A new diabatization scheme for direct quantum dynamics: Procrustes diabatization

Gareth W. Richings[a] and Scott Habershon[b]
Department of Chemistry and Centre for Scientific Computing, University of Warwick, Coventry, CV4 7AL, UK

We present a new scheme for diabatizing electronic potential energy surfaces, for use within the recently implemented direct-dynamics grid-based (DD-GB) class of computational nuclear quantum dynamics methods (DD-SM and DD-MCTDH), called Procrustes diabatization. Calculations on the well-studied molecular systems LiF and the butatriene cation, using both Procrustes diabatization and the previously implemented propagation and projection diabatization schemes, have allowed detailed comparisons to be made which indicate that the new method combines the best features of the older approaches; it generates smooth surfaces which cross at the correct molecular geometries, reproduces interstate couplings accurately and hence allows the correct modelling of non-adiabatic dynamics.

I. INTRODUCTION

In recent years, simulating the non-adiabatic, nuclear quantum dynamics (QD) of molecules following photo-excitation has become an increasingly popular and powerful tool to complement experimental studies of, for example, organic chromophores,[1–5] organometallic complexes,[6–8], DNA bases,[9–11], and more.

A range of different non-adiabatic dynamics simulation methods are now available, with examples including trajectory surface hopping (TSH),[10,12–22] *ab initio* multiple spawning (AIMS),[11,23–25], multi-configuration Ehrenfest,[5,26,27] multi-configuration time-dependent Hartree (MCTDH),[1–4,28] multi-layer MCTDH[29–31] and variational multi-configuration Gaussian (vMCG).[32,33] These methods share the common goal of trying to determine the time evolution of a mixed electronic and/or nuclear wavefunction according to the time-dependent Schrödinger equation, but differ in the details of the representation of the wavefunction and its time-propagation, depending upon the rigour with which they treat quantum-mechanical effects such as tunnelling, zero-point energy conservation, and non-adiabatic transition between different electronic states. However, a feature which all share is the need to know the potential energy operator in the Hamiltonian, which is given in multi-state problems by a set of interacting potential energy surfaces (PESs); for QD methods where the nuclear wavefunction is expanded in a localised basis (*e.g.* TSH, AIMS and vMCG) only localised knowledge of the PESs is required, but for fully quantum-mechanical methods such as MCTDH, a global form of the PESs is required for all degrees-of-freedom under consideration.[1–4,28]

In traditional schemes for solving the time-dependent Schrödinger equation, a great deal of effort is required to generate a mathematically compact representation of the global PESs before one performs nuclear dynamics simulation.[1–4,7,8,34] However, much recent effort has gone into developing on-the-fly QD methods where PESs are generated during the propagation of the nuclear wavefunction[35–39] Over the last few years we have implemented this idea in the arena of grid-based QD methods, particularly MCTDH and the simpler standard method (SM), allowing the performance of fully quantum dynamics without the hassle of pre-fitting the PESs (termed DD-GB (grid-based) methods of which DD-SM and DD-MCTDH are variants) and have applied this approach to systems such as malonaldehyde,[40,41] salicylaldimine,[40–42] pyrazine,[41,42] butatriene[41,43] and molecular sunscreens.[44] Our on-the-fly fitting procedure uses kernel ridge regression (KRR)[45–56] to generate representations of the PES using electronic energies calculated at molecular geometries selected according to the location of the nuclear wavefunction; this strategy generates a global PES which is periodically updated as the wavepacket evolves.

Methods such as TSH[10,12–22] and AIMS[11,23–25] use the adiabatic representation of the relevant PESs, such that the PESs are energy-ordered at all points in configuration space, and transfer of nuclear wavepacket population between them is governed by vectorial non-adiabatic coupling terms (NACTs). As explained later, the adiabatic PESs are not smooth and the couplings are infinite at points where different electronic states become degenerate, so, to provide smooth surfaces amenable to a global fit with smooth functions, our strategy is to transform to the diabatic representation of the PESs prior to the generation of a global PES fit using KRR. Unfortunately, there is no way to perform this diabatization transformation exactly (such that all NACTs are exactly zero everywhere[57,58]), so we must instead resort to approximate diabatization methods which give smooth surfaces whilst reducing the troublesome NACTs as much as is possible.

Many diabatization schemes have been proposed to address this problem, with representative schemes based on integration of the NACTs,[43,59–69] block-diagonalization of the Hamiltonian matrix,[57,70–74] minimising the change in some selected molecular property of the states between geometries *e.g.* dipole moments[64] (plus quadrupoles[75]) or

[a] Electronic mail: G.Richings@warwick.ac.uk
[b] Electronic mail: S.Habershon@warwick.ac.uk

atomic transition charges,[76] or the state itself[77–80], the fourfold way,[81–83] or construction of regularized diabatic states.[84,85] There are also QD methods where diabatization is performed locally in configuration space before transforming back to the adiabatic representation in order to get NACTs.[86].

