Nudged elastic band method in the CRYSTAL code: Theory and Applications

Nudged Elastic Band Method in the CRYSTAL code. Theory and Applications

Andreha Gelli,[1,2] Silvia Casassa,[1] Albert Rimola,[2] and Chiara Ribaldone[1]
[1)] *Dipartimento di Chimica, University of Torino, via Giuria 5, 10125, Torino, Italy*
[2)] *Department de Química, Universitat Autònoma de Barcelona, E-08193, Catalonia, Spain*

(*Electronic mail: chiara.ribaldone@unito.it)

The nudged elastic band (NEB) method is a widely used algorithm for determining minimum energy paths and transition states in chemical reactions and phase transitions. In this work, we present the implementation of different NEB algorithm schemes in the CRYSTAL code, a quantum mechanical ab initio program for the calculation of electronic properties of condensed matter systems, based on Hartree-Fock and density functional theory. The use of a set of localized Gaussian-type functions to expand the wavefunction permits a very efficient evaluation of the exact exchange series for the hybrid exchange-correlation functionals. Therefore, our implementation allows an accurate characterization of transition states in both molecular and condensed phase systems, at the hybrid functional level of theory. The theoretical framework, including the force projection scheme, tangent estimation, and optimization strategies, and the extensions of the method with climbing image and variable spring constant variants, is recalled. Then, the NEB algorithm is validated through a series of benchmark tests: two molecular reactions (a collinear proton transfer process and the keto-enol tautomerization in formamide) and a proton exchange process in a periodic chabazite zeolite. Our results are in excellent agreement with experimental and previous theoretical data, confirming the accuracy and applicability of the implementation. This work opens the possibility for future studies of complex reactive processes on extended periodic systems, using hybrid functionals.

I. INTRODUCTION

The calculation of energy barriers and transition states for processes such as surface-catalyzed heterogeneous reactions, ion trapping and migration within solid matrices, defect formation and propagation, as well as phase transitions, is a current challenge in condensed matter physics and theoretical chemistry. Despite the great advances in the computational community for the development of novel methods and algorithms to accelerate molecular dynamics simulations,[1–4] such phenomena are typically many orders of magnitude slower than vibrations of the atoms, making direct ab initio dynamical simulations an impractical way to calculate transition rates. At the same time, this separation of time scales allows the use of statistical approaches such as the harmonic approximation to transition state theory (HTST)[5,6] or Kramers theory.[7] Within these methods, the activation energy of a transition, representing the main factor that determines thermal stability, is defined as the energy difference between the local minimum of the potential energy surface corresponding to the initial state and the first-order saddle point along the path connecting the minima of the initial and final states. Among all the families of theoretical methods and algorithms that have been developed for tackling the problem of saddle points searching on potential energy surfaces, the nudged elastic band (NEB) approach, introduced by Jónsson and Henkelman,[8–13] has been established, since its early days, as one of the most efficient and accurate. Both its basic[8] and improved versions[10,11] are based on interpolation between the initial and final states of transition, converting the search for a first-order saddle point in configuration space to an optimization problem in a discretized path space. In this way, only the evaluation of the first order energy derivative is required, differently from surface walking approaches[14–16] or transition states localization techniques.[17] Thanks to its efficiency, the NEB method is widely used in conjunction with electronic structure calculations such as Hartree-Fock or density functional theory (DFT), which allows for an accurate evaluation of the forces among atoms, as well as a precise estimation of the potential energy surface (PES). In particular, in the framework of DFT, semilocal density functionals can often be insufficient to achieve a good description of electronic properties associated with the PES, especially for localized defects or systems with d and f orbitals. The over-delocalization of the electrons, due to the intrinsic self-interaction error in the semilocal functionals, as well as the associated underestimation of the band gap, can be mitigated by introducing hybrid density functionals, which incorporate a fraction of nonlocal exact exchange. Unfortunately, when studying extended systems, the evaluation of the exact exchange terms can become very demanding in terms of computational time, especially for plane waves-based codes. In the CRYSTAL code, the use of localized Gaussian functions as basis set for expanding the wavefunction has allowed the development of very efficient algorithms for the evaluation of the exact exchange series.[18,19] Therefore, the current implementation, combining the accuracy of the saddle-point searching technique via NEB with the efficiency of the program in evaluating the exchange series, can pave the way for a high-accurate description of transitions states at the DFT level, by means of hybrid density functionals.

The article is organized as follows. In Section II, the main theoretical background underlying the nudged elastic band method is briefly recalled and discussed. In Section III, with reference to selected test cases, (i) the reliability of the nudged elastic band method is assessed through comparisons with experimental data and (ii) its efficiency is analyzed in relation to other ab initio quantum mechanical programs. Finally, in Section V, the conclusions are outlined, together with some future perspectives.

II. THEORY

Although the theoretical framework and methodological aspects are well known and consolidated, a dedicated section to presenting the current implementation of the NEB method is presented, so that users can get the most out of it, having documented the meaning and effect of the few parameters to be set and the specific of the various options. For clarity, a flowchart of the algorithm is provided in Section 6 of the supplementary material.

Consider a given nuclear configuration as a collection of N atoms in a d dimensional space, described by a Nd dimensional vector $\boldsymbol{q}_i \in \mathbb{R}^{Nd}$ where each component of the vector identifies a position of an atom in a given spatial direction. Using the concept of PES, a map can be established between each nuclear configuration $\boldsymbol{q}_i$ and its ground state energy $E(\boldsymbol{q}_i)$, computed using Hartree-Fock or DFT.
Starting from an initial state $\boldsymbol{q}_1$, the NEB method defines a scheme to find the minimum energy path connecting the initial state $\boldsymbol{q}_1$ with a specific final state $\boldsymbol{q}_m$, by sampling the PES with a set of $(m-2)$ intermediate nuclear configurations. The total set of m configurations thus defined, called images, can be indicated as $\boldsymbol{q} = [\boldsymbol{q}_1, \boldsymbol{q}_2, ..., \boldsymbol{q}_m]$, where the endpoints $\boldsymbol{q}_1$ and $\boldsymbol{q}_m$ are fixed and correspond to two local minima of the PES. The positions of the $(m-2)$ intermediate images, which ultimately identify the minimum energy path (MEP), are found by minimizing the objective function

$$S(\boldsymbol{q}) = \sum_{i=1}^{m} E(\boldsymbol{q}_i) + \frac{1}{2}\sum_{i=2}^{m} k_{i(i-1)}\,(\boldsymbol{q}_i - \boldsymbol{q}_{i-1})^2 \qquad (1)$$

This function represents a physical system consisting of an elastic chain formed by m objects, each connected to its neighbor by a spring with elastic constant $k_{i(i-1)}$, with the endpoints $\boldsymbol{q}_1$ and $\boldsymbol{q}_m$ held fixed. Therefore, the first term in equation (1) is the sum of the ground state energies of each nuclear configuration, while the second term is the elastic potential energy due to the presence of fictitious springs connecting the images.

A. Generation of the initial path

Firstly, the question arises of how to define the initial guess for the intermediate nuclear configurations $[\boldsymbol{q}_2, ..., \boldsymbol{q}_{m-1}]$ which enter in the objective function (1) and represent the initial path between the endpoints in the PES. A simple yet effective solution is to initialize the intermediate images by means of a linear interpolation between the initial $\boldsymbol{q}_1$ and final $\boldsymbol{q}_m$ nuclear configurations. For completeness, the adopted algorithm is detailed in Section 1 of the supplementary material.

