Supporting Information: Dynamically correct intrinsic reaction rate constants in heterogeneous catalysis

Massimo Bocus,[*,†] Lukas Baldauf,[‡] Titus S. van Erp,[‡] and Veronique Van Speybroeck[*,†]

†*Center for Molecular Modeling, Ghent University, Technologiepark 46, 9052 Zwijnaarde, Belgium*

‡*Department of Chemistry and Biomedical Science, Norwegian University of Science and Technology, 7941 Trondheim, Norway*

E-mail: massimo.bocus@ugent.be; veronique.vanspeybroeck@ugent.be

Contents

1 Extended computational details

1.1 Bias walls

Assume one would like to simulate the full time evolution of a zeolite crystallite undergoing a reaction with a mobile molecule, such as ethene in some of the considered reactions. This would mean that there are diffusive regions, where ethene is moving mostly freely in the crystal, and reactive regions, where ethene is considered in interaction with the active site. Obviously, this is a fictitious division and the choice should ultimately not matter. A larger reactive region would reduce the computed reaction rate since the molecule has now more space to move around the active site, but it would also shorten the time required to diffuse from reactive region to reactive region. In practice, however, limiting the molecule movement in the zeolite unit cell has a major impact on converging the simulations. If let free to fluctuate, especially in a relatively large cell of H-ZSM-5, an ethene molecule could easily spend nanoseconds away from the active site, making the simulations unfeasible to converge in a reasonable amount of time. Additionally, even assuming convergence was possible, one would get a reaction rate including diffusion for a very specific H-ZSM-5 crystal in which the Si/Al=95 and the Al are all perfectly arranged in the crystal. For this reason, it can be more sensible to include a bias potential that restrains the molecule in proximity of the active site.

The movement of ethene was restrained in proximity of the active site for the protonation and methylation reaction. For US, the bias was imposed by PLUMED[1,2] in addition to the umbrella bias, while for ∞RETIS we implemented a custom calculator in the atomic simulation environment[3] (ASE). The applied walls are summarized in Table S1. Note that the wall on the ethene methylation reaction is applied on the same CV used to characterize the reaction itself. Therefore, the wall is not needed in umbrella sampling since the CV range is already restricted by the presence of the umbrella biasses.

While the isobutene protonation reaction did not require any wall to restrict the movement of the adsorbate molecule (the CHA cages are relatively small and isobutene does not diffuses through the gate in the simulations timescale), we added a wall to prevent the transfer of an hydrogen atom to oxygen atoms not in the first coordination sphere of the Al defect (CN(C;H) + CN(O;H) < 7.0, in which O includes the four O atoms in the first coordination sphere of the Al defect). These were only needed in REUS as no side reaction was observed in regular US or ∞RETIS. Similarly, we restrained the ethene methylation reaction to prevent proton transfers from ethene to the zeolite (CN(O_{zeo};H_{et}) $-$ CN(C_{et};H_{et}) > -3.0), and to prevent the formation of surface-bound species (CN(C_{et};O_{zeo}) > 0.3). This type of walls, which does not restrain the movement of the molecule, can be seen as simple guard rails to stop the molecules from jumping

off the transition state region because of unfavorable thermal fluctuations and, if chosen appropriately, should not affect the resulting rate constant. It is interesting to notice that in ∞RETIS the walls are not needed because the system is not restrained in the high-energy region of the TS, but it rather crosses it in a more 'natural' fashion towards the product state.

Table S1: Repulsive walls applied in the simulations. The wall is defined as $W(\xi) = \frac{\kappa}{2}(\xi - \xi_0)^2$ if $\xi > \xi_0$ and 0 otherwise if it is an upper wall, while the opposite for a lower wall. In the collective variables definition, d stands for distance, com for center of mass, and CN for coordination number.

Reaction	**Collective variable**	**REUS**	**∞RETIS**	**type**	ξ_0 (Å)	κ (kJ·mol^{-1}·ξ_{unit}^{-2})
Ethene protonation	d(com(C_{et}); Al)	✓	✓	upper	6.5	3000
Ethene methylation	d(C_{me}; O_{me}) - d(com(C_{et}); C_{me})	✗	✓	lower	-5.0	3000
	CN(O_{zeo};H_{et}) - CN(C_{et};H_{et})	✓	✗	upper	-3.0	3000.0
	CN(C_{et};O_{zeo})	✓	✗	upper	0.3	3000.0
Isobutene protonation	CN(C;H) + CN(O;H)	✓	✗	lower	7.0	3000.0

1.2 Replica-exchange umbrella sampling

Differently from regular US, REUS requires all simulations to run synchronously as, at each time step, a swap can be attempted. The acceptance probability for the swap between two umbrellas is proportional to $\exp\{-\beta[(\mathcal{U}_j(\mathbf{i}) + \mathcal{U}_i(\mathbf{j})) - (\mathcal{U}_i(\mathbf{i}) + \mathcal{U}_j(\mathbf{j}))]\}$, in which $\mathcal{U}_i$ and $\mathcal{U}_j$ are two hamiltonians (in our case, two umbrellas with different centers) and **i** and **j** the current configuration in simulations i and j, respectively. We set the swapping frequency such that, on average, a swap move is attempted every two simulation steps.

The umbrella centers and force constants selection was done heuristically based on our previous experience with the three reactions under investigation.[4–6] The number of umbrellas was also limited by the synchronous execution requirement of replica-exchange which, in the current i-PI[7] implementation, practically reflects on the fact that all simulations must run on the same computing node. This limited us to a maximum of 40 umbrellas per simulation, which are nevertheless more than sufficient to satisfactorily sample the reaction and lead to a good amount of replica-exchange swaps. We provide a complete overview of the umbrella sampling parameters,

both for the REUS simulation as well as the additional tighter umbrella atop the transition state (to generate initial RF structures, see main text), in Table S2.

Table S2: The umbrella centers ($\boldsymbol{\xi}_0$), force constants (κ) and simulation time for the REUS simulations of the three reactions under investigation. For the centers, we adopt the notation $[\xi_i : \xi_j]_N$ to indicate N equally-spaced values between ξ_i and ξ_j, included.

Reaction		$\boldsymbol{\xi}_0$	$\boldsymbol{\kappa}$	**simulation time** (ps)
Isobutene protonation	umbrellas	$[5.5 : 7.5]_{40}$	3000 kJ·mol^{-1}	100
	TS umbrella	6.56	5000 kJ·mol^{-1}	200
Ethene methylation	umbrellas	$[-5.0 : 0.0]_{40}$	1000 kJ·mol^{-1}·Å^{-2}	100
	TS umbrella	-0.137	3000 kJ·mol^{-1}·Å^{-2}	200
Ethene protonation	umbrellas	$[1.7 : 4.7]_{40}$	3000 kJ·mol^{-1}	100
	TS umbrella	3.529	4000 kJ·mol^{-1}	200

1.3 ∞RETIS initialization

To start, an ∞RETIS simulation requires the location of the various interfaces $\{\xi_0, \xi_1, ..., \xi_N\}$ and at least one valid path per ensemble. Since deciding the right interface location is non-trivial and initial paths are not readily available, we adopted the `inf-init` procedure implemented within the `infretis` library. First, a 1 ns long MD simulation was performed in the basins of attraction representing the reactant (A) and product (B) states. This allows to define reasonable values for $\xi_0 \equiv \xi_{\mathrm{A}}$ and $\xi_N \equiv \xi_{\mathrm{B}}$ based on the criterion that the interface should be crossed every few ps at most, preventing excessively long paths in the $[0^-]$ ensemble. To summarize, we hereafter focus on the reactant-to-product reaction but the same applies to the backward process. 16 random frames are extracted from the reactant MD simulation and propagated backward and forward in time to generate two paths in the $[0^-]$ and $[0^+]$ ensembles. One of the $[0^-]$ paths is randomly chosen as initial path for the ensemble. The $[0^+]$ paths of the remaining 15 simulations are ranked based on the maximal value reached by the order parameter and 15 interfaces $\{\xi_1, \xi_2, ..., \xi_{15}\}$ are placed such that every path is valid in one of them. Of course, coming from a short MD simulation the interfaces will be very close to each other and to ξ_0.

A short ∞RETIS simulation is now run starting from these initial paths. At its end, the crossing probability is estimated via WHAM analysis and the interfaces are moved such that the probability of crossing ξ_{i+1} given that we crossed ξ_i first is 0.3. If necessary, additional interfaces can also be automatically added. The interesting aspect of this procedure is that the WF moves in the last ensemble will likely push the system forward, up the reaction barrier. Hence, after few iterations, one should start to observe reactive paths being sampled. This is clearly visible

in Figure S1, where the maximum order parameter reached by each path and the interfaces location is shown for 6 `inf-init` iterations applied to the ethene methylation reaction. In general, for our systems, we performed 5–10 `inf-init` iterations starting from ~ 200 paths for the first iteration up to ~ 2000 for the last one (always excluding the first $\sim 20\%$ of the paths as equilibration). This provided a well converged set of interfaces as well as an initial path for each ensemble, which were then used to initialize the production ∞RETIS simulations.