In this Article, we introduce a new diabatization scheme which is based on the notion of directly minimising the change in the electronic wavefunctions across configuration space using the solution to the orthogonal Procrustes problem;[76,87–91] our scheme is directly compatible with our recent development of direct-dynamics methods, but is not limited to this application. To assess the new diabatization scheme, we perform a detailed analysis of the properties of the generated diabatic states in the case of LiF[64,77,92–94] and the butatriene cation,[37,41,43,66,95–98] both of which have well-known state crossings in the diabatic representations. We also confirm that the wavepacket dynamics generated by the new diabatization scheme is consistent with previous results for the butatriene cation. To aid in this assessment, we also perform simulations using two recently-developed diabatization schemes which have also been used in the context of DD-GB methods, namely projection diabatization and propagation diabatization. We conclude that the new diabatization scheme developed here combines several advantages of these previously-used diabatization schemes; overall, the combination of the new Procrustes diabatization scheme with our on-the-fly DD methods is a promising tool for highly-accurate non-adiabatic simulations of molecular quantum dynamics.

II. METHODOLOGY

A. Grid-Based Quantum Dynamics: Standard Method

The standard grid-based method for QD is very well established,[28] so we only give a brief summary relevant to the calculations discussed herein. Considering a molecular system with f degrees-of-freedom, the nuclear wavefunction may be represented as a linear expansion of products of *time-independent* basis functions, the coefficients of each product, $C_{j_1,\cdots,j_f}$, being time-dependent and complex. The basis functions used in this work are in the highly localised and orthonormal discrete variable representation (DVR),[28] hence the description of the expansion as a grid. A nuclear wavepacket moving on electronic state, s, has the form

$$\begin{aligned}\Psi^{(s)}(q_1,\cdots,q_f,t) &= \sum_{j_1=1}^{N_1}\cdots\sum_{j_f=1}^{N_f} C^{(s)}_{j_1,\cdots,j_f}(t)\prod_{\kappa=1}^{f}\chi^{(\kappa)}_{j_\kappa}(q_\kappa)\\ &= \sum_{J} C^{(s)}_{J}(t)\, X_J(\mathbf{q})\end{aligned} \tag{1}$$

where, for clarity, we have introduced a compound index, $J{=}j_1,\cdots,j_f$. For a system with N_s orthonormal electronic states, the total wavefunction is defined as

$$|\Psi\rangle = \sum_{s=1}^{N_s}|\Psi^{(s)}\rangle|s\rangle. \tag{2}$$

The total Hamiltonian may be written as

$$\hat{H} = \sum_{su}^{N_s}|s\rangle\hat{H}^{(su)}\langle u|, \tag{3}$$

so by use of the Dirac-Frenkel variational principle,[99,100] a set of coupled equations-of-motion can be derived for the expansion coefficients:

$$i\hbar\dot{C}^{(s)}_J = \sum_{u=1}^{N_s}\sum_{L}\langle X_J|\hat{H}^{(su)}|X_L\rangle C^{(u)}_L. \tag{4}$$

Given an appropriate set of initial conditions, these equations-of-motion are integrated so as to follow the time evolution of the wavefunction. Using a well-chosen basis of sufficient size, a numerically exact propagation of the wavepacket according to the time-dependent Schrödinger equation is obtained for the given Hamiltonian.

B. PESs for Grid-Based Wavefunction Propagation: Kernel Ridge Regression

To create PESs on-the-fly during wavefunction propagation, we use the machinery of KRR. Our overall simulation strategy has been explained in detail in earlier work,[40–44] so we give only brief, relevant details here.

Given a set of M molecular geometries, $\{\mathbf{q}_k\}$, at which the *diabatic* electronic energies and couplings have been calculated, we define a kernel in the degrees-of-freedom under consideration (*e.g.* normal modes) as

$$k(\mathbf{q},\mathbf{q}_i) = \exp(-\alpha|\mathbf{q}-\mathbf{q}_i|^2) \tag{5}$$

where α is a parameter defining the width. Using KRR, each element of the diabatic PES matrix can be represented as

$$V^{(su)}(\mathbf{q}) \approx \sum_{i=1}^{M} w_i^{(su)} k(\mathbf{q},\mathbf{q}_i). \tag{6}$$

The expansion weights, $\{w_i^{(su)}\}$, are determined by solving the linear equation

$$\mathbf{K}\mathbf{w}^{su} = \mathbf{b}^{su}, \tag{7}$$

where the weights are included in the vector $\mathbf{w}^{su}$, the elements of the covariance matrix, $\mathbf{K}$, are[46]

$$K_{ij} = k(\mathbf{q}_i,\mathbf{q}_j) + \gamma^2\delta_{ij}, \tag{8}$$

with γ^2 being a small, positive regularization parameter (set here to 10^{-8}) and

$$b_i^{(su)} = V^{(su)}\left(\mathbf{q}_i\right), \tag{9}$$

the actual value of the potential at $\mathbf{q}_i$. Using a DVR basis in a quantum dynamics simulation, we can exploit the useful property that

$$\langle X_J|\hat{V}^{(su)}|X_K\rangle \approx V^{(su)}\left(\mathbf{q}_J\right)\delta_{JK}. \tag{10}$$

