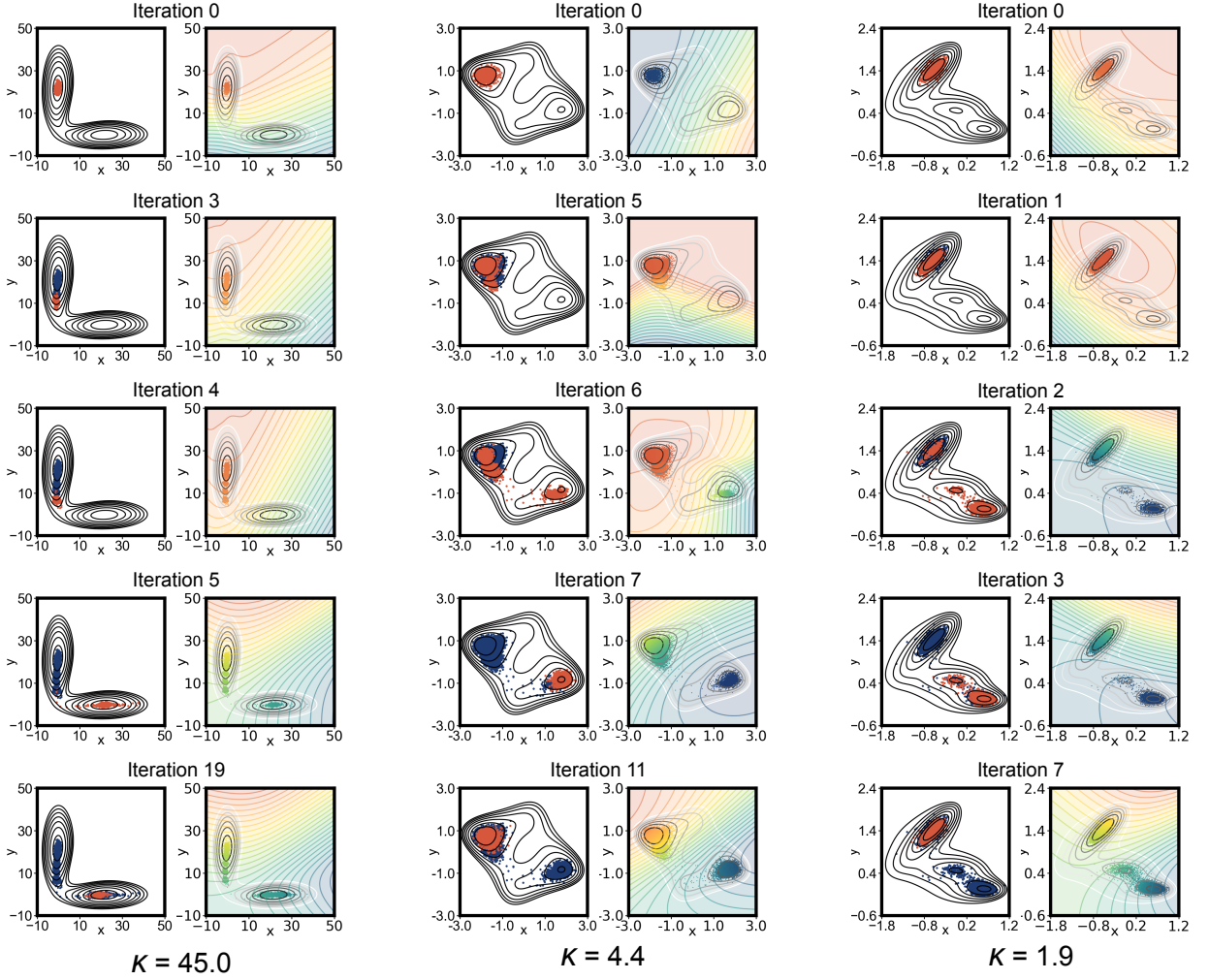


FIG. S4. **Progression of sampling in three model potential cases.** Each left subcolumn displays sampling from the initial iteration without loxodynamic bias, three consecutive iterations with state transitions in all model PESs, and the final iteration at simulation termination. In each left subcolumn, the blue points represent samples from the incremental dataset, while red points correspond to samples from the current iteration. In all cases, the wall, determined by the trained CV, guided the system toward the saddle point of the PES. The transition from one minimum to another occurred in the subsequent iteration of the simulation. Each right subcolumn shows contour plots of the current latent CV projected onto the xy plane. As the simulation progresses, the latent space evolves into an effective CV that captures pathways connecting all metastable states and saddle points. Lighter regions correspond to higher values in the latent CV space, whereas darker regions represent lower values.

1. Comparison of Different $\mathcal{L}_{\text{skew}}$

We assess the performance of the Skewencoder upon modifying the skewness-related loss term, $\mathcal{L}_{\text{skew}}$, and motivate the adopted formulation in view of its impact on the exploration phase. To benchmark performance, we focus on the first model PES shown in Figure S3-(1), which features two elongated, highly skewed minima oriented at 90° with respect to one another. This model is particularly representative as it can be regarded as a prototypical PES for an S_N2 -type reaction and poses nontrivial challenges to exploration algorithms.

Loxodynamics simulations were initiated from the upper-left minimum, aiming to reach the lower-right minimum. A systematic campaign of explorations was performed by varying the harmonic stiffness constant κ from 25 to 80. For each value of κ , four independent loxodynamics explorations (comprising multiple iterations) were carried out, testing different skewness-based loss terms, along with the loss function simply without the skewness-based loss terms, within the Skewencoder objective. Specifically, we compared:

1. the smooth “softplus” of $-\gamma^2$, which is the functional form adopted in our calculations, $\mathcal{L}_{\text{skew}}^{(1)} = \ln(1 + e^{-\gamma^2})$,

2. a simpler Gaussian-like penalty, $\mathcal{L}_{\text{skew}}^{(2)} = e^{-\gamma^2}$,
3. the direct negative squared skewness of the latent CV, $\mathcal{L}_{\text{skew}}^{(3)} = -\gamma^2$.
4. no skewness-related loss term, i.e. the Skewencoder loss function in Equation (2) is simply reduced to a loss function for minimizing Autoencoder reconstruction error (i.e. Equation (3))

All terms were incorporated into the Skewencoder loss function with the same hyperparameter weight $\alpha = 0.1$, as defined in Equation 2.

A well-performing loss function is expected to facilitate the discovery of new metastable states in the smallest possible number of iterations for a given choice of κ . This criterion is used to compare the performance of the different formulations and to motivate our particular choice, which favors both a minimal number of iterations and a low κ . As a further justification for its suitability, we also consider a representative case in which all loss functions succeed in reaching the new metastable state within the same number of iterations. In this situation, we compare the quality of the projected CVs on the PES, where a well-performing loss function is expected to generate a proper foliation of the space, resembling the effect of an ideal path CV.

In Figure S5, we report the number of iterations required to exit the initial basin and reach the bottom-right basin via the transition region as a function of the stiffness κ of the applied harmonic bias. For convenience, the exploration is terminated after a maximum of n_{iter} iterations (here $n_{\text{iter}} = 20$) if no new state is reached. At the lowest stiffness shown, $\kappa = 28$, only the simulation using $\mathcal{L}_{\text{skew}}^{(1)}$ reaches the target basin, doing so within 8 iterations. Increasing to $\kappa = 30$, both $\mathcal{L}_{\text{skew}}^{(1)}$ and $\mathcal{L}_{\text{skew}}^{(2)}$ succeed. As κ grows further, the number of iterations decreases, as expected under a stronger bias. Nevertheless, at lower κ values, $\mathcal{L}_{\text{skew}}^{(1)}$ consistently requires fewer iterations than the Gaussian-like $\mathcal{L}_{\text{skew}}^{(2)}$. In contrast, using the bare $-\gamma^2$ term ($\mathcal{L}_{\text{skew}}^{(3)}$) fails to induce transitions unless very large stiffnesses are employed—values that, in realistic systems, would produce unphysically large atomic forces, uncontrolled exploration, and potentially simulation instabilities. This observation led us to discard $\mathcal{L}_{\text{skew}}^{(3)}$. Similarly, omitting the skewness-related term altogether leads to transitions only at very high stiffnesses (e.g., $\kappa \gtrsim 70$), which again may render simulations unstable in practice.

Regarding the remaining choices, $\mathcal{L}_{\text{skew}}^{(1)}$ and $\mathcal{L}_{\text{skew}}^{(2)}$ perform similarly overall, with a slight advantage for $\mathcal{L}_{\text{skew}}^{(1)}$ at low κ . Beyond this quantitative edge, the softer gradients of the softplus formulation tend to stabilize Skewencoder training and provide a smoother, more informative signal about the local basin geometry (an effect difficult to prove rigorously but supported by our observations).

To further probe qualitative differences, we selected the lowest stiffness for which the four loss functions methods succeed to reach the new states and projected the isovalues of the final latent-space CV onto the (x, y) plane of the model PES (top panels of Figure S5). The CV obtained with $\mathcal{L}_{\text{skew}}^{(1)}$ yields a markedly better foliation, closely resembling that of a path CV along the minimum-energy path. These results motivated our adoption of $\mathcal{L}_{\text{skew}}^{(1)}$ as the default in practical loxodynamics applications, while retaining $\mathcal{L}_{\text{skew}}^{(2)}$ as a viable alternative in the current code release. $\mathcal{L}_{\text{skew}}^{(3)}$ fails to learn a sufficiently detailed and informative foliation of the space, yielding a non-representative CV that steers biasing forces in enhanced-sampling schemes (e.g., metadynamics or OPES) along suboptimal directions. Finally, omitting the skewness term produces at best a moderately meaningful foliation, however it requires substantially larger stiffness, which makes it ill-suited to the primary aim of Loxodynamics: efficient exploration.

2. Multiple reaction pathways and selectivity with loxodynamics

The stiffness parameter κ governs the pathway (mechanism) that Loxodynamics preferentially follows during sampling. We systematically examine this dependence on the second model PES (Figure S3-(2)), which exhibits multiple reaction routes. Two distinct pathways are identified (sketched in panel B of Figure S6), each associated with a different saddle point on the PES: one near $(x, y) \approx (-0.8, -0.7)$ with energy $E \approx 16$, and another near $(x, y) \approx (0.6, 0.5)$ with energy $E \approx 19$.