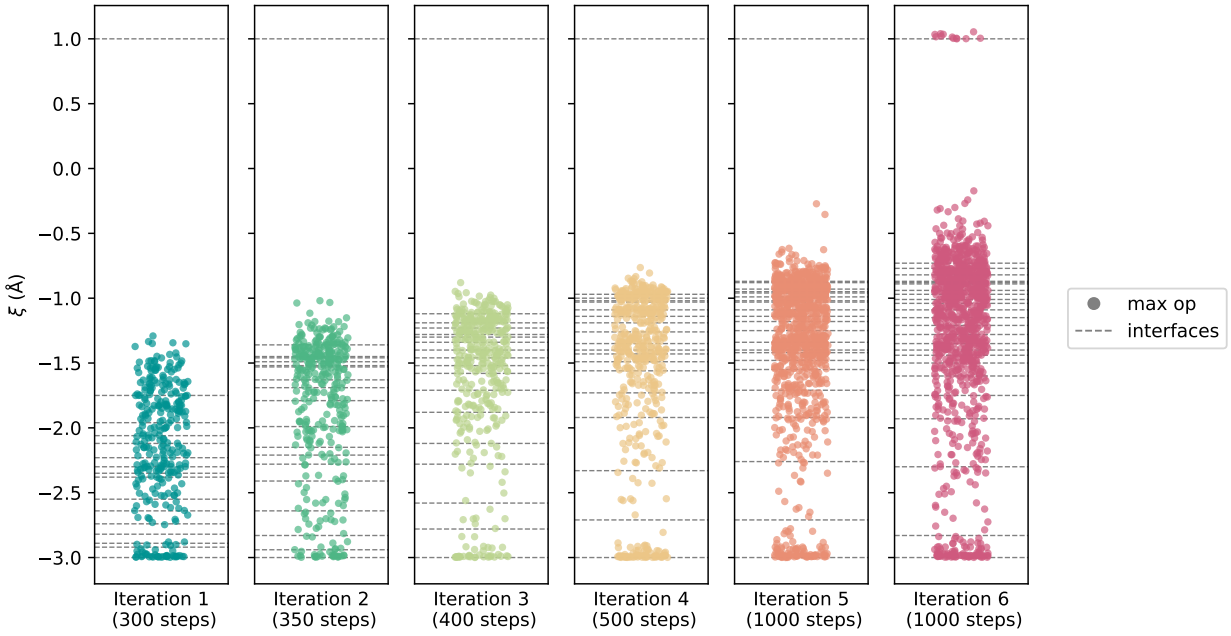

Figure S1: The 6 `inf-init` loops applied to the ethene methylation reaction. The dots represent the maximum order parameter reached by each path, while the grey dashed lines the current interfaces. It can be seen how, by the sixth iteration, some paths are reaching ξ_B at $\xi = 1.0$ Å.

1.4 Error estimation

Error estimate on the ∞RETIS rates are computed with the recursive block error procedure proposed in ref. 8. An estimate of the true error is obtained by fitting the block uncertainty with the empirical expression $BE = TE\sqrt{B/(B+\tau-1)}$, in which BE is the error of a block size B, TE the true error, and τ an estimate of the correlation size. The last two parameters are obtained from the fit. Since the conditional free energies are used exclusively for a qualitative comparison with (RE)US, they are computed from all the available data without any uncertainty estimation.

The error estimate on the REUS rates is more complex. As mentioned in the main article, the Bennett-Chandler rate equation is:

$$k_{\mathrm{A}\to\mathrm{B}}(t) = \left\langle \dot{\xi}\left(\mathbf{r}_0\right) h\left(\xi\left(\mathbf{r}_t\right) - \xi^{\ddagger}\right)\right\rangle_{\xi(\mathbf{r}_0)=\xi^{\ddagger}} \frac{e^{-\beta F(\xi^{\ddagger})}}{\int_{-\infty}^{\xi^{\ddagger}} e^{-\beta F(\xi)}\mathrm{d}\xi} \tag{1}$$

We divide the error calculation in two parts, one for the thermodynamics term and one for the kinetics term. For the former, we use `ThermoLIB`[9] to estimate the statistical inefficiency of each CV time series by means of autocorrelation functions. The statistical inefficiency is used to subsample the CV time series and decorrelate the samples, which allow to obtain an uncertainty estimate of the FEP by the maximum-likelihood estimate method implemented in `ThermoLIB`. For the kinetic part, the approach changes between TST and RF. For TST, the rate prefactor $\frac{1}{2}\left\langle\left|\dot{\xi}\left(\mathbf{r}_0\right)\right|\right\rangle_{\xi(\mathbf{r}_0)=\xi^{\ddagger}}$ is computed for all structures in a small interval $\delta\xi$ around $\xi^{\ddagger}$. Since these structures come from the same trajectory, they are correlated and, therefore, the uncertainty is obtained with block averaging. In reactive flux, we get either $\dot{\xi}\left(\mathbf{r}_0\right)$ or zero depending on the basin in which the simulations end up. Since every structure is given randomized velocities, we can consider the entries to be uncorrelated—but not normally distributed. The uncertainty estimate is therefore obtained via bootstrapping. We obtain a set of values for $k_{\mathrm{A}\to\mathrm{B}}$ by taking the computed values of $\left\langle \dot{\xi}\left(\mathbf{r}_0\right) h\left(\xi\left(\mathbf{r}_t\right) - \xi^{\ddagger}\right)\right\rangle_{\xi(\mathbf{r}_0)=\xi^{\ddagger}}$ and random samples from the uncertainty estimate of $F(\xi)$ for the thermodynamic part. We then performed bootstrapping on the $k_{\mathrm{A}\to\mathrm{B}}$ values and report the uncertainty as twice the standard deviation on the bootstrapping values. Note that in all cases the error bars are reported as a 95% confidence interval (i.e., twice the standard deviation).

To sum things up, the error on k is assumed to be normally distributed when obtained from ∞RETIS, log-normally distributed when obtained from TST (since the error on $F(\xi)$ is assumed to be normally distributed and the rate is dominated by $\exp(-\beta F(\xi))$) and, finally, normally distributed for RF coming from bootstrapping.

1.5 MLP parameters

The MLP parameters for the isobutene reaction were kept identical as our previous work.[6] For the ethene methylation and ethene protonation reactions, some hyperparameters had different values, which are summarized in Table S3. We do not provide any particular recommendation concerning the hyperparameters choice. The main difference between the two models can be summarized in a smaller rotation order (from 2 to 1) but a larger number of channels (16 to 32). We believe that both choices are valid and provide similar results (both in terms of accuracy and efficiency) and, therefore, the choice should be heuristic and based on the actual accuracy which is desired.

For isobutene protonation, the new model was trained on 4491 structures and 499 additional

structures for validation. The final model reaches a root mean squared error (RMSE) of 0.3 meV per atom on the energies and 29.9 meV·Å^{-1} on the forces. For ethene methylation and ethene protonation, the model was trained on 5931 structures and 660 additional structures for validation. The final model reaches a RMSE of 0.5 meV per atom on the energies and 33.9 meV·Å^{-1} on the forces. For both models the training metrics are more than satisfactory enough and known by now to result in quantitative reaction free energies. As our overarching goal is not to obtain exact reaction rates (the DFT level of theory used to train the model is known to significantly underestimate reaction barriers in zeolite catalysis,[10,11] way beyond any inaccuracy of the trained model) we did not perform additional validations of the model accuracy, aside from ensuring that no unphysical states were generated during the dynamics.

Table S3: Summary of the hyperparameters adopted for the model trained on the ethene protonation and ethene methylation reactions. All additional hyperparameters were kept at their default value.

Hyperparameter	**Value**
cutoff radius (`r_max`)	6.5 Å
maximum rotation order (`max_L`)	1
number of channels (`num_channels`)	32
number of radial basis functions (`num_radial_basis`)	8
batch size (`batch_size`)	8
learning rate (`lr`)	0.01
scheduler patience (`scheduler_patience`)	10
patience	20
maximum number of epochs (`max_num_epochs`)	1,000
start stochastic weight averaging (`start_swa`)	400
energy weight (`energy_weight`)	100
swa energy weigth (`swa_energy_weight`)	200

1.6 Rate constants from unbiased MD simulations

If a certain number of transitions between reactants and products can be observed during a regular, unbiased MD simulation, one can compute an estimate of the reaction rate constant. As this was the case for the isobutene protonation reaction, we proceeded as follow: first, we computed the CV $\xi = \mathrm{CN(C;H)} - \mathrm{CN(O;H)}$ along the simulation trajectory. We then defined two interfaces at CV values of 5.8 and 7.2. States with $\xi < 5.8$ are attributed to isobutene, while states with $\xi > 7.2$ to the *tert*-butyl cation. The trajectory is divided in chunks each time that the system enters in one of the basins while coming from the other one (or, in other words, when a complete reactive event occurs, see Figure S5). Each chunk is attributed to the basin in which the

simulation had entered. The reaction rate constant $k_{\mathrm{A}\to\mathrm{B}}$ can then be estimated as the number of A→B transitions divided by the cumulative simulation time of the chunks belonging to A.[12]

2 Extended results

2.1 NPT equilibrations

The unit cell of H-ZSM-5 was first equilibrated in the NPT ensemble for both the ethene methylation (1 atm, 623 K) and ethene protonation (1 atm, 573 K) reactions. For the former, we considered only the reactant state (ethene + methanol) while for the latter both reactant (ethene) and product (surface ethoxide species) were simulated. For every basin of attraction we performed 4 independent 1 ns long MD simulations in the NPT ensemble, relying on a Martyna-Tuckerman-Tobias-Klein barostat [13,14] to regulate the pressure. From each simulation, we extracted the average cell parameters to assess their variability and we then further average over all of them for the final simulations. The results are depicted graphically in Figure S2 and Figure S3 for ethene methylation and ethene protonation, respectively.