As suggested in the Introduction, in order to fit global PESs on-the-fly we must sample relevant geometries at which to evaluate the electronic energies. This is done by Sobol sampling[41,101–103] within a given distance of the wavepacket centre along all degrees-of-freedom on each state at pre-defined time intervals. To reduce the computational effort we evaluate the KRR variance function[46]

$$\sigma^2\left(\mathbf{q}\right) = k(\mathbf{q},\mathbf{q}) + \gamma^2 - \mathbf{k}^{\mathrm{T}}\mathbf{K}^{-1}\mathbf{k} \tag{11}$$

at each sampled geometry, where $\mathbf{k}$ contains the covariances, $k_i = k(\mathbf{q},\mathbf{q}_i)$, between the point of interest and all other points previously sampled. If the variance, $\sigma^2\left(\mathbf{q}\right)$, is less than a user-defined threshold parameter, the PES fits at $\mathbf{q}$ are assumed to be sufficiently accurate, so further electronic energy evaluations are not required. If $\sigma^2\left(\mathbf{q}\right)$ exceeds the threshold parameter, the new state energies are evaluated and added to the database of points, increasing the accuracy of the global fit in the region.

Using KRR allied with the SM we can efficiently sample configuration space to generate an accurate fit of the PESs in regions visited by the wavepacket. We also note that all energies and geometries are saved in a database, so can be re-used (and augmented) in subsequent calculations.

C. Diabatic States

In the discussion above, we have assumed that the KRR fitting, and eventual wavepacket propagation, is performed in the diabatic electronic state representation, giving rise to PES elements $V^{(su)}(\mathbf{q})$. However, the electronic energies and wavefunctions obtained from standard *ab initio* electronic structure codes such as Molpro[104] are ordered by energy and form an adiabatic representation of the PESs, with a series of non-crossing surfaces. The coupling between these states is a result of their non-zero gradients with respect to the nuclear coordinates; for adiabatic states ψ_i and ψ_j, where $i \neq j$, this NACT is a vector of the following form:

$$\begin{aligned}\mathbf{F}_{ij} &= \frac{\langle\psi_i|\nabla\hat{H}|\psi_j\rangle}{V_{jj}^A - V_{ii}^A} \\ &= \langle\psi_i|\nabla\psi_j\rangle \\ &= -\langle\nabla\psi_i|\psi_j\rangle,\end{aligned} \tag{12}$$

where V_{ii}^A and V_{jj}^A are the adiabatic energies of the corresponding states and ∇ is the gradient operator with respect to the nuclear coordinates. From the first line of Eq. (12), it is clear that the vector NACT has a pole at nuclear geometries where $V_{ii}^A = V_{jj}^A$, meaning that the magnitude of the coupling becomes infinite at such points of degeneracy. From the second and third lines of Eq. (12) we see that the infinite coupling is equivalently caused by a discontinuity in the gradient of both states at that point, leading to the cusp-like shape of the PESs at conical intersections.

When performing grid-based QD calculations, a global fit to the PES is required (whether fitted prior to, or during, the running of the dynamics); PESs with infinite couplings and discontinuous derivatives are not easily represented with a reasonable number of well-behaved functions, so it is desirable to transform to a diabatic representation (although we note that machine learning approaches have been used to fit adiabatic states in the context of TSH dynamics[53–56]). Specifically, the aim of diabatization is to transform the adiabatic states into a set of diabatic states in which the vector NACTs are eliminated. The resulting, diabatic PESs can now cross one another smoothly at points of degeneracy, with the interstate couplings being represented by finite, potential-like terms. As a result, diabatic PESs and couplings lend themselves well to expansion in terms of well behaved analytic functions, such as those used in KRR.

To fit a diabatic PES on-the-fly, it is necessary first to transform the adiabatic energies calculated at each sampled geometry to the diabatic representation. The key part of this process is determining the unitary,[59] coordinate-dependent adiabatic-diabatic transformation (ADT) matrix, $\mathbf{A}$, at each point, $\mathbf{q}$, such that

$$\mathbf{V}^{\mathrm{D}}(\mathbf{q}) = \mathbf{A}^{\mathrm{T}}(\mathbf{q})\mathbf{V}^{\mathrm{A}}(\mathbf{q})\mathbf{A}(\mathbf{q}), \tag{13}$$

where $\mathbf{V}^A$ and $\mathbf{V}^D$ are the adiabatic and diabatic energy matrices respectively (the former being diagonal, the latter not diagonal). Once $\mathbf{V}^D$ has been determined, its elements can be used in fitting the PESs and diabatic couplings (the off-diagonal terms in $\mathbf{V}^D$) using the KRR procedure outlined in section II B.