Loxodynamics simulations were initiated from the upper-left local minimum with the objective of reaching the lower-right minimum. We conducted a systematic series of explorations by varying the harmonic stiffness κ , starting from $\kappa = 4.4$ —the minimal value required to induce state transitions (second column of Figure S4). For consistency, each exploration was limited to a maximum of $n_{\text{iter}} = 20$ iterations. As κ increased, sampling began to access path 2 (panel B of Figure S6) in addition to primarily following path 1. Notably, state transitions were observed exclusively during iterations dominated by path 1 (see the progression from iteration 5 to iteration 7 in the second column of Figure S4). Upon increasing κ to 6.7, sampling became predominantly aligned with path 2 while still enabling

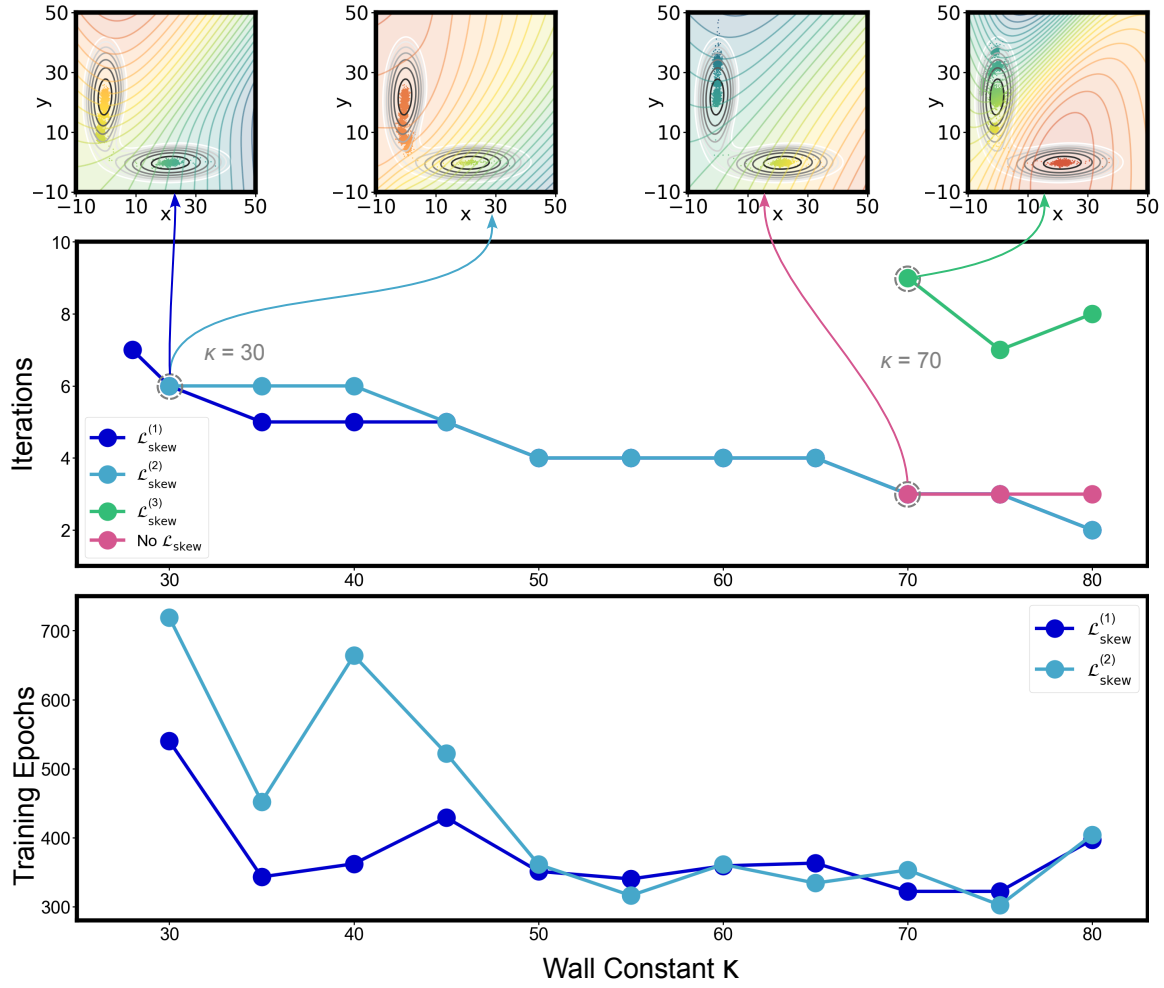


FIG. S5. **Comparison of different loss function.** Lower panel: The number of iterations required for state transition as a function of κ . The starting point of each line indicates the smallest κ and the corresponding number of iterations needed for a state transition in each model. Upper panel: For $\kappa = 30$, the latent space CV contour plots are shown for $\mathcal{L}_{skew}^{(1)} = \ln(1 + e^{-\gamma^2})$ (left) and $\mathcal{L}_{skew}^{(2)} = e^{-\gamma^2}$ (right) at iteration 6, when the state transition occurred.

transitions. From these explorations, we selected representative cases in which (i) path 1 predominates and (ii) path 2 predominates, and we analyzed the shape and performance of the learned CVs during enhanced sampling. The contour plots in Figure S6 indicate that CV1 varies preferentially along path 1, whereas CV2 exhibits stronger directionality along path 2.

To evaluate the selectivity of the two CVs, we performed Metadynamics simulations to test whether each CV preferentially explores a specific pathway. Langevin dynamics followed the protocol described in Section S2 A. Metadynamics was implemented with bias factor $\gamma = 80$, hills deposited every 50 time steps, an initial hill height $W = 2.0$, and a Gaussian width $\sigma = 0.2$.

Selectivity was quantified by counting samples assigned to each pathway during state transitions. We first restricted attention to configurations whose CV values fall within a prescribed interval (pink-shaded region on the left side of panel A in Figure S6), chosen to capture transition-region information while excluding neighborhoods of local minima. We then further filtered out configurations located far from both paths or approximately equidistant from them, thereby isolating points in the vicinity of the saddle region of either path while maintaining comparable sample sizes for the two CVs. In this framework, samples near a given saddle are deemed to “belong” to that path. Four sets were thus defined (panel B of Figure S6): CV1-path1 (deep blue), CV1-path2 (light blue), CV2-path1 (light green), and CV2-path2 (deep green).

As shown in panel B of Figure S6, more samples are associated with path 1 than with path 2, irrespective of whether CV1 or CV2 is used. This trend is expected: although CV2 is designed to follow path 2, Metadynamics primarily enhances fluctuations within basins, so the lower-energy pathway is explored more frequently. Notably, when sampling

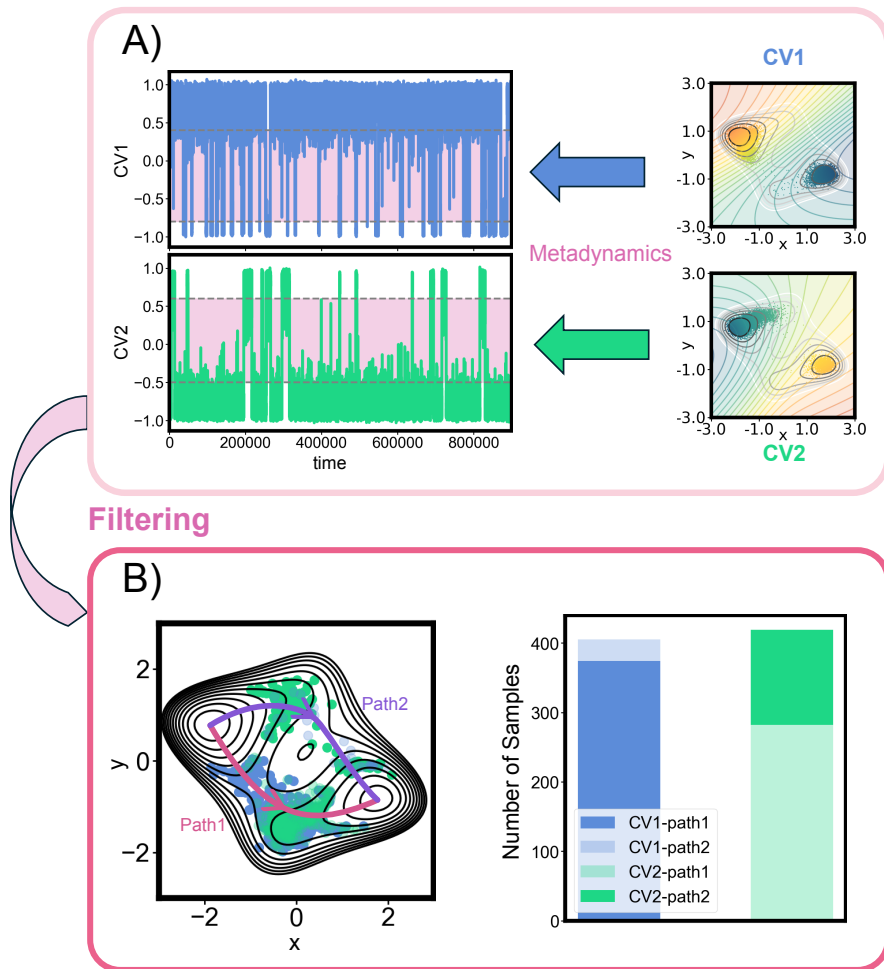


FIG. S6. **Discussion on exploring multiple pathways in Loxodynamics simulation.** **a**, Results of Metadynamics simulations employing two distinct CVs. **b** Distribution statistics of samples collected during state transitions, categorized by pathway and associated CV.

with CV2 a substantially larger fraction of reactive events proceeds via path 2 than with CV1. Thus, while CV2 does not ensure perfect mechanistic selectivity, its geometry captures a greater share of transitions along path 2. It is worth noting that the limited selectivity observed with Metadynamics follows from the considerations above. It is crucial to emphasize that, unlike Metadynamics—a CV-based flooding technique—Loxodynamics is a directional-bias method. Therefore, selectivity along the preferential directions is ensured, as backtracking along the pathway is discouraged.