B. Forces acting on the intermediate images

Once the initial images are set, the minimization of the objective function (1), that is, the search for its stationary points, provides the equation of motion for the displacement of each image towards the MEP. Taking the derivative of equation (1) with respect to a given nuclear configuration, the form of the force acting on each nuclei in that image can be devised as

$$\bar{\mathbf{F}}_i^{\mathrm{NEB}} = -\frac{\partial S(\boldsymbol{q})}{\partial \boldsymbol{q}_i} = \mathbf{F}_i + \mathbf{F}_i^s \qquad (2)$$

Following this prescription, the nuclear forces (2) will include two terms, one related to the real nuclear forces $\mathbf{F}_i = -\nabla E(\boldsymbol{q}_i)$, and the other one associated to the elastic forces $\mathbf{F}_i^s$ due to the spring interactions. Hence, the minimization of the total forces (2) acting on each image leads the initial chain of images to converge to the MEP, whose first-order saddle points are assigned with transition states, corresponding to the highest energy points on the MEP connecting the initial $\boldsymbol{q}_1$ and final $\boldsymbol{q}_m$ nuclear configurations. However, this formulation implies two well-known problems: (*i*) the component of the real forces parallel to the path causes the images to move away from high energy regions towards the energy minima, thereby decreasing the density of the images in the neighborhood of the saddle points; (*ii*) the component of the elastic forces perpendicular to the path tends to prevent the band from following a curved trajectory, thus avoiding the retrieval of the true saddle point. The first issue, called the *sliding down* problem, becomes relevant for small values of the elastic constant, and can be avoided considering only the component $(\mathbf{F}_i)|_\perp$ of the real forces perpendicular to the path. The second challenge, known as *corner cutting* effect, mostly arises for high values of the elastic constant, and can be overcome by taking into account only the component $(\mathbf{F}_i^s)|_\parallel$ of the elastic forces parallel to the path, thus leading to the following redefinition of the total forces acting on the i-th image

$$\begin{aligned}\mathbf{F}_i^{\mathrm{NEB}} &= (\mathbf{F}_i)|_\perp \;+\; (\mathbf{F}_i^s)|_\parallel \\ &= -\nabla E(\boldsymbol{q}_i)|_\perp \;+\; (\mathbf{F}_i^s)|_\parallel \end{aligned} \qquad i = 2, ..., (m-1) \qquad (3)$$

In order to evaluate the components of the forces with respect to the path, a unitary local tangent vector $\hat{\boldsymbol{\tau}}_i$ has to be defined for each i-th image, so that the real and elastic forces can be projected according to equation (3). In this way, the component of the real forces perpendicular to the path can be obtained by subtracting the component parallel to the path to the total real forces vector, that is

$$\begin{aligned}(\mathbf{F}_i)|_\perp &= \mathbf{F}_i - [\mathbf{F}_i \cdot \hat{\boldsymbol{\tau}}_i]\,\hat{\boldsymbol{\tau}}_i \\ &= -\nabla E(\boldsymbol{q}_i) + [\nabla E(\boldsymbol{q}_i) \cdot \hat{\boldsymbol{\tau}}_i]\,\hat{\boldsymbol{\tau}}_i \end{aligned} \qquad (4)$$

while the component of the elastic forces parallel to the path can be computed using the well-known formula[10]

$$(\mathbf{F}_i^s)|_\parallel \;=\; k(\|\boldsymbol{q}_{i+1} - \boldsymbol{q}_i\| - \|\boldsymbol{q}_i - \boldsymbol{q}_{i-1}\|)\,\hat{\boldsymbol{\tau}}_i \qquad (5)$$

Applying equations (4) and (5) for $i = 2, ..., (m-1)$, that is, for all the intermediate images, and then using the formula (3), the total nudged elastic band forces can be evaluated and minimized, thus finding the MEP.

C. Estimation of the tangent to the path

Initially, the calculation of $\hat{\tau}_i$ was based on a simple estimation of the unit tangent at the i-th image as the normalized line segment between two configurations adjacent the i-th image.[8] Later, the evaluation of this quantity was made more effective, leading to better localization of the saddle points and avoiding the formation of kinks in the path, which cause instabilities and slow down convergence behavior, by defining it in relation to the energy of the first closest images, in the following way[10]

$$\begin{aligned}&\boldsymbol{\tau}_i = \begin{cases} \boldsymbol{\tau}_i^+ & \text{if } E_{i+i} > E_i > E_{i-1} \\ \boldsymbol{\tau}_i^- & \text{if } E_{i+i} < E_i < E_{i-1} \end{cases} \qquad i = 2, ..., (m-1) \\ &\text{with} \quad \boldsymbol{\tau}_i^+ = \boldsymbol{q}_{i+1} - \boldsymbol{q}_i \quad \text{and} \quad \boldsymbol{\tau}_i^- = \boldsymbol{q}_i - \boldsymbol{q}_{i-1} \end{aligned} \qquad (6)$$

where $E_i \equiv E(\boldsymbol{q}_i)$ is the ground state energy of the i-th image. The previous equation (6) describes the case where adjacent images have monotonically increasing or

decreasing energies. For cases where the i-th image is at a minimum ($E_{i+1} > E_i < E_{i-1}$) or at a maximum ($E_{i+1} < E_i > E_{i-1}$), then the tangent is estimated to be a linear combination of the vectors defined in equation (6), leading to the expression[10]

$$\boldsymbol{\tau}_i = \begin{cases} \boldsymbol{\tau}_i^+ \Delta E_i^{max} + \boldsymbol{\tau}_i^- \Delta E_i^{min} & \text{if } E_{i+i} > E_{i-1} \\ \boldsymbol{\tau}_i^+ \Delta E_i^{min} + \boldsymbol{\tau}_i^- \Delta E_i^{max} & \text{if } E_{i+i} < E_{i-1} \end{cases} \tag{7}$$

where the weight coefficients are given by

$$\begin{aligned} \Delta E_i^{max} &= \max(|E_{i+1} - E_i|, |E_{i-1} - E_i|) \\ \Delta E_i^{min} &= \min(|E_{i+1} - E_i|, |E_{i-1} - E_i|) \end{aligned} \tag{8}$$

for $i = 2, ..., (m-1)$. This definition only plays a role at stationary points along the path, and it serves to smoothly switch between the two possible tangents $\boldsymbol{\tau}_i^+$ and $\boldsymbol{\tau}_i^-$ defined in equation (6). Otherwise, when one image becomes higher in energy than another, the tangent changes abruptly, which can result in convergence issues.
Finally, the tangent vector associated with each image needs to be normalized, namely

$$\hat{\boldsymbol{\tau}}_i = \frac{\boldsymbol{\tau}_i}{\|\boldsymbol{\tau}_i\|} \qquad i = 2, ..., (m-1) \tag{9}$$

With this modified tangent, the elastic band is well behaved and it converges reliably to the MEP, provided that a sufficient number of images are included in the path. In the cases of images with degenerate energies, the simple definition of the tangent as the normalized line segment connecting the previous and the next images to a given i-th image in the chain is recovered. Moreover, if the tangent vector has to be estimated also for the initial and the final fixed nuclear configurations, the previous definition becomes ambiguous, and a specific choice has to be done. All technical aspects are reported in Section 2 of the supplementary material.

D. Path minimization methods

At each step of the NEB method, after the calculation of the tangent at each image, the total forces vector can be computed, by means of equation (3), for each image along the path. At this point, the total forces vector has to be minimized for all the intermediate images. Two main minimization methods have been implemented in the CRYSTAL code to find the MEP, starting from the initial nudged elastic band path: (*i*) the steepest descent (SD), and (*ii*) the conjugate gradient (CG) optimization methods. Details about the equations involved and their implementation are reported in Section 3 of the supplementary material, for the three considered path minimization methods.
To consider the minimization process convergent towards the minimum energy path, two criteria must be satisfied for each intermediate image, both associated with the real nuclear force vector orthogonal to the path: (*i*) its Euclidean norm and (*ii*) the root-mean-square of its Cartesian components must be less than given tolerances

$$\begin{aligned} &(i)\ \|(\mathbf{F}_i)|_\perp\| < \Omega_n \\ &(ii)\ \text{RMS}\left[(\mathbf{F}_i)|_\perp\right] < \Omega_r \end{aligned} \qquad \forall\ i = 2, ..., (m-1) \tag{10}$$

These requirements are based on the fact that, for an image located in the energy minimum, the real nuclear force component perpendicular to the path vanishes. The default values for the two tolerances in (10) are $\Omega_n = 0.05$ eV/Å $= 0.001$ Ha/Bohr and $\Omega_r = \Omega_n/2$.

E. Beyond regular nudged elastic band method

The main goal of the NEB method is the search for transition states, corresponding to saddle points in the PES. However, there are cases in which the resolution of the MEP near the saddle point is poor and the estimate of the activation energy obtained from the interpolation is subject to large uncertainty. Aiming at a better detection and characterization of the saddle points, two main variants of the basic NEB have been developed,[11] as described below.