It can be seen that the four replicas produce very close results, with variations in the order of 10^{-2} Å and 10^{-1} degrees. Interestingly, it appears that going from the mobile ethene to the surface-anchored ethoxide species has a non-negligible impact on the unit cell parameters, as visible by the clustering of the yellow and blue dots in Figure S3. While this is interesting to note, for all practical purposes the differences are very small and, therefore, we opted for continuing the production runs in the NVT ensemble with the overall average cell parameters, which are summarized in Table S4.

Table S4: The average H-ZSM-5 unit cell parameters for the ethene methylation and ethene protonation reactions.

Reaction	**T (K)**	**P (atm)**	**a (Å)**	**b (Å)**	**c (Å)**	α **(deg)**	β **(deg)**	γ **(deg)**
Ethene methylation	623	1.0	20.143	19.972	13.417	90.05	90.09	90.04
Ethene protonation	573	1.0	20.150	19.967	13.417	89.93	89.92	90.04

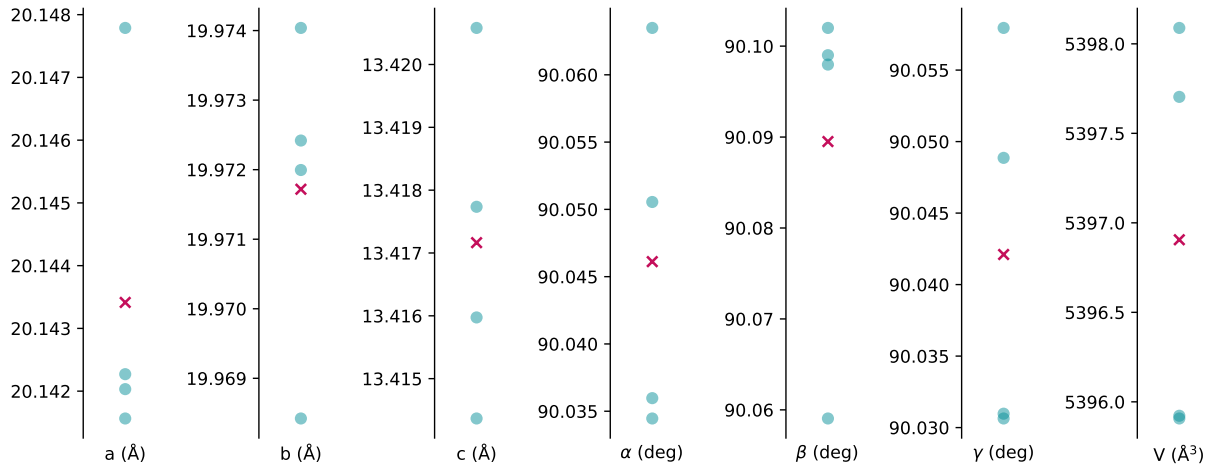

Figure S2: The average unit cell parameters and volume at 1 atm and 623 K from 4 independent NPT MD simulations of ethene + methanol in the H-ZSM-5 zeolite. The overall average is indicated by a red cross.

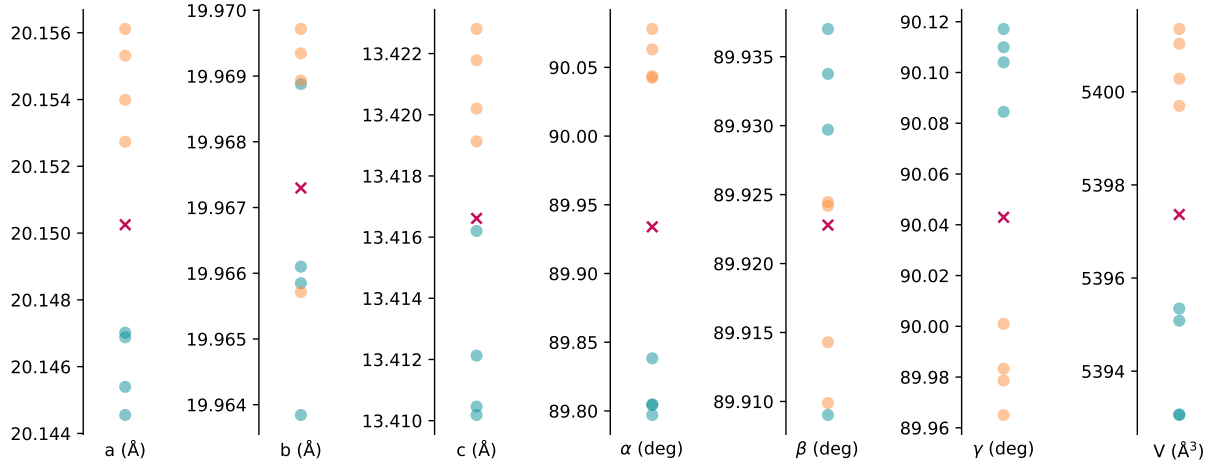

Figure S3: The average unit cell parameters and volume at 1 atm and 573 K from 4 independent NPT MD simulations of ethene (yellow dots) and the surface ethoxide species (blue dots) in the H-ZSM-5 zeolite. The overall average is indicated by a red cross.

2.2 Isobutene protonation

Both regular US and REUS were used to investigate the reaction. A discrete number of swaps was observed in the REUS simulations (Figure S7), although not all trajectories visited all the umbrellas—which would be a strong indication of convergence. To evaluate the advantage of using replica-exchange, we recomputed the reaction FEP at various simulation times and evaluated the mean absolute error (MAE) compared to the final, converged FEPs (Figure S8). As expected, REUS exhibits a (almost) monotonically decreasing error, while US can have quite some important oscillations depending on the phase space regions that the system happens to be visiting. Hence, aside from the more complex algorithmic implementation, REUS is in general recommended over regular US.

The results from the regular MD simulations are summarized in Figure S5 and Figure S6. Figure S9 reports the transmission coefficient from the reactive flux simulation. Figure S10 and Figure S11 plot a set of diagnostic properties obtained from the ∞RETIS simulations. Finally, a 2-dimensional deprojection of the original 1-dimensional FEP along the two CNs that constitute the final CV can be seen in Figure S12.

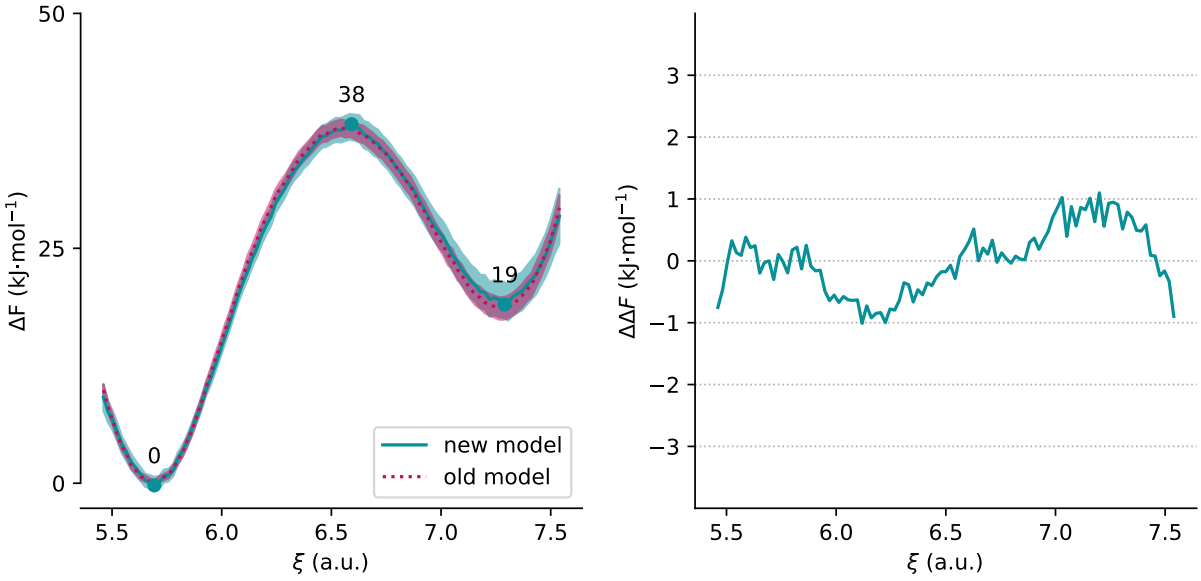

Figure S4: The newly trained model produces an identical free energy surface as the old one from ref. 6. Indeed, the two curves differ at most of 1 kJ·mol^{-1}, which is well within the error bars.