3. Warmstarting efficiency

We evaluate the performance and efficiency of Skewencoder using a warm-start strategy, compared with retraining from scratch at each iteration. For the sake of clarity, and without loss of generality, we focus the benchmark on the first two model PESs shown in Figure S3.

For the first PES model (Figure S3), we performed a systematic exploration by varying the harmonic stiffness κ at 4.8, 5.1, and 5.4. For the second PES model (Figure S3), simulations were carried out with κ values starting at 25. For each selected κ , two independent Loxodynamics runs were executed—one using a warm-start training strategy and one trained from scratch—to compare their performance. Each exploration was capped at $N_{\max} = 20$ iterations if no new state was detected.

The warmstart strategy is intended to maintain an appropriate search direction with minimal computational effort, particularly by reducing training costs. Specifically, the warmstart approach aims to prevent substantial alterations in the structure of the latent space between consecutive iterations. When a state transition does not occur, most samples remain near the initial state, resulting in the accumulation of similar information characterizing this region during

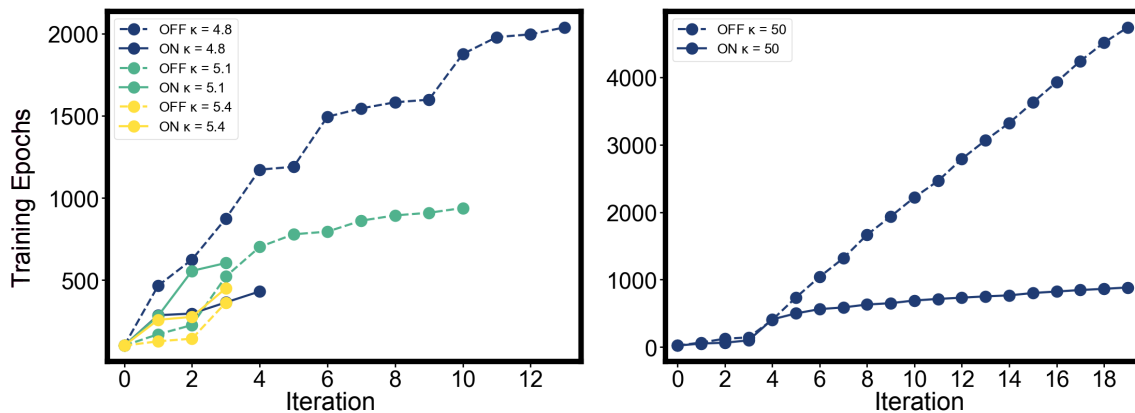


FIG. S7. Training epochs needed to minimize Skewencoder loss function with respect to Loxodynamics iterations upon changing bias wall stiffness with and without warmstarting. Left panel: Comparison on PES Figure S3-(2). Right panel: Comparison on PES Figure S3-(1).

exploration. In such cases, initializing training from previously optimized NN parameters allows for more efficient utilization of existing knowledge about the neighborhood of the initial state, thereby decreasing the computational burden associated with autoencoder reconstruction within Skewencoder. Consequently, both a reduction in the number of required training epochs and a decrease in the number of Loxodynamics iterations necessary to identify a new state can be anticipated.

Figure S7 presents aggregated totals of training epochs after each iteration for all simulations considered. The left panel displays how training epoch counts evolve across six simulations on the PES depicted in Figure S3-(2), comparing scenarios with and without warmstart activation. In all cases shown, the second state was identified within 20 iterations. Notably, when Skewencoder was trained without warmstart at $\kappa = 4.4$ —the minimum value found to permit state transitions on this PES—the simulation failed to reach another state within 20 iterations. Therefore, our discussion is confined to simulations that successfully located an alternative state within this threshold; each run was terminated at the iteration corresponding to successful state transition.

As illustrated in Figure S7 (left panel), employing warmstart at $\kappa = 4.8$ led to significant reductions in both total training epochs and required iterations for achieving a state transition. For $\kappa = 5.1$, this trend persisted: simulations initialized from scratch demanded substantially more iterations before reaching another state compared to those using warmstart initialization. However, as κ increased further, it was observed that during early iterations, warmstart procedures sometimes necessitated more training epochs than runs without warmstart activation.

As discussed in Section S2 A 2, this model PES contains two distinct "reaction" pathways; increasing κ compels Loxodynamics sampling toward higher-energy paths. Under these conditions, applying warmstart may amplify noise effects arising from samples along secondary pathways, requiring additional corrective training steps. Importantly, however, the use of warmstart reduces sensitivity of latent space CVs to variations in κ , especially during initial simulation stages—as evidenced by relatively stable epoch counts over early iterations despite changes in κ . This property helps stabilize search directions at simulation onset.

The right panel of Figure S7 presents a comparative analysis for a single pair of simulations conducted on the second model PES in Figure S3, with $\kappa = 50$. It is noteworthy that all simulations performed with κ values below 50 did not succeed in identifying new states when the warmstart training strategy was not employed. In these cases, sampling remained localized near the initial state, and exploration trajectories exhibited spiral-like distortions. As illustrated in Figure S5, a value of $\kappa = 50$ constitutes a relatively large spring constant, imposing a strong biasing wall that induces aggressive exploration, often bypassing thorough sampling along the MEP.

Although only modest improvements are observed upon activating the warmstart strategy at $\kappa = 50$ for this particular PES, Loxodynamics simulations without warmstart can still occasionally identify alternative states due to the substantial magnitude of κ . To more clearly demonstrate potential advantages conferred by the warmstart approach in this context, the total number of training epochs following successful state transition is depicted in the right panel of Figure S7. During state discovery, warmstart continues to offer significant reductions in training effort; conversely, disabling this protocol can lead to renewed random searches concentrated around the initial state—particularly evident when using smaller κ values—which ultimately diminishes exploration efficiency.

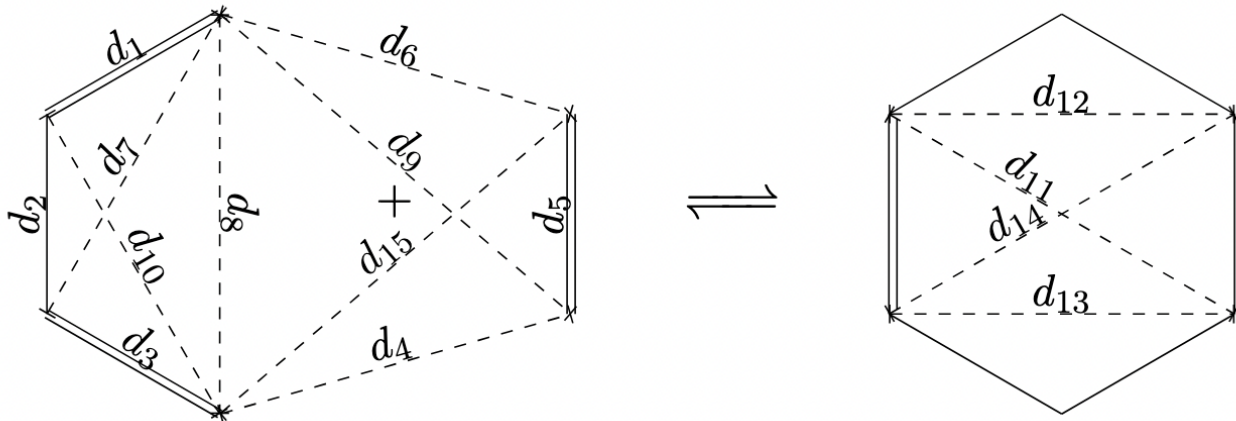


FIG. S8. Distances among all the heavy (carbon) atoms used to explore the DA reaction.

B. Gas Phase Chemical reactions

1. S_N2 Reaction

We used the two atom-pair distances, d_F and d_{Cl} , as descriptors for the Skewencoder input. Since our MD simulations were performed under non-periodic boundary conditions, dissociation channels could arise if either d_F or d_{Cl} increased significantly. To mitigate this effect and maintain sample distribution consistency, we applied two harmonic walls at $d_F = 4.0 \text{ \AA}$ and $d_{Cl} = 4.0 \text{ \AA}$ to restrain the Cl^- and F^- ions, ensuring proper compensation for dissociation. Additionally, we introduced a constant harmonic wall at a distance of $1.7 \text{ \AA}$ between each ion and any hydrogen atom to prevent the formation of $H-Cl$ or $H-F$ bonds.

Regarding the NN structure, since the Skewencoder input remained 2-D, as in the model system, we retained the same architecture. All distance values in the training set were transformed into mean-free data with a standard deviation of 1.

For bond detection, we set the parameters to $p = 0.25$, $m = 12$, and $n = 6$. The equilibrium bond distance threshold d_0 was set to $1.8 \text{ \AA}$ for the C-F bond and $2.5 \text{ \AA}$ for the C-Cl bond, approximately 1.2 times the mean lengths of the respective bonds.

2. Diels-Alder Reaction

Figure S8 depicts the distances among all carbon atoms in the system, denoted as d_1 to d_{15} , which serve as the input for our architecture to sample the reaction. To prevent excessive separation of the two fragments in the reactant state, we applied harmonic walls at $d_4 = 5.0 \text{ \AA}$ and $d_6 = 5.0 \text{ \AA}$.

Regarding the NN structure, the number of layers increases as the input dimensionality grows. However, all distance values are standardized to mean-free data with a standard deviation of 1.

To validate the generalization of our method, we tested two different NN architectures, as discussed in the section on default parameter values. The first architecture is shallower, with the number of nodes in each encoder layer set to $[15, 40, 20, 5, 1]$, while the second is deeper and wider, with nodes configured as $[15, 90, 40, 2, 20, 1]$. For bond detection, we set $p = 0.25$, $m = 8$, $n = 6$, and $d_0 = 1.7 \text{ \AA}$, approximately 1.2 times the bond length of C-C in cyclohexene.