An important point to note is that strict diabatic states, where all NACTs are completely eliminated, can only be found in special cases, for example for diatomic molecules.[58] For polyatomic systems one can eliminate the couplings by using a crude adiabatic basis where the electronic basis is taken as being the set of adiabatic states at a single reference geometry. This basis is then used to expand the total wavefunction at all geometries and, because it is a constant basis, the NACTs vanish. However, as the adiabatic functions do change with geometry, a larger and more computationally-demanding crude adiabatic basis is needed to represent the PES than if a more flexible basis were used.[105] The crude adiabatic basis has been used successfully in a local sense with trajectory-based dynamics methods[106–109], the adiabatic states at the geometry at the beginning of the timestep for each trajectory being used as the quasi-diabatic basis in which quantities are expanded for the rest of the

timestep, at the end of which a new crude adiabatic basis is defined. Useful as these methods are, however, their locality means that they cannot be used in our case where the full expansion of the nuclear wavefunction requires a global diabatic representation with a single gauge.

Alternatively to the crude adiabatic basis a complete manifold of states is needed for polyatomic systems,[58] but this condition means that an infinite number of states would need to be included in order to remove all NACTs;[58] this is clearly impossible in practice. As such, the two formal ways of exactly diabatizing PESs for polyatomics are impractical in our case, so we have to accept that we will be unable to reduce the NACTs to exactly zero, with the remainder terms being non-removable couplings. Fortunately, around points of degeneracy, the singular coupling is removable[110] and the non-removable parts are insignificant.[111] Overall we are then left with a set of finite, residual NACTs which are usually safely ignored along the lines of the Born-Oppenheimer approximation due to the terms in the transformed Schrödinger equation being multiplied by a coefficient inversely proportional to the nuclear masses.[110,112] The upshot is that the states we obtain by calculating the ADT matrix are only *quasi-diabatic*; however, for brevity we will refer to the states we calculate simply as diabatic states.

To date, our DD-GB simulation approach has employed two different diabatization schemes to enable on-the-fly non-adiabatic simulations; in the following sections, we will recap these two diabatization schemes, and highlight some of their weaknesses. We then describe the new Procrustes diabatization scheme, and how it offers advantages over the previous approaches.

1. Propagation diabatization

The first diabatization scheme implemented for the DD-GB class of methods using KRR-fitted PESs was the propagation diabatization method,[43] which was originally used by one of us in the *Quantics* package to diabatize states for use with the DD-vMCG approach.[66,98] This earlier implementation used modified Shepard interpolation to fit the diabatic PES as a weighted sum of local harmonic approximations, and hence required diabatic gradients and Hessians (transformed from the adiabatic quantities available from many electronic structure packages). However, our KRR fitting procedure only uses energy information so we use a modified algorithm which is outlined in the supporting information (SI) for completeness. We return to this difference after a brief description of the theory behind the method.

As the name suggests, the propagation diabatization method relies on the propagation of the ADT matrix[65] from molecular configuration to molecular configuration. This method uses the approximate differential equation[58–60]

$$\nabla \mathbf{A} \approx -\underline{\mathbf{F}}\mathbf{A}, \tag{14}$$

where $\underline{\mathbf{F}}$ is a square matrix, the elements of which are the NACTS, $\mathbf{F}_{ij}$. For a complete set of states this expression is an exact equality, the inexactness here reflecting the presence of non-removable couplings. However, as the main aim of this diabatization scheme is to smooth out the PESs in the region around points of degeneracy where the removable couplings dominate, we pragmatically treat Eq. (14) as an equality in order to determine approximate ADT matrices. Doing so, and assuming we know $\mathbf{A}$ at some geometry, $\mathbf{q}$, then the ADT matrix at some new point, $\mathbf{q}+\boldsymbol{\Delta}\mathbf{q}$, is[59,65]

$$\mathbf{A}\left(\mathbf{q}+\boldsymbol{\Delta}\mathbf{q}\right) = \exp\left(-\int_{\mathbf{q}}^{\mathbf{q}+\boldsymbol{\Delta}\mathbf{q}} \underline{\mathbf{F}}\cdot \mathrm{d}\mathbf{q}\right)\mathbf{A}\left(\mathbf{q}\right). \tag{15}$$

Evaluation of the integral in Eq. (15) allows us to propagate $\mathbf{A}$ from configuration to configuration.