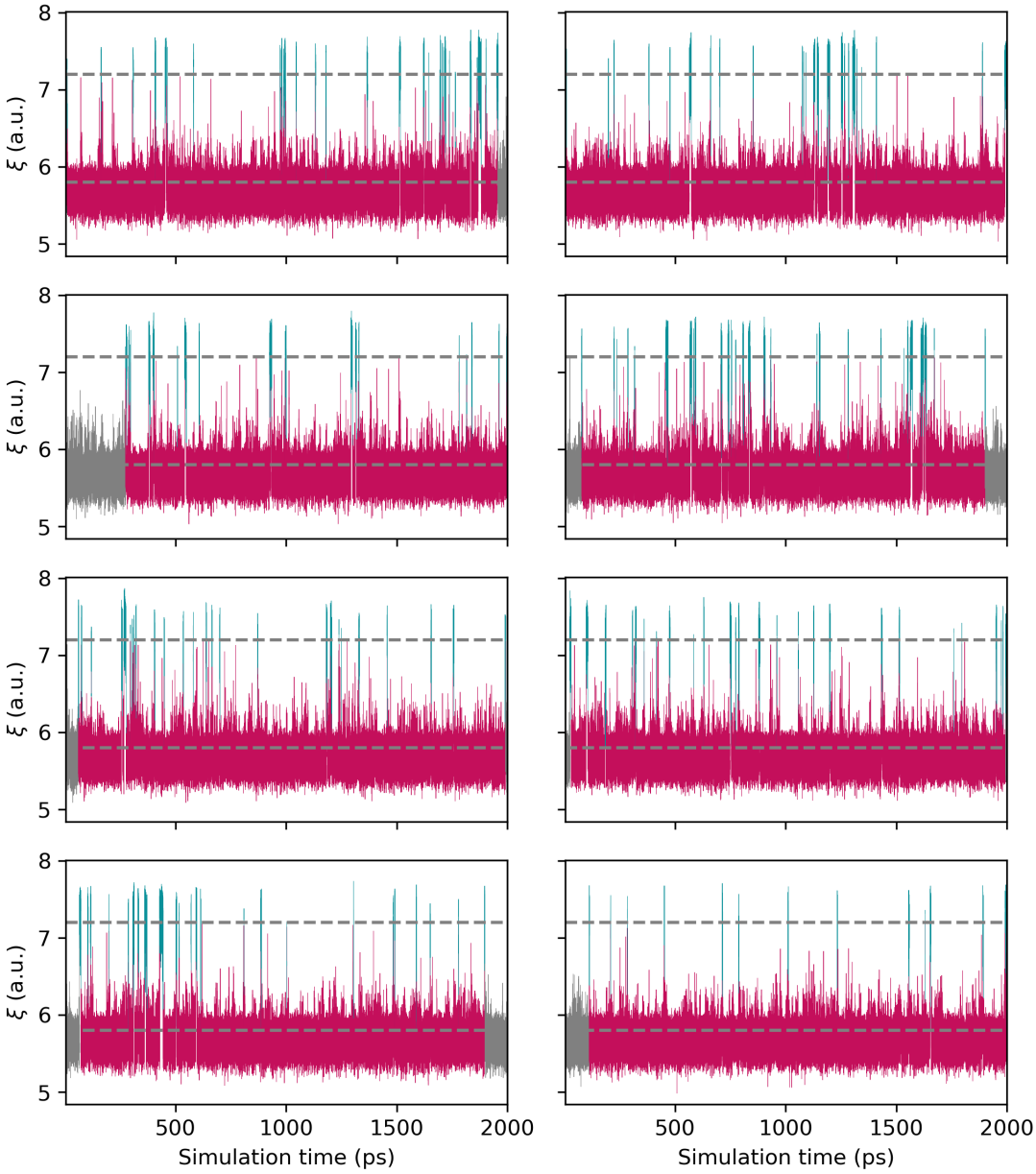

Figure S5: Time evolution of $\xi = \mathrm{CN(C;H)} - \mathrm{CN(O;H)}$ during 8 independent MD simulations. The horizontal dotted lines are possible boundaries for the isobutene ($\xi < 4.8$) and *tert*-butyl cation ($\xi > 7.2$) basins, though more combinations were also tried (see Figure S6). The trajectory is colored to illustrate how the rate calculation is performed. Isobutene is given a red color while the *tert*-butyl cation a green color. Each time that the system enters in one of the stable states, the line gets tinted with the respective color. As an example, starting from a state for which $\xi < 4.8$, the line will be red and will change color only when crossing the second interface at $\xi = 7.2$. Excursions in the central 'no man's land' do not change the line color. The rate of the isobutene to *tert*-butyl cation reaction is then equal to the number of times that a red-to-green color change was observed divided by the total time that the line was red (and vice-versa for the opposite reaction). The gray regions are excluded from the calculation as they are truncated by the beginning and end of the simulation.

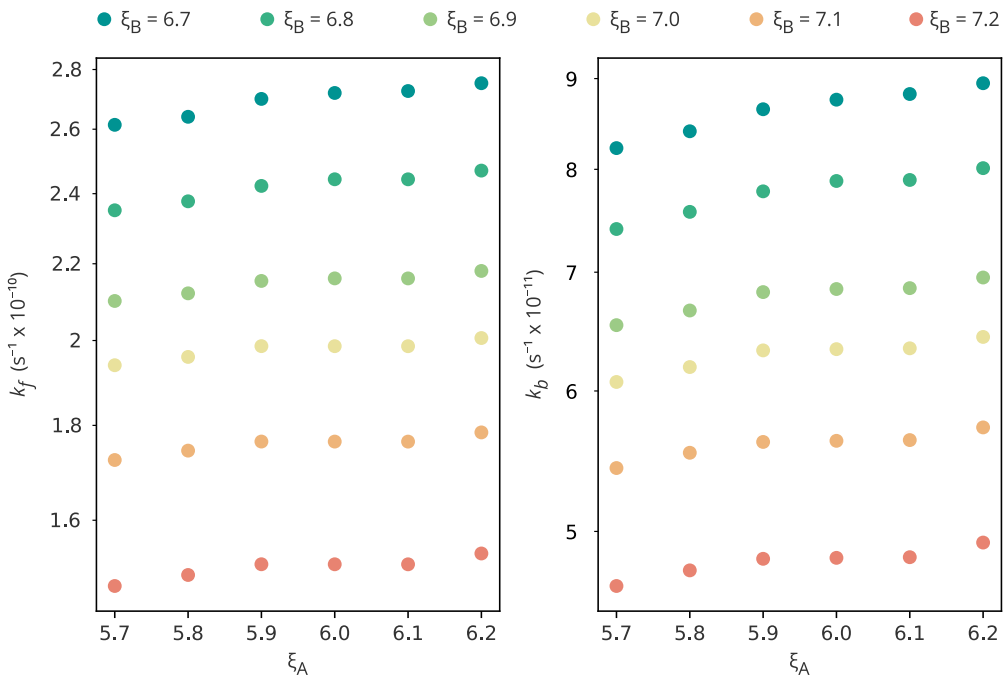

Figure S6: The computed values of k_f and k_b for the isobutene protonation reaction from 2 ns long, unbiased MD simulations, as function of the arbitrary location of the ξ_{A} and ξ_{B} values.

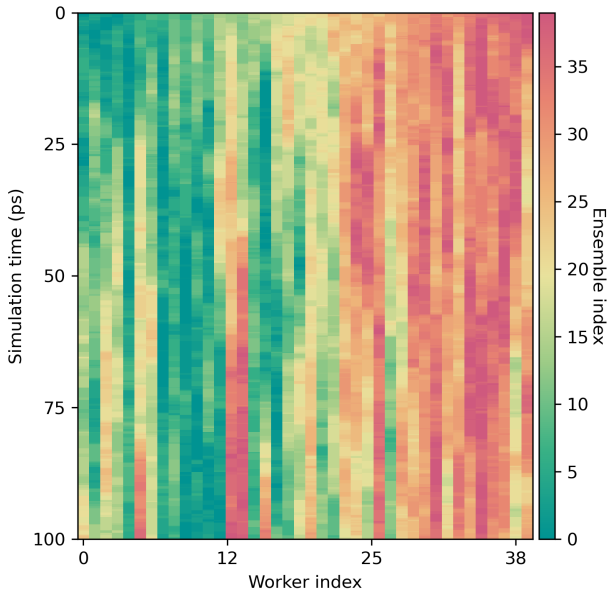

Figure S7: Schematic depiction of the swaps observed during the REUS of the isobutene protonation reaction. A *worker*, indexed on the horizontal axis, is a continuous MD simulation that can swap between *ensembles*, represented by different colors, here intended as umbrellas with a different center. So at the beginning walker i begins its journey from umbrella i but, in time, the swaps bring it to explore other umbrellas, shown by the change in color vertically, as time progresses.

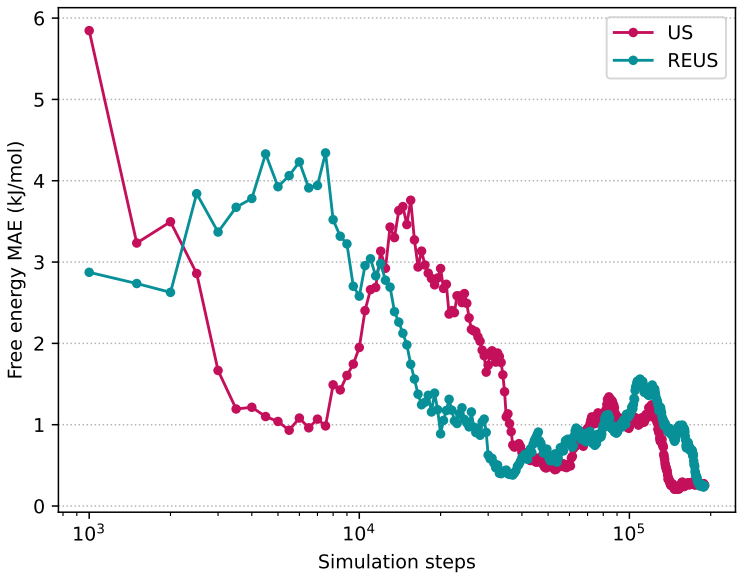

Figure S8: The mean absolute error (MAE) on the REUS and US free energy profiles computed after a certain amount of simulation steps. The reference free energy is the mean between REUS and US and does therefore not represent an actual “target".