The results of the deeper NN confirm that our method is valid in both reaction directions. Figure S9 demonstrates that even with a simpler NN structure, our approach remains effective. Initially, learning the symmetric and asymmetric properties of the reaction is more challenging. However, with the aid of the skewness-based loss term and the biased wall, a rough approximation of the reaction path is established, enabling the system to quickly reach alternative states due to the stronger wall.

Notably, in the shallower NN case, the system transitioned back to the original state. This behavior was not observed in the test with the deeper architecture when $\kappa = 350$, suggesting that the product state had lower free energy than the reactant. With a sufficiently large κ , this energy barrier can be overcome bidirectionally.

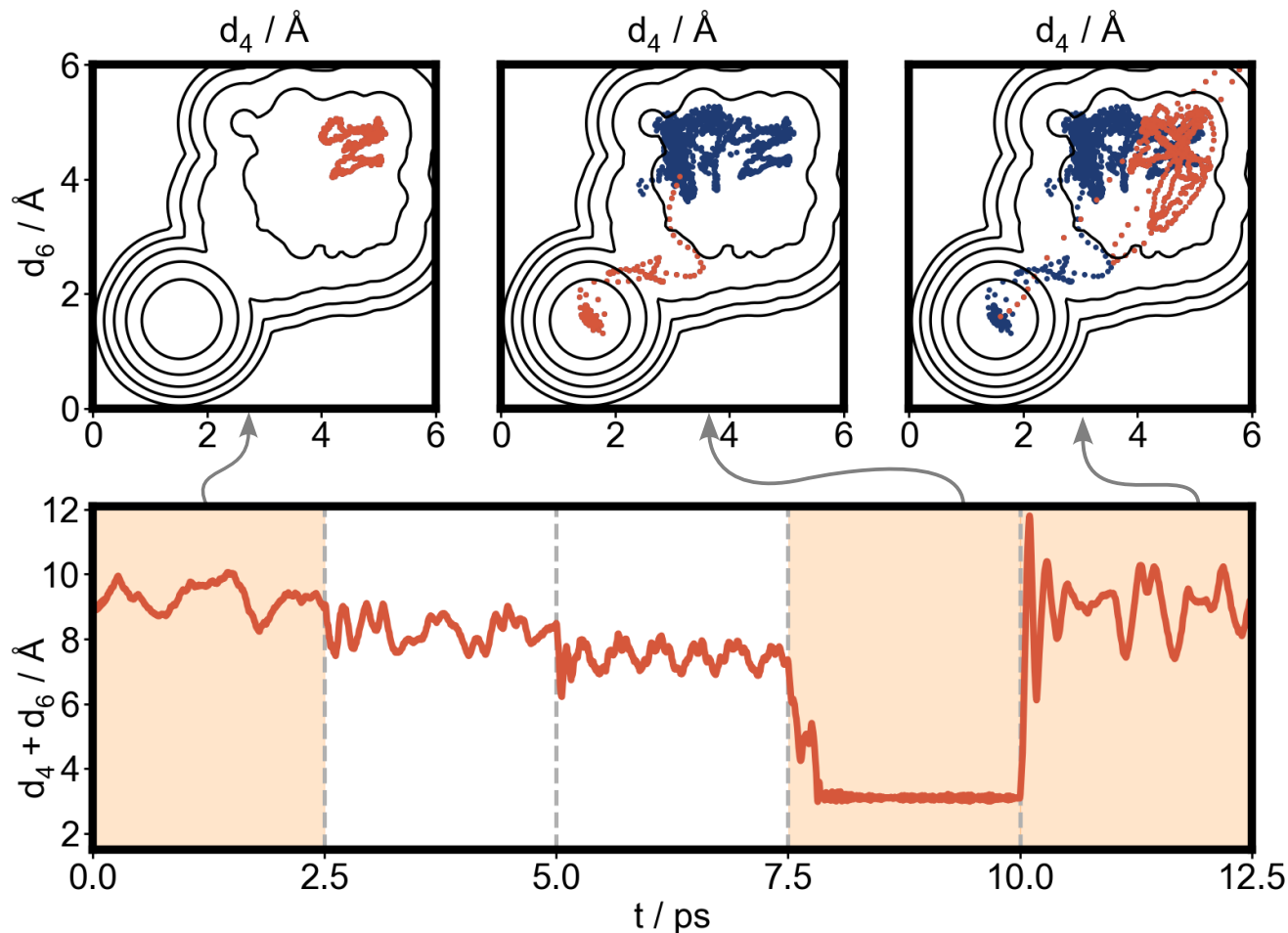


FIG. S9. **Simulation results of searching for (meta)stable states in the Diels-Alder reaction with a shallower NN.** The final κ in the shallower NN is 470, which is sufficiently large to enable recrossing of the energy barrier between the two states.

We also aimed to gain further physical insight into the process through a feature importance analysis performed on the NN at each iteration. Similar to Bonati et al.¹¹, we rank the features by computing the modulus of the derivatives of the latent space variable s_{LS} with respect to each input, averaged over all points in the training set, and weighted by the standard deviation of the corresponding input in the training set. The feature importance ranking is calculated as shown in Equation (S3):

$$r_k = \frac{\sum_{j=1}^n \left| \frac{\partial s_{LS}}{\partial d_k} \right|_j \sigma(d_k)}{\sum_{k=1}^m \sum_{j=1}^n \left| \frac{\partial s_{LS}}{\partial d_k} \right|_j \sigma(d_k)} \quad (\text{S3})$$

where d_k are the descriptors, s_{LS} is the latent space variable, m is the number of inputs, n is the number of samples, and $\sigma(d_k)$ is the standard deviation of d_k in the training set. The ranking r_k is normalized so that the sum of all rankings equals one.

We tested both architectures using different randomized Torch seeds, obtaining 10 successful simulations with the deeper NN structure and 15 with the shallower one, repeating this analysis in each iteration. In some successful tests, bonds d_9 and d_{15} formed instead of d_4 and d_6 , aligning with our expectations. In each iteration, we computed the mean (colored lines in Figure S10) and standard deviation (shaded areas in Figure S10) of r_k across all simulations.

Intermolecular distances were paired based on geometric symmetry, and the rankings of the descriptors within each pair, as well as all intermolecular heavy-atom pair distances, were summed. As the simulation progressed, the rankings of relevant pairs increased slightly, while the total ranking sum of these descriptors rose sharply to 0.6, highlighting their significance. This sum remained constant in later iterations due to the symmetric geometry of the reactant basin.

However, the warm-start strategy ensured that these pairs consistently maintained the highest rankings among all descriptors. Notably, the significance of the pairs d_4 - d_6 and d_9 - d_{15} converged, consistent with chemical intuition.

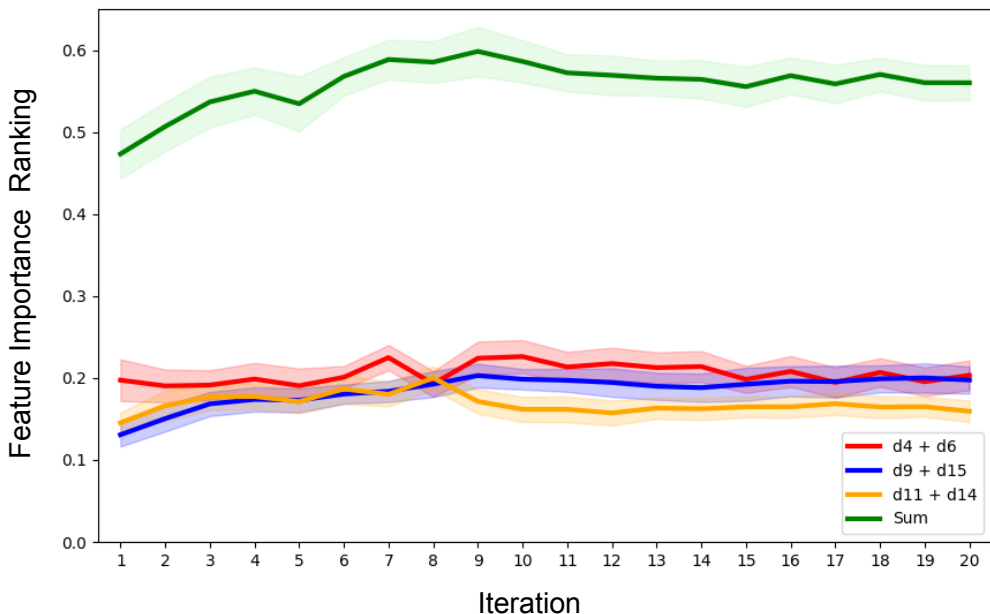


FIG. S10. **Evolution of feature importance rankings for three pairs of geometrically symmetric intermolecular distances over multiple iterations.** The rankings are computed based on the magnitude of the derivative of the latent space variable s_{LS} with respect to the descriptors $d_1 - d_{15}$ and are normalized such that their sum equals 1. The green line represents the total feature importance ranking sum across all distance pairs.

These results demonstrate that our method effectively identifies key features of a reaction system from a large descriptor set without *a priori* knowledge of the reaction pathway. However, the importance of these pairs is not dominant, as the concerted elongation and contraction of distances d_1 , d_2 , d_3 , and d_5 also influence the reaction process¹².

3. Assessment of Skewencoder CVs in Conventional Enhanced Sampling Methods

This section evaluates the performance of CVs obtained during Loxodynamics simulations. Skewencoder CVs are iteratively trained and refined throughout the simulation process. As illustrated in Figure S4, the CV evolves towards a quasi-optimal representation that closely approximates the reaction pathway on the two-dimensional model potential. Our analysis demonstrates that, when Loxodynamics successfully identifies the target products, the resulting Skewencoder CV serves as an effective surrogate for the reaction coordinate and is suitable for use in conventional enhanced sampling techniques to generate free energy surfaces (FES). Specifically, we employed the CV obtained at iteration 4 of the reversed S_N2 reaction (corresponding to the final window shown in Figure 4c, see main text), as well as the CV from iteration 9 of the forward Diels-Alder reaction, to compute FES via Metadynamics simulations for these two systems.