The assumption that we know $\mathbf{A}$ at the initial geometry is justified by considering Eq. (14). $\mathbf{A}$ is a unitary matrix, dependent on the molecular geometry but we are free to choose any constant, unitary matrix, $\mathbf{A}_0$, and make the replacement in Eq. (14), $\mathbf{A} \rightarrow \mathbf{A}' = \mathbf{A}\mathbf{A}_0$, and the (approximate) equality still holds. This being so, we are free to choose a reference geometry, $\mathbf{q}_0$, where $\mathbf{A}(\mathbf{q}_0) = \mathbf{I}$ (*i.e.* formally $\mathbf{A}_0 = \mathbf{A}^\dagger(\mathbf{q}_0)$). In practice we choose $\mathbf{q}_0$ as the centre of the initial wavepacket and because we know $\mathbf{A}$ at one point, we can calculate it everywhere using Eq. (15).

We should also note other sources of error in this approach here; we are assuming that Eq. (14) is an equality for practical purposes, but as this is not actually so, the integral in Eq. (15) is path dependent[58,59,62] meaning that $\mathbf{A}(\mathbf{q})$ depends on the path taken to reach it. This means $\mathbf{A}(\mathbf{q})$ can be altered by the choice of the shape of path between $\mathbf{q}$ and $\mathbf{q}+\boldsymbol{\Delta}\mathbf{q}$ (as noted in the SI, we choose a straight line in configuration space to take the shortest path and hence minimise the error in the numerical integral), and also by the choice of the previous points in the propagation. This means that $\mathbf{A}(\mathbf{q})$ is the result of a series of propagation steps starting from $\mathbf{q}_0$: from $\mathbf{q}_0$ we propagate to $\mathbf{q}_a$, then to $\mathbf{q}_b$ and so on before reaching $\mathbf{q}$. However, propagating from $\mathbf{q}_0$ to $\mathbf{q}_{a'}$ to $\mathbf{q}_{b'}$ etc. will, in general, give a different $\mathbf{A}(\mathbf{q})$. In the SI we give the method for choosing this path uniquely, but briefly, the path is chosen by finding the nearest point to $\mathbf{q}$, which is also nearer to $\mathbf{q}_0$ than $\mathbf{q}$ is (and indeed may be $\mathbf{q}_0$ itself). In this way the propagation radiates out from the reference geometry, $\mathbf{q}_0$. This choice of path does not necessarily mean that the path from $\mathbf{q}_0$ to $\mathbf{q}$ is the shortest possible (*i.e.* a straight line between the two), but means that the path passes through other points where $\underline{\mathbf{F}}$ has been calculated exactly rather than having to be approximated for the whole path between the end points.

This brings us to the second, potential source of inaccuracy; the need to numerically integrate the NACTs in Eq. (15). We use the NACTs in the form of the first line in Eq. (12), interpolating the numerator and denominator separately. The energies in the denominator are

fitted by a KRR of all available adiabatic energies (including from points previously diabatized or those awaiting diabatization). The numerator is expected to be well behaved[110] and so is simply linearly interpolated between the NACTs calculated at $\mathbf{q}$ and $\mathbf{q}+\mathbf{\Delta q}$. Using these two fits, a trapezium rule integration is used in Eq. (15). We have already stated that one of the reasons for using the diabatic representation is to allow an accurate fit of the PES in terms of the smooth functions used in the KRR fit so as to avoid fitting around the cusp-like shape of the adiabatic PES at points of degeneracy, which we are doing here. It follows that the accuracy of the integral will be most affected in the region where we most need it, around conical intersections. As noted in our earlier work, a linear interpolation of the energy would be even worse.[41]

It is also necessary to carefully account for the signs of the NACTs; because the overall signs of all coupled states are arbitrary, it is possible for the sign of a NACT to suddenly change when comparing the vectors at $\mathbf{q}$ and $\mathbf{q}+\Delta\mathbf{q}$ (rather than a natural, smooth change as a function of coordinate). This has potential to yield completely incorrect results for the integral of Eq. (15). In our implementation of the propagation diabatization scheme, we overlap the electronic wavefunctions at the two geometries and consider the resulting sign; if the overlap of the states is less than zero, then the sign of the state and all of its associated NACTS is changed. The method of overlapping the wavefunctions is described in the SI (Section I) as is the exact method for determining whether a state has changed sign (Section II).

2. Projection diabatization

In an attempt to overcome the deficiencies in the propagation diabatization scheme, we recently implemented[41] a second approach, namely the projection diabatization method of Robertson *et al.*[77]. The central idea underpinning this approach is to mix adiabatic states in such a way that the rate of change of the states with respect to the nuclear coordinates is minimised, and hence the NACTs are minimised. This idea has a long pedigree,[78–80] reflecting the intuitive picture that a diabatic representation is one where the character of the states (*e.g.* ionic, covalent *etc.*) does not change as a function of nuclear coordinate. As mentioned above, this may be trivially achieved by using a crude adiabatic basis where the electronic basis is taken as being the set of adiabatic states at a single reference geometry.

The projection diabatization algorithm has been explained in detail in the two earlier works[41,77] and its implementation is unchanged here. We will, however, recount some of the underlying assumptions which illustrate the differences between the projection and propagation diabatization schemes.

The projection diabatization method relies on using configuration interaction (CI) expansions of the electronic wavefunction, practically CASSCF and RASSCF (complete and restricted active-space self-consistent field, respectively), where the adiabatic states are expanded in terms of a set of Slater determinants (usually a ground state determinant plus a set representing electronic excitations) In other words, for state s, we have

$$|\psi_s\rangle = \sum_{i=1}^{N_{\text{Config.}}} c_i^s |\varphi_i\rangle, \tag{16}$$

where $\{c_i^s\}$ are the coefficients corresponding to each configuration represented by the determinant, φ_i. The propagation diabatization scheme is not, in principle, restricted to CI-type electronic structure methods, but, in practice, it is limited by the necessity to calculate the NACTs. As noted in Section II C 1, we also rely on being able to overlap the electronic wavefunctions at neighbouring geometries so as to maintain constant phases for the NACTs. The first major advantage of the projection diabatization scheme, then, is that the issue of the NACTs' phases disappears for the simple reason that NACTs are not required. This has the more significant advantage that the calculation of NACTS can be computationally demanding, particularly for larger molecules and/or when a large number of states.

In the projection diabatization scheme, a reference geometry, $\mathbf{q}_0$, is chosen where the adiabatic and diabatic states are taken to be one and the same, or, in other words $\mathbf{A}(\mathbf{q}_0) = \mathbf{I}$. At $\mathbf{q}_0$, a calculation is performed, including the number of states of interest (*e.g.* using state-averaged (SA-) CASSCF),[104,113,114] whilst at all other sampled geometries, $\{\mathbf{q}\}$, two calculations are performed. The first is the same as at $\mathbf{q}_0$, including the number of states required. The active space molecular orbitals (MOs) at $\mathbf{q}$ are then rotated so as to maximise their overlaps with the corresponding orbitals at $\mathbf{q}_0$ using Molpro[104] hence diabatizing the MOs. A second calculation is then performed, using these diabatized orbitals (which are not optimized during the calculation; only the CI part is carried out) in order to get the CI coefficients for a larger number of states (much greater than the number of states of interest). With CI vectors for all states at both $\mathbf{q}_0$ and $\mathbf{q}$ being in terms of sets of maximally coincident MOs, the final step of the projection diabatization proceeds with reference only to the CI vectors. To do so, a projector onto the space of the CI vectors corresponding to the states in the diabatic manifold (the adiabats at $\mathbf{q}_0$) is constructed; this is then used to extract those components of the CI space at $\mathbf{q}$ which are in the reference space. The Procrustes method (to be described in detail in the next section) is then used to rotate the resulting vectors into maximum coincidence with those at $\mathbf{q}_0$. This final operation is necessary to ensure that, not only do the vectors at $\mathbf{q}$ span the same space as those at the reference geometry, but that the individual basis vectors most closely represent the reference states. In this way we get the ADT matrix needed to obtain the diabatic energies at $\mathbf{q}$ using Eq. (13).

From the summary above, it is clear that there are two steps required to bring the states at $\mathbf{q}$ into maximal, diabatic alignment with those at the reference point, $\mathbf{q}_0$, namely the orbital and CI vector transformations. The reason for this can be seen by expanding the NACTs in terms of the expansion in Eq. (16). Using Eq. (12), for states a and b

$$\mathbf{F}_{ab} = \sum_{ij}^{N_{\text{Config.}}} c_i^{a*}\, c_j^b \langle \varphi_i | \nabla \varphi_j \rangle + \sum_{i=1}^{N_{\text{Config.}}} c_i^{a*}\, \nabla c_i^b \qquad (17)$$

where we have used the orthonormality of the Slater determinants. From this expression it is clear that the NACTs arise from two contributions: the coordinate dependence of the MOs (consisting of the MO coefficients and the atomic orbitals (AOs)) through the gradient of the Slater determinants, and the CI coefficients.[115] By transforming the MOs so that they are maximally overlapped between $\mathbf{q}$ and $\mathbf{q}_0$, the coordinate dependence on the MOs (and hence on the determinants) is minimised. In the original implementation,[77] the MOs were overlapped using the AO overlap matrix at $\mathbf{q}$ along with the MO coefficients at the respective geometries, and then diabatized using the Procrustes method (to be outlined in section II C 3), whilst in the implementation used in the *Quantics* package[41] this was modified to using the Jacobi rotation based method in Molpro.[104] In both cases the AO contribution to the diabatization is not treated explicitly but implicitly by the rotations of the MOs; the quality of the results gained by both methods indicates that this is not a significant problem. Assuming the orbital transformation has been done and the second CI calculation performed to get a new set of CI coefficients, we realise that the first term in the equation has been minimised, albeit the total wavefunctions and hence values of the NACTs are unchanged, but with as much of their magnitude a possible now being contained in the second term, so we are left with minimising the second term *i.e.* bringing the CI vectors at $\mathbf{q}$ into alignment with those at the reference point. It is this, second, transformation which is used to convert the adiabatic energies to the diabatic representation.

Returning briefly to the propagation diabatization method, the decomposition of the NACT into MO and CI components in Eq. (17) reveals the reason it was not necessary to diabatize the orbitals in that scheme before determining the ADT matrix. Specifically, the integration of the NACTs in Eq. (15) involves an integration over both components. This means that the ADT matrix obtained in the propagation scheme can be decomposed into contributions due to the MO components and CI coefficients, and has the necessary flexibility to rotate the total wavefunctions of the states in a single step.

The projection method was implemented in the *Quantics* package and shown to work successfully in the case of the dynamics of pyrazine and butatriene.[41] Additionally, calculations were performed along one degree-of-freedom of butatriene, the $14A_g$ stretching mode,[41] and for LiF[77], demonstrating that the diabatization correctly deals with surface crossings along one dimension, a problem with which the propagation diabatization struggles unless there is gradient information available. However, less success was found when using the projection diabatization method in the dynamics of some selected sunscreen molecules,[44] where larger scale vibrational motions were involved in the dynamics. This latter case shows that the projection diabatization method is very robust in keeping track of the diabats defined at the reference geometry, even if, during the course of the dynamics, that diabat rises in energy so that the wavepacket cannot follow it. It also means that diabats which were high in energy at $\mathbf{q}_0$, but which become lower in energy in regions of configuration space that the wavepacket visits, are not included in the set of reference states; a subsequent calculation would need to be performed to include such a state in the reference.

3. Procrustes diabatization

As the previous two sections have shown, the two diabatization schemes we have previously implemented for the DD-GB set of methods both successfully generate smooth, quasi-diabatic PESs, removing the discontinuities in the gradients of the PESs and infinite couplings at points of degeneracy. However, they both have different strengths and weaknesses; here, we propose a scheme which attempts to combine the positives (and remove some of the negatives) of both methods. We refer to this new method as Procrustes diabatization hereafter, as it simply relies on two applications of the solution to the orthogonal Procrustes problem.[89,90] We acknowledge that the Procrustes method has been used to diabatize states before, although not necessarily with that name,[76,77,87,88] and has been shown to be equivalent to the block diagonalization of the Hamiltonian.[87,116] However, to the best of our knowledge, the diabatization strategy presented here has not been described before.

For those unfamiliar with the orthogonal Procrustes problem, we briefly summarise the method here. The basic idea of the problem is: given two matrices $\mathbf{P}$ and $\mathbf{Q}$, we seek the orthogonal matrix, $\mathbf{R}$, which transforms $\mathbf{P}$ so that the result most closely matches the target, $\mathbf{Q}$. In other words, we seek

$$\mathbf{R} = \arg\min_{\mathbf{\Omega}} ||\mathbf{P}\mathbf{\Omega} - \mathbf{Q}||_{\text{F}}, \qquad (18)$$

where $||\cdot||_{\text{F}}$ indicates the Frobenius norm, such that $\mathbf{\Omega}^T\mathbf{\Omega} = \mathbf{I}$. The solution to this problem is

$$\mathbf{R} = \mathbf{U}\mathbf{V}^T \qquad (19)$$

where the matrices $\mathbf{U}$ and $\mathbf{V}^T$ are obtained by performing a singular value decomposition (SVD) on the product of $\mathbf{P}^T$ and $\mathbf{Q}$

$$\mathbf{P}^T\mathbf{Q} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \qquad (20)$$

with $\mathbf{\Sigma}$ being the matrix of singular values.

The utility of the Procrustes method for diabatization is apparent; we seek a unitary matrix which transforms the adiabats at some geometry, $\mathbf{q}$, so that they best match those at the reference geometry. This method has been used previously to diabatize electronic states using atomic transition charges as the property of the states which defines the diabats,[76] and also with CASSCF wavefunctions.[77,87,88]

However, key to our approach is the fact that we do not use molecular properties as the characteristics of each diabat, but the wavefunctions themselves. In the spirit of the projection diabatization scheme we are looking to transform the adiabats at $\mathbf{q}$ so that they most closely resemble those at $\mathbf{q}_0$. In other words we wish the overlap matrix of the diabats at the two geometries to be as close to the unit matrix as possible, so that we must minimise $||\mathbf{S}^D - \mathbf{I}||_\mathrm{F}$ where $S^D_{ij} = \langle\psi^D_i|\psi^D_j\rangle$, the superscript indicating that we are considering the diabatic states. It follows that if we know the adiabatic states and the ADT matrix at some geometry, $\mathbf{q}_a$, we also know the adiabatic states at $\mathbf{q}$, and can evaluate the overlap between these two sets of adiabats, $\mathbf{S}^A$, then, considering the Procrustes problem, we have

$$\mathbf{P} = \mathbf{A}(\mathbf{q}_a)\mathbf{S}^A \tag{21a}$$

$$\mathbf{Q} = \mathbf{I}. \tag{21b}$$

The necessary SVD is

$$\mathbf{P}^T\mathbf{Q} = (\mathbf{S}^A)^T\mathbf{A}^T(\mathbf{q}_a) = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{22}$$

and thus the required ADT matrix is

$$\mathbf{A}(\mathbf{q}) = \mathbf{U}\mathbf{V}^T. \tag{23}$$

We will only be using CI-type wavefunctions here so, by the same arguments made in section II C 2 in relation to Eq. (17), we need to bring both the MOs and CI coefficients at $\mathbf{q}$ into alignment with those at the reference point in order to minimise the NACTs.

We provide a detailed algorithm for the Procrustes diabatization scheme in the SI (Section III) but briefly, given the adiabatic wavefunction, diabatic MOs and ADT matrix at $\mathbf{q}_a$, we calculate the N_s electronic energies and wavefunctions at $\mathbf{q}$. The MOs at both geometries are overlapped (diabatic at $\mathbf{q}_a$ and adiabatic at $\mathbf{q}$) and the Procrustes procedure used to transform the MO coefficients at $\mathbf{q}$ to the diabatic representation (in the case of CASSCF-type wavefunctions, the active and closed space orbitals are transformed separately to ensure the invariance of the electronic energies to the transformations). A second CI only calculation is then performed at $\mathbf{q}$ using these new orbitals to get the CI coefficients in terms of the diabatic Slater determinants. The overlaps of the wavefunctions at both geometries are then evaluated, giving $\mathbf{S}^A$ and the Procrustes method used to get the required ADT matrix, $\mathbf{A}(\mathbf{q})$, which is then used to transform the adiabatic energies to the diabatic using Eq. (13).

With reference to Eq. (17), the first, MO transformation deals with the first term in the NACTs, namely that arising directly from the Slater determinants. We noted above that the MO contribution to the NACTs could be further decomposed into a part due to geometry dependent changes in the MO coefficients and a part due to the change of the AO basis.[115] The Procrustes transformation brings the orbitals at $\mathbf{q}$ into best alignment with those at $\mathbf{q}_a$ so the coefficient part is certainly dealt with, but a question remains over the AO part. In some earlier applications of the Procrustes method to diabatization[87,88] there was no transformation of the MOs, only of the total wavefunction, resulting in a strategy which was effectively the same as our second step. We initially tried a method where only the second step was performed and found that it was insufficient to replicate the diabatic states and couplings in the butatriene cation (see Section III B for the final results); the initial transformation of the MO overlaps was necessary to qualitatively reproduce the proper surfaces.[80] In those earlier papers[87,88] there is discussion of the need to deal with the AO non-adiabatic coupling, with the solution being to perform the wavefunction overlaps using the AO basis at only one of the geometries, hence avoiding the problem. Here we use the AOs at both $\mathbf{q}_a$ and $\mathbf{q}$ to get the exact overlap, so there may be residual coupling due to the movement of the AOs with their associated atoms. We note that it would be far from straightforward to achieve a usable diabatization of the AOs in the context of our method, which is meant to automatically diabatize wavefunctions using information obtainable from commercial electronic structure codes; generation and input of a transformed basis in such codes seems to be extremely tricky, if possible at all. Instead, we will just make the following observations to justify our neglect of AO diabatization: all sampled molecular geometries in our calculations are transformed to the Eckart frame given by the normal modes at the reference geometry, $\mathbf{q}_0$, where the ADT matrix is taken to be the unit matrix, hence the changes in the atomic positions between sampled geometries is minimised. The AOs therefore change as little as possible, reducing the residual coupling. We could, of course, have chosen a single basis with which to evaluate the overlaps but wanted to maximise the accuracy of the wavefunction overlaps; that this may be at the expense of having to accept some residual coupling due to the AO basis is something we choose to accept, and the results presented in the following section confirm this to be a reasonable compromise. Furthermore, it has been noted previously that the AO component of the NACTs do not contribute to the induction of electronic transitions.[117]

By using the idea of configurational conformity to diabatize the states, the Procrustes method is similar to the projection diabatization scheme, which in its original form also used the Procrustes method to rotate both the

MOs and final wavefunction, but used an intermediate projection operation which is not used here.[77] There is also no need to calculate the expensive NACTs needed in the propagation diabatization method. Unlike the projection diabatization scheme, the Procrustes method does not use a single reference to which all others are compared, but uses a propagation-like series of steps radiating out from $\mathbf{q}_0$ (where $\mathbf{A} = \mathbf{I}$); in other words, the states at $\mathbf{q}$ are diabatized in comparison to some point $\mathbf{q}_a$, which is the closest point to $\mathbf{q}$ which is also closer to $\mathbf{q}_0$ (and has hence already been diabatized). The aim of this is to better follow the gain and loss of relevant electronic states in different regions of configuration space, as is achieved by the propagation diabatization method but not by the projection diabatization. