"On-the-fly" *Crystal*: How to reliably and automatically characterize and construct potential energy surfaces

Mahdi Aarabi, Ankit Pandey, and Bill Poirier[a)]

Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, Texas, USA

(Dated: October 2, 2023)

In this work, the ***Crystal*** code, developed previously by the authors to find "holes" as well as legitimate transition states in existing potential energy surface (PES) functions [*JPC Lett.* **11**, 6468 (2020)], is retooled to perform on-the-fly "direct dynamics"-type PES explorations, as well as automatic construction of new PES functions. In all of these contexts, the chief advantage of ***Crystal*** over other methods is its ability to globally map the PES, thereby determining the most relevant regions of configuration space quickly and reliably—even when the dimensionality is rather large. Here, ***Crystal*** is used to generate a uniformly-spaced grid of density functional theory (DFT) or *ab initio* points, truncated over the relevant regions, which can then be used to either: (a) hone in precisely on PES features such as minima and transition states, or; (b) create a new PES function automatically, via interpolation. Proof of concept is demonstrated via application to three molecular systems: water (H_2O), methane (CH_4), and methylene imine (CH_2NH).

[a)]Electronic mail: Bill.Poirier@ttu.edu

I. INTRODUCTION

The chemical dynamics "story" is told through an unfolding of potential energy surface (PES) features, taking us from reactant valleys and/or minima, through transition states (TSs) and reaction intermediates, and ultimately to products. Just as with any story, knowing the beginning and ending is often less important than learning about the journey in between. For small molecular systems with say, less than five atoms, there are not so many dynamical possibilities, and it may be feasible to explore them all. As molecules grow in size, however—and also as the energy increases—the size of the dynamical space can grow astronomically.

In order to find our way around in such vast spaces, we typically rely on a very powerful tool called "chemical intuition." To be sure, this has led to many successful predictions of chemical structures and reaction paths (RPs). However, there have also been many failures. Simply put, "chemical intuition" is far from fool-proof; we thus need tools that can *reliably* find all relevant RPs—whether they match our preconceived expectations or not. Indeed, this challenge has been deemed "the most important task in the study...of chemical reactions".[1]

Generally speaking, the dynamically most important RPs are those with the lowest-lying energy barriers. Since straightforward tools (e.g., steepest descent[2]) exist for determining RPs from their corresponding TS barriers, the bottleneck challenge is thus to compute *all* of the TSs (i.e., first-order saddle points) lying below a given energy $E_{\max}$, in a guaranteed and automated fashion that does not rely a priori on "chemical intuition." This is the heart of what we propose to call "chemical dynomics"—in analogy with genomics, proteomics, etc.

As a crucial first step in the chemical dynomics direction, the authors recently introduced a new software tool—i.e., our ***Crystal*** code[3–6]—to completely characterize *existing* PES functions as previously published in the literature. In this context, ***Crystal*** has already been applied to existing PESs with as many as 48 degrees of freedom (dofs).[4] In addition to finding all legitimate TSs up to the desired $E_{\max}$, ***Crystal*** also finds all PES "holes"—i.e., unphysical saddle points typically leading to valleys of infinitely negative energy, which can wreak havoc on dynamical calculations. Holes are notoriously difficult to find, because they can occur literally anywhere, and will therefore not succumb at all to chemical intuition.

Despite best intentions, holes invariably arise in PES fitting when the dimensionality is large enough, for the simple reason that PES developers have traditionally lacked the tools needed to globally map their PES functions. Indeed, until recently, the conventional wisdom had many believing that "there are no general methods...guaranteed to work" beyond about

three atoms.[1] ***Crystal*** has changed all of that, however, making global PES mapping feasible for much larger systems and higher energies than previously thought possible.[4,5] This has already led to the aforementioned discovery of new legitimate TSs—and regrettably, also, a good number of heretofore unexpected holes—in previously published PES functions.

While using ***Crystal*** to "fix" existing PES functions is no doubt a useful enterprise, it will undoubtedly prove better, going forward, to construct new global PES functions that are devoid of such problems in the first place. Specifically what we mean by this are PES functions that, up to some predetermined $E_{\max}$:

(a) include all energetically accessible minima, TSs, and RPs;

(b) contain no accessible PES holes.

To this end, we develop in this paper an "on-the-fly" version of ***Crystal***, that operates in the absence of any preexisting PES function.

As with previous ***Crystal*** implementations, the new on-the-fly version constructs a uniformly spaced grid of points over the desired configuration space (either full-dimensional or restricted) that is truncated to include only those points with potential energy value less than $E_{\max}$. The main difference is that, rather than call a preexisting PES function routine to evaluate the energy at a given point, either a density functional theory (DFT) or *ab initio* calculation is instead performed. In this fashion, a grid of points over the relevant region of configuration space is efficiently constructed on the fly.

The above constitutes the first phase of on-the-fly ***Crystal***. In the second phase, there are at least two alternative directions that may be pursued, depending on the desired goal. The first option is *on-the-fly characterization* of the PES—meaning that all minima, TSs, and RPs are accurately determined up to $E_{\max}$, using procedures analogous to those used previously.[4,5] Of course, there are no PES holes, since there is no fitting; only legitimate PES structures will be discovered—subject, of course, to the whims of the particular electronic structure method being used.

The second option is *on-the-fly construction* of a new PES function, based on precomputed grid point values. Of course, to this end, any of a number of well established strategies may be employed.[7–10] Historically, these may be divided as to whether they are primarily either fitting or interpolation strategies. Whereas the former is traditionally preferred, here we will advocate for one specific approach that is interpolation-based, due to the following two considerations. First, consider one of the main reasons why fitting is often used over

interpolation—i.e., because it can (for some functional forms) lead to "global" PES functions that are well-behaved over the entire configuration space, even well away from the grid points. In contrast, interpolation is notorious for producing poor quality PES functions in the "extrapolation" region. However, our philosophy here is to include as global a region of configuration space as possible, but *only up to* $E_{\max}$. The grid that ***Crystal*** produces automatically places points throughout exactly this region; hence, there is no need nor desire to extrapolate beyond it. The second consideration is the highly pragmatic one of simplicity; interpolation is straightforward to perform when structured, uniformly-spaced grids are employed (as is the case here), and the results are reasonably accurate. In contrast, global PES fitting typically requires a lot more effort and finesse to implement effectively.

Interpolation with ***Crystal*** thus provides a ready avenue for so-called "automatic" PES construction—which, moreover, is guaranteed to result in PES functions that satisfy conditions (a) and (b) above. In contrast, most fitting strategies will not necessarily satisfy these conditions, and in addition, are far more tricky to implement in practice. That said, there are a number of other "automatic" PES construction strategies that have been developed in recent years,[11–14] some using neural networks or other AI-inspired ideas,[15–18] and all more sophisticated than the simple strategy we describe here. Some of these approaches tend to center around optimal placement of the $n+1$'th point, based on the current grid of n points. As such, they necessarily lead to a non-uniform grid, not well suited for interpolation strategies.

In this paper, we shall focus on on-the-fly PES construction using ***Crystal***. A representative test suite of applications is considered, spanning a range of system dimensionalities, electronic structure methodologies (***Gaussian 16*** is used here[19]), and other factors. We begin with the well-studied triatomic molecule, water (H_2O). The full vibrational space of H_2O has only three dofs; however, to make it interesting, we have chosen our energy cut-off value of $E_{\max} = 32,000$ cm^{-1} to lie well above the linear inversion barrier TS (near 11,000 cm^{-1}) that leads to the second equivalent bent minimum structure. The resultant PESs constructed by on-the-fly ***Crystal*** are therefore global, symmetric double-well PESs. Accordingly, the 3D PES grids are chosen to respect this symmetry, as well as the structure of the two equivalent global minima. For this system, too, we have constructed separate PESs using both *ab initio* and DFT electronic structure methods.

For our second application, we consider the restricted configuration space describing a single CH bond in methane. Keeping the methyl group frozen, the remaining H atom is

allowed to move in all three Cartesian directions, (x, y, z). This PES, constructed using DFT calculations on the grid points, is thus 3D. Again, the energy cutoff of $E_{\max} = 28,000$ cm^{-1} is quite high—although, as in the case of H_2O, well below dissociation. The restricted CH_4 PES thus serves as a useful benchmark, especially in comparison with the above H_2O calculations. Note that both PESs have the same number of dofs, and the two molecules are also isoelectronic.

The third application is the most difficult—i.e., methylene imine (CH_2NH), computed in full vibrational dimensionality (9 dofs). Completing all requisite grid point calculations is quite challenging in this case, despite the fact that only DFT-level electronic structure calculations are performed. This numerical challenge reflects both the increased number of grid points due to the larger dimensionality, as well as the greater computational (CPU) cost needed to perform each such electronic structure calculation. One consequence of this increased numerical difficulty is the lower energy cutoff used for this system, i.e. $E_{\max} = 5300$ cm^{-1}. Nevertheless, with almost no "human oversight," a viable 9D CH_2NH PES function was constructed, which has no holes within the specified energy range, and is otherwise well behaved. Proof of concept of the method is thereby established.

Note that we were able to perform the many on-the-fly PES grid calculations needed for CH_2NH in a timely fashion—but only by virtue of a new parallel implementation of the ***Crystal*** code. This parallel implementation is designed to be highly efficient whenever the CPU bottleneck of the calculation is the evaluation of the PES energy at the grid points. This condition is certainly satisfied for the present on-the-fly applications, but may also be true more generally. For instance, some PES function evaluations (e.g. using the method of Dawes[11]) are notoriously slow to evaluate. In any event, given its broader scope, the parallel version of ***Crystal*** shall serve as the focus of a concurrent publications[6] although a brief description is also provided here.

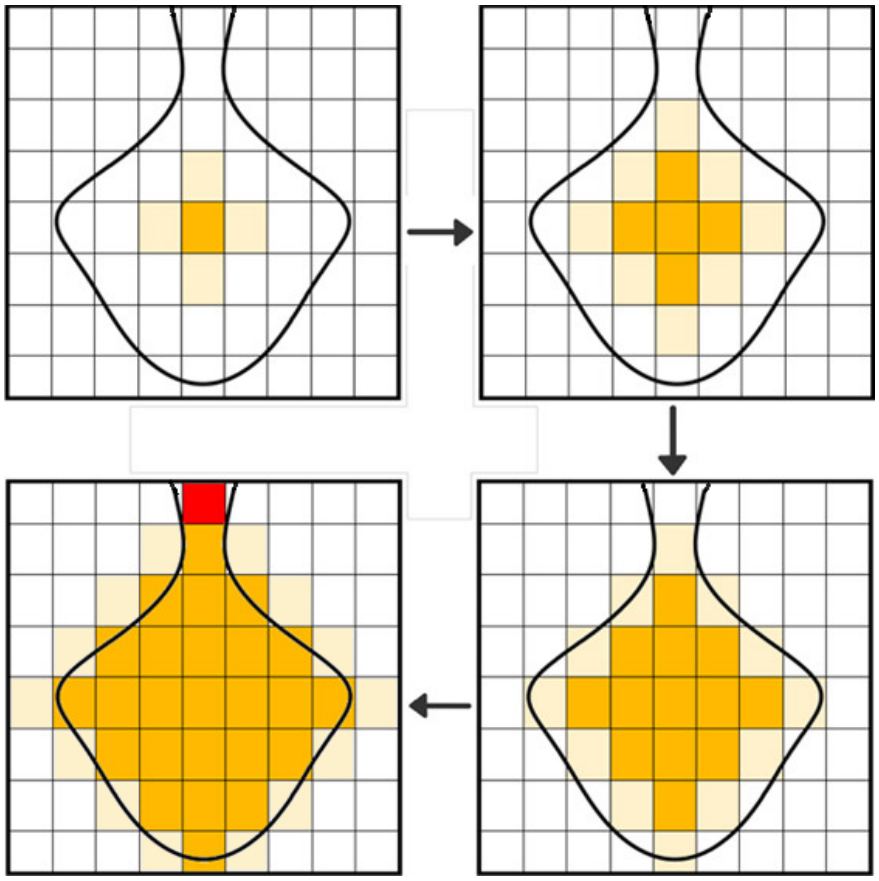
II. THEORY

A. The *Crystal* code

Crystal is an open source code developed and maintained by the authors. The basic structure of the ***Crystal*** algorithm is described in Fig 1. As shown, ***Crystal*** performs an outward expansion into a volume space of interest; for our current purpose, this volume is the region enclosed by a PES contour at the desired energy cutoff, $E_{\max}$. As seen in

Fig 1, each shell (collection of light-yellow squares, whose centers are the lattice points) is generated by the lattice points of the previous (i.e. next innermost) shell. Duplicates are removed. In addition, the lattice points with higher energy than a predetermined cutoff are also removed. The shell-generation process is repeated until no new points are generated. Of course, there are many other technical details that must be considered to make the scheme feasible in practice;[3–5] nevertheless, the basic operation of ***Crystal***, as conveyed in Fig 1, is very simple.

Figure 1: Operation of the ***Crystal*** algorithm. Light yellow squares denote lattice points for the current shell, whereas dark yellow squares denote current set of lattice points In each ***Crystal*** step, lattice points for the current shell are determined as the nearest neighbors to all dark yellow lattice points. PES evaluations for all such points are then conducted, either via function evaluation, or (for on-the-fly ***Crystal***) via a call to external *ab initio*/DFT software. Those shell lattice points whose energy lies below the cutoff $E_{\max}$ (denoted by the black contour curve) are retained and added to the current set; the rest are discarded. In the example indicated, the red lattice point signifies the avenue to a second well region (not indicated)—and therefore heralds the presence of a TS nearby.



Despite its remarkable simplicity, the basic ***Crystal*** algorithm has been found to be very suitable for a wide range of purposes, especially involving molecular systems with up to tens of dofs. It has formerly been utilized not only to find TSs and RPs, but also to

calculate configuration (or phase) space volumes (e.g. for computing partition functions), as well as to select basis set functions for performing quantum dynamics calculations for vibrational spectroscopy[3–5]. As discussed in Sec. I, in the current paper, we focus on the new PES-construction aspect of our ***Crystal*** code, which we have developed to automate the determination of uniformly-spaced lattice points—and commensurate electronic structure calculations—that are required for PES construction. In particular, not only is the new PES constructed automatically (via interpolation over the lattice points), but the new PES function is *guaranteed* not to have any holes up to the predetermined energy cutoff $E_{\max}$.

One notable strength of ***Crystal*** is that it can be readily parallelized. This is accomplished by distributing the PES evaluations for the current shell of lattice points across as many computational (CPU) nodes as desired (presumably up to the total number of lattice points in the shell). There is some communication required to avoid duplicate calculations, but this is minimal in comparison with the CPU cost of the electronic structure calculations themselves. As mentioned earlier, a much deeper analysis of the parallel ***Crystal*** code may be found in the concurrent paper,[6] and in future publications. However, we note that parallel ***Crystal*** is likely to allow for reliable exploration of much larger volumes—and in higher dimensional spaces—than what has been considered in the literature up to this point, including our own previous ***Crystal*** papers.

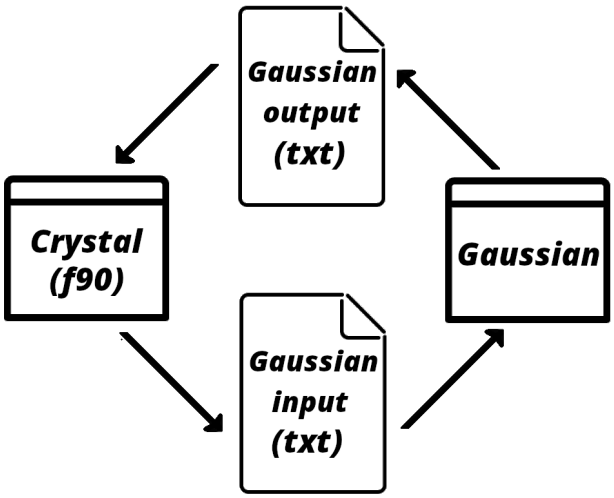
In the present study, the parallel version of ***Crystal*** was applied only to the largest application, i.e. methylene imine, for which it was found to be highly efficient. In particular, nearly perfect parallel scalability was achieved, in that the computational speedup observed with respect to the number of CPU nodes used was nearly linear. As the methylene imine calculations are not extremely numerous, we distributed the lattice points of each current shell across up to eight nodes in total. Each node was then responsible for computing a single PES evaluation at a time. Note that the computational clusters that we used (see Acknowledgments) have many available cores per node. Accordingly, intranode parallelization across multiple cores was employed at the level of the electronic structure codes (i.e., ***Gaussian 16*** in our case), but not distributed memory parallelization across multiple nodes.

B. *Crystal* with on-the-fly PES evaluations

Next, we address the on-the-fly version of ***Crystal***. Although all calculations presented in this study were performed using on-the-fly ***Crystal***, the calculations for the water and methane applications were performed without parallelization, and so we address this simpler

strategy first. For these applications, the PES evaluations inside every shell were performed serially—i.e., one lattice point after another—which, at least up until now, has been the standard ***Crystal*** implementation. The new wrinkle, of course, is the on-the-fly aspect of how these serial lattice-point PES evaluations are evaluated, which is described below.

Figure 2: Schematic of on-the-fly implementation of the ***Crystal*** code, as explained in the text. The ***Crystal*** algorithm is implemented as a continuously running FORTRAN 90 executable, which communicates with a sequence of ***Gaussian 16*** jobs via input/output .txt files, all coordinated through a shell script.



The basic operation involves running one ***Crystal*** executable job throughout the calculation, together with a sequence of individual ***Gaussian 16*** jobs, each designed to perform a single PES evaluation for a given lattice point. For the initial lattice-point PES evaluation, a ***Gaussian 16*** input file, corresponding to the initial reference molecular geometry, is first created by the ***Crystal*** code. A ***Gaussian 16*** job is then executed to read in this input file, compute a single PES evaluation at the specified geometry, and produce an output file which contains the molecular energy at this geometry, which is then fed back into ***Crystal***. Thereafter, this feedback loop is repeated, as illustrated in Fig. 2: ***Crystal*** generates new shells based on the prior ***Gaussian 16*** energy outputs, and creates new Gaussian input files for subsequent evaluation. Simultaneously, ***Gaussian 16*** produces output files, based upon which, Crystal generates the next set of input files, etc. Communication is straightforward enough as to preclude the need for sophisticated tools such as MPI; instead, simple shell scripts and built-in time delays suffice for this purpose.

For methylene imine, a parallel version of on-the-fly ***Crystal*** was used, as follows. First, the ***Crystal*** codes themselves were partitioned into two separate executables. The first is the "primary" ***Crystal*** code, which performs all of ***Crystal***'s functions that do *not* involve

PES evaluations—such as lattice-data collection, distribution, and storage. Throughout the parallel on-the-fly calculation, there is just a single primary code that is executed. In addition, however, there are multiple "secondary" ***Crystal*** executables, one running on each node used in the parallel calculation (i.e., eight in all, in the case of methylene imine). Each of these secondary jobs runs in a feedback loop with its own set of ***Gaussian 16*** jobs, to perform a sequence of PES evaluations on a given node, one at a time. After each such PES evaluation is completed, the corresponding secondary ***Crystal*** job communicates with the primary job, to receive a new lattice point for PES evaluation.

Typically, a ***Crystal*** shell may be expected to contain hundreds to thousands of new lattice points, for systems with as many dofs as methylene imine. Moreover, this number increases polynomially quickly as a function of system dimensionality. Since the communication cost is very limited in comparison to the PES evaluation cost, the parallel load sharing of PES evaluations is expected to scale nearly perfectly, until the number of nodes or secondary jobs starts to approach the shell size. Therefore, we also have the desirable situation that systems with higher dimensionality—i.e., those that are larger and more challenging—are also expected to exhibit better parallel scalability.

C. Electronic structure calculations

For all three of the molecular systems considered in this work, i.e. water, methane, and methylene imine, DFT methods have been used to compute structures and PES functions, via on-the-fly ***Crystal***. In addition, for H_2O, a second on-the-fly PES has also been constructed using an *ab initio* electronic structure method. For the DFT calculations, after examining several options, we selected B3LYP/aug-cc-pVTZ as a suitable choice of functional/basis set, that would take proper account of all the atoms in our investigated molecules. In the *ab initio* context, coupled-cluster methods are widely used to generate PESs, with CCSD(FULL) a fairly accurate standard choice available in ***Gaussian 16***. Accordingly, CCSD(FULL) was used to construct our *ab initio* PES for H_2O. Note that for these CCSD calculations, we also used the aug-cc-pVTZ basis set.

Regarding the above choices, it should be noted that our primary goal here is not to construct new state-of-the-art PESs to compete with others already on the market (although we may well have achieved this result for CH_2NH[20]). Rather, it is to demonstrate the viability of our on-the-fly ***Crystal*** PES construction scheme, for realistic calculations. Accordingly, our choices above reflect what are some of the most commonly used methods in electronic

structure, as opposed to what might be optimal for the specific molecular systems considered. Note that for this first-ever, proof-of-concept effort, we are not constructing any PESs with TSs, apart from the H_2O system. For this system, we wish to construct fully global double-well PES functions, that include the linear TS structure, as well as the two equivalent bent minima. Hence, this seems the appropriate system for which to construct an *ab initio* PES, in addition to the DFT PES, in order to validate the quality of the latter—especially in the vicinity of the TS structure. In the event, the two H_2O PESs, DFT and CCSD, are found to be reasonably similar, as discussed in Sec. IV A.

All electronic structure calculations presented in this work were performed using the ***Gaussian 16*** program.[19] References for the B3LYP DFT functional, CCSD(FULL) method, and basis sets used here, may be found in the ***Gaussian 16*** Users Reference—as well as in standard electronic structure textbooks.[21,22]

D. Interpolation

In most of its incarnations, the ***Crystal*** algorithm generates a uniformly-spaced lattice of grid points, over the configuration space of interest. For PES construction, the next and final stage is to use this lattice to construct a PES function, that can be evaluated at any desired point in the space (at least within the energy range of interest). To this end, regular grids of the sort produced by ***Crystal*** are very well suited to interpolation methods—which are preferred over fitting methods for this work, as discussed. However, although the lattice *itself* is regular; the *truncation* of lattice grid points as determined by ***Crystal*** is most decidedly *not* regular. The region fleshed out by the grid points is anything but rectangular, in other words. Indeed, this is the point; if a fully rectangular grid of lattice points were to be employed, the number of such points would grow exponentially with system dimensionality, which is what ***Crystal*** is striving to circumvent.

In principle, a truncated regular grid does not pose insurmountable difficulties for interpolation. However, for the present paper, we rely on the canned *Interpolation* routine from ***Mathematica***, which does indeed require a fully rectangular (i.e. non-truncated) grid of lattice points in order to interpolate effectively. At present, we deal with this obstacle in the most obvious manner possible—by simply extending our ***Crystal***-truncated grid into a fully rectangular lattice, and assigning a constant energy value to all of the additional lattice points. This procedure is straightforward and fast—even when the fully rectangular grid contains many times the number of points as the truncated grid. (As a technical point, we

note that in practice, better results are obtained when the assigned energy value is chosen to be somewhat *higher* than $E_{\max}$ itself.)

Of course, the interpolation was handled in the above manner solely for expediency. Although it works well enough for the molecular systems considered here (up to 9D), this approach and its exponential scaling (albeit with a very small prefactor) is not a tenable one for larger systems. Going forward, there are a number of customized interpolation approaches that could be relied upon. Or conversely, if interpolation should prove too onerous, PES fitting is always an option—albeit one that requires more finesse, and is therefore less well-suited for "automatic" PES construction.

III. COMPUTATIONAL DETAILS

A. Water

As described in Sec. I, on-the-fly ***Crystal*** can be used to characterize structures such as TSs, as well as to construct PESs. For on-the-fly minima, it is generally more convenient to compute these directly via the electronic structure code itself, rather than using on-the-fly ***Crystal***, although the reverse may be true for on-the-fly TSs. Conventional (i.e. "non-on-the-fly") ***Crystal*** can, of course, also be used to find both minima and TSs for existing PES functions. For all of the molecular systems considered here, the grid of lattice points is chosen so as to always include the global minimum (or minima) explicitly. Since interpolation is used, this ensures that the PES function values are *exact* at the true PES minima (subject to the vagaries of the electronic structure methods employed...). Thus, there is never a need to use either conventional or on-the-fly ***Crystal*** explicitly to find these structures.

In the case of H_2O, there is, in addition, the TS structure representing the inversion barrier to isomerization. However, because this TS structure has a known linear symmetric geometry, it is once again easier to compute this structure via constrained optimization within the electronic structure code itself, rather than using on-the-fly ***Crystal***. That said, however, we do use conventional ***Crystal*** to compute the TS structures of the resultant interpolated H_2O PES functions. In any event, the various ground state minima and inversion barrier TS structures we computed for H_2O, as described above, may be found in Tables I and IV, respectively.

For H_2O PES construction purposes, rather than work directly in valence bond coordinates, we found it convenient to create a grid of points in the Cartesian-like internal

Table I: Global minimum structures for H_2O, as computed using DFT and CCSD electronic structure methods, and incorporated into corresponding interpolated PES functions. Structures are reported in three different coordinate systems: Cartesian, internal, and valence bond. Cartesian and radial coordinate values are presented in units of Å; angular coordinates are in radians. Some additional results from the literature are also provided.

Minimum geometries	Atom/Coord	DFT	CCSD
Our result	O	(0.000, 0.000, 0.000)	(0.000, 0.000, 0.000)
(Cartesian)	H_1	(-0.763, 0.585, 0.000)	(-0.756, 0.586, 0.000)
	H_2	(0.763, 0.585, 0.000)	(0.756, 0.586, 0.000)
Our result	H_{1x}	0	0
(internal)	H_{2x}	0	0
	H_y	± 0.58487	± 0.58490
Our result	r_1	0.9617	0.9562
(valence)	r_2	0.9617	0.9562
	θ	1.834	1.824
Ref. 23	r_1		0.958
(valence)	r_2		0.958
CCSD(T)/aug-cc-pVTZ	θ		1.821
Ref. 24	r_1		0.9572
(valence)	r_2		0.9572
Experiment	θ		1.824

coordinates, (H_{1x}, H_{2x}, H_y), indicated in Fig. 3. Note that H_{1x} and H_{2x} are defined as displacements relative to equilibrium, whereas H_y=0 corresponds to linear geometries with bend angle $\theta = \pi$. The (H_{1x}, H_{2x}, H_y) coordinate system can be easily converted to the more standard valence bond coordinate system, (r_1, r_2, θ), as indicated in Eqs. 1 to 3.

$$r_1 = \sqrt{(H_{1x} - x_0)^2 + H_y^2} \tag{1}$$

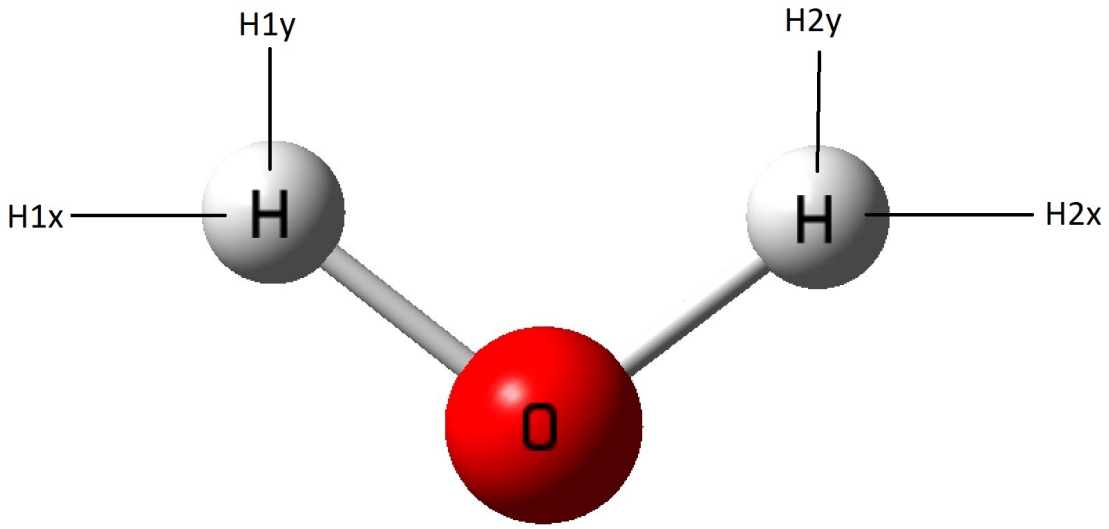
$$r_2 = \sqrt{(H_{2x} + x_0)^2 + H_y^2} \tag{2}$$

$$\begin{aligned} \theta &= \pi - \Phi_1 - \Phi_2, \\ \text{where} \quad \Phi_1 &= \arctan[-H_y/(H_{1x} - x_0)] \\ \text{and} \quad \Phi_2 &= \arctan[H_y/(H_{2x} + x_0)] \end{aligned} \tag{3}$$

In the equations above, $\pm x_0$ represents the equilibrium value of the Cartesian x coordinate—e.g., $x_0 = 0.756370$, in the case of CCSD (as indicated in Table I). In any event, all 3D PES lattice grids for H_2O were constructed using on-the-fly ***Crystal*** in conjunction with ***Gaussian 16***. Lattice spacings and maximum extents in each coordinate direction are as

indicated in Table III.

Figure 3: Internal coordinates used to construct lattice grids for both DFT and CCSD on-the-fly ***Crystal*** PESs for H_2O. Following the orientation of the figure, x and y refer to horizontal and vertical planar motions of the two H atoms, H_1 and H_2. The body-fixed frame is defined by constraining H_{1y} and H_{2y} to move together; we thus replace these with the single coordinate H_y.



Note that for all molecular systems, the coordinates and grids are defined such that the global minimum is a grid point. For H_2O, for which we construct symmetric double-well PES functions, *both* equivalent global minima are grid points. The geometry representing the origin of the (H_{1x}, H_{2x}, H_y) coordinate system is taken to be the *average* of the two minima structures, as represented in Cartesian coordinates. This is a linear symmetric structure, in terms of which the two minima are described in internal coordinates by (0, 0, $\pm$ $H_y^{\rm min}$), and the TS by (-$H_x^{\rm TS}$, $H_x^{\rm TS}$, 0). Note that $H_y^{\rm min}$ and $H_x^{\rm TS}$ values may be readily obtained from Tables I and IV, respectively, where the minima and TS structures are expressed in internal as well as valence bond coordinates.

In order to ensure that both minima are included as explicit lattice points in our grid, it is necessary to choose an H_y grid spacing Δ_y that divides into 2 $H_y^{\rm min}$. For both H_2O grids considered here, we choose $\Delta_y = H_y^{\rm min}/7$, which provides a good number of grid points in the H_y direction (i.e., 15 just from minimum to minimum). For the symmetric H_{1x} and H_{2x} directions, a sufficiently small grid spacing Δx was chosen so as to ensure good sampling of the PES in these directions, as indicated in Table III.

In comparing Δx grid spacing values between the CCSD and DFT lattices, precise values were finessed a bit, in order to ensure that the maximum extent of both grids (as measured by the number of lattice points in H_{1x} or H_{2x}) is the same. For the same reason, we also used slightly different energy cutoffs for the two lattices, i.e. $E_{\max} = 32{,}000$ cm^{-1} for the DFT lattice, and $E_{\max} = 33{,}000$ cm^{-1} for the CCSD lattice. (Note that the DFT dissociation threshold energy for H_2O is around 41,600 cm^{-1}.) With these choices, nearly 14,000 lattice points were obtained, for both the CCSD and DFT lattices.

B. Methane

Our second application system is a prototype for CH bond activation, or the "predissociation" dynamics in methane, i.e. $CH_4 \rightarrow CH_3 + H$. A 3D restricted PES is constructed, by allowing one of the hydrogens to move along all three Cartesian directions, (x, y, z). Note that this coordinate system is defined such that the C atom lies at the origin, and the moving H atom lies along the positive z axis when CH_4 is in its global minimum configuration. The other three H atoms (in the methyl group subunit) thus lie parallel to and below the x–y plane, with one of these lying in the x–z plane. Due to tetrahedral symmetry, the global minimum geometry is uniquely determined by specifying a CH bond length—which is found to be 1.089 Å, at the B3LYP/aug-cc-pVTZ level of theory. (For comparison, the experimental value is 1.0870 Å.[25]) Note that for simplicity, the methyl group subunit is frozen at its CH_4 minimum geometry, even as the mobile H atom is displaced, which is slightly unrealistic.

This system was chosen for several reasons. Firstly, as discussed in Sec. I, we envision on-the-fly ***Crystal*** being used to construct full-dimensional as well as restricted PESs, and so it is important to have at least one example of the latter. Secondly, we wish to demonstrate the applicability of on-the-fly ***Crystal*** across a variety of coordinate system types, so we wanted to consider a Cartesian coordinate calculation. Finally, the PES dimensionality for this system is the same as for the H_2O PES, and the electronic structure calculations are also isoelectronic. So this system serves as a good comparison with H_2O, to determine whether CPU costs are widely different.

In fact, they turn out to be comparable, although the lattice size is somewhat smaller, because CH_4 is a single-well system. Note that we do not extend the PES lattice all the way to the dissociation channel, for which the threshold energy is 36,700 cm^{-1} (according to the B3LYP/aug-cc-pVTZ level of theory). Instead, we truncate the lattice at $E_{\max} =$

28,000 cm^{-1}—which is very far into the anharmonic region, if still below dissociation. In any event, this choice gives rise to around 11,000 lattice points in all, which enables a very detailed sampling of the PES region of interest. Lattice spacings and maximum extents in each coordinate direction are as indicated in Table III. Other computational details are as discussed in Secs. II and III A.

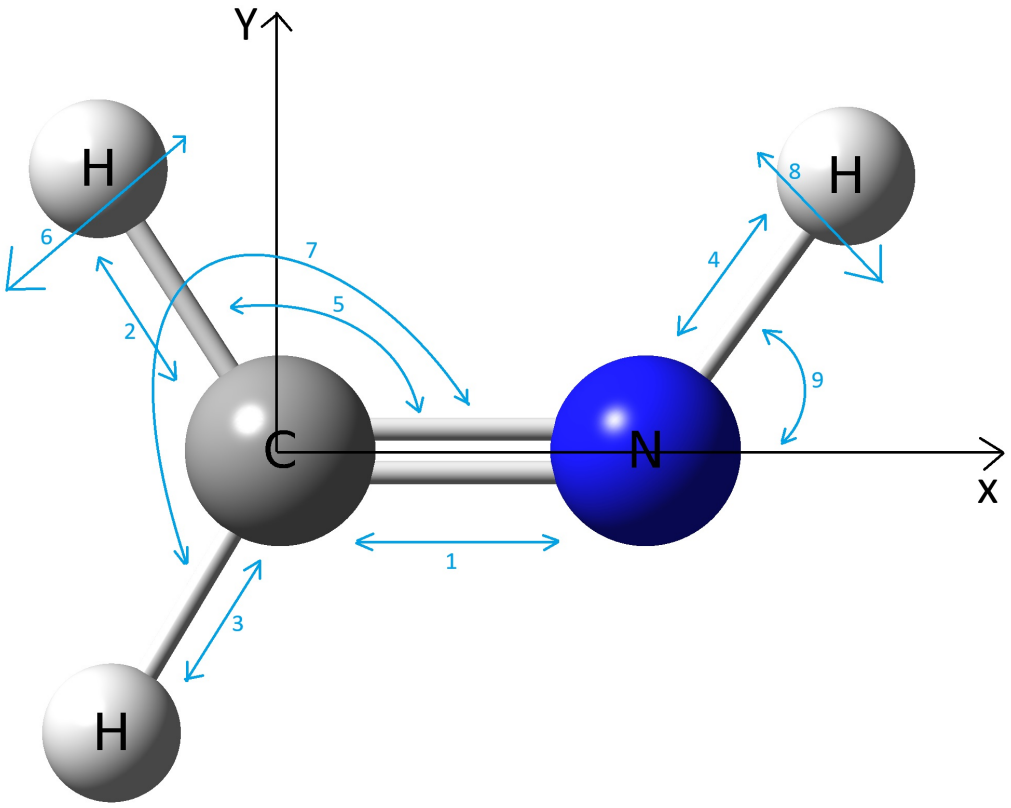
C. Methylene imine

Our largest and most challenging molecular application is methylene imine, CH_2NH, studied here in full dimensionality. Of course, the individual electronic structure calculations for CH_2NH are much more computationally expensive than for the other applications, and in addition, many more such calculations are required to flesh out the 9D lattice. In addition to these difficulties, the CPU cost of the PES interpolation itself also becomes non-trivial in this regime. Nevertheless, using DFT with on-the-fly ***Crystal***, and the other computational techniques as discussed, we were able to construct a viable 9D PES for this system—which is moreover, guaranteed to be hole-free, up to the chosen energy cutoff.

The global minimum structure of CH_2NH, as computed at the B3LYP/aug-cc-pVTZ level of theory, is presented in Table II, in both Cartesian and internal coordinates. This structure, together with the nine internal coordinates used to construct the 9D lattice and PES, is also indicated in Fig. 4. Four of these internal coordinates are radial stretch coordinates (i.e. q_1–q_4), three are essentially valence bond bend angle coordinates (i.e., q_5, q_7, and q_9), and the remaining two (i.e., q_6 and q_8) are Cartesian out-of-plane motions. Note that since the lower HCN subunit defines the body-fixed x–y plane, only the upper two H atoms, H_1 and H_3, may move out of plane. In any event, q_6 and q_8 are defined to be along the perpendicular z direction, rather than as true torsional angles.

Because of the aforementioned computational difficulty of this system, it was not possible to maintain a similar energy cutoff value as for the previously considered systems, while still keeping the same highly detailed level of PES region sampling. Consequently, $E_{\max}$ for this system was reduced down to 5,300 cm^{-1}, giving rise to a lattice of over 80,000 points. This is still sufficiently energetically high as to enable a detailed investigation of the most relevant vibrational dynamics of this system. In comparison, the only other published CH_2NH PES function of which we are aware[20] has a hole at just 2800 cm^{-1} above the ground state minimum![26] Additionally, the number of lattice points, or explicit electronic structure calculations, is greater here than what is typically found in the literature for

Figure 4: Internal coordinates used to construct 9D lattice for DFT PES for CH_2NH. Following the orientation of the figure, x and y refer to horizontal and vertical planar motions, so that z refers to out-of-plane motion. The methylene imine structure indicated is the ground state minimum. The body-fixed frame is defined by constraining the H-C-N subunit to lie in the x-y plane. Accordingly, only the upper two H atoms, i.e. H_1 (left) and H_3 (right), may move out of plane—corresponding to coordinates q_6 and q_8, respectively.



many constructed PES functions of similar system dimensionality. To this end, of course, the parallel version of ***Crystal*** was immensely helpful, reducing the CPU time that would otherwise be required by about an order of magnitude.

All of that said, in point of fact the computational bottleneck for this system was not the ***Crystal***/***Gaussian 16*** calculation per se, but rather, the subsequent PES interpolation procedure. This is discussed in more detail in the next subsection.

Table II: Global minimum structure for CH_2NH, as computed using DFT, and incorporated into corresponding interpolated PES function. Structure is reported in Cartesian and internal coordinates, in Col. IV. Cartesian and radial coordinate values are presented in units of Å; angular coordinates are in radians. Atoms and internal coordinates are defined in Cols. II and III.

Methylene imine	Atom No.	Atom	Coordinates
Minimum geometry (Cartesian)	1	C	(0.000, 0.000, 0.000)
	2	N	(1.264, 0.000, 0.000)
	3	H_1	(-0.624, 0.899, 0.000)
	4	H_2	(-0.528, -0.953, 0.000)
	5	H_3	(1.637, 0.950, 0.000)
	Coord.	Description	Value at minimum
Minimum geometry (internal)	q_1	r (1,2)	1.2636
	q_2	r (1,3)	1.0938
	q_3	r (1,4)	1.0892
	q_4	r (2,5)	1.0205
	q_5	θ (2,1,3)	2.1774
	q_6	z (3)	0
	q_7	$2\pi - \theta$ (2,1,4)	4.2066
	q_8	z (5)	0
	q_9	$\pi - \theta$ (1,2,5)	1.1966

Table III: Lattice spacings and coordinate ranges obtained in the construction of PES grids for all three molecular systems: water, methane, methylene imine. Units are as indicated in Col. II. Coordinate values for the grid points corresponding to ground state minima are also indicated, in Col. III.

Molecule	Coord (unit)	Value at minimum	Spacing	Min	Max
Water	H_{1x} (Å)	0.000	0.042	-0.798	0.84
	H_{2x} (Å)	0.000	0.042	-0.84	0.798
	H_y (Å)	$\pm$ 0.585	0.084	-1.169	1.169
Methane	x (Å)	0.000	0.10	-1.2	1.2
	y (Å)	0.000	0.10	-1.2	1.2
	z (Å)	1.089	0.05	0.389	1.939
Methylene imine	q_1 (Å)	1.264	0.05	1.164	1.414
	q_2 (Å)	1.094	0.05	0.944	1.294
	q_3 (Å)	1.089	0.05	0.939	1.289
	q_4 (Å)	1.020	0.05	0.920	1.220
	q_5 (rad)	2.177	0.18	1.817	2.537
	q_6 (Å)	0.000	0.225	-0.675	0.675
	q_7 (rad)	4.207	0.18	3.847	4.567
	q_8 (Å)	0.000	0.225	-0.675	0.675
	q_9 (rad)	1.197	0.18	0.657	1.557

D. Interpolation

As discussed in Sec. II D, we are currently using a canned *Interpolation* routine that requires extending our ***Crystal*** grid out into a fully rectangular lattice. The cost to evaluate the (constant) energy value for each new lattice point is trivial compared to each ***Crystal*** lattice point PES evaluation. Nevertheless, the *number* of such additional points grows exponentially with system dimensionality, such that before too long, the interpolation itself becomes the computational bottleneck. It is worth exploring this in some detail for our three application systems.

Starting with H_2O, we have seen that ***Crystal*** constructs a truncated grid of nearly 14,000 lattice points. However, upon extending the grid out to encompass the fully rectangular 3D lattice, the number of points increases to over 48,000. Similar behavior is observed for the 3D CH_4 PES, for which the total number of lattice points grows from ∼11,000 to ∼20,000—although the ratio here is a bit smaller than for H_2O, because the PES itself is more regular.

For 9D methylene imine, the grid limits in each coordinate direction are as indicated in Table III. Multiplying all nine 1D grid sizes together yields the total number of fully rectangular lattice points—approximately 3.5 million in this case. This represents an enormous increase from the original ∼81,000 truncated grid points—so much so that interpolation now becomes the bottleneck of the entire computation.

IV. RESULTS

A. Water

Two different H_2O PESs were constructed, one using DFT and one using CCSD. Both are symmetric double-well PESs that incorporate the linear inversion barrier to isomerization. To satisfy symmetry, the internal coordinate system origin is accordingly placed at a linear geometry (albeit not the TS), as discussed in Sec. III A.

In Fig. 5, we present 1D "slices" of the DFT H_2O PES, along each of the internal coordinate directions, (H_{1x}, H_{2x}, H_y). As expected, the H_{1x} and H_{2x} plots are mirror images of each other. Note that the flat regions on the right and left sides, respectively, of these two plots, arise because of the additional constant-energy points, added to form the fully rectangular lattice for interpolation purposes. The H_y plot is also fully symmetric, as expected, with both equivalent PES wells and global minima evident. Note that whereas the

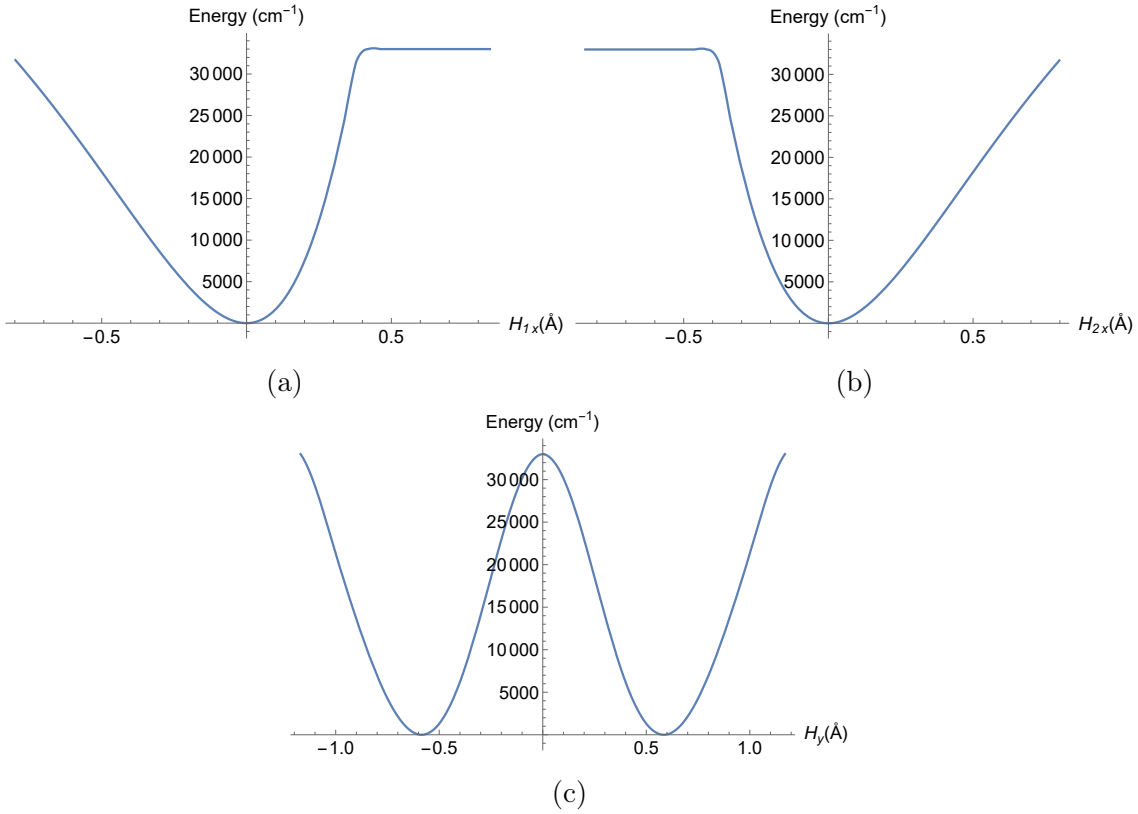


Figure 5: 1D "slice" plots of the DFT H_2O PES, taken along each of the three internal coordinate directions, keeping the other two coordinates fixed at their global minimum values: (a) H_{1x}; (b) H_{2x}; (c) H_y.

minima here really are the true global minima, the barrier at H_y=0 is by no means the true inversion barrier TS, but rather, the aforementioned origin of the coordinate system. The true inversion barrier height is therefore much lower than what might otherwise be inferred from this plot.

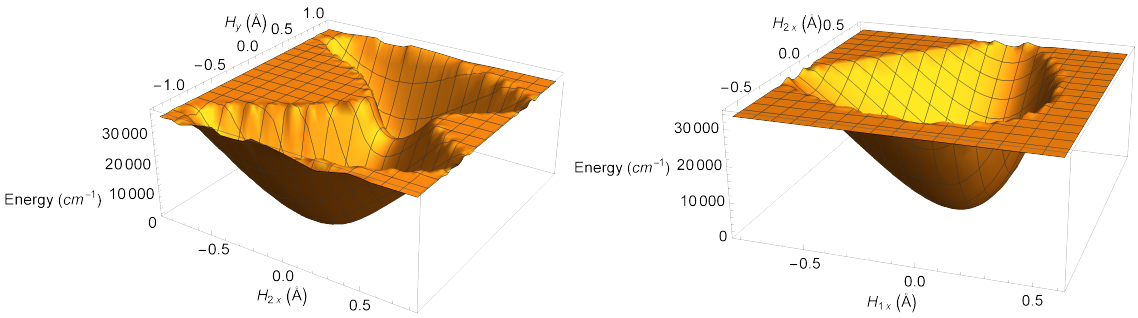


Figure 6: Surface plots of the DFT H_2O PES, taken along two of the three internal coordinate directions, keeping the third coordinate fixed at its global minimum value: (a) varying (H_{2x}, H_y) at fixed H_{1x}=0; (b) varying (H_{1x}, H_{2x}) at fixed $H_y = H_y^{\min}$=0.58487 Å.

To round out our characterization of the DFT H_2O PES, we also present several 3D plots. Fig. 6 presents surface plots for which one of the three internal coordinates is held constant at its equilibrium value. In particular, the left panel, corresponding to fixed $H_{1x} = 0$, shows both PES wells with very pronounced anharmonicity. Note that here too, the apparent TS, while lower than in Fig. 6, is not the true inversion barrier TS either.