For both reactions, the initial MD simulation configurations were prepared as described in Section General Workflow of the main text. In the S_N2 reaction, a well-tempered Metadynamics simulation was performed with a bias factor γ , hill deposition interval $\tau_G = 25$ fs (corresponding to every 50 time steps), and an initial hill height $W = 5.0$ kJ/mol. The hill width was set to $\sigma = 0.2$. Restraining harmonic walls were applied to prevent inefficient sampling in dissociative regions. For the Diels-Alder reaction, Metadynamics was conducted with identical bias factor and deposition interval, but with an initial hill height $W = 4.0$ kJ/mol and a narrower hill width $\sigma = 0.01$. Similar restraining walls were implemented to confine exploration within relevant regions of configuration space. As the Skewencoder CV incorporates all 15 carbon-atom-pair distances as input descriptors, additional restraining biases were imposed on non-reaction pair distances to avoid unintended exploration of alternative states during Metadynamics simulations.

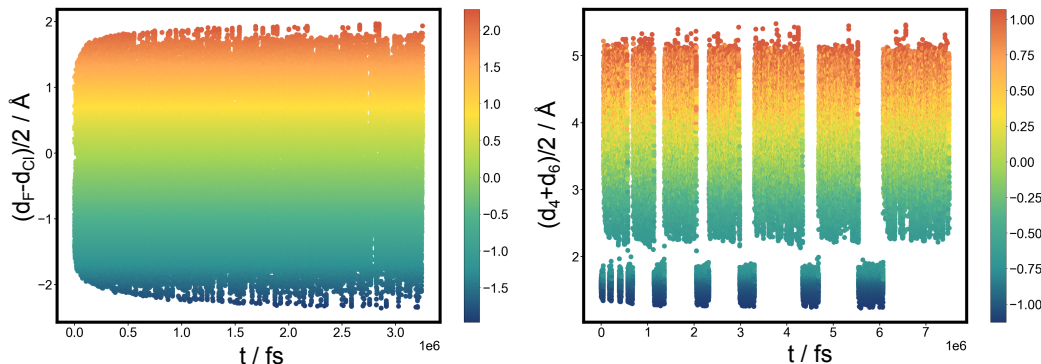


FIG. S11. **Metadynamics simulation based on Skewencoder CV.** To indicate the change of states during sampling, time evolution of heuristic CVs based on chemical intuition for the S_N2 (left panel) and Diels-Alder (right panel) reactions during the Metadynamics run is illustrated. The color indicates the value of the Skewencoder CV along which the bias is deposited. A copy of this figure using a perceptually uniform color palette is available in Section S3 of the Supplementary Information.

Figure S11 presents the trajectories from two simulations utilizing Skewencoder-derived CVs, alongside results obtained using conventional CVs—specifically, $\frac{d_F - d_{Cl}}{2}$ for the S_N2 reaction and $\frac{d_4 + d_6}{2}$ for the Diels-Alder reaction. These distance-based variables are defined in accordance with previous sections. Analysis of transition rates reveals that, although only scalar distances serve as input descriptors for Skewencoder—in particular, several correlated and unrelated distances are included in the Diels-Alder case—the neural network successfully constructs a nonlinear combination capable of distinguishing and facilitating transitions between states in both systems. Thus, the learned CV effectively guides the system between relevant basins and serves as a reliable proxy for the internal reaction coordinate.

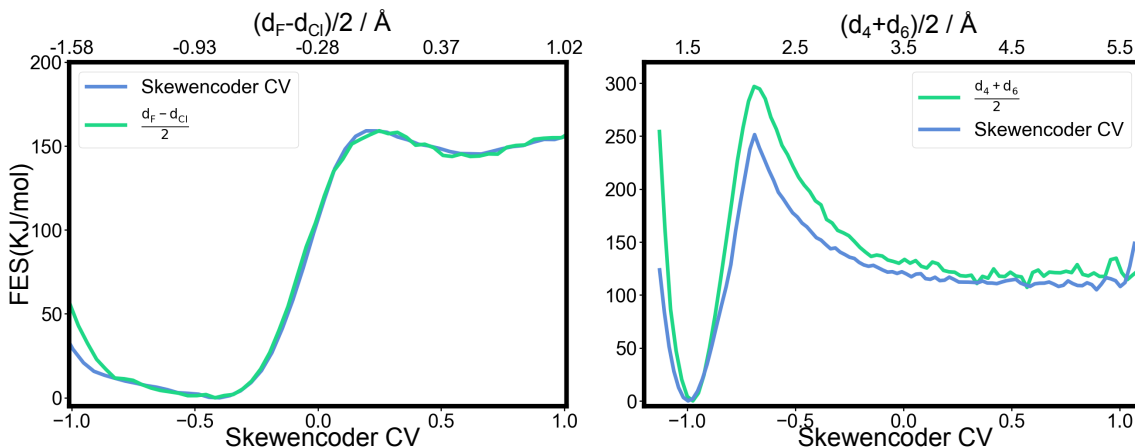


FIG. S12. **FES obtained from Metadynamics on chemically-inspired heuristic and Skewencoder CVs.** The FES are reconstructed via statistical reweighting for the S_N2 reaction (left panel) and for the Diels-Alder reaction (right panel).

Figure S12 displays the free energy profiles obtained by reweighting along the Skewencoder CVs, in comparison with those generated and reweighted using the conventional CVs described above for both reactions. A consistent hill width of $\sigma = 0.01$ was employed throughout all reweighting procedures. For the S_N2 reaction, the FES and associated energy differences derived from the Skewencoder CV closely match those obtained with chemically-inspired CV $\frac{d_F - d_{Cl}}{2}$. The free energy differences between reactants and products were calculated by integrating the probability densities within the respective basins. The dividing boundary between basins was set to -0.28 for the chemically-inspired CV and 0.0 for the Skewencoder CV. The chemically-inspired CV simulation yielded a free energy difference of 143.177 kJ/mol, while the Skewencoder CV simulation resulted in 143.204 kJ/mol, demonstrating excellent agreement between the two simulation results. In contrast, for the Diels-Alder reaction, minor deviations are observed in the positions of the broader basin minima and barrier heights. To mitigate unintended exploration of alternative states, additional restraining biases were applied to non-reaction pair distances. Given that the mapping between input descriptors and latent space CV is not injective, imposing bias on non-reactive bond distances is in fact imposing bias in the latent

space, which can promote sampling along the desired pathway, which may contribute to reduced free energy differences, compared with the simulation based on chemically-inspired CV $\frac{d_4+d_6}{2}$. Still, we calculated the free energy difference by integrating over basins defined by a threshold of -2.2 (chemically-inspired CV) and -0.7 (Skewencoder CV). The resulting values were 106.223 kJ/mol and 102.342 kJ/mol, respectively, indicating reasonable agreement between the two simulation results.

C. Ethanol Dehydration in Acidic Chabazite under Operando Conditions

1. Simulation Setup

We constructed a 37-dimensional input descriptor set with minimal chemical intuition. Instead of relying solely on heavy-atom pairs, we also incorporated hydrogen–oxygen distances to demonstrate the flexibility of our approach. The 37 input distances include:

- All distances between any of the four oxygen atoms near the aluminum site in H-CHA and any hydrogen atom in the system, totaling 28 distances.
- All distances between the oxygen atom in the hydroxyl group of ethanol and any hydrogen atom in the system, totaling 7 distances.
- Distances between each of the two carbon atoms in ethanol and the oxygen atom in the hydroxyl group.

As the input dimensionality increases, the number of nodes in each layer must also increase. Therefore, the NN structure consists of an encoder with layer sizes set to [37, 90, 40, 20, 5, 1]. For bond detection, we set $p = 0.3$, $m = 12$, $n = 6$, and d_0 to 1.2 Å for OH bonds and 1.7 Å for CO bonds.

The spring constant of the biased wall $\kappa = 280$ in the concerted mechanism and $\kappa = 300$ in the stepwise mechanism. The difference in κ corresponds to the difference of energy barriers in both reaction mechanisms¹³.

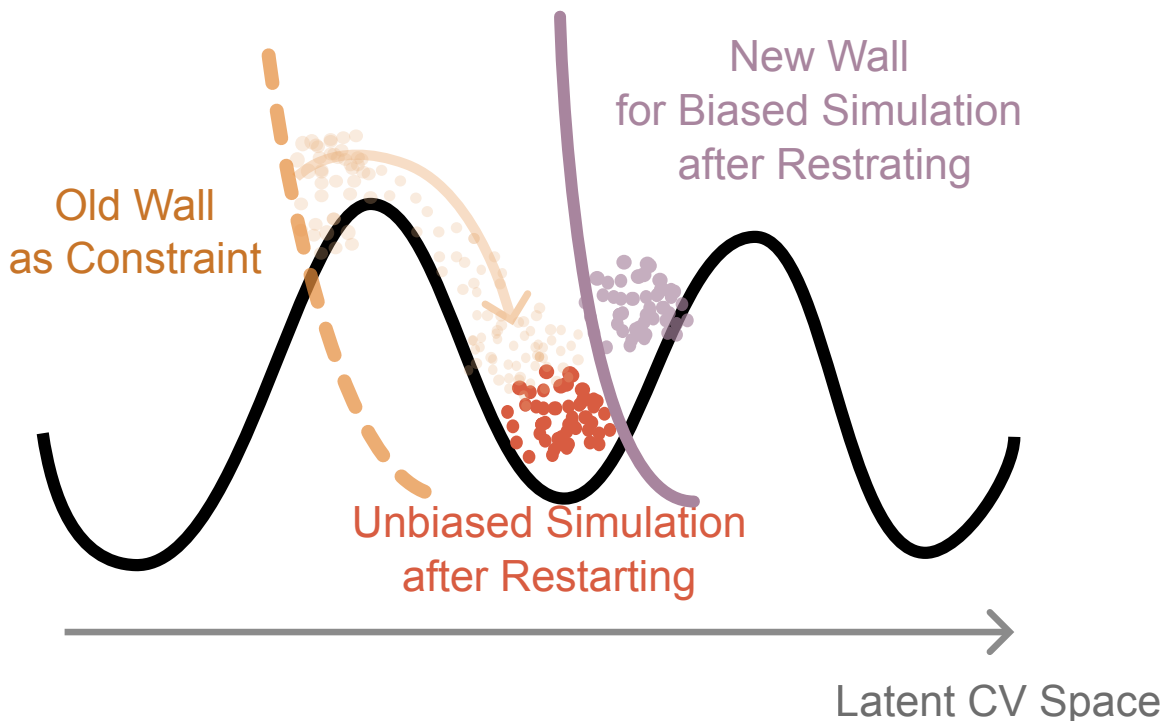


FIG. S13. **Restarted Loxodynamics protocol applied during exploration of the stepwise mechanism.** The previously applied wall that facilitated the transition to the intermediate state is retained as a constraint to prevent reversion to the original state. A new locally "unbiased" simulation is subsequently performed in the presence of this constraint. The Loxodynamics procedure then continues according to the workflow depicted in Figure S1.

2. Restart Strategy for Exploring Multiple Potential Wells

It should be emphasized that the Loxodynamics simulation was restarted upon identification of a key intermediate—specifically, the formation of an ethoxy group bonded to a BAS oxygen atom. The procedure for restarting Loxodynamics is illustrated in Figure S13. In the restarted simulation, the previously applied wall that enabled the transition to the ethoxy intermediate is maintained to prevent reversion to the original reactant state. To offset potential distortions introduced by this constraint, the wall spring constant was increased to 400, thereby promoting progression from the intermediate towards the final product state. Successful exploration of the stepwise mechanism using this restart protocol demonstrates the capability of Loxodynamics to investigate sequential reaction pathways and highlights its potential for mapping complete and complex reaction networks.

3. Validation of the Reaction Mechanism Identified by Loxodynamics Simulations

In this section, we assess the quality of MD trajectories generated by Loxodynamics simulations, with particular attention to how closely the system dynamics align with the minimum energy pathway (MEP) for ethanol dehydration in acidic Chabazite. Validation was performed using the Climb-Image Nudged Elastic Band (CI-NEB) method¹⁴ applied to all three cases. CI-NEB calculations were carried out with CP2K, employing convergence criteria of 0.003 for maximum force, 0.002 for maximum displacement, and 0.005 for both root mean square (RMS) displacements and forces. Frame alignment was enabled at the start of each calculation, and replica rotation was permitted throughout to improve the accuracy of the resulting energy profiles. Energy and forces were also computed with the semiempirical GFN-XTB.

We now outline the refined NEB workflow applied to all trajectories. For each of the three reactions—namely, the concerted pathway, the initial step of the stepwise mechanism leading to formation of the ethoxy intermediate, and the subsequent transformation from this intermediate—a sequence of frames corresponding to state transitions was extracted from the MD trajectories. The geometries in both the initial and final frames were optimized to serve as endpoints for the NEB calculations. Each trajectory segment contained over 100 frames, which were used as input for CI-NEB analysis. A spring constant of 0.005 was employed in all cases to ensure convergence along these high-resolution NEB pathways. From each set, 14 representative images were manually selected, predominantly from regions near energy maxima and minima along the path. This preliminary filtering enabled identification of MD frames approximating saddle points within the original Loxodynamics trajectories, thus providing accurate estimates of MEPs near transition states. Subsequently, CI-NEB calculations were rerun using these refined image selections, resulting in improved precision while reducing computational cost.

Subsequently, 14 representative MD frames were extracted from each trajectory generated by Loxodynamics. These frames, manually selected from the original trajectories, were used to optimize the initial and final configurations, which served as endpoints for a refined NEB calculation. The optimized endpoints, along with the remaining frames, were provided as input for a new CI-NEB computation with a spring constant set to 0.005. The resulting CI-NEB pathways are presented in the left column of Figure S14. The computed NEB energy profiles closely match the energies of stable states and transition states reported by Kim et al.¹³. Moreover, the highest-energy geometries along these NEB paths (also shown in Figure S14) exhibit minimal deviation from their corresponding initial configurations in the Loxodynamics MD trajectories. These observations demonstrate that Loxodynamics is effective in guiding system exploration along MEPs within MD simulations, irrespective of the level of theory employed.

D. 1-Butanol Dehydration in Acidic Chabazite

We constructed a 28-dimensional input descriptor set with minimal chemical intuition. Likewise in Section S2 C 1, we also incorporated hydrogen–oxygen distances to demonstrate the flexibility of our approach. The 28 input distances include:

- All distances between any of the five oxygen atoms (four near the aluminum site in H-CHA and the rest one belongs to the hydroxyl group of 1-butanol) and any carbon atom in the system, totaling 20 distances.
- All distances between the four oxygen atoms near the aluminum site in H-CHA and 2 hydrogen atom in the system, totaling 8 distances.

Regarding the (hyper-)parameters of the Skewencoder, the NN structure consists of an encoder with layer sizes set to [28, 50, 30, 10, 3, 1]. For bond detection, we set $p = 0.3$, $m = 12$, $n = 6$, and d_0 to 1.2 Å for OH bonds and 1.7 Å for CO bonds.

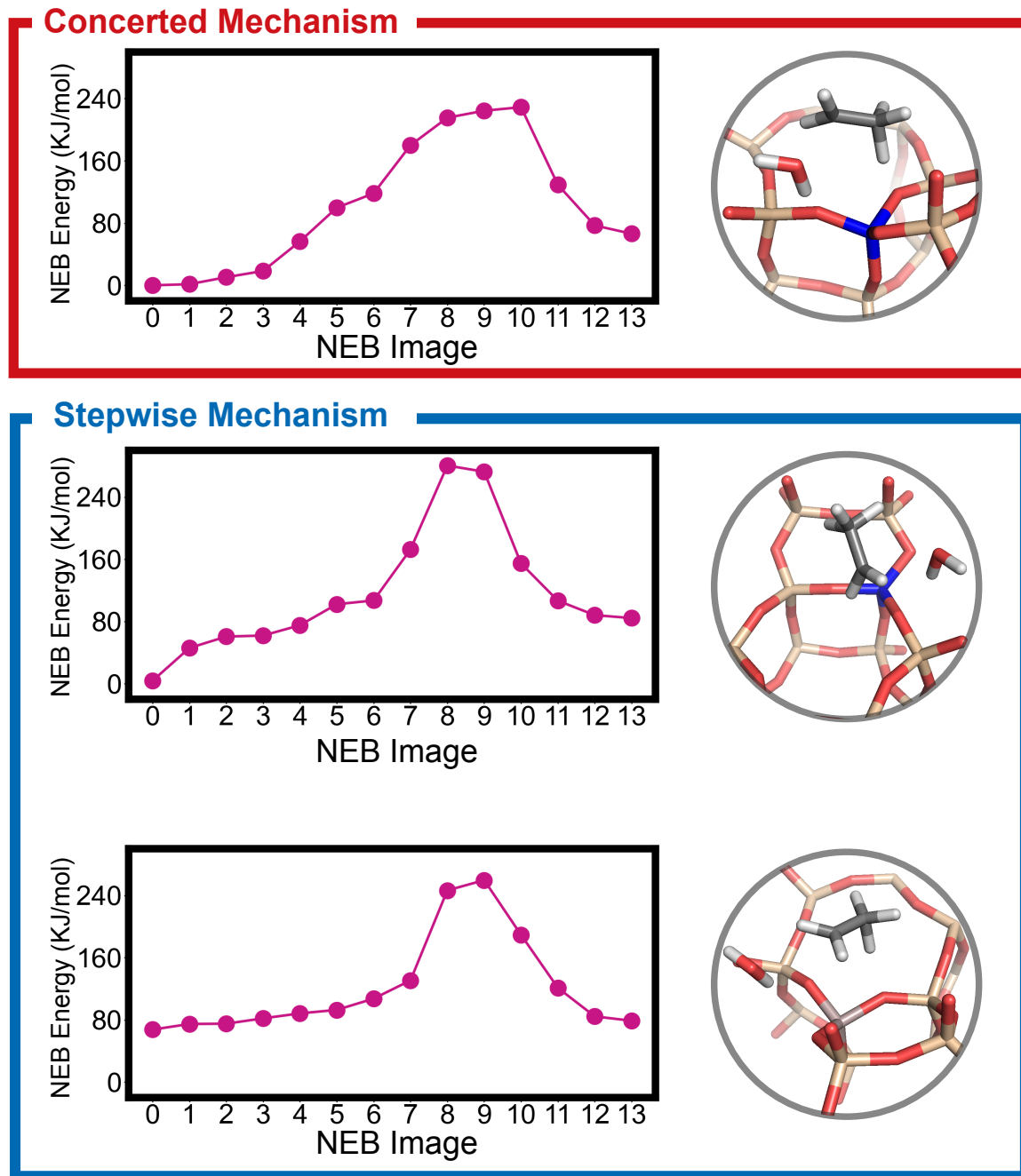


FIG. S14. **NEB energy profile of ethanol dehydration in Acidic Chabazite** Upper panel: NEB energy profile corresponding to the concerted reaction mechanism. Lower panel: NEB energy profile associated with the stepwise mechanism. Right column: transition state geometries determined via the NEB method, representing system topologies at the highest-energy NEB images for each pathway.

In the concerted mechanisms (shown in Figure S15), the spring constant of the biasing wall was set to $\kappa = 310$ for

simulations yielding 1-butene as the final product, while values of $\kappa = 350$ and $\kappa = 370$ were applied for identifying trans- and cis-2-butene conformers, respectively. The specific value of κ associated with each conformer reflects its relative thermochemical stability. For the stepwise mechanism involving detection of the butoxy intermediate, a restart simulation protocol analogous to that described in Section S2 C 2 was employed to prevent recurrence to the original reactant configuration. Accordingly, we first apply $\kappa = 320$ to search the intermediate, then a spring constant of $\kappa = 420$ was used to facilitate exploration of the final product state (1-butene), compensating for possible distortions introduced by constraints from previous state transitions.