1. *Climbing Image method*

The climbing image approach constitutes a small modification to the basic NEB scheme.[11] After a few iterations with the regular NEB, the image i_{max} with the highest energy is identified. The force on this image is then computed using a modified version of equation (3), given by

$$\begin{aligned} \mathbf{F}_i^{\text{NEB}} &= (\mathbf{F}_i)|_\perp - (\mathbf{F}_i)|_\parallel \\ &= \mathbf{F}_i - 2(\mathbf{F}_i \cdot \hat{\boldsymbol{\tau}}_i)\, \hat{\boldsymbol{\tau}}_i \end{aligned} \qquad \text{for} \quad i = i_{max} \tag{11}$$

Equation (11) represents the full force due to the real potential, but with the component along the elastic band inverted. By using this force expression for the maximum energy image, it becomes completely unaffected by the spring forces, so that the spacing between adjacent images will be different on each side of the climbing image. As it moves towards the saddle point, the images on one side will become compressed, while those on the other side will spread out. Qualitatively, the climbing image moves up the potential energy surface along the elastic band and down the potential surface perpendicular to the band. The only challenge remained at this point is ensuring there are enough images near the climbing image to obtain an accurate estimate of the reaction coordinate, since this determines the climbing direction.

2. *Variable Spring Constants method*

Since the saddle point is the most important point along the MEP, it is preferable to have more resolution in the neighborhood of the saddle point than near the endpoints. Indeed, as the images are moved closer to the saddle point, the approximation of the tangent will become more accurate. In such cases, it is more efficient to distribute the images unevenly along the path. This can be accomplished by using higher values of the elastic constants near the saddle point. On the base of Ref.[11], a scheme can be used where the spring constant depends linearly on the energy of the images, in such a way that images with low energy get connected by a weaker spring constant. This can be accomplished using the formula

$$k_i = \begin{cases} k_{\max} - \Delta k \left(\dfrac{E_{\max} - E_{ik}}{E_{\max} - E_{\text{ref}}} \right) & \text{if } E_{ik} > E_{\text{ref}} \\ k_{\max} - \Delta k = k_{\min} & \text{if } E_{ik} \leq E_{\text{ref}} \end{cases} \tag{12}$$

with $\Delta k = k_{\max} - k_{\min}$ and for $i = 1, ..., m$. The energy E_{ik} is defined as

$$E_{ik} = \begin{cases} \max\{E_i, E_{i+1}\} & \text{for} \quad i = 1, ..., m-1 \\ \max\{E_i, E_{i-1}\} & \text{for} \quad i = m \end{cases} \tag{13}$$

and it represents the highest energy between the two images connected by the i-th spring, while $E_{\max}$ is the maximum value of E_i for the whole elastic band. Finally, $E_{\rm ref}$ is a reference value for the energy, defining a minimum value of the spring constant. In Ref.[11], $E_{\rm ref}$ is taken to be the energy of the highest energy endpoint in the path, that is

$$E_{\rm ref} = \max\{E_{\rm IS}, E_{\rm FS}\} \tag{14}$$

where $E_{\rm IS}$ and $E_{\rm FS}$ are the energies of the initial and final states, respectively. This choice ensures that the density of images is roughly equal near the two endpoints, even for highly asymmetric paths. The spring constant is, therefore, linearly scaled from a maximum value of $k_{\max}$ for highest energy images to a minimum value of $k_{\max} - \Delta k = k_{\min}$ for the images with energy equal to $E_{\rm ref}$ or lower.
Using the variable spring constants method, the parallel component of the spring forces for each image is defined in a slightly different way with respect to equation (5), namely

$$(\mathbf{F}_i^s)|_{\parallel} = (k_i\|\boldsymbol{q}_{i+1} - \boldsymbol{q}_i\| - k_{i-1}\|\boldsymbol{q}_i - \boldsymbol{q}_{i-1}\|)\,\hat{\boldsymbol{\tau}}_i \tag{15}$$

for $i = 2, ..., (m-1)$, so that a unique value of the elastic constant is assigned to each spring that describes the interaction among images. In equation (15), the value of the elastic constant associated with the spring connecting the i-th and the $(i+1)$-th images only depends on the value k_i of the elastic constant conferred to the i-th image, estimated by means of equation (12).

F. Minimum energy path interpolation

The analysis of the results of a NEB calculation requires an interpolation between the images, in order to get estimates of the coordinates of atoms and the energy at maxima and minima along the MEP. For a good interpolation scheme, it is better to include the components of the real forces parallel to the path in the definition of the interpolation function, that can lead to the identification of extrema in the MEP, which could not be seen if only the energy data are taken into account.[10]
In the CRYSTAL code, a cubic polynomial to represent the MEP between each pair of adjacent images is used. The polynomial describing the curve between two points (images) in the range $[\boldsymbol{q}_i, \boldsymbol{q}_{i+1}]$, being $\boldsymbol{q}_i$ the starting point (image) and $\boldsymbol{q}_{i+1}$ the final point (image) of the i-th line, can thus be written as

$$p(\boldsymbol{q}) = a_i\,\|\boldsymbol{q} - \boldsymbol{q}_i\|^3 + b_i\,\|\boldsymbol{q} - \boldsymbol{q}_i\|^2 + c_i\,\|\boldsymbol{q} - \boldsymbol{q}_i\| + d_i \tag{16}$$

valid for all points $\boldsymbol{q} \in [\boldsymbol{q}_i, \boldsymbol{q}_{i+1}]$. The equations describing the coefficients of the polynomial in equation (16) can be derived by imposing four conditions for the matching of both the energy and the force at the two endpoints of the interval. That is, given a starting point $\boldsymbol{q}_i$ with ground state electronic energy $E_i \equiv E(\boldsymbol{q}_i)$ and real forces component parallel to the path provided by the projection

$$F_i \equiv \mathbf{F}_i \cdot \hat{\boldsymbol{\tau}}_i = -\nabla E(\boldsymbol{q}_i) \cdot \hat{\boldsymbol{\tau}}_i \tag{17}$$

as well as a endpoint $\boldsymbol{q}_{i+1}$ with ground state electronic energy $E_{i+1} \equiv E(\boldsymbol{q}_{i+1})$ and component of the real forces parallel to the path defined following equation (17), the four conditions are

$$p(\boldsymbol{q}_i) = E_i \qquad\qquad p(\boldsymbol{q}_{i+1}) = E_{i+1} \tag{18}$$

$$[\nabla p(\boldsymbol{q})]\big|_{\boldsymbol{q}=\boldsymbol{q}_i} = -F_i \qquad [\nabla p(\boldsymbol{q})]\big|_{\boldsymbol{q}=\boldsymbol{q}_{i+1}} = -F_{i+1} \tag{19}$$

By imposing these conditions on the cubic polynomial (16), the four parameters for the interpolation curve between two adjacent images have the form

$$a_i = \frac{2(E_i - E_{i+1})}{\|\boldsymbol{q}_{i+1} - \boldsymbol{q}_i\|^3} - \frac{F_i + F_{i+1}}{\|\boldsymbol{q}_{i+1} - \boldsymbol{q}_i\|^2} \tag{20}$$

$$b_i = \frac{3(E_{i+1} - E_i)}{\|\boldsymbol{q}_{i+1} - \boldsymbol{q}_i\|^2} + \frac{2F_i + F_{i+1}}{\|\boldsymbol{q}_{i+1} - \boldsymbol{q}_i\|} \tag{21}$$

$$c_i = -F_i \tag{22}$$

$$d_i = E_i \tag{23}$$

The cubic polynomial interpolation discussed above is typically quite smooth.[10] It is worth noting that the cubic interpolation introduced above does not guarantee the continuity of second derivatives at the images. However, adopting interpolation polynomial of higher orders, although ensuring continuous second derivatives, can add spurious oscillations between couples of images.[10] Thus, if a continuous second derivative is not essential, the straightforward cubic interpolation outlined above is likely the preferred choice.

III. RESULTS AND DISCUSSION

To assess the accuracy and the general applicability of the NEB implementation in the CRYSTAL code, a set of test cases covering different chemical environments and reaction types were selected. These include elementary gas-phase reactions and more complex processes, relevant in solid-state chemistry. In particular, we examined two molecular reactions, namely, the collinear proton transfer reaction,

$$\mathrm{H_2 + H \rightarrow H + H_2} \tag{R1}$$

and the keto-enol tautomerization of formamide,

$$\mathrm{HC(=O)NH_2 \rightarrow HC(\text{-}OH)NH} \tag{R2}$$

These tests are discussed in Section III A and provide a well-balanced set of systems in terms of complexity and chemical diversity, allowing for a comprehensive validation of the NEB algorithms in the molecular regime. In addition, a bulk process involving a proton transfer between two oxygen atoms in a chabazite zeolite was simulated to test the method in a fully periodic condensed phase environment. The results and the analysis of this process are discussed in Section III B. Together, these tests aim to demonstrate the flexibility and reliability of the NEB algorithm in the CRYSTAL code across different length scales and chemical contexts.

A. Gas Phase reactions

1. *The collinear proton transfer reaction* (R1)

The collinear proton transfer reaction (R1) in the framework of Hartree-Fock and DFT has been extensively studied by B. S. Jursic,[20,21] using GAUSSIAN 94 computational

package.[22] In particular, using a cc-pVQZ basis set, the activation barrier for the reaction (R1) has found to be equal to 16.8 kcal/mol ≈ 0.7285 eV and 2.4 kcal/mol ≈ 0.1041 eV, respectively using Hartree-Fock Hamiltonian and BLYP exchange-correlation functional.[21] The experimental activation barrier used in literature to be compared with theoretical results has been computed by W. R. Schulz *et al.*[23] from reaction rate versus temperature data over a temperature range of about (330, 444) K,[23] for which the simple Arrhenius form was suitable and for which reaction rates were available, and it has a value of 9.70 kcal/mol ≈ 0.4206 eV.[23] However, it has been observed[24] that the experimentally deduced activation barriers depend on the temperature range used for the Arrhenius fit and that this complicates a direct comparison with reaction barriers computed quantum mechanically. Indeed, the previously mentioned ab initio calculations by B. S. Jursic,[20,21] together with more recent theoretical results,[24,25] confirm the finding that neither hybrid nor gradient-corrected density functionals methods are capable of reproducing the experimental reaction barrier found by W. R. Schulz *et al.*,[23] and only a coupled-cluster theory with single, double, and perturbative triple substitutions [CCSD(T)] approach,[26] can reach activation energy values comparable to the experimental one.[24,25]

In this work, avoiding the debated and controversial comparison with experimental data, for the sake of testing our implementation against other theoretical methods and quantum mechanical packages, we employed unrestricted Hartree-Fock (UHF) Hamiltonian, using an all electron basis set consisting of 6 atomic orbitals for each hydrogen atom,[27] and we compare our results with those obtained using the ORCA code,[28] within the same computational setup and NEB methods. The total number of images along the path is taken equal to $m = 9$, including the initial and final states of the reaction. The initial and final states of the reaction are shifted with respect to the corresponding center of mass of the system, before starting the NEB calculation. The nuclear positions for the initial, final and intermediate states at each NEB step for all the calculations listed in Table I are made available through a Zenodo repository.[29] The initial reaction path is obtained through linear interpolation between the Cartesian nuclear coordinates of the atoms in the initial and final states, while the minimization of the NEB path towards the MEP is analyzed using two algorithms, namely, steepest descent (SD) method, with an optimization step equal to $ds = 1.0$ Bohr^2/Ha, and the conjugate gradient (CG) technique, with $ds = 0.8$ Bohr^2/Ha. The calibration of the optimization step for the path minimization methods is detailed in Section 8 and Table 3 of the supplementary material.

The results obtained using the different NEB techniques, in conjunction with the three procedures for path optimization, are listed in Table I, together with the corresponding results computed with the ORCA code.[28] The energy barriers, at the UHF level, are in striking agreement with the ones found by the ORCA code, as well as with the other theoretical references. For all the calculations, the vibrational frequencies of the transition state structure have one imaginary eigenvalue, whose value is reported in Table I, thus confirming the detection of a saddle point in the PES.

Examples of the symmetric form of the MEP computed with the basic, climbing image (CI-NEB) and variable spring constants (VARK-NEB) methods, using the steepest descent minimization algorithm, are reported in Figure 1, superimposed to the corresponding minimum energy paths obtained using the same computational setup within the ORCA code, showing a perfect agreement between the results of the two programs.

Reaction (R1) with $m = 9$ images - UHF method

Method		n_{opt}	E_a [eV]	ν [cm^{-1}]
Basic NEB	SD	58	0.721139	2267.5720 i
	CG	76	0.721134	2266.9278 i
	ORCA-LBFGS	20	0.721145	2231.0023 i
CI-NEB	SD	58	0.721139	2267.5719 i
	CG	76	0.721134	2266.9285 i
	ORCA-LBFGS	20	0.721145	2246.6890 i
VARK-NEB	SD	58	0.721139	2267.5727 i
	CG	76	0.721134	2266.9284 i
	ORCA-LBFGS	27	0.721145	2217.8591 i
B. S. Jursic[21]			0.72852	2476 i
B. G. Johnson *et al.*[25]			0.721145	2296 i

TABLE I. Number of optimization steps n_{opt} required to find the MEP, activation energy E_a and imaginary frequency ν of the transition state. Different NEB methods (see Section II) and various optimization schemes (see Section II D) are compared with the results obtained with the ORCA code,[28] adopting the same computational setup. The optimization steps in the SD and CG path minimization algorithms are reported in the main text of Section III A 1.

2. *The keto-enol tautomerization of formamide* (R2)

The tautomerization reaction (R2), describing the proton transfer from the hydroxyl group to the amino nitrogen in formimidic acid to produce the amide form of formamide, has been extensively investigated in literature by several first principle electronic structure studies. Among these, the pioneering work by Schlegel *et al.*[30] already highlighted the energetic preference of the amide form over the enolic tautomer, with an estimated tautomerization enthalpy of about 12 kcal/mol ≈ 0.52 eV, including zero-point energy corrections and MP2-level electron correlation. Later, more refined benchmark calculations conducted by G. Fogarasi[31] confirmed this result, reporting a best-estimate energy difference of about 10.6 kcal/mol (0.46 eV) between the two forms at the CCSD(T)//MP2/aug-cc-pVTZ level of theory. C. Adamo and coworkers[32] further analyzed this reaction using B3LYP functional and post-Hartree-Fock methods, finding a barrier for direct intramolecular proton transfer in gas phase equal to 46.2 kcal/mol ≈ 2.00 eV when computed at the B3LYP/6-31G(d,p) level, the same value also computed by G. Fogarasi[31] using the same level of theory. In the present work, we investigate reaction (R2) using the B3LYP exchange-correlation functional and an all-electron basis set consisting of 57 atomic orbitals (see Section IV for further details). The number of images for the NEB calculations was set equal to $m = 9$, including the initial and final states of the reaction. The minimization of the NEB path towards the MEP is computed using two different algorithms, namely, the steepest descent (SD) method, with an optimization step equal to $ds = 1.0$ Bohr^2/Ha, and the conjugate gradient (CG) technique, with $ds = 0.8$

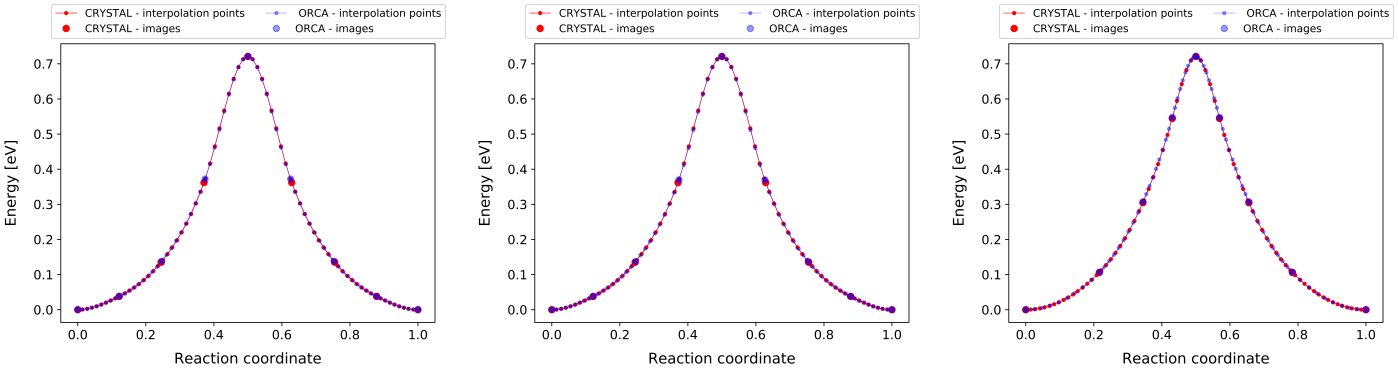

FIG. 1. Minimum energy paths obtained with the basic NEB (left panel), the climbing image (CI-NEB, central panel) and the variable spring constants (VARK-NEB, right panel) methods, using the steepest descent (SD) algorithm with an optimization step of $ds = 1.0$ Bohr^2/Ha to converge the reaction path towards the minimum energy path. Details about the activation barrier, as well as about the vibrational frequencies for the transition state configurations, are reported in Table I. The interpolation of the minimum energy paths is done using 72 points. The method used for the interpolation is described in Section II F.

Bohr^2/Ha. For a detailed discussion on the computational setup, refer to (*i*) Section 8 and Table 4 of the supplementary material for the choice of the optimization step for the path minimization procedures, as well as to (*ii*) Section IV for the climbing image and the variable spring constants methods.

The nuclear positions for the initial, final and intermediate states at each NEB step for all the calculations listed in Table II are made available through a Zenodo repository.[29]

The bond lengths of the optimized structure for the reactant and the product of the reaction (R2), used as initial and final states in the NEB calculations, as well as the geometry configuration lengths of the transition states obtained through the NEB methods listed in Table II, are reported in Table 2 of the supplementary material, together with the experimental values[33] for the initial and final geometry structures, and the theoretical results obtained by C. Adamo and coworkers.[32]

A picture of the initial, final, and transition state configurations, the latter obtained by means of the CI-NEB method in conjunction with the steepest descent path minimization, is reported in Figure 2.

The computed activation barriers, summarized in Table II, range from 1.95 eV to 2.00 eV (depending on the NEB method and the reaction path minimization scheme employed), which are in excellent agreement with the theoretical value of 46.2 kcal/mol $\approx$ 2.00 eV reported by C. Adamo *et al.*[32] as well as by G. Fogarasi,[31] obtained using B3LYP functional in conjunction with the 6-31G(d,p) basis set. Furthermore, the calculated tautomerization energy obtained from the energy difference between the reactant (formamide) and product (formamidic acid) states amounts to $\Delta E_r \approx 0.55$ eV, in good agreement with the best estimates from previous studies.[30,31] The transition state has been characterized by a vibrational frequency analysis, identifying the presence of a single imaginary eigenvalue, with values (reported in Table II) consistent with a proton transfer phenomenon, thus confirming the identification of a first order saddle point. Moreover, our calculated values are in good agreement with the imaginary frequency of $1894\,i$ cm^{-1} reported by G. Fogarasi,[31] calculated at the MP2/cc-pVTZ level of theory. The MEPs obtained using the three NEB methods (i.e., basic, CI-NEB and VARK-NEB), in conjunction with the steepest descent path optimization, are reported in Figure 3.

As in the collinear proton transfer case of Section III A 1, we observe that the three NEB variants, combined with the selected minimization schemes, lead to consistent results, and all the transition state geometries exhibit a well-defined imaginary frequency associated with the reaction coordinate.

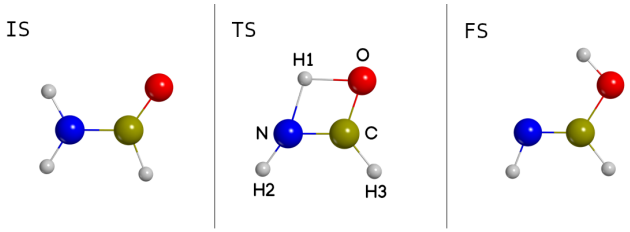

FIG. 2. Initial (IS), transition (TS) and final (FS) states for the keto-enol tautomerization of formamide reaction. Dark olive, blue, red and white spheres represent carbon, nitrogen, oxygen and hydrogen atoms, respectively. The transition state is computed using the CI-NEB method in conjunction with the SD path minimization scheme. The associated MEP is reported in the central panel of Figure 3.

Reaction (R2) with $m = 9$ images - B3LYP functional

Method		n_{opt}	E_a [eV]	ν [cm^{-1}]
Basic NEB	SD	117	1.949339	$1825.8207\,i$
	CG	149	1.949700	$1826.4288\,i$
CI-NEB	SD	111	2.000514	$1905.1026\,i$
	CG	143	2.000514	$1905.1072\,i$
VARK-NEB	SD	75	1.986986	$1870.2894\,i$
	CG	98	1.986956	$1870.1877\,i$
C. Adamo *et al.*[32]			2.00	–
G. Fogarasi[31]			2.00	–

TABLE II. Number of optimization steps n_{opt} required to find the MEP, activation energy E_a and imaginary frequency ν of the transition state, resultant from different NEB methods (see Section II) and various optimization schemes (see Section II D) for the search of the minimum energy path. The optimization steps in the SD and CG path minimization algorithms are reported in the main text of Section III A 2.

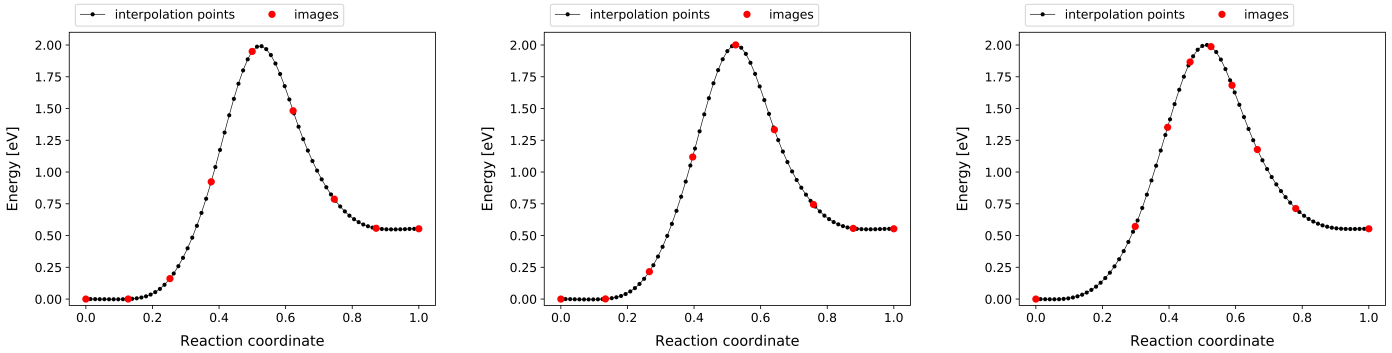

FIG. 3. Minimum energy paths obtained with the basic NEB (left panel), the climbing image (CI-NEB, central panel) and the variable spring constants (VARK-NEB, right panel) methods, using the steepest descent (SD) algorithm with an optimization step of $ds = 1.0$ Bohr^2/Ha to converge the reaction path towards the minimum energy path. Details about the activation barrier, as well as about the vibrational frequencies for the transition state configurations, are reported in Table II. The interpolation of the minimum energy paths is done using 72 points. The method used for the interpolation is described in Section II F.

B. Proton exchange process in chabazite bulk

As a validation test for the implementation of the NEB method in the CRYSTAL code in three-dimensional periodic systems, the proton jump between two Brønsted sites in the acidic chabazite was analyzed as a test case.

The acidic chabazite zeolite consists of a network of double six-membered alumino-silicate rings, connected by four-membered rings. The unit cell contains 37 atoms, with formula unit $HAlSi_{11}O_{24}$. This system is widely studied in the literature as an exemplary case of reactivity in zeolites. M. Sierka *et al.*[34] have studied the proton jump in several proton-exchanged alumino-silicates, including chabazite. In their work, periodic calculations were carried out adopting the QM-Pot approach, in which the QM region, (involving the reaction site) was computed at a B3LYP level combining an Ahlrich's Gaussian double-zeta polarized (for H, Si and Al atoms) with a triple-zeta polarized (for O atoms) basis set, whereas the long range contribution was handled by molecular mechanics embedding.[34] The energy barrier computed in this way, related to the proton transfer between oxygen sites O1 and O2 in dry chabazite (see Figure 4) is equal to 17.6 kcal/mol ≈ 73.6 kJ/mol.[34] At the same time, A. Rimola[35] and coworkers have implemented in the CRYSTAL code the distinguished reaction coordinate (DRC) method for the location of the transition states, using a set of redundant internal valence coordinates, with a localization of the saddle point based on the calculation of the Hessian matrix of the initial guess structure.[35] For the proton-exchanged process in the acidic chabazite zeolite bulk, modeled with B3LYP exchange-correlation functional, they obtained an energy barrier of 15.0 kcal/mol ≈ 62.8 kJ/mol using the basic DRC method, a value of 18.3 kcal/mol ≈ 76.6 kJ/mol adopting the DRC method with fixed cell approach, and of 17.1 kcal/mol ≈ 71.5 kJ/mol using the basic DRC method with a triple zeta polarized basis set.[35]

In order to analyze the same proton-exchange process using the NEB method implemented in the CRYSTAL code, a path with $m = 9$ images is considered, including initial and final states. The values of the elastic constants used for the basic, CI-NEB and VARK-NEB methods are reported in Section IV, together with other details about the NEB calculation setup. The geometry of the initial image was fully optimized, allowing relaxation of both atomic positions and lattice parameters, in the absence of symmetry constraints. On the contrary, the structure of the final image has been optimized keeping the volume and the lattice parameters constant and equal to those obtained by optimizing the initial image. This corresponds to a fixed cell approach, as all the images have the same reference cell, with $a = 9.4445$ Å, $b = 9.3940$ Å, $c = 9.3711$ Å, $\alpha = 93.83°$, $\beta = 94.47°$, $\gamma = 94.79°$. The search of the MEP was performed using two minimization algorithms, namely, the steepest descent (SD) and the conjugate gradient (CG), both with a constant optimization step of $ds = 0.8$ Bohr^2/Ha.

The initial, final and the transition state geometrical configurations in the surrounding of the proton exchange center, obtained using the CI-NEB method in conjunction with the steepest descent path minimization scheme, are reported in Figure 4. The nuclear positions for the initial, final and intermediate states at each NEB step for all the calculations listed in Table II are made available through a Zenodo repository.[29] The values of the distances between the atoms involved in the proton exchange process for the transition state configuration, obtained with the three different NEB techniques, are reported in Table IV, together with the corresponding values of other theoretical works previously mentioned. The values for the activation barriers, as well as the imaginary frequencies associated with the transition state found using different NEB schemes within our current implementation, together with the theoretical literature values, are reported in Table V.

Configuration	d(O1-H)	d(O2-H)	d(O1-Al)	d(O2-Al)
IS	0.9694	2.5645	1.8988	1.7028
	0.969 [a]	2.565 [a]	1.899 [a]	1.703 [a]
FS	3.3976	0.9713	1.7062	1.8801
	3.398 [a]	0.971 [a]	1.877 [a]	1.705 [a]

TABLE III. Distances [Å] between the relevant atoms involved in the proton exchange process for the transition state configuration of the acidic chabazite, in the initial (IS) and final (FS) structures used as first and last images for the nudged elastic band calculations. Results labeled with [a] are taken from A. Rimola and coworkers.[35] The labels of the atoms refer to Figure 4. All the calculations reported are performed using the hybrid B3LYP exchange-correlation functional.

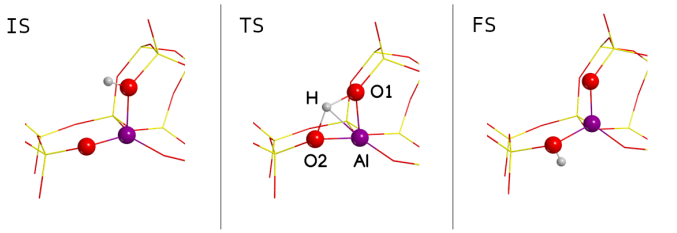

FIG. 4. Initial (IS), transition (TS) and final (FS) states for the proton exchange process in the acidic chabazite bulk. Yellow, purple, red and white spheres represent silicon, aluminum, oxygen and hydrogen atoms, respectively. The transition state is computed using the CI-NEB method in conjunction with the SD path minimization scheme, as implemented in the CRYSTAL code. The associated MEP is reported in the central panel of Figure 5. Details about the computational setup are given in Section III B of the main text.

Transition state (TS) configuration
Number of images $m = 9$ - B3LYP functional

Method		d(O1-H)	d(O2-H)	d(O1-Al)	d(O2-Al)
Basic NEB	SD	1.6115	1.0399	1.7679	1.8673
	CG	1.6113	1.0399	1.7679	1.8673
CI-NEB	SD	1.2316	1.2647	1.8280	1.8193
	CG	1.2316	1.2647	1.8280	1.8193
VARK-NEB	SD	1.2113	1.2876	1.8314	1.8158
	CG	1.2115	1.2874	1.8313	1.8158
DRC		1.222 [a]	1.245 [a]	1.830 [a]	1.820 [a]

TABLE IV. Distances [Å] between the relevant atoms involved in the proton exchange process for the transition state configuration of the acidic chabazite, resultant from different NEB methods (see Section II) and various optimization schemes (see Section II D) for the search of the minimum energy path, compared with other theoretical results. Results labeled with [a] are taken from Rimola *et al.*,[35] with [b] from Sierka and coworkers.[34] The labels of the atoms refer to Figure 4.

While the basic NEB method is not capable of reproducing the bond lengths nor the activation barriers found in other theoretical studies, both the CI-NEB and the VARK-NEB approaches lead to values in very good agreement with the existing literature.

The inability of the basic NEB method to capture the correct transition state is evident when looking at the profile of the MEP reported in the left panel of Figure 5, as the number of images is not sufficient to correctly locate the transition state. On the other hand, both the CI-NEB and the VARK-NEB methods, characterized by an increase in the resolution around the saddle point (see central and right panels of Figure 5, respectively) provide a good description of the transition state, even with $m = 9$ images along the whole path.

In order to better characterize the transition state in the case of the basic NEB method, a greater number of images has been used to discretize the reaction path. The results of these calculations are outlined in Table VI, and the corresponding minimum energy paths obtained using the steepest descent path minimization method are reported in Figure 3 of the supplementary material, highlighting an increasing in the accuracy for the calculation of the activation energy as the number of the images increases. However, even with a value of $m = 25$, the activation energy and especially the imaginary frequency associated to the transition state found by the basic NEB method are not in good agreement with the results obtained with the climbing image and the variable spring constants approaches, as well as with the theoretical literature values. In this case, the basic NEB method proves to be less accurate and less efficient in locating the transition state with respect to the other two improved methods, that are the climbing image and the variable spring constants strategies, with a much greater number of images required to correctly characterize the transition state.

Number of images $m = 9$ - B3LYP functional

Method		n_{opt}	E_a [eV]	ν [cm^{-1}]
Basic NEB	SD	1038	0.475143 (0.364577)	—
	CG	1035	0.475451 (0.364885)	—
CI-NEB	SD	508	0.799548 (0.688982)	1343.1208 i
	CG	511	0.799556 (0.688990)	1343.1214 i
VARK-NEB	SD	532	0.797144 (0.686577)	1322.5652 i
	CG	532	0.797199 (0.686633)	1322.9498 i
DRC			0.6504 [a]	1218 i [a]
DRC fixed cell			0.7935 [a]	1334 i [a]
DRC triple ζ			0.7415 [a]	1308 i [a]
QM-Pot			0.7628 [b]	1151 i [b]

TABLE V. Number of optimization steps n_{opt} required to find the MEP, activation energy E_a and imaginary frequency ν of the transition state. Different NEB methods (see Section II) and various optimization schemes (see Section II D) for the search of the minimum energy path are compared with other theoretical results. Results labeled with [a] are taken from Rimola *et al.*,[35] with [b] from Sierka and coworkers.[34] The optimization steps in the SD and CG path minimization algorithms are reported in the main text of Section III B.

Basic NEB - B3LYP functional

Images	Method	n_{opt}	E_a [eV]	ν [cm^{-1}]
$m = 9$	SD	1038	0.475143 (0.364577)	—
	CG	1035	0.475451 (0.364885)	—
$m = 12$	SD	1760	0.697654 (0.587087)	441.0805 i
	CG	1766	0.697345 (0.586779)	438.2428 i
$m = 15$	SD	565	0.649948 (0.539381)	414.8259 i
	CG	538	0.644330 (0.533763)	388.3860 i
$m = 18$	SD	539	0.704545 (0.593979)	744.4183 i
	CG	544	0.705515 (0.594949)	749.0149 i
$m = 25$	SD	537	0.749974 (0.639408)	1050.4820 i
	CG	541	0.748394 (0.637828)	1040.1668 i

TABLE VI. Number of optimization steps n_{opt} required to find the MEP, activation energy E_a and imaginary frequency ν of the transition state, obtained with the basic NEB method (see Section II) using different number of images m and various optimization schemes (see Section II D) for the search of the minimum energy path. The optimization steps in the SD and CG path minimization algorithms are reported in the main text of Section III B.

The excellent agreement between our results and the other theoretical calculations, at the B3LYP level, confirms the reliability of all the NEB algorithms, implemented within

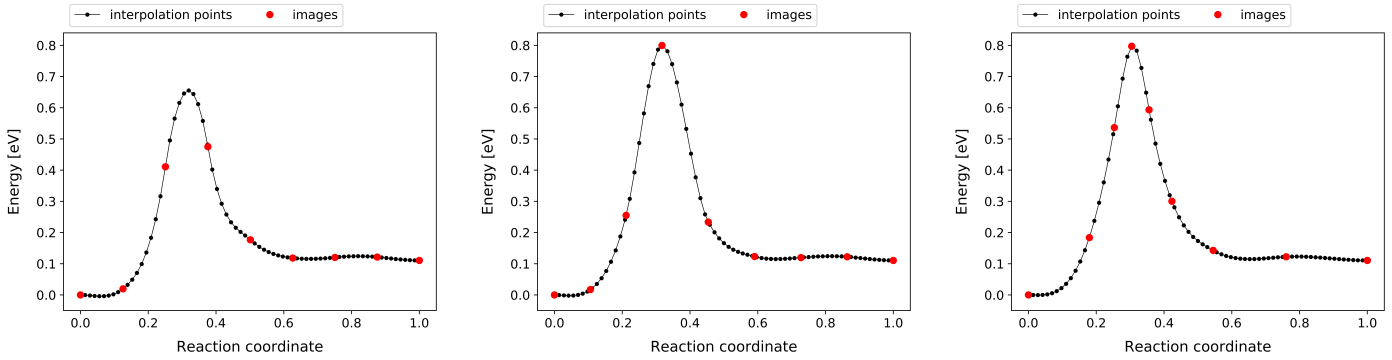

FIG. 5. Minimum energy paths obtained with the basic NEB (left panel), the climbing image (CI-NEB, central panel) and the variable spring constants (VARK-NEB, right panel) methods, using the steepest descent (SD) algorithm with an optimization step of $ds = 0.8$ Bohr^2/Ha to converge the reaction path towards the minimum energy path. Details about the activation barrier, as well as about the vibrational frequencies and the geometrical distances among the atoms involved in the proton exchange process for the transition state configurations are reported in Tables IV and V. The interpolation of the minimum energy paths is done using 72 points. The method used for the interpolation is described in Section II F.

the CRYSTAL code, in computing the activation energy and finding the transition state for local transitions in periodic crystalline systems.

IV. COMPUTATIONAL DETAILS

The new implementation of the NEB algorithm described in this work has been performed within a beta version of the CRYSTAL23 package for ab initio quantum chemistry and physics of solid state.[19]

Geometry optimizations, for which the default convergence thresholds were adopted,[36] are performed using analytical gradients with respect to atomic coordinates, within a quasi-Newtonian algorithm combined with the Broyden-Fletcher-Goldfarb-Shanno (BFGS) formula.[37–41] The default convergence thresholds were adopted.[19,36]

Harmonic frequencies, computed at the $\mathbf{\Gamma}$ point, are obtained through the diagonalization of the mass-weighted Hessian in Cartesian coordinates, where the second derivatives of the energy with respect to atomic displacements are evaluated numerically by means of a central-difference formula.[42,43] In the calculation of frequencies for the transition states, any residual symmetry has been removed, to ensure consistency with the NEB calculations performed within the P_1 space group. The self-consistent field energy threshold was set equal to 10^{-12} Ha per unit cell, for all the systems.

The NEB calculations for the reaction (R1) were performed using an unrestricted Hartree-Fock Hamiltonian. The Coulomb and exchange series, summed in direct space, are truncated using overlap criteria, with thresholds given by $[20, 20, 20, 20, 20]$.[36] The threshold for the self-consistent field algorithm convergence on the total energy is chosen equal to 10^{-8} Ha. An all electron basis set, consisting of contracted Gaussian-type atomic orbitals (AOs) functions, named 5-11 G*, was used.[27] Details on the NEB parameters are reported in Section III A 1.

The NEB calculations for the reaction (R2) were performed in the framework of DFT, using the B3LYP exchange-correlation functional. The Coulomb and exchange series, summed in direct space, are truncated using overlap criteria with thresholds set to $[20, 20, 20, 20, 20]$. The threshold for the self-consistent field algorithm convergence on the total energy is fixed at 10^{-8} Ha. The all-electron basis set by C. Gatti *et al.*[44] were adopted for the different elements, resulting in 57 AOs for the whole molecular system.

The NEB calculations for the proton exchange process in the acidic chabazite were performed in the framework of DFT, using B3LYP exchange-correlation functional. The DFT exchange-correlation contribution was evaluated by numerical integration over the unit cell volume, using a pruned grid,[45,46] consisting of 75 radial points and a maximum number of 974 angular points in regions relevant for chemical bonding.[36] The diagonalization of the Fock matrix and the integration over the reciprocal space is carried out using the Monkhorst-Pack mesh,[47] consisting in a grid of 14 $\boldsymbol{k}$ points in the irreducible part of the first Brillouin zone (IBZ). The Coulomb and exchange series, summed in direct space, are truncated using overlap criteria, with thresholds given by $[6, 6, 6, 6, 14]$.[36] The threshold for the self-consistent field algorithm convergence on the total energy is chosen equal to 10^{-8} Ha. A set of 5-11 G* type all-electron basis were used for the different elements, with a total of 557 AOs per unit cell.[27]

A default values of the few parameters controlling the NEB algorithm have been defined and used in all tests. In the basic and CI-NEB approaches, the value of the elastic constant was set to $k = 0.1$ Ha/Bohr^2. The CI-NEB technique is activated after 10 NEB iterations, i.e. starting from the eleventh NEB step. The minimum and maximum values of the spring constants in the VARK-NEB approach are set equal to $k_{min} = 0.1$ Ha/Bohr^2 and $k_{max} = 0.5$ Ha/Bohr^2, respectively, and the method is active from the beginning of the simulation. The values of the optimization steps adopted for each simulation is reported in Section III, for all the analyzed systems. The thresholds for the path minimization regarding (i) the Euclidean norm and (ii) the root-mean-square of the components of the real nuclear force vectors orthogonal to the path are 10^{-3} Ha/Bohr and $5{\cdot}10^{-4}$ Ha/Bohr, respectively.

V. CONCLUSIONS

This work documents the implementation of the NEB method within the CRYSTAL code, a quantum mechanical ab initio program, capable of modeling systems with dif-

ferent periodicity, based on localized Gaussian wavefunctions as basis set. This algorithm allows to compute reaction paths for molecular and periodic systems, using hybrid density functionals, opening the possibility to obtain accurate and efficient determination of minimum energy paths as well as activation barriers and transition state configurations in solid state reactions, surface processes, and chemical-physical transformations in molecular and extended periodic environments. To validate the implementation, we analyzed the reliability and efficiency of different NEB variants (namely, basic, CI-NEB and VARK-NEB) on representative molecular and bulk reactions. The results demonstrate that: (*i*) the implementation is reliable, robust, versatile and general purpose; (*ii*) the parameter associated with the step for the path minimization has a strong impact on the convergence speed, while maintaining a good and consistent accuracy up to a reasonable high value; (*iii*) the CI-NEB and the VARK-NEB methods proved to be the most efficient in identifying the correct transition state, and (*iv*) the steepest descent algorithm for finding the MEP outperforms the conjugate gradient method. The integration of the NEB procedure into the CRYSTAL code, which is able to perform quantum mechanical hybrid density functional calculations on periodic systems at a reasonably low cost, lays the foundation for future investigations of complex chemical-physical processes in condensed phase systems, at a high level of theory.

[1] J. M. Herbert and M. Head-Gordon, "Accelerated, energy-conserving Born–Oppenheimer molecular dynamics via Fock matrix extrapolation," Phys. Chem. Chem. Phys. **7**, 3269–3275 (2005).

[2] P. Steneteg, I. A. Abrikosov, V. Weber, and A. M. N. Niklasson, "Wave function extended Lagrangian Born-Oppenheimer molecular dynamics," Phys. Rev. B **82**, 075110 (2010).

[3] É. Polack, G. Dusson, B. Stamm, and F. Lipparini, "Grassmann Extrapolation of Density Matrices for Born–Oppenheimer Molecular Dynamics," J. Chem. Theory Comput. **17**, 6965–6973 (2021).

[4] F. Pes, É. Polack, P. Mazzeo, G. Dusson, B. Stamm, and F. Lipparini, "A Quasi Time-Reversible Scheme Based on Density Matrix Extrapolation on the Grassmann Manifold for Born-Oppenheimer Molecular Dynamics," J. Phys. Chem. Lett. **14**, 9720–9726 (2023).

[5] E. Wigner, "The transition state method," Trans. Faraday Soc. **34**, 29–41 (1938).

[6] G. H. Vineyard, "Frequency factors and isotope effects in solid state rate processes," J. Phys. Chem. Solids **3**, 121–127 (1957).

[7] H. Kramers, "Brownian motion in a field of force and the diffusion model of chemical reactions," Physica **7**, 284–304 (1940).

[8] H. Jónsson, G. Mills, and K. W. Jacobsen, "Nudged elastic band method for finding minimum energy paths of transitions," in *Classical and Quantum Dynamics in Condensed Phase Simulations* (1998) pp. 385–404.

[9] G. Henkelman and H. Jónsson, "A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives," J. Chem. Phys. **111**, 7010–7022 (1999).

[10] G. Henkelman and H. Jónsson, "Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points," J. Chem. Phys. **113**, 9978–9985 (2000).

[11] G. Henkelman, B. P. Uberuaga, and H. Jónsson, "A climbing image nudged elastic band method for finding saddle points and minimum energy paths," J. Chem. Phys. **113**, 9901–9904 (2000).

[12] D. Sheppard, R. Terrell, and G. Henkelman, "Optimization methods for finding minimum energy paths," J. Chem. Phys. **128**, 134106 (2008).

[13] D. Sheppard, P. Xiao, W. Chemelewski, D. D. Johnson, and G. Henkelman, "A generalized solid-state nudged elastic band method," J. Chem. Phys. **136**, 074103 (2012).

[14] C. J. Cerjan and W. H. Miller, "On finding transition states," J. Chem. Phys. **75**, 2800–2806 (1981).

[15] A. Banerjee, N. Adams, J. Simons, and R. Shepard, "Search for stationary points on surfaces," J. Phys. Chem. **89**, 52–57 (1985).

[16] J. Baker, "An algorithm for the location of transition states," J. Comput. Chem. **7**, 385–395 (1986).

[17] A. Rimola, C. M. Zicovich-Wilson, R. Dovesi, and P. Ugliengo, "Search and Characterization of Transition State Structures in Crystalline Systems Using Valence Coordinates," J. Chem. Theory Comput. **6**, 1341–1350 (2010).

[18] R. Dovesi, A. Erba, R. Orlando, C. M. Zicovich-Wilson, B. Civalleri, L. Maschio, M. Rérat, S. Casassa, J. Baima, S. Salustro, and B. Kirtman, "Quantum-mechanical condensed matter simulations with CRYSTAL," Wiley Interdiscip. Rev. Comput. Mol. Sci. **8**, e1360 (2018).

[19] A. Erba, J. K. Desmarais, S. Casassa, B. Civalleri, L. Donà, I. J. Bush, B. Searle, L. Maschio, L. Edith-Daga, A. Cossard, C. Ribaldone, E. Ascrizzi, N. L. Marana, J.-P. Flament, and B. Kirtman, "CRYSTAL23: A Program for Computational Solid State Physics and Chemistry," J. Chem. Theory Comput. **19**, 6891–6932 (2023).

[20] B. S. Jursic, "An accurate evaluation of activation barriers for hydrogen abstraction reactions with Becke's 88 density functional theory and high-level G1 and G2 ab initio methods," Int. J. Quantum Chem. **65**, 75–82 (1997).

[21] B. S. Jursic, "Exploring the potential energy surface for the H2 + H → H + H2 reaction with ab initio and density functional theory methods," J. Mol. Struct. THEOCHEM **427**, 117–121 (1998).

[22] M. J. Frisch, G. W. Trucks, H. B. Schlegel, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, M. H. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. Zeng, J. A. V. M. Bloino, J. L. Ghysels, and B. J. F. J. de Vera, "Gaussian 94, Revision B.3," (1995), Gaussian, Inc., Pittsburgh, PA, USA.

[23] W. R. Schulz and D. J. Le Roy, "Kinetics of the Reaction H + p-H_2 = o-H_2 + H," J. Chem. Phys. **42**, 3869–3873 (1965).

[24] B. Temelso, C. D. Sherrill, R. C. Merkle, and R. A. Freitas, "High-Level ab Initio Studies of Hydrogen Abstraction from Prototype Hydrocarbon Systems," J. Phys. Chem. A **110**, 11160–11173 (2006).

[25] B. G. Johnson, C. A. Gonzales, P. M. Gill, and J. A. Pople, "A density functional study of the simplest hydrogen abstraction reaction. effect of self-interaction correction," Chemical Physics Letters **221**, 100–108 (1994).

[26] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, "A fifth-order perturbation comparison of electron correlation theories," Chem. Phys. Lett. **157**, 479–483 (1989).

[27] R. Dovesi, C. Ermondi, E. Ferrero, C. Pisani, and C. Roetti, "Hartree-Fock study of lithium hydride with the use of a polarizable basis set," Phys. Rev. B **29**, 3591–3600 (1984).

[28] F. Neese, "Software update: The ORCA program system—Version 5.0," Wiley Interdiscip. Rev.: Comput. Mol. Sci. **12**, e1606 (2022).

[29] A. Gelli, S. Casassa, A. Rimola, and C. Ribaldone, "Nudged Elastic Band Tests using a develop version of the CRYSTAL code," `https://www.zenodo.org/10.5281/zenodo.17463172` (2025).

[30] H. B. Schlegel, P. Gund, and E. M. Fluder, "Tautomerization of formamide, 2-pyridone, and 4-pyridone: an ab initio study," J. Am. Chem. Soc. **104**, 5347–5351 (1982).

[31] G. Fogarasi, "Studies on tautomerism: Benchmark quantum chemical calculations on formamide and formamidine," J. Mol. Struct. **978**, 257–262 (2010).

[32] C. Adamo, M. Cossi, and V. Barone, "Catalytic and bulk solvent effects on proton transfer: Formamide as a case study," J. Comput. Chem. **18**, 1993–2000 (1997).

[33] R. D. Brown, P. D. Godfrey, and B. Kleibömer, "The conformation of formamide," J. Mol. Spectrosc. **124**, 34–45 (1987).

[34] M. Sierka and J. Sauer, "Proton Mobility in Chabazite, Faujasite, and ZSM-5 Zeolite Catalysts. Comparison Based on ab Initio Calculations," J. Phys. Chem. B **105**, 1603–1613 (2001).

[35] A. Rimola, C. M. Zicovich-Wilson, R. Dovesi, and P. Ugliengo, "Search and characterization of transition state structures in crystalline systems using valence coordinates," J. Chem. Theory Comput. **6**, 1341–1350 (2010).

[36] R. Dovesi, V. R. Saunders, C. Roetti, R. Orlando, C. M. Zicovich-Wilson, F. Pascale, B. Civalleri, K. Doll, N. M. Harrison, I. J. Bush, P. D'Arco, M. Llunell, M. Causá, Y. Nöel, L. Maschio, A. Erba, M. Rérat, S. Casassa, B. G. Searle, and J. Desmarais, "online CRYSTAL23 user's manual," www.crystal.unito.it/include/manuals/crystal23.pdf , **2022**.

[37] C. G. Broyden, "The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations," IMA J. Appl. Math. **6**, 76–90 (1970).

[38] C. G. Broyden, "The Convergence of a Class of Double-rank Minimization Algorithms 2. The New Algorithm," IMA J. Appl. Math. **6**, 222–231 (1970).

[39] R. Fletcher, "A new approach to variable metric algorithms," Comput. J. **13**, 317–322 (1970).

[40] D. Goldfarb, "A family of variable-metric methods derived by variational means," Math. Comput. **24**, 23–26 (1970).
[41] D. Shanno, "Conditioning of Quasi-Newton Methods for Function Minimization," Math. Comput. **24**, 647–656 (1970).
[42] F. Pascale, C. M. Zicovich-Wilson, F. López Gejo, B. Civalleri, R. Orlando, and R. Dovesi, "The calculation of the vibrational frequencies of crystalline compounds and its implementation in the CRYSTAL code," J. Comput. Chem. **25**, 888–897 (2004).
[43] C. M. Zicovich-Wilson, F. Pascale, C. Roetti, V. R. Saunders, R. Orlando, and R. Dovesi, "Calculation of the vibration frequencies of α-quartz: The effect of Hamiltonian and basis set," J. Comput. Chem. **25**, 1873–1881 (2004).
[44] C. Gatti, V. R. Saunders, and C. Roetti, "Crystal field effects on the topological properties of the electron density in molecular crystals: the case of urea," J. Chem. Phys. **101**, 10686–10696 (1994).
[45] A. D. Becke, "A multicenter numerical integration scheme for polyatomic molecules," J. Chem. Phys. **88**, 2547–2553 (1988).
[46] M. D. Towler, A. Zupan, and M. Causà, "Density functional theory in periodic systems using local gaussian basis sets," Comput. Phys. Commun. **98**, 181–205 (1996).
[47] H. J. Monkhorst and J. D. Pack, "Special points for Brillouin-zone integrations," Phys. Rev. B **13**, 5188–5192 (1976).