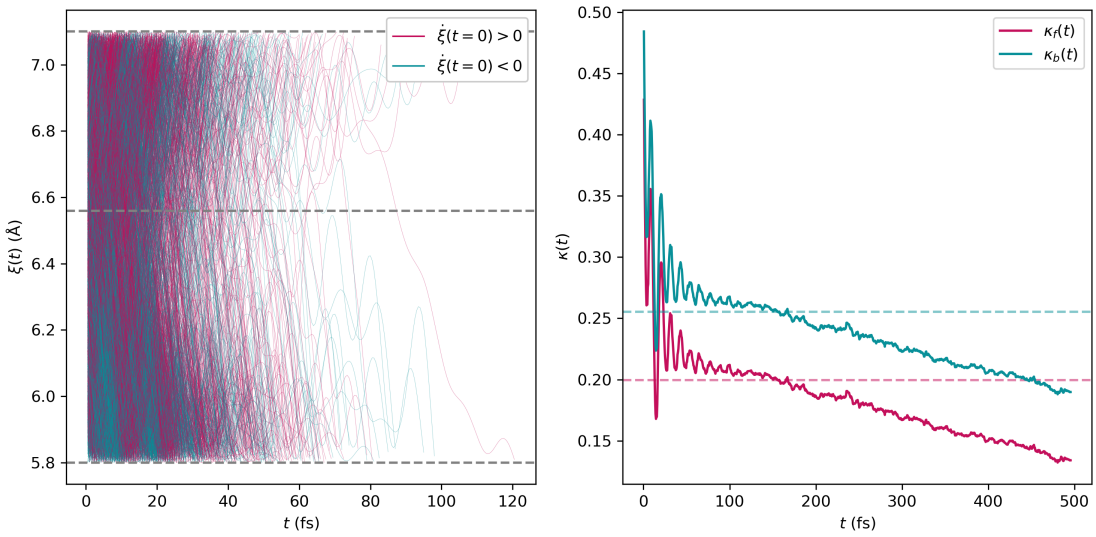

Figure S9: The reactive flux trajectories used to compute the transmission coefficient $\kappa(t)$ of the isobutene protonation reaction, which is reported on the right for the forward and backward reactions. Remark that $\kappa(t)$ should be the same for both directions and the observed deviations are due to the statistical noise. The trajectories are colored based on the initial velocity of the system along the CV ($\dot{\xi}(t=0)$). It is interesting to note that $\kappa(t)$ does not reach a plateau due to the low free energy barrier, so the final transmission coefficient used to compute the RF rate constant was taken as the average value between 100 and 200 fs (dotted horizontal lines).

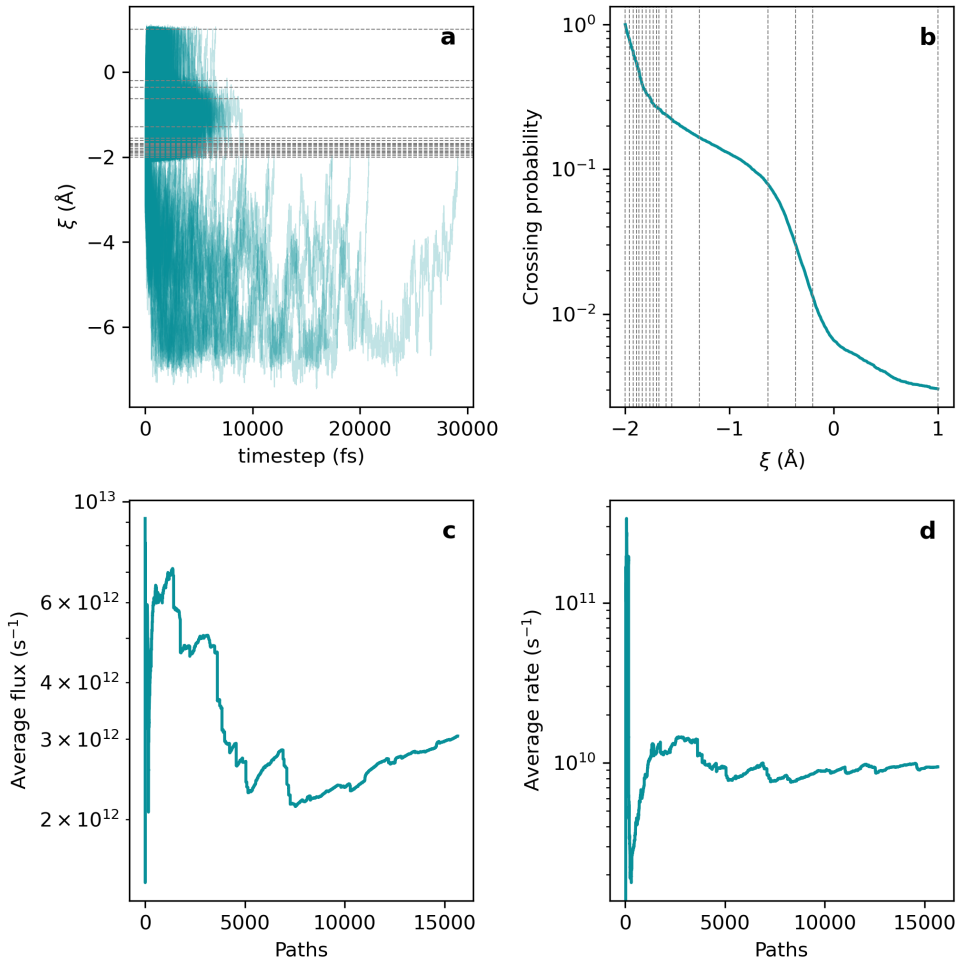

Figure S10: A summary of the ∞RETIS results for the forward isobutene protonation reaction (isobutene to *tert*-butyl cation). **a**. The 15682 accepted paths (5000 initial paths are excluded as equilibration). **b**. The crossing probability as function of the order parameter. **c**. The running average of the flux through the ξ_0 surface. **d**. The running average of the predicted rate.

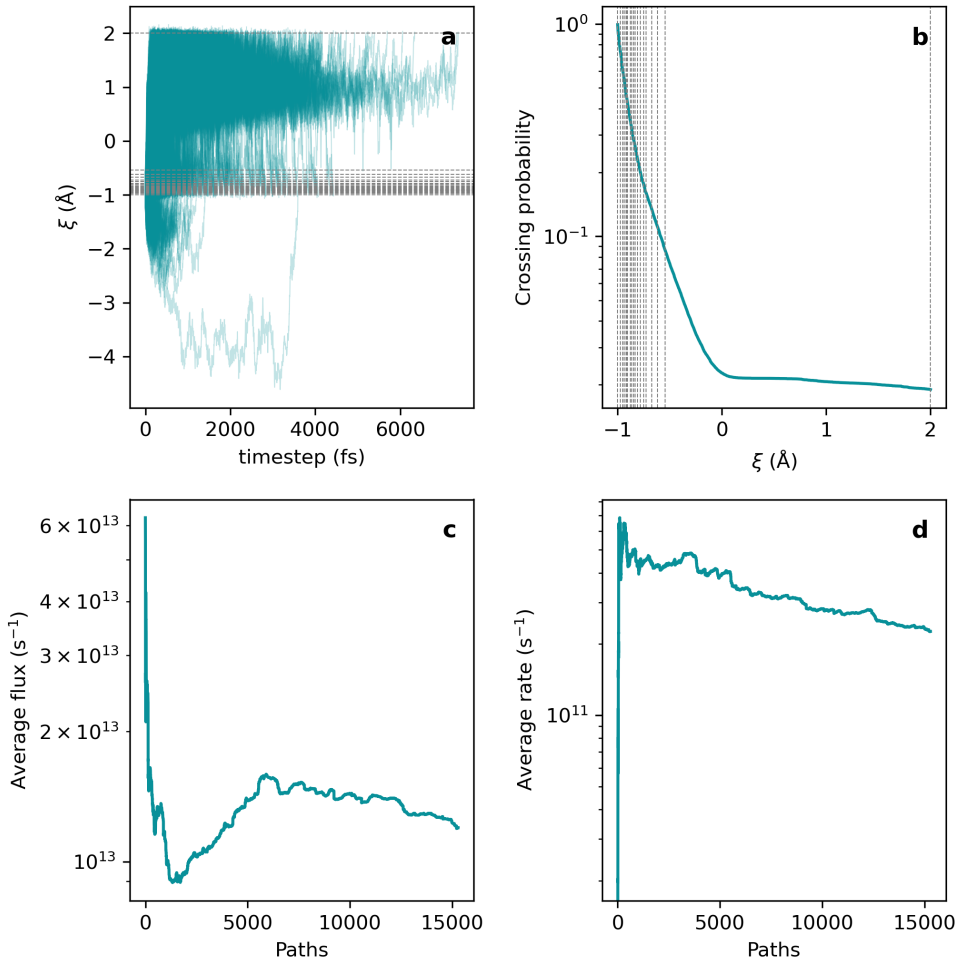

Figure S11: A summary of the ∞RETIS results for the backward isobutene protonation reaction (*tert*-butyl cation to isobutene). **a**. The 15316 accepted paths (5000 initial paths are excluded as equilibration). **b**. The crossing probability as function of the order parameter. **c**. The running average of the flux through the ξ_0 surface. **d**. The running average of the predicted rate.

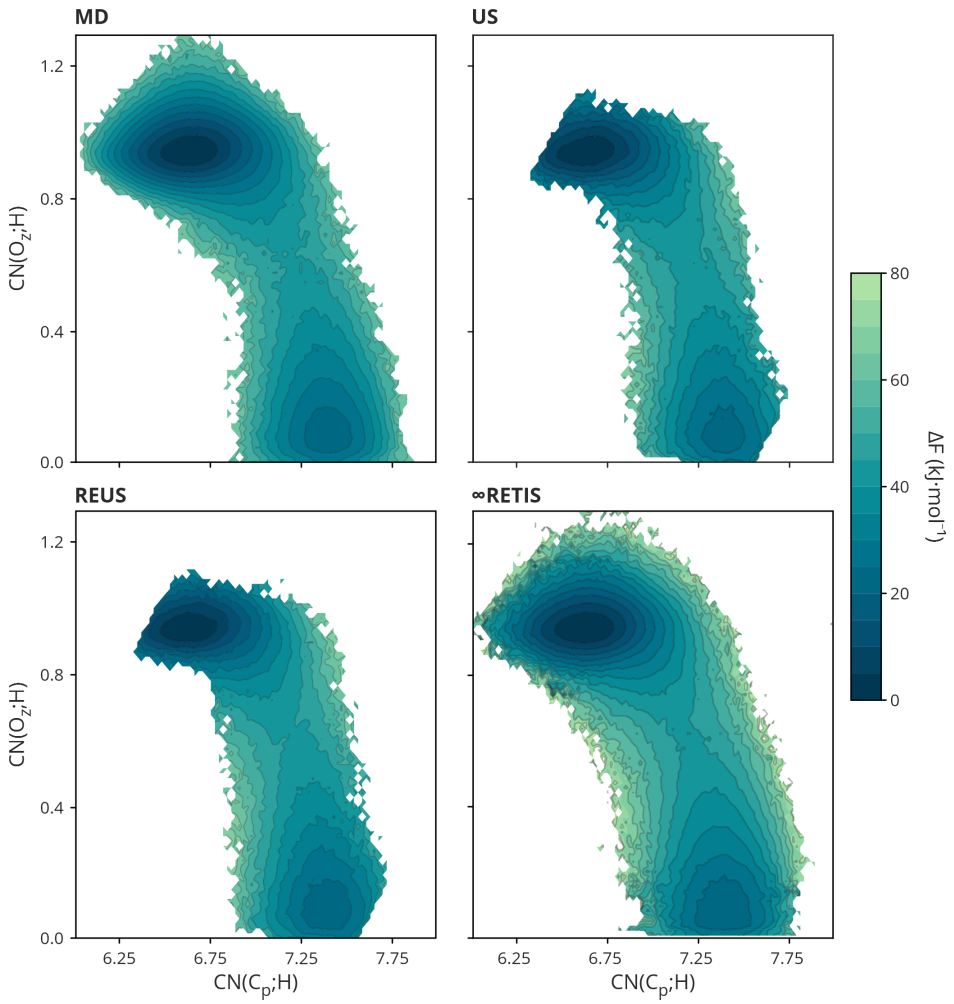

Figure S12: Free energy surfaces for the isobutene protonation reaction, shown along the CN between the primary carbon atoms and all hydrogen atoms (CN(C_p;H)) and the CN between the four zeolite oxygen atoms in the first coordination sphere of the Al defect and all hydrogen atoms (CN(O_z;H)). The 1-dimensional CV used in US is the difference between the two. The free energies are obtained from regular MD simulations (simple probability histogram), from US and REUS (deprojection with conditional probability), and from ∞RETIS (weighted histogram of forward and backward conditional free energies).

2.3 Ethene methylation

A graphic depiction of the swaps observed during the REUS simulation of ethene methylation in H-ZSM-5 is shown in Figure S13, where it can be seen that a good degree of scrambling is present (though not all simulations visit all the ensemble in the allocated simulation time). Figure S14 shows the transmission coefficient obtained from the reactive flux simulation and Figure S15 reports some diagnostic properties from the ∞RETIS simulation.

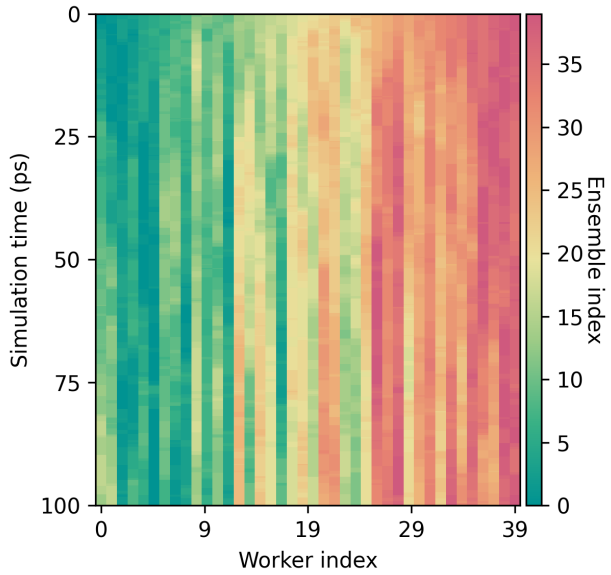

Figure S13: Schematic depiction of the swaps observed during the REUS simulation of ethene methylation. A *worker*, indexed on the horizontal axis, is a continuous MD simulation that can swap between *ensembles*, represented by different colors, here intended as umbrellas with a different center. So at the beginning walker i begins its journey from umbrella i but, in time, the swaps bring it to explore other umbrellas, shown by the change in color vertically, as time progresses.

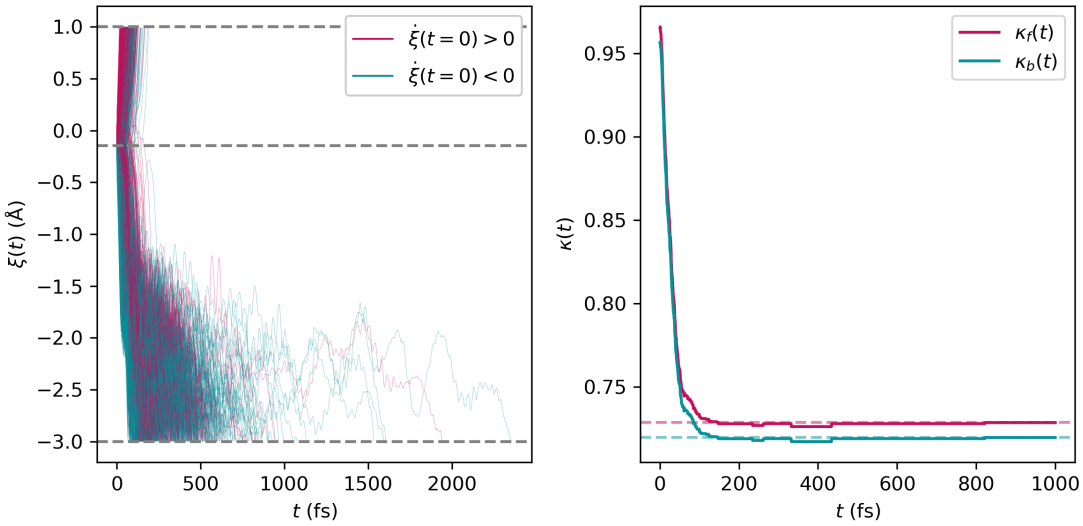

Figure S14: The reactive flux trajectories used to compute the transmission coefficient $\kappa(t)$ of the ethene methylation reaction, which is reported on the right for the forward and backward reactions. In this case, the backward reaction is ignored. Remark that $\kappa(t)$ should be the same for both directions and the observed deviations are due to the statistical noise. The trajectories are colored based on the initial velocity of the system along the CV ($\dot{\xi}(t = 0)$).

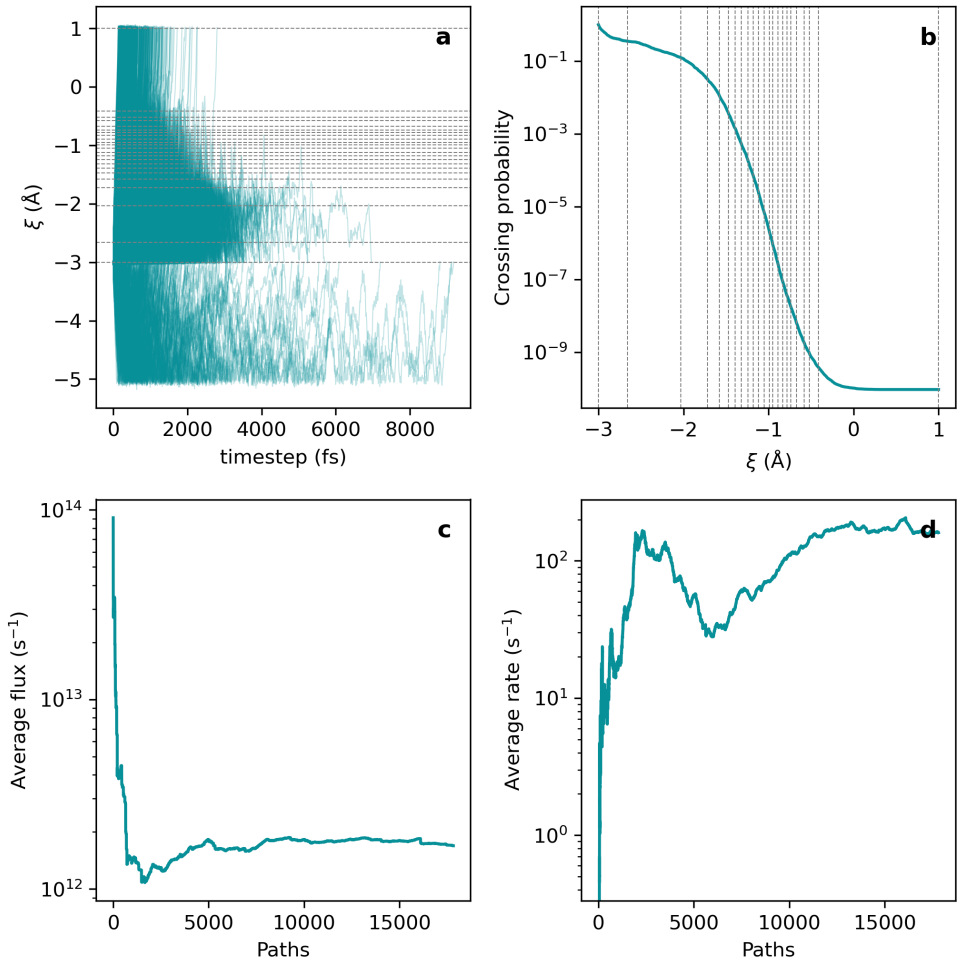

Figure S15: A summary of the ∞RETIS results for the ethene methylation reaction. **a**. The 17843 accepted paths (7500 initial paths are excluded as equilibration). **b**. The crossing probability as function of the order parameter. **c**. The running average of the flux through the ξ_0 surface. **d**. The running average of the predicted rate.

2.4 Ethene protonation

The protonation of ethene to a surface ethoxide species is studied in H-ZSM-5 at 573 K. US, REUS, and ∞RETIS are used to investigate the reaction. The number of RE swaps is decent but far from a complete shuffling of the replicas due to the hysteresis in the FES (Figure S16). We also report the CV histogram of the REUS simulation and the sampled values of the rate prefactor in the TS simulation (Figure S17). The transmission coefficient from the reactive flux simulation is shown in Figure S18, and some diagnostic metric from the ∞RETIS simulations are reported in Figure S19 and Figure S20.

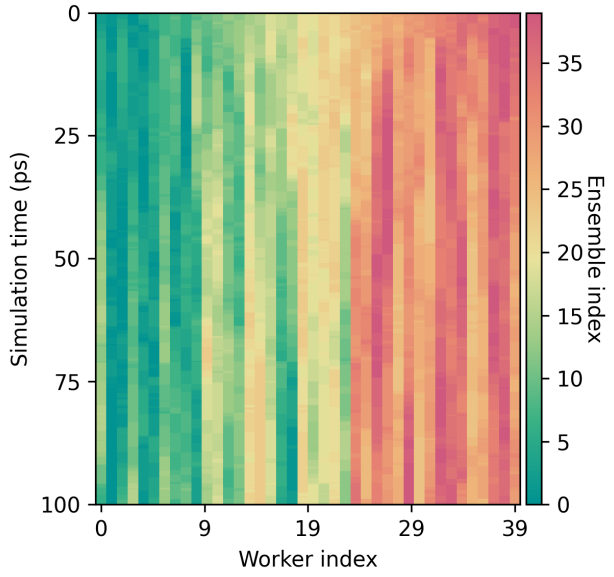

Figure S16: Schematic depiction of the swaps observed during the REUS simulation of ethene protonation. A *worker*, indexed on the horizontal axis, is a continuous MD simulation that can swap between *ensembles*, represented by different colors, here intended as umbrellas with a different center. So at the beginning walker i begins its journey from umbrella i but, in time, the swaps bring it to explore other umbrellas, shown by the change in color vertically, as time progresses.

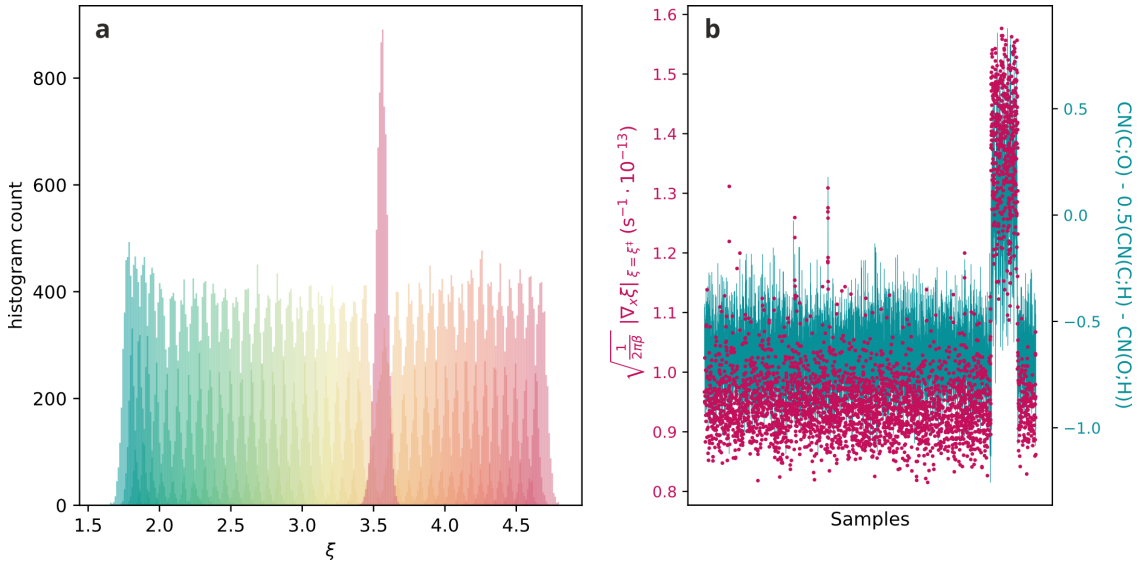

Figure S17: a. The CV histogram of the REUS simulation of the ethene protonation reaction. The larger, red histogram in the center corresponds to the 200 ps long TS umbrella used to evaluate the TST rate and to gather initial structures for RF. It can be noticed how, just by looking at the histogram, there is no evident sign of a pathological sampling deficiency. b. The rate prefactor $\frac{1}{2}|\dot{\xi}(\mathbf{r}_0)| = \sqrt{\frac{1}{2\pi\beta}}|\nabla_x \xi|$ sampled within a ± 0.03 Å interval around $\xi^{\ddagger}$. In the formula, ∇_x is the derivative with respect to the system's mass-weighted coordinates, which is obtained from $\dot{\xi}$ via analytical integration of the momenta. It can be seen how two clearly different regions are sampled. *A posteriori*, this behavior can be attributed to the simulation jumping briefly from one side to the other of the transition state, which is visible by plotting the values of the orthogonal CV to ξ (CN(C;O) $-$ 0.5(CN(C;H) $-$ CN(O;H))).

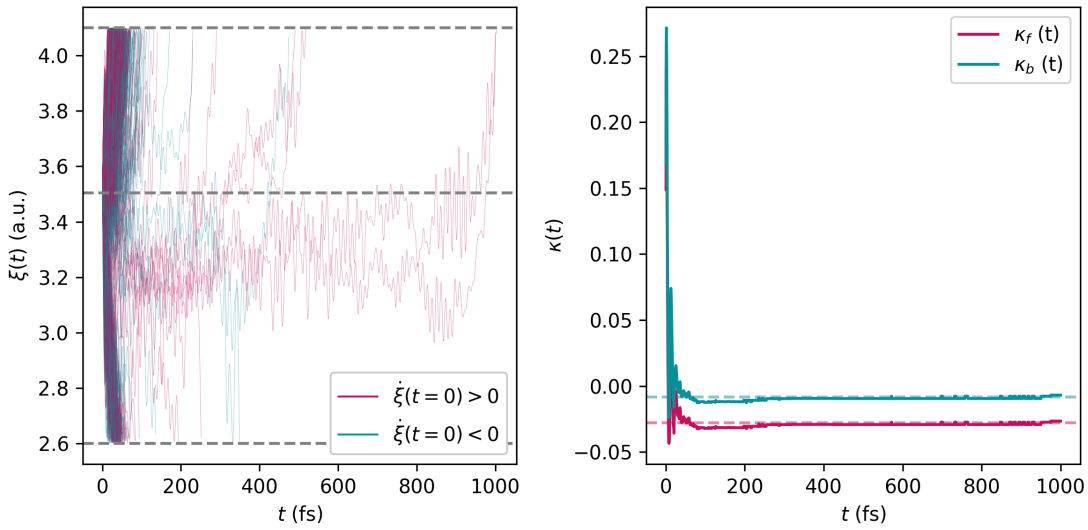

Figure S18: The reactive flux trajectories used to compute the transmission coefficient $\kappa(t)$ of the ethene protonation reaction, which is reported on the right for the forward and backward reactions. The trajectories are colored based on the initial velocity of the system along the CV ($\dot{\xi}(t = 0)$). Remark that $\kappa(t)$ should be the same for both directions and should always be greater than zero; the observed deviations are due to the statistical noise. The decay to near-to-zero values indicates a fundamental lack of correlation between the initial velocity of the system and the basin of attraction it ends up falling into—a symptom of hidden hysteresis in the free energy landscape.

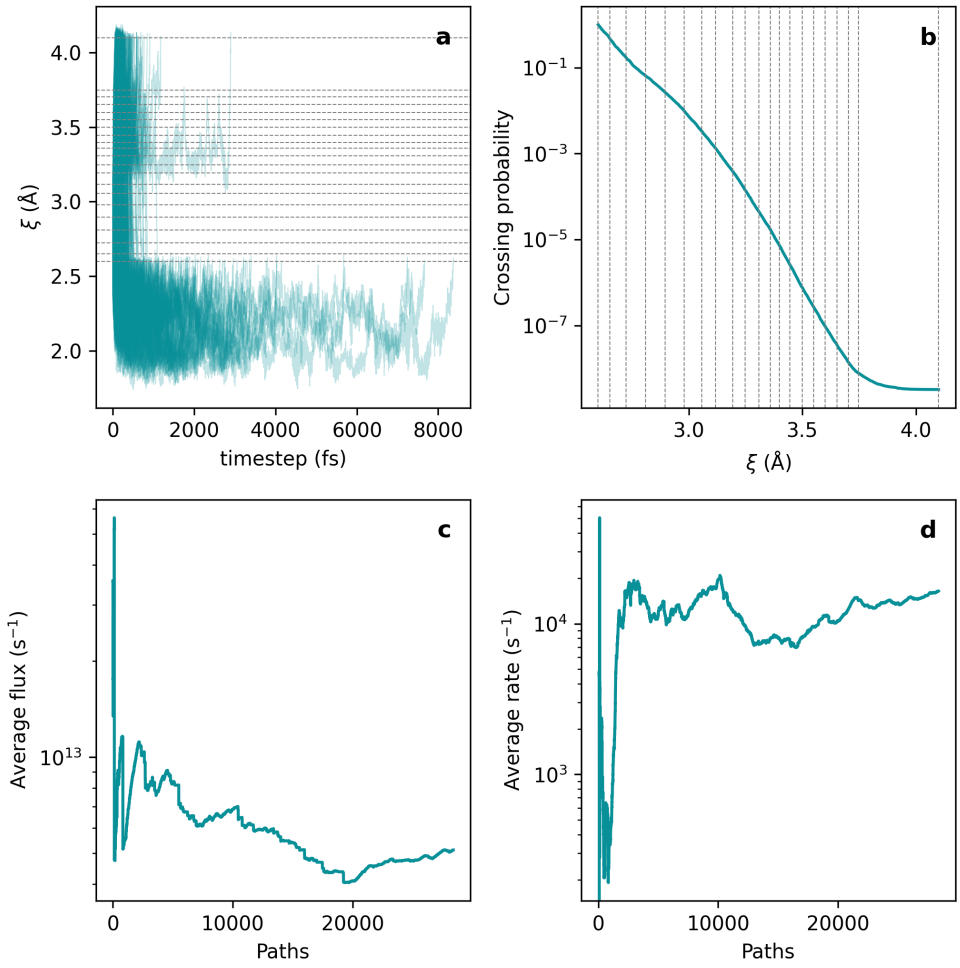

Figure S19: A summary of the ∞RETIS results for the forward ethene protonation reaction (ethene to surface ethoxide species). **a**. The 28420 accepted paths (20000 initial paths are excluded as equilibration). **b**. The crossing probability as function of the order parameter. **c**. The running average of the flux through the ξ_0 surface. **d**. The running average of the predicted rate.

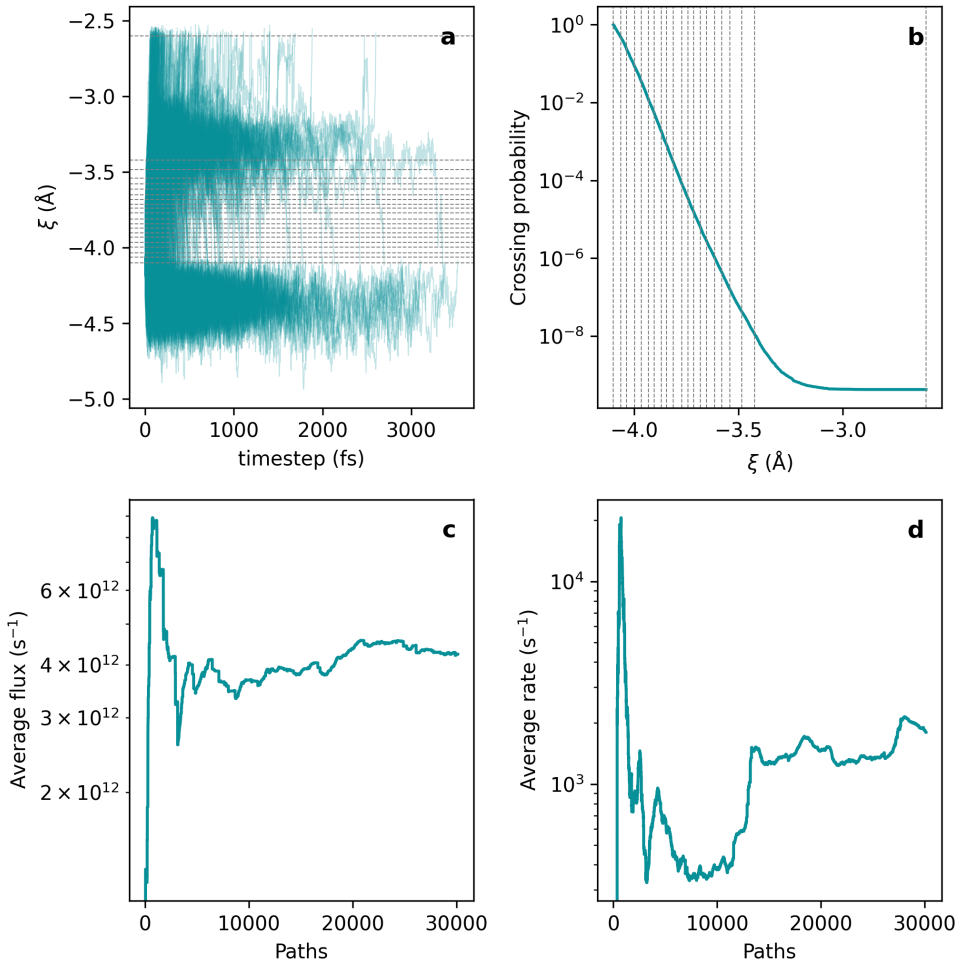

Figure S20: A summary of the ∞RETIS results for the backward ethene protonation reaction (surface ethoxide species to ethene). **a**. The 30147 accepted paths (20000 initial paths are excluded as equilibration). **b**. The crossing probability as function of the order parameter. **c**. The running average of the flux through the ξ_0 surface. **d**. The running average of the predicted rate.

References

(1) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Computer Physics Communications* **2014**, *185*, 604–613, https://doi.org/10.1016/j.cpc.2013.09.018.

(2) The PLUMED consortium Promoting transparency and reproducibility in enhanced molecular simulations. *Nature Methods* **2019**, *16*, 670–673, https://doi.org/10.1038/s41592-019-0506-8.

(3) Larsen, A. H.; Mortensen, J. J.; Blomqvist, J.; Castelli, I. E.; Christensen, R.; Dułak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C.; others The atomic simulation environment—a Python library for working with atoms. *Journal of Physics: Condensed Matter* **2017**, *29*, 273002, https://doi.org/10.1088/1361-648X/aa680e.

(4) Bailleul, S.; Dedecker, K.; Cnudde, P.; Vanduyfhuys, L.; Waroquier, M.; Van Speybroeck, V. Ab initio enhanced sampling kinetic study on MTO ethene methylation reaction. *Journal of Catalysis* **2020**, *388*, 38–51, https://doi.org/10.1016/j.jcat.2020.04.015.

(5) Bocus, M.; Vanduyfhuys, L.; De Proft, F.; Weckhuysen, B. M.; Van Speybroeck, V. Mechanistic characterization of zeolite-catalyzed aromatic electrophilic substitution at realistic operating conditions. *JACS Au* **2022**, *2*, 502–514, https://doi.org/10.1021/jacsau.1c00544.

(6) Bocus, M.; Vandenhaute, S.; Van Speybroeck, V. The Operando Nature of Isobutene Adsorbed in Zeolite H- SSZ- 13 Unraveled by Machine Learning Potentials Beyond DFT Accuracy. *Angewandte Chemie International Edition* **2025**, *64*, e202413637, https://doi.org/10.1002/anie.202413637.

(7) Litman, Y.; Kapil, V.; Feldman, Y. M.; Tisi, D.; Begušić, T.; Fidanyan, K.; Fraux, G.; Higer, J.; Kellner, M.; Li, T. E.; others i-PI 3.0: A flexible and efficient framework for advanced atomistic simulations. *The Journal of Chemical Physics* **2024**, *161*, https://doi.org/10.1063/5.0215869.

(8) Vervust, W.; Zhang, D. T.; Ghysels, A.; Roet, S.; van Erp, T. S.; Riccardi, E. PyRETIS 3: Conquering rare and slow events without boundaries. *Journal of Computational Chemistry* **2024**, *45*, 1224–1234, https://doi.org/10.1002/jcc.27319.

(9) Bocus, M.; Vanduyfhuys, L. ThermoLIB – A Python Library for Constructing and Post-Processing Free Energy Surfaces to Extract Thermodynamic and Kinetic Properties. 2026; https://arxiv.org/abs/2601.23071.

(10) Sauer, J. Ab initio calculations for molecule–surface interactions with chemical accuracy. *Accounts of chemical research* **2019**, *52*, 3502–3510, https://doi.org/10.1021/acs.accounts.9b00506.

(11) Plessow, P. N.; Studt, F. How accurately do approximate density functionals predict trends in acidic zeolite catalysis? *The Journal of Physical Chemistry Letters* **2020**, *11*, 4305–4310,

https://doi.org/10.1021/acs.jpclett.0c01240.

(12) van Erp, T. S. How far can we stretch the timescale with RETIS? *Europhysics Letters* **2023**, *143*, 30001, https://doi.org/10.1209/0295-5075/ace9f6.

(13) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics* **1994**, *101*, 10–1063, https://doi.org/10.1063/1.467468.

(14) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. Explicit reversible integrators for extended systems dynamics. *Molecular Physics* **1996**, *87*, 1117–1157, https://doi.org/10.1080/00268979600100761.