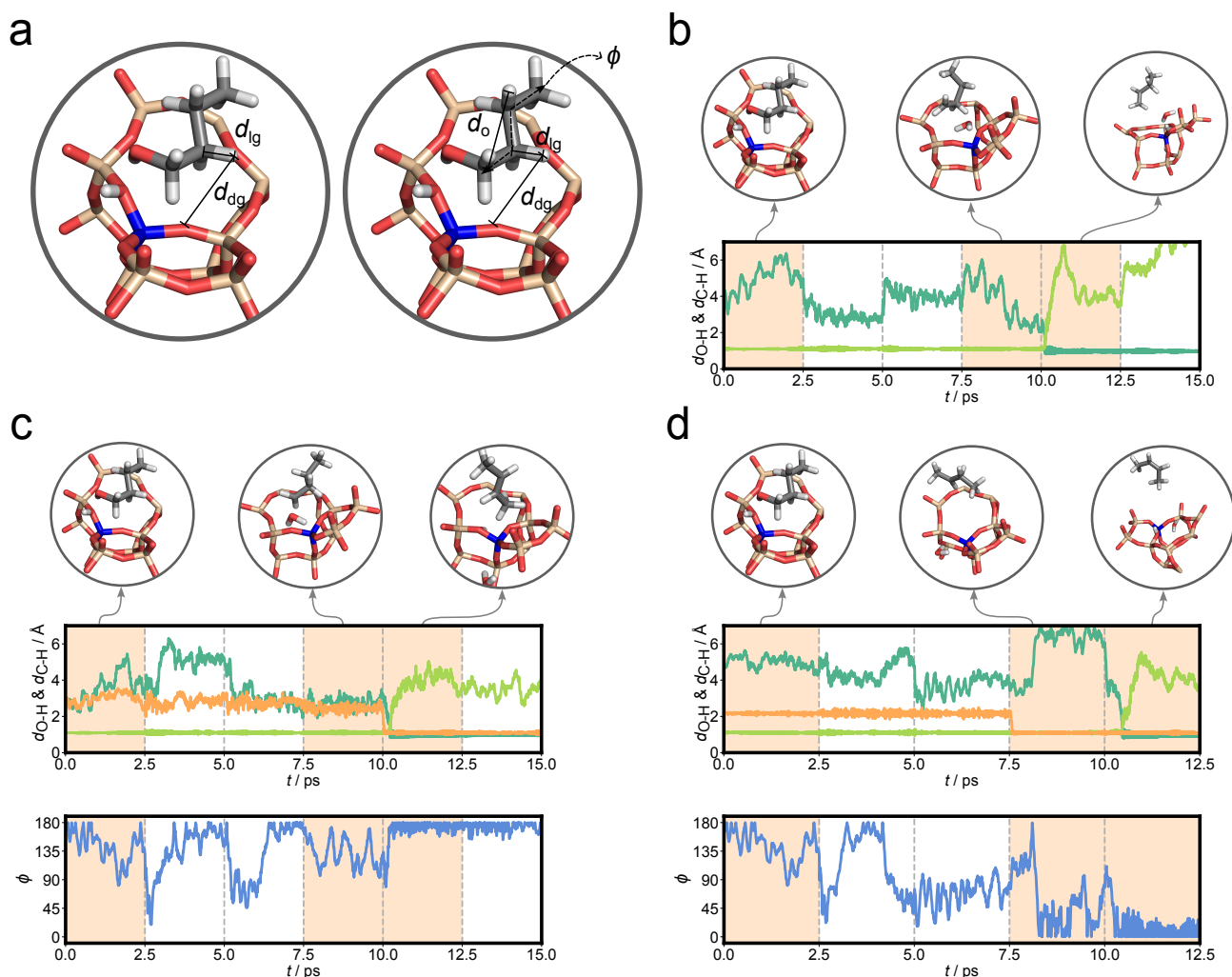


FIG. S15. **Concerted mechanisms of 1-butanol dehydration in acidic chabazite.** **a**, Definition of heuristic collective variables (CVs) for the 1-butene pathway and the cis-/trans-butene pathways. **b-d**, Time evolution of Loxodynamics-accelerated simulations yielding 1-butene (**b**), trans-butene (**c**), and cis-butene (**d**). Geometric parameters are color-coded: d_{ig} (light green) and d_{dg} (dark green) correspond to C-H and O-H distances tracking the interaction with the chabazite host. d_o (orange) denotes the intramolecular C-H distance associated with hydrogen migration, while ϕ (blue) is the dihedral angle distinguishing cis- and trans-butene.

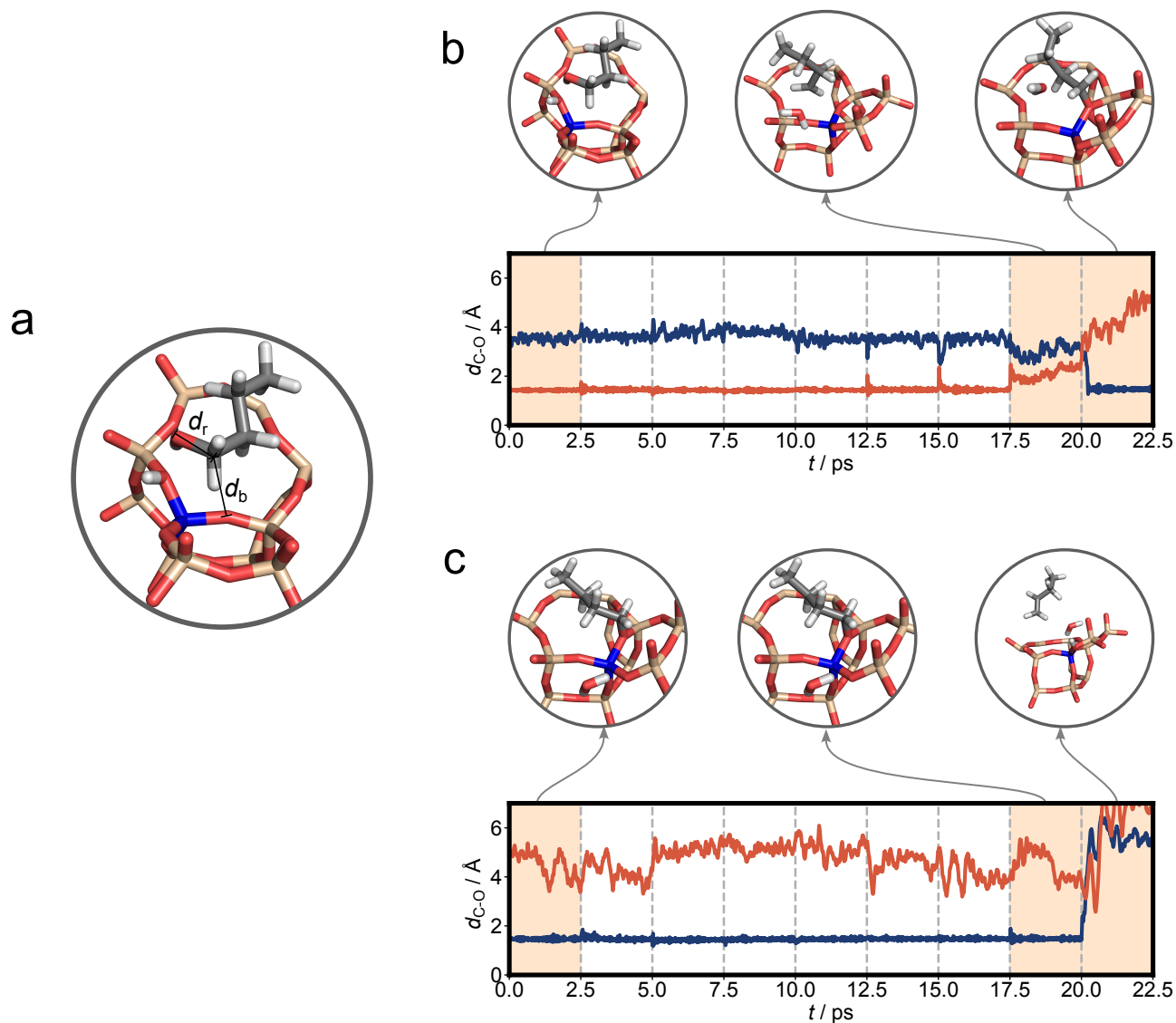


FIG. S16. **Stepwise mechanism of 1-butanol dehydration in acidic chabazite.** **a**, Definition of heuristic collective variables (CVs) for the two-step pathways. **b**, **c**, Time evolution of Loxodynamics-accelerated simulations leading to the formation of a surface-bound butoxyl intermediate (**b**) and the final 1-butene product (**d**). Geometric parameters are color-coded: d_r (red) and d_b (blue) correspond to the two C-O distances tracking the interaction with the chabazite host.

S3. PERCEPTUALLY UNIFORM VISUALIZATIONS

To guarantee an accurate interpretation of the data gradients, we provide here replicas of Figure 3a (main text), Figure S3, and Figure S11 rendered using a perceptually uniform color palette (Viridis). These visualizations are included to eliminate artifacts associated with non-uniform color maps and ensure a faithful representation of the underlying values.

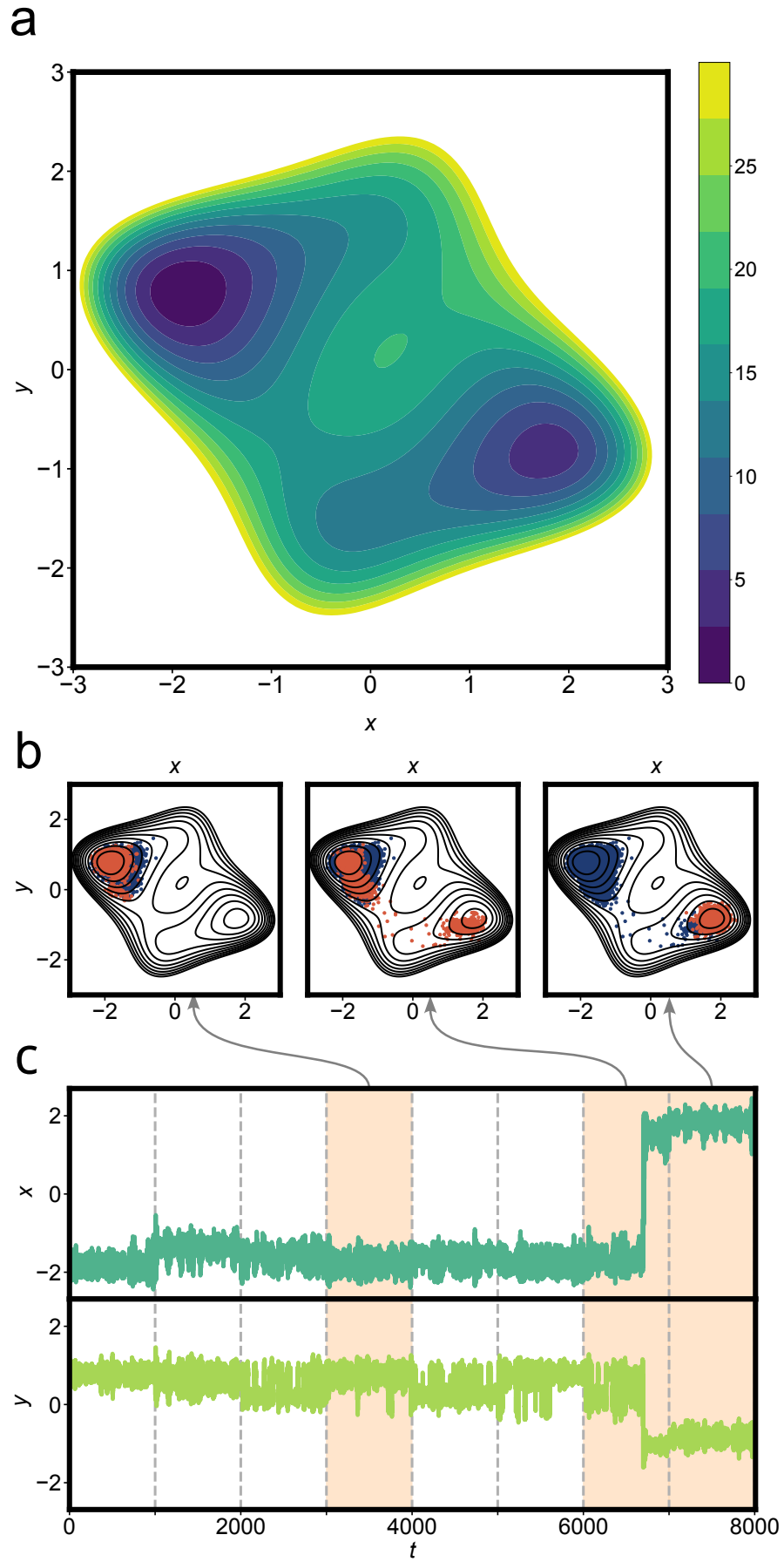


FIG. S17. Copy of Figure 3 (main text) using a perceptually uniform color palette

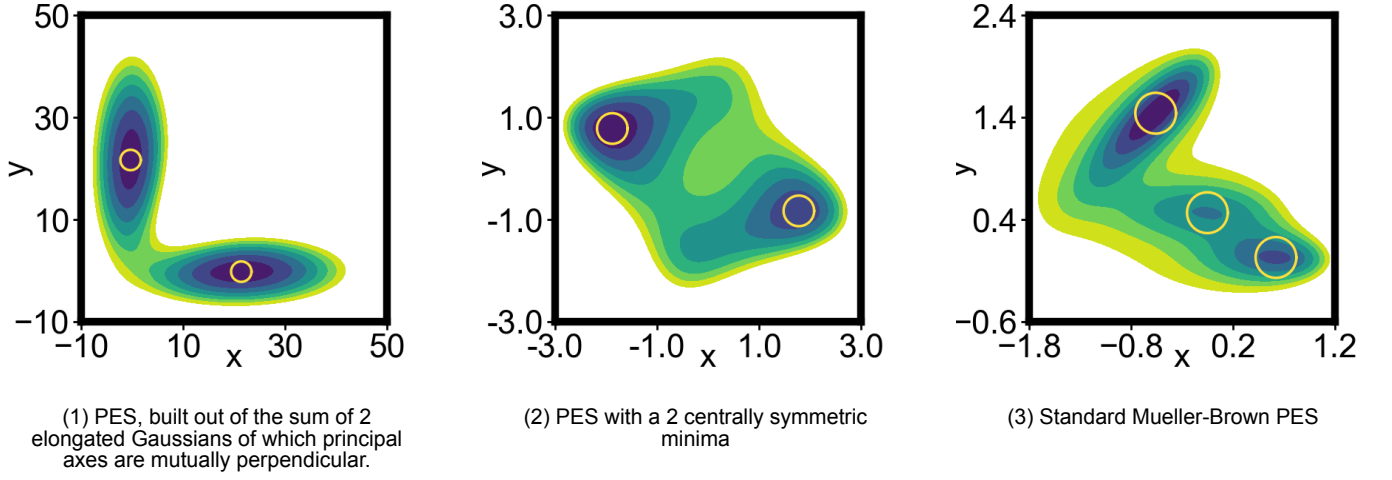


FIG. S18. Copy of Figure S3 using a perceptually uniform color palette

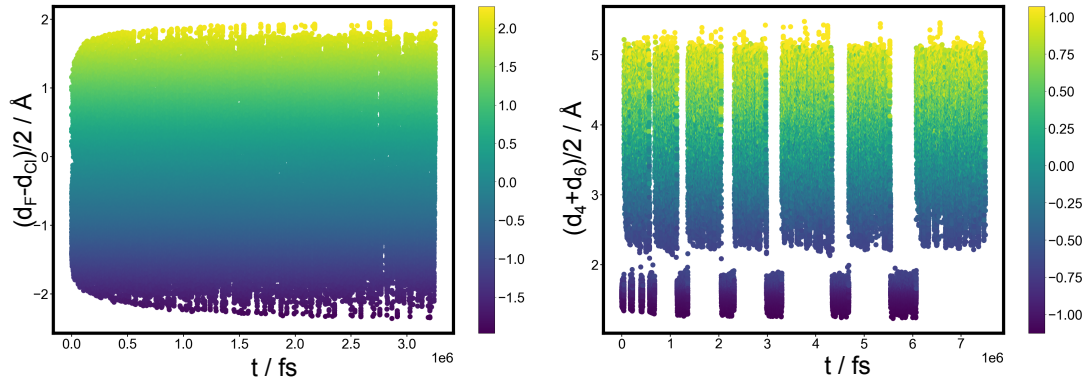


FIG. S19. Copy of Figure S11 using a perceptually uniform color palette

REFERENCES

- ¹He, X. *et al.* Practical lessons from predicting clicks on ads at facebook. In *Proceedings of the Eighth International Workshop on Data Mining for Online Advertising*, ADKDD'14, 1–9 (ACM, New York, NY, USA, 2014). URL <https://doi.org/10.1145/2648584.2648589>. 2
- ²Ash, J. & Adams, R. P. On warm-starting neural network training. *Adv. Neural Inf. Process. Syst.* **33**, 3884–3894 (2020). 2
- ³Ramkumar, V. R. T., Arani, E. & Zonooz, B. Learn, unlearn and relearn: An online learning paradigm for deep neural networks (2023). URL <https://arxiv.org/abs/2303.10455>. 2303.10455.
- ⁴Zhou, H., Vani, A., Larochelle, H. & Courville, A. Fortuitous forgetting in connectionist networks (2022). URL <https://arxiv.org/abs/2202.00155>. 2202.00155. 2
- ⁵Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning* (MIT Press, 2016). <http://www.deeplearningbook.org>. 3
- ⁶Schütt, K. T., Sauceda, H. E., Kindermans, P.-J., Tkatchenko, A. & Müller, K.-R. Schnet – a deep learning architecture for molecules and materials. *J. Chem. Phys.* **148**, 241722 (2018). URL <https://doi.org/10.1063/1.5019779>. https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.5019779/16655678/241722_1_online.pdf. 3
- ⁷Wright, D. B. & Herrington, J. A. Problematic standard errors and confidence intervals for skewness and kurtosis. *Behav. Res. Methods* **43**, 8–17 (2011). 5
- ⁸Bonati, L., Trizio, E., Rizzi, A. & Parrinello, M. A unified framework for machine learning collective variables for enhanced sampling simulations: mlcolvar. *J. Chem. Phys.* **159**, 014801 (2023). URL <https://doi.org/10.1063/5.0156343>. https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/5.0156343/18281524/014801_1_5.0156343.pdf. 6
- ⁹Invernizzi, M. & Parrinello, M. Best of a bad lot: A strategy for improving collective variables. *J. Chem. Theory Comput.* **15**, 2187–2194 (2019). 7
- ¹⁰Müller, K. & Brown, L. D. Location of saddle points and minimum energy paths by a constrained simplex optimization procedure. *Theor. Chim. Acta* **53**, 75–93 (1979). 7
- ¹¹Bonati, L., Rizzi, V. & Parrinello, M. Data-driven collective variables for enhanced sampling. *J. Phys. Chem. Letters* **11**, 2998–3004 (2020). URL <https://doi.org/10.1021/acs.jpclett.0c00535>. PMID: 32239945, <https://doi.org/10.1021/acs.jpclett.0c00535>. 14
- ¹²Mendels, D., Piccini, G. & Parrinello, M. Collective variables from local fluctuations. *J. Phys. Chem. Lett.* **9**, 2776–2781 (2018). URL <https://doi.org/10.1021/acs.jpclett.8b00733>. PMID: 29733652, <https://doi.org/10.1021/acs.jpclett.8b00733>. 15
- ¹³Kim, S., Robichaud, D. J., Beckham, G. T., Paton, R. S. & Nimlos, M. R. Ethanol dehydration in hzsm-5 studied by density functional theory: Evidence for a concerted process. *J. Phys. Chem. A* **119**, 3604–3614 (2015). URL <https://doi.org/10.1021/jp513024z>. PMID: 25802969, <https://doi.org/10.1021/jp513024z>. 17, 18
- ¹⁴Henkelman, G., Uberuaga, B. P. & Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of Chemical Physics* **113**, 9901–9904 (2000). URL <https://doi.org/10.1063/1.1329672>. https://pubs.aip.org/aip/jcp/article-pdf/113/22/9901/19259681/9901_1_online.pdf. 18