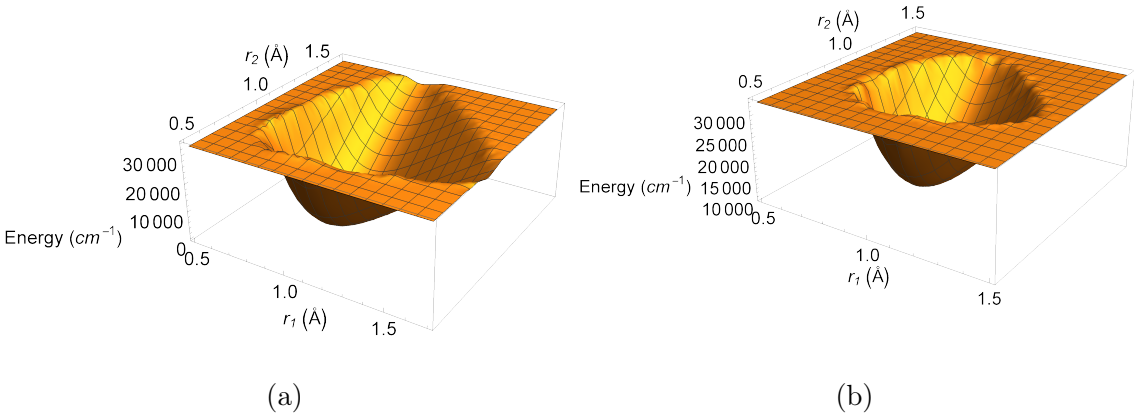


Figure 7: Surface plots of the DFT H_2O PES, taken along the r_1 and r_2 valence bond coordinate directions, keeping the bend angle fixed: (a) at the global minimum value, $\theta = 105.08°$; (b) at the TS value, $\theta = 180°$.

Finally, in Fig. 7, we present 3D plots in valence bond coordinates—specifically, in the two OH bond lengths, keeping the bond angle θ fixed. In the left panel, θ is set to the global minimum value, whereas in the right panel, θ is set to the TS value of 180°. The minimum in the latter plot therefore corresponds to the true inversion barrier TS.

As discussed, we have also computed a CCSD PES for H_2O. It is instructive to compare the DFT and CCSD PESs, in order to better comprehend the validity of the former—particularly in the barrier regions. Fig. 8 presents 1D "slice" plot comparisons in the internal coordinates, as in Fig. 5. To be sure, the overall trends between the two PESs are consistent, with the DFT and CCSD H_{1x} plots as presented in the upper part of the figure appearing very nearly identical. However, these plots are on slightly different scales, with the CCSD curve consistently higher than the DFT curve throughout. In the bottom panel, the two H_y plots are presented. Here, the rescaling is quite evident—although interestingly, the DFT curve becomes *higher* than the CCSD curve in the barrier region. Furthermore, the agreement appears no worse in the barrier region than elsewhere, and is in fact rather better here than in the asymptotic regions at comparable energy.

Although the apparent minima in Fig. 8 are true minima, the local maximum at H_y=0 in

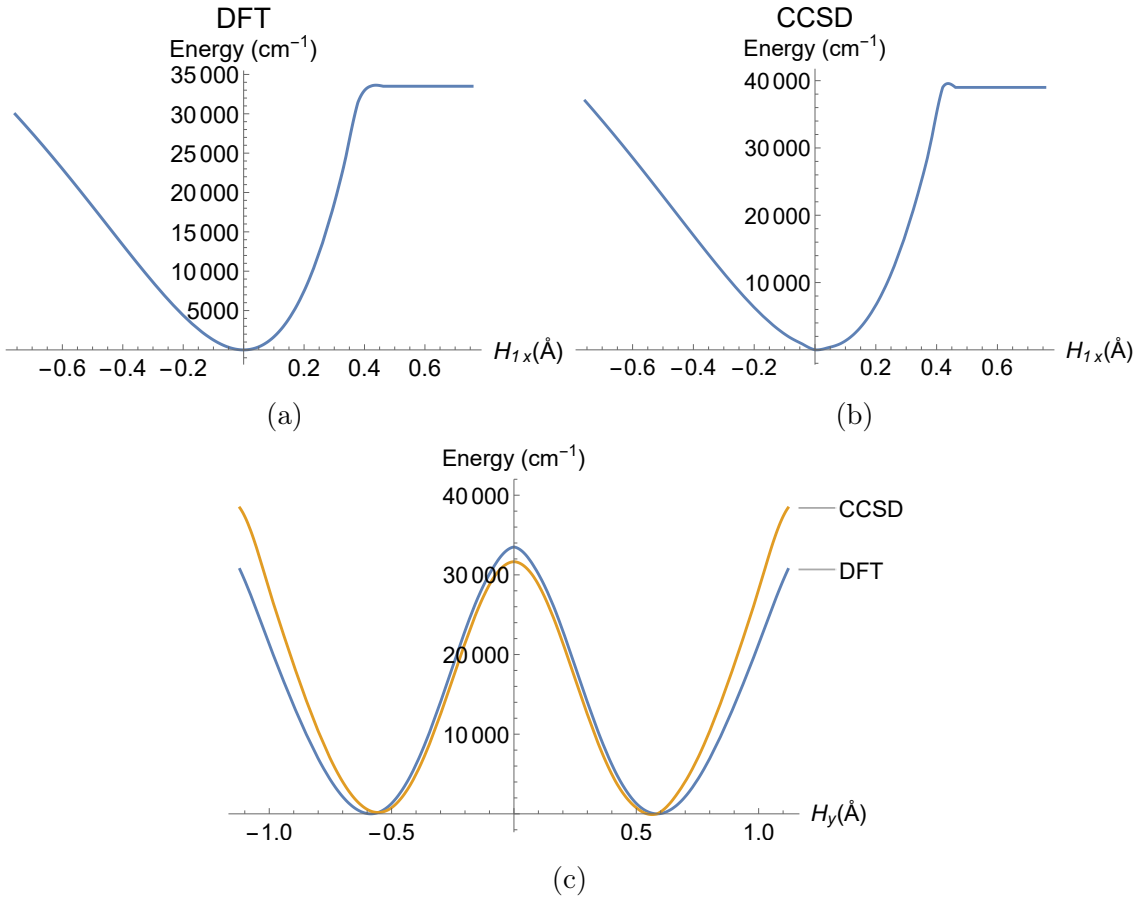


Figure 8: Comparison of DFT and CCSD H_2O PESs, as observed through 1D "slice" plots taken along one internal coordinate direction, keeping the other two coordinates fixed at their global minimum values: (a) H_{1x} plot for DFT PES; (b) H_{1x} plot for CCSD PES; (c) H_y plot for both DFT and CCSD PESs.

the bottom panel is *not* the true inversion barrier TS, as discussed. Of course, it is important to compare the true structures—not only as computed via DFT vs. CCSD directly in ***Gaussian 16***, but also with respect to the respective interpolated PES functions. By the same token, since H_2O is such a well-studied system, it is also worthwhile to compare our results with previous results from the literature.[23,24,27,28]

As discussed, all interpolated PES functions include the true minima as explicit lattice points; these structures are presented in Table I, and seen to show very close agreement between DFT and CCSD, although the CCSD bond lengths are a bit shorter. The table also presents minimum geometry data as obtained from a previous CCSD(T)/aug-cc-pVTZ calculation,[23] and from experiment.[24] Bond lengths for the former agree with our CCSD results extremely well—i.e., to less than .002 Å—although it should be noted that these results were obtained using a very similar level of electronic structure theory. Even so,

agreement between our results and experiment is even better—with bond lengths to within 0.001 Å, and the computed bend angle within .001 radians of the experimental value.

Table IV: Comparison of TS structures and energies for H_2O, as computed using DFT and CCSD electronic structure methods within ***Gaussian 16***, and compared with TSs for corresponding interpolated PESs functions. TS structures are reported in both internal and valence bond coordinates. Cartesian and radial coordinate values are presented in units of Å; angular coordinates are in radians; energies are in cm^{-1} relative to the corresponding ground state minima. Some additional results from the literature are also provided.

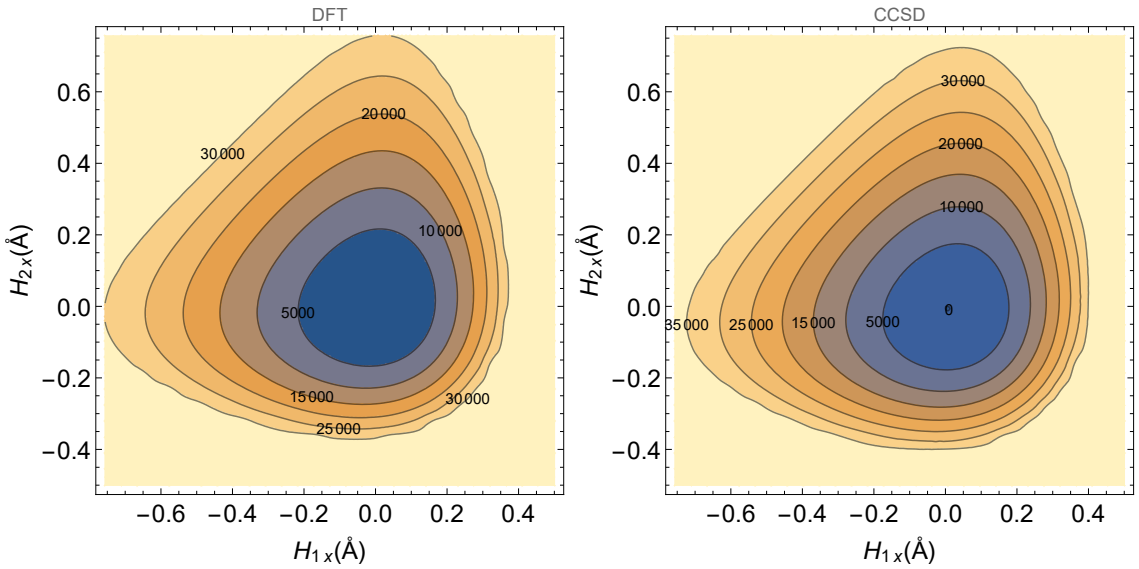
Inversion barrier TS geometries	Coord	DFT	CCSD
Our result	H_{1x}	-0.174	-0.166
(internal)	H_{2x}	0.174	0.166
Gaussian 16	H_y	0.000	0.000
Our result	r_1	0.937	0.932
(valence)	r_2	0.937	0.932
Gaussian 16	θ	3.142	3.142
Our result	H_{1x}	-0.174	-0.166
(internal)	H_{2x}	0.174	0.166
Interpolated PES	H_y	0.000	0.000
Our result	r_1	0.937	0.922
(valence)	r_2	0.937	0.922
Interpolated PES	θ	3.142	3.142
Ref. 23	r_1		0.934
(valence)	r_2		0.934
CCSD(T)/aug-cc-pVTZ	θ		3.142
Energy (Ours, ***Gaussian 16***)		10425.9	11254.7
Energy (Ours, Interpolated PES)		10424.9	11247.0
Energy (Ref. 23, CCSD(T)/aug-cc-pVTZ)			11269
Energy (Ref. 23, best)			11127

As for the inversion barrier TS structures, both the ***Gaussian 16*** and interpolated PES structures, for both DFT and CCSD, are presented in Table IV, in internal and valence bond coordinates. Here, too, our ***Gaussian 16*** geometries are quite similar, differing by only .005 Å (again, with CCSD a tiny bit shorter). The corresponding interpolated PES TS structures are interesting, in that the DFT case exhibits almost *perfect* agreement with the ***Gaussian 16*** result, whereas on the CCSD side, the agreement is good (i.e., to .01 Å) but certainly not perfect. That being the case, the spectacular DFT agreement may well be a fluke.

Ample theoretical data from the prior literature also exists for comparison of inversion barrier TS structures and energies—with Ref. 23 serving as one definitive resource on this topic. In Table IV, we present data on a previous CCSD(T)/aug-cc-pVTZ calculation, which is the closest we could find to our current CCSD calculation. As was the case for the similar

comparison with global minimum structures, the bond lengths agree to .002 Å. Barrier height energies are also in excellent agreement, to within 15 cm^{-1}. That said, our energy result is actually closer to the "best" theoretical result from Ref. 23 (also listed in the table), which incorporates extrapolation, core correlations, relativistic, diagonal Born-Oppenheimer, and other corrections.

We conclude this section with a contour plot comparison between the DFT and CCSD H_2O PESs, in the coordinates H_{1x} and H_{2x}, with H_y set to its global minimum value. The two contour plots, with labeled contours, are presented in Fig. 9. Once again, the two plots appear nearly indistinguishable, apart from a systematic rescaling of energy. All in all, we are pleased with the level of agreement between our DFT and CCSD interpolated PES functions, particularly given the global nature of the PESs, which incorporate extensive energy ranges, multiple equivalent global minima, inversion barrier TSs, and isomerization pathways (albeit not dissociation). The agreement with previous computational and experimental data is also very satisfactory.



(a) (b)

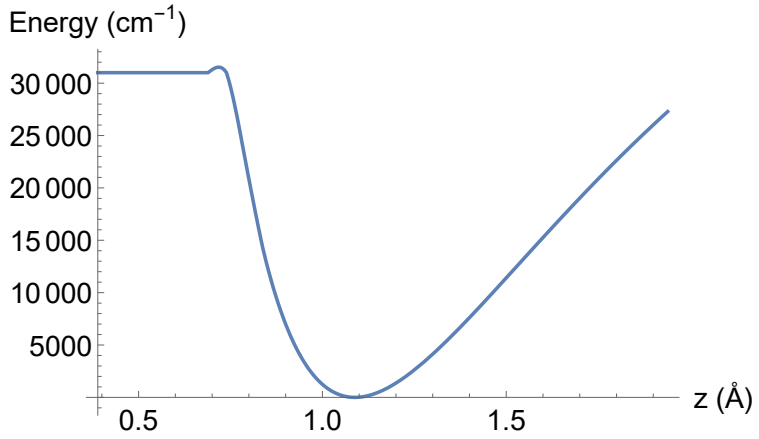
Figure 9: Comparison of DFT and CCSD H_2O PESs, as observed through contour plots taken along the H_{1x} and H_{2x} directions, keeping H_y fixed at its global minimum value: (a) contour plot for DFT PES; (b) contour plot for CCSD PES.

B. Methane

As discussed in Sec. III B, we have also constructed one restricted-coordinate DFT PES, describing activation/predissociation of a single CH bond in methane. Cartesian coordinates (x, y, z) are used for the construction of the lattice grid and interpolated PES function, with z representing the dissociation reaction coordinate. Although $E_{\max} = 28,000$ cm^{-1} was chosen somewhat below the dissociation threshold, it still permits a very broad latitude of motion for the mobile H atom. Moreover, the use of ∼11,000 lattice points for this relatively smooth 3D PES function ensures a very well-behaved and accurate interpolated PES function.

In Fig. 10 we present a 1D slice of our CH_4 PES, taken along the reaction coordinate z. As in the case of H_2O, the flat region at the smaller z values is due to the fully-rectangular extension of the lattice grid, needed for interpolation. As is evident from the figure, the minimum energy occurs around 1.089 Å, with the PES reaction profile exhibiting a typical bond-breaking softening pattern for larger z values. The large z end of the interpolated region signifies an energy an excess of 75% of the dissociation threshold.

Figure 10: 1D "slice" plot of the CH_4 PES, taken along the dissociation reaction coordinate, z, keeping $x = y = 0$ fixed at their global minimum values.



In Fig. 11 we present contour plots of the CH_4 PES in the "perpendicular " directions, (x, y), for several different fixed z values. For the most part, these show quite isotropic behavior, as might be expected. However, as the value of z decreases, the contour plots increasingly deviate from circular symmetry, becoming a bit more triangular in nature. Of course, this shift represents the increasing proximity to the methyl group subunit H atoms. In particular, the presence of one such H atom in the x-z plane implies the highest energies in the directions 0°, 120°, and 240°, and lowest energies in the directions 60°, 180°, and 300°. Indeed, this behavior is clearly evident in the third panel of Fig. 11.

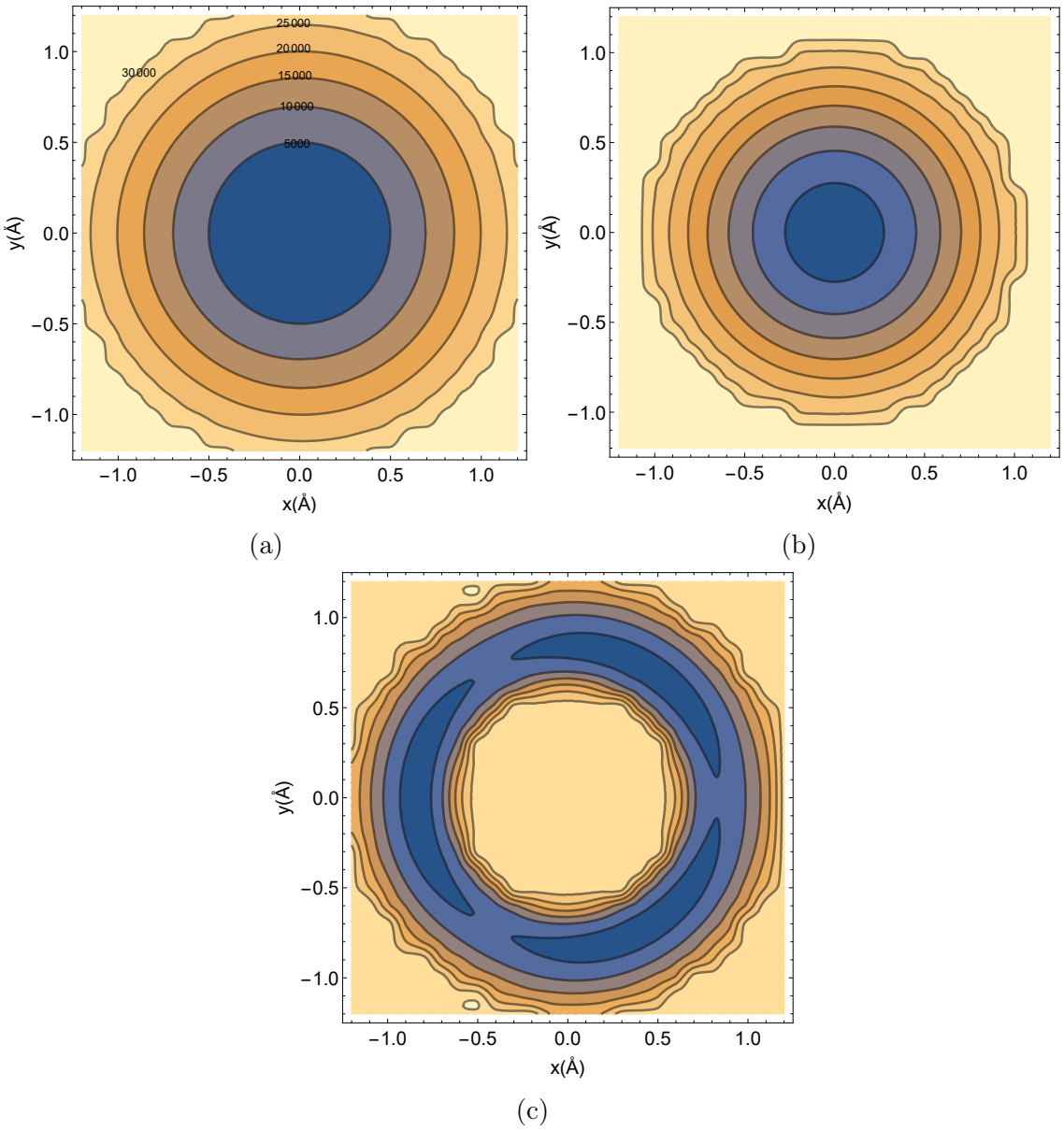
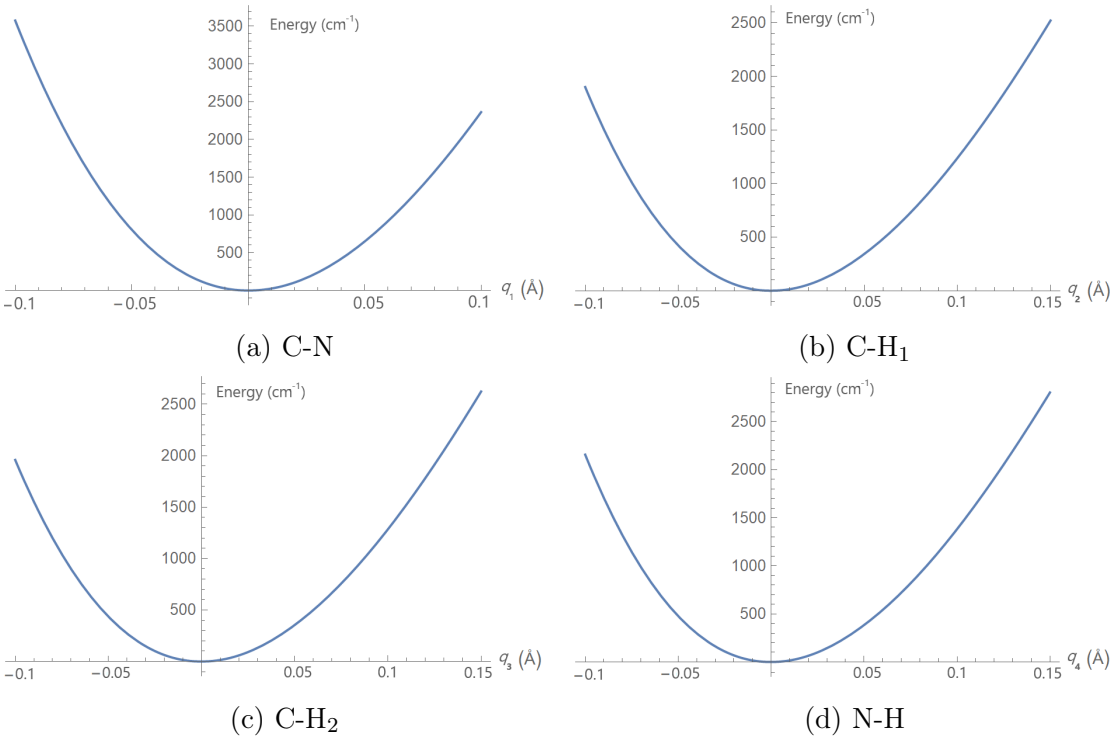


Figure 11: Contour plots of the CH_4 PES, taken along the perpendicular directions x and y, for several different values of the dissociation reaction coordinate, z: (a) $z = 1.089$ Å$= z^{\min}$; (b) $z = 1.489$ Å$> z^{\min}$; (b) $z = 0.589$ Å$< z^{\min}$.

C. Methylene imine

Methylene imine (CH_2NH) is the largest and the most challenging system considered in this study. Nevertheless, using the very straightforward methods discussed, we were able to construct a fully automated, hole-free, full-dimensional (9D) PES for this system. As a reminder, the coordinate system and ranges, lattice spacing, and minimum geometry, are presented in Fig. 4 and in Tables II and III. As discussed, over 80,000 truncated lattice points were used in the construction of the PES (i.e., prior to the extension to a fully rectangular

lattice). This enables quite decent sampling and smooth interpolation over the energy range considered—which is lower, and therefore less anharmonic, than for the other molecular systems of this work.



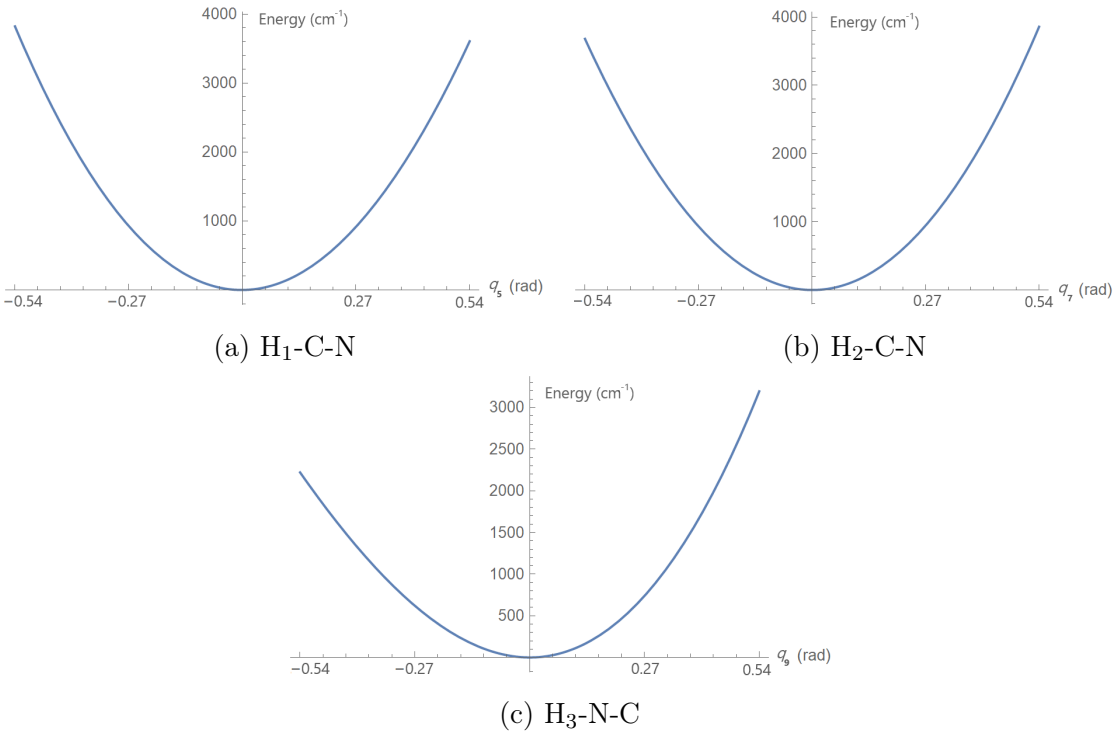
(a) C-N (b) C-H_1

(c) C-H_2 (d) N-H

Figure 12: 1D "slice" plots of the CH_2NH PES, taken along each of the four radial valence bond stretch coordinates, q_1–q_4, keeping all other internal coordinates fixed at their global minimum values. For plotting purposes, coordinates have been shifted so that the global minimum appears at the origin: (a) q_1 (C-N stretch); (b) q_2 (C-H_1 stretch); (c) q_3 (C-H_2 stretch); (d) q_4 (N-H stretch).

In Fig. 12, we present 1D slice plots of the CH_2NH PES, along each of the four radial valence bond stretch coordinates, i.e. q_1–q_4. Note that for plotting purposes, these coordinates have been shifted by their equilibrium values, so that the true minimum always appears at the origin. Note also that the apparent energy limits vary a bit—especially on the left or small-value side of the plots, where the PES curves are steeper. This is due to the vagaries of discrete grid truncation, which is more pronounced for larger system dimensionalities. Of course, in practice, we could extend all of these curves outward somewhat by extrapolating our interpolated PES function. This procedure will generate a smooth and well-behaved extension, and is reliable up to a point—presumably up to $E_{\max}$, anyhow—before running into the types of problems discussed in Sec. I. For the figure, however, we decided to present only "pure" curves, that end at the edge of the interpolation grid. From the figure, the latter three curves, which all describe H atom bond lengths, look very similar. Indeed, the

q_2 (C-H_1) and q_3 (C-H_2) curves are nearly identical, to within a few tens of cm^{-1} or so.



(a) H_1-C-N (b) H_2-C-N

(c) H_3-N-C

Figure 13: 1D "slice" plots of the CH_2NH PES, taken along each of the three valence bond bend-type angular coordinates, q_5, q_7, and q_9, keeping all other internal coordinates fixed at their global minimum values. For plotting purposes, coordinates have been shifted so that the global minimum appears at the origin: (a) q_5 (H_1-C-N bend); (b) q_7 (H_2-C-N bend); (c) q_9 (H_3-N-C bend).

Fig. 13 depicts 1D slice plots along the three valence bond bend-type angle coordinates, i.e. q_5, q_7, and q_9. Here too, the angles have been shifted to their equilibrium values for plotting purposes. From the figure, it is clear that the H_1 and H_2 bend profiles (i.e., the H atoms bound to the C atom) are much more symmetric than the H_3 bend profile—and also stiffer. Apart from an expected mirror reflection, these first two are again nearly identical to each other, to within 10 cm^{-1} or so—even closer than what was observed for the H_1 (q_2) and H_2 (q_3) stretch profiles already discussed. Finally, we have the out-of-plane z motions for H_1 (q_6) and H_3 (q_8), for which 1D slice plots are presented in Fig. 14. Note that both of these plots are perfectly symmetric, in their respective coordinates. However, the H_3 (q_8) profile is considerably stiffer than that of H_1 (q_6).

It is also worth examining mode- or coordinate-coupling, within our full 9D CH_2NH PES. Of course, it is only possible to plot the PES as a function of two coordinates at a time. For a 9D PES, there are 36 such coordinate pairs in all. We have made 3D and contour plots of all 36 coordinate pairs, in an attempt to identify those that are most strongly coupled

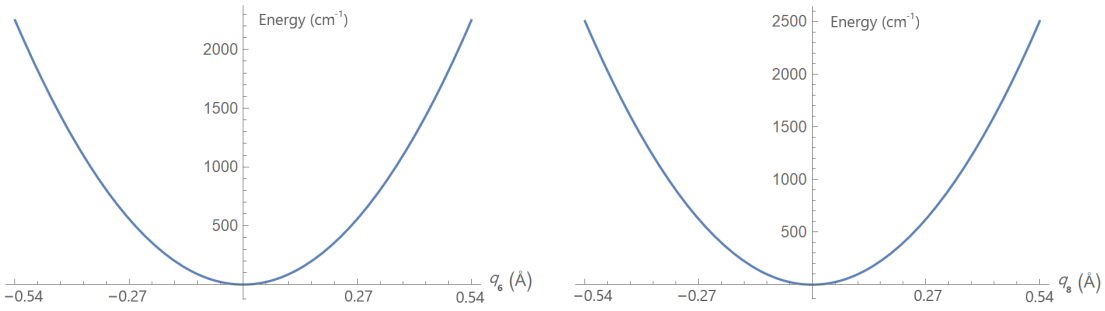
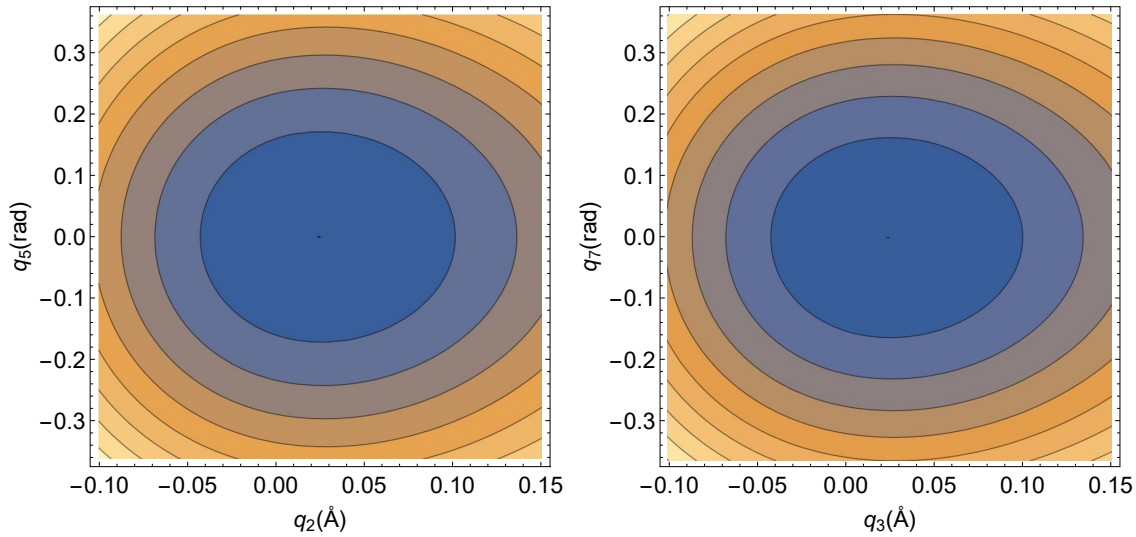


Figure 14: 1D "slice" plots of the CH_2NH PES, taken along each of the two Cartesian out-of-plane coordinates, q_6 and q_8, keeping all other internal coordinates fixed at their global minimum values. For plotting purposes, coordinates have been shifted so that the global minimum appears at the origin: (a) q_6 (H_1 z motion); (b) q_8 (H_3 z motion).

over the range of energies considered. Note that, whereas we might have used Hessians to determine the strongest mode couplings, in principle this procedure is only reliable in the vicinity of a minimum, and not throughout a given energy range.

Our approach, though it relies on qualitative visual analysis, is in any event unambiguous for the present CH_2NH PES. For this system, the most coupled coordinate pairs are clearly (q_2, q_5) and (q_3, q_7). In other words, the bend and stretch coordinates for each of the C-bound H atoms are the most strongly coupled. In Fig. 15(a), we present a contour plot for the CH_2NH PES with respect to the H_1 (q_2, q_5) coordinates, keeping all other coordinates fixed at their global minimum values. As can be seen from the figure, the "egg-shaped" contours are clear evidence of significant anharmonicity and/or coupling, which become more pronounced at higher energies. Fig. 15(b) presents a similar, and in fact nearly identical, contour plot for the H_2 (q_3, q_7) coordinates.



(a) (b)

Figure 15: Contour plots of the CH_2NH PES, taken along the valence bond bend and stretch coordinates for each of the two C-bound H atoms, keeping all other internal coordinates fixed at their global minimum values. For plotting purposes, coordinates have been shifted so that the global minimum appears at the origin: (a) H_1 atom, (q_2, q_5); (b) H_2 atom, (q_3, q_7).

V. CONCLUSIONS

On the time scale of the theoretical chemistry enterprise, the art of constructing potential energy surface (PES) functions for molecular systems may be regarded as "ancient." Well-known PES models such as Morse,[29] Eckart[30] and LEPS,[31] go back to nearly the earliest days of quantum mechanics, and predate even the development of *ab initio* electronic structure as a standalone discipline. For some decades, PES construction usually meant human beings fitting intuitive functional forms to experimental or theoretical data for three- and four-atom systems. Although such applications may be far from trivial, the configuration spaces (i.e., 6 dofs or fewer) are small enough so that any "surprises" that might arise can generally be dealt with relatively easily.

No more. The current research climate demands PES characterization and construction for much larger systems, and for novel environments and contexts, where traditional "chemical intuition" may be a far less reliable guide. Above all, the increased system dimensionalities of interest nowadays imply exponentially larger configuration spaces that can wreak havoc in all sorts of ways. To begin with, the *legitimate* chemical structures and pathways can become far more intricate and numerous, motivating the development of the "chemical dynomics" enterprise, as discussed in Sec. I. Additionally however, and even more

insidious, is the greatly increased likelihood that the constructed PES functions exhibit pathologically unphysical features—undetected and unsuspected by their creators. Indeed, one of the lessons learned from the early ***Crystal*** papers—now that the community has a reliable means of characterizing existing PES functions—is the ubiquitous presence of low-lying "holes," even in well-established PES functions.[4,5]

Whereas previous ***Crystal*** work has focused on characterizing existing PES functions—and in particular, on discovering previously undetected low-lying holes—the present work represents a significant departure, in that we now aim to characterize PESs, or construct new PES functions *without* recourse to an existing PES function. Instead, on-the-fly PES evaluations are conducted, during the course of the ***Crystal*** code execution itself. As with all on-the-fly methods, individual PES evaluations are typically far more costly than they would be with a preexisting function; consequently, viability demands some careful means of optimizing the choice of points to be evaluated. While many strategies exist,[11–14,17] ***Crystal*** is unique in offering a global exploration of precisely the region of configuration space that is needed (up to energy $E_{\max}$). As a consequence, the resultant automatically-generated PES is guaranteed to capture all possible dynamical pathways (whether envisioned *a priori* or not), while also avoiding the undesirable creation of spurious hole features.

The results of the present study indicate the viability of the on-the-fly ***Crystal*** PES construction approach across a variety of contexts—e.g., varying dimensionalities (up to 9D), multiple isomerization minima, DFT vs. *ab initio* electronic structure methods, restricted vs. full dimensionality, different coordinate types, etc. In all such contexts, the operation of on-the-fly ***Crystal*** was found to be sufficiently straightforward that the resultant PESs may truly be deemed to have been "automatically" generated. Although a large number of PES evaluations is typically required, especially for the larger system dimensionalities, the external electronic structure codes (***Gaussian 16*** in this case) are called automatically, without the direct involvement of a human user.

Likewise, the use of interpolation over fitting renders the final PES function construction fully automatic as well—although slightly more sophisticated methods than ***Mathematica***'s *Interpolation* routine will be needed going forward, in order to avoid the need for fully rectangular grids. Indeed, without the rectangular constraint, 9D is by no means the limit. For comparative purposes, we point out that serial ***Crystal*** has already been used to characterize existing PES functions up to 48D—with the parallel version expected to extend such calculations to perhaps ∼100D or so. So it is likely that creating new PES functions using

parallel on-the-fly ***Crystal***, without fully rectangular grids, will scale effectively up to at least several tens of dofs.

Of course, one important factor that must not be ignored in such assessments is the increasing cost with increasing system size of the electronic structure methods used—which are well known to follow very different scaling laws, depending on the level of theory. In particular, higher-level multi-reference methods may be desired, depending on one's objectives. Even for the relatively cheap methods used in this paper, we observe, e.g., that our 9D CH_2NH ***Crystal*** grid calculations would have taken over a week to complete using our serial code. The parallel version reduces this time to less than a single day, using only 8 nodes. As discussed, parallel scalability is nearly perfect, and increases quickly with system dimensionality. In theory, this parallel performance will be maintained so long as the overall CPU cost bottleneck is the PES evaluations. Consequently, we envision no practical limit, at present, on the size of the molecular systems that may be considered, provided that sufficient computational resources are available.

Many future developments are planned for on-the-fly ***Crystal***, both in terms of algorithms and applications. On the algorithmic front, in addition to the truncated-grid interpolation scheme already mentioned, we will seek to better exploit the natural symmetry of a given problem into the calculation itself. We also aim to pursue several ideas pertaining to still-better parallelization schemes, as discussed in the concurrent[6] and future papers. Additionally, more code development work is required to beef up the PES *characterization* (as opposed to construction) features of on-the-fly ***Crystal***. Several schemes are presented in Ref. 6, which emphasizes the PES characterization aspect of ***Crystal*** much more than this paper. There, too, we discuss another exciting new algorithmic development—enabling ***Crystal*** to characterize features of conical intersection seams.

Insofar as applications are concerned, we are currently identifying a set of molecular and chemical systems for which chemical intuition does *not* likely provide all the answers, and a more chemical dynomics approach is needed. This can be the case, either because the nature of the interactions or environment is non-standard, or simply because the number of available pathways is large. Rare gas clusters, with many equivalent minima and low-lying isomerization barriers, exemplify both criteria, and otherwise serve as an interesting and naturally scalable testbed.[32–37] Proton transfer is important in many biological processes,[38] yet is often difficult to model accurately; not only are there typically many potential reaction pathways to contend with, but quantum effects such as tunneling can play a role,

implying that higher-than-classical TS barriers must be considered. Regarding restricted PESs, hydrogen ligand dynamics and reactivity *within* transition metal polyhydride complexes present another interesting application area, relevant for catalysis and hydrogen fuel storage.[39–46] There are also many interesting applications from photochemistry as well. In the concurrent paper,[6] we explore photoswitches,[47,48] notably Photostatin-1[47], for which we will allow the methoxy groups to rotate in future work. Other photoswitches such as Donor Accepter Sternhouse Adducts[48] exhibit the kind of complex, multistep chemistry that would seemingly lend itself very well to a ***Crystal*** analysis.

The above, of course, represents nalbut a tiny sampling of possible application areas. Should the reader find himself or herself coming up with ideas of his or her own, we remind the reader that the latest ***Crystal*** source code and user's manual are always available from the authors on request. We also make ourselves available to offer our assistance in utilizing these resources effectively.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (CHE-1665370), and the Robert A. Welch Foundation (D-1523). The authors also gratefully acknowledge the Texas Advanced Computing Center for use of the Lonestar6 and Frontera clusters.

REFERENCES

[1]S. Maeda, K. Ohno, and K. Morokuma, "Systematic exploration of the mechanism of chemical reactions: the global reaction route mapping (GRRM) strategy using the ADDF and AFIR methods," Physical Chemistry Chemical Physics **15**, 3683–3701 (2013).

[2]F. Jensen, *Introduction to Computational Chemistry* (Wiley, Chichester, UK, 1999).

[3]A. Pandey and B. Poirier, "Using phase-space Gaussians to compute the vibrational states of $OCHCO^+$," The Journal of Chemical Physics **151**, 014114 (2019).

[4]A. Pandey and B. Poirier, "An algorithm to find (and plug) "holes" in multi-dimensional surfaces," J. Chem. Phys. **152**, 214102 (2020).

[5]A. Pandey and B. Poirier, "Plumbing potentials for molecules with up to tens of atoms: How to find saddle points and fix leaky holes," J. Phys. Chem. Lett. **11**, 6468–6474 (2020).

[6]A. Pandey, B. Poirier, and R. Liang, "Global exploration of conical intersection seam

space using crystal algorithm coupled with multireference electronic structure methods," J. Chem. Theory Comput. , (submitted) (2023).

[7]R. P. A. Bettens and M. Collins, J. Chem. Phys **111**, 816 (1999).

[8]M. A. Collins and D. H. Zhang, J. Chem. Phys **111**, 9924 (1999).

[9]T.-S. Ho and H. Rabitz, "RKHS interpolation," J. Chem. Phys. **104**, 2584–2597 (1996).

[10]T.-S. Ho, H. Rabitz, and G. Scoles, "RKHS interpolation," J. Chem. Phys. **112**, 6218–6227 (2000).

[11]R. Dawes, D. L. Thompson, Y. Guo, A. F. Wagner, and M. Minkoff, "Interpolating moving least-squares methods for fitting potential energy surfaces: Computing high-density potential energy surface data from low-density ab initio data points," The Journal of Chemical Physics **126**, 184108 (2007).

[12]E. Quintas-Sánchez and R. Dawes, "Autosurf: A freely available program to construct potential energy surfaces," J. Chem. Inf. Model. **59**, 262–271 (2019).

[13]J. Ischtwan and M. A. Collins, "Molecular potential energy surfaces by interpolation," J. Chem. Phys. **100**, 8080–8088 (1994).

[14]T. Győri and G. Czakó, "Automating the development of high-dimensional reactive potential energy surfaces with the robosurfer program system," J. Chem. Theory Comput. **16**, 51–66 (2020).

[15]J. Behler and M. Parrinello, "Generalized neural-network representation of high-dimensional potential-energy surfaces," Physical Review Letters **98**, 146401 (2007).

[16]S. Manzhos and T. Carrington Jr, "Using neural networks to represent potential surfaces as sums of products," The Journal of Chemical Physics **125**, 194105 (2006).

[17]B. Jiang and H. Guo, "Permutation invariant polynomial neural network approach to fitting potential energy surfaces," The Journal of Chemical Physics **139**, 054112 (2013).

[18]C. Petty, R. F. K. Spada, F. B. C. Machado, and B. Poirier, "Accurate rovibrational energies of ozone isotopologues up to j = 10 utilizing artificial neural networks," The Journal of Chemical Physics **149**, 024307 (2018).

[19]https://gaussian.com/gaussian16/.

[20]C. Pouchan and K. Zaki, J. Chem. Phys. **107**, 342 (1997).

[21]R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford Science Publications, UK, 1999).

[22]A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to advanced electronic structure theory* (Courier Corporation, 2012).

[23]G. Tarczay, A. G. Császár, W. Klopper, V. Szalay, W. D. Allen, and I. Schaefer, Henry F., "The barrier to linearity of water," The Journal of Chemical Physics **110**, 11971–11981 (1999).

[24]M. Huš and T. Urbic, "Strength of hydrogen bonds of water depends on local environment," The Journal of Chemical Physics **136**, 144305 (2012).

[25]E. Hirota, "Anharmonic potential function and equilibrium structure of methane," J. Mol. Spect. **77**, 213–221 (1979).

[26]T. Halverson and B. Poirier, J. Chem. Phys. **140**, 204112 (2014).

[27]T. Ho, T. Hollebeek, H. Rabitz, L. B. Harding, and G. C. Schatz, "A global H_2O potential energy surface for the reaction $O(1D)+H_2 \rightarrow OH+H$," The Journal of Chemical Physics **105**, 10472–10486 (1996).

[28]J. Sarka, B. Poirier, V. Szalay, and A. G. Császár, "On neglecting Coriolis and related couplings in first-principles rovibrational spectroscopy: Considerations of symmetry, accuracy, and simplicity. II. Case studies for H_2O isotopologues, H_3^+, O_3, and NH_3," Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy **250**, 119164 (2021).

[29]N. Rosen and P. Morse, Phys. Rev. **42**, 210 (1932).

[30]C. Eckart, Phys. Rev. **35**, 1303 (1930).

[31]S. Sato, J. Chem. Phys. **23**, 592, 2465 (1955).

[32]R. L. Johnston, *Atomic and Molecular Clusters* (Taylor and Francis, New York, 2002).

[33]S. Sugano and H. Koizumi, *Microcluster Physics* (Springer-Verlag, Berlin, 1998).

[34]D. M. Leitner, R. S. Berry, and R. M. Whitnell, "Quantum chaos of Ar_3: Statistics of eigenvalues," J. Chem. Phys. **91**, 3470 (1989).

[35]D. M. Leitner and J. D. Doll and R. M. Whitnell, "Quantum mechanics of small Ne, Ar, Kr, and Xe clusters," J. Chem. Phys. **94**, 6644 (1991).

[36]B. Yang, W. Chen, and B. Poirier, "Rovibrational bound states of neon trimer: Quantum dynamical calculation of all eigenstate energy levels and wavefunctions," J. Chem. Phys. **135**, 094306 (2011).

[37]J. Sarka, C. Petty, and B. Poirier, "Exact bound rovibrational spectra of the neon tetramer," J. Chem. Phys. **151**, 174304 (2019).

[38]E. by P. Ädelroth, "Proton Transfer Reactions in Biological Systems," Biochim. et Biophys. Acta–Bioenerg. **1757**, 867–1064 (2006).

[39]"Hydrogen storage," http://energy.gov/eere/fuelcells/hydrogen-storage.

[40]F. Maseras, M. Duran, A. Lledos, and J. Bertran, "Molecular hydrogen complexes with

a hydride ligand. An ab initio study on the iron hydride, $[Fe(PR_3)_4H(H_2)]^+$, system," J. Am. Chem. Soc. **113**, 2879–2884 (1991).

[41]L. S. Van Der Sluys, J. Eckert, O. Eisenstein, J. H. Hall, J. C. Huffman, S. A. Jackson, T. F. Koetzle, G. J. Kubas, P. J. Vergamini, and K. G. Caulton, "An Attractive "Cis-Effect" of Hydride on Neighbor Ligands: Experimental and Theoretical Studies on the Structure and Intramolecular Rearrangements of $Fe(H_2)(\eta^2\text{-}H_2)(PEtPh_2)_3$," J. Am. Chem. Soc. **112**, 4831–4841 (1990).

[42]N. Došlić, V. Gomzi, M. Mališ, I. Matanović, and J. Eckert, "Fluxionality of Hydrogen Ligands in $Fe(H)_2(H_2)(PEtPh_2)_3$," Inorg. Chem. **50**, 10740–19747 (2011).

[43]R. H. Morris, "Dihydrogen, dihydride and in between: NMR and structural properties of iron group complexes," Coord. Chem. Rev. **252**, 2381–2394 (2008).

[44]G. J. Kubas, R. R. Ryan, B. I. Swanson, P. J. Vergamini, and H. J. Wasserman, "Characterization of the first examples of isolable molecular hydrogen complexes, $M(CO)_3(PR_3)_2(H_2)$ (M = Mo, W; R = Cy, i-Pr). Evidence for a side-on bonded H_2 ligand." J. Am. Chem. Soc. **106**, 451–452 (1984).

[45]M. E. Gonzalez, J. Eckert, A. J. A. Aquino, and B. Poirier, "A quantum dynamical study of the rotation of the dihydrogen ligand in the $Fe(H)_2(H_2)(PEtPh_2)_3$ coordination complex," J. Chem. Phys. **148**, 154303 (2018).

[46]M. Aarabi, J. Sarka, A. Pandey, R. Nieman, A. J. A. Aquino, J. Eckert, and B. Poirier, "Quantum dynamical investigation of dihydrogen–hydride exchange in a transition-metal polyhydride complex," The Journal of Physical Chemistry A **127**, 6385–6399 (2023).

[47]M. Borowiak *et al.*, "Photoswitchable inhibitors of microtubule dynamics optically control mitosis and cell death," Cell **162**, 403–411 (2015).

[48]U. Raucci, D. M. Sanchez, T. J. Martinez, and M. Parrinello, "Enhanced sampling aided design of molecular photoswitches," J. Am. Chem. Soc. **144**, 19265–19271 (2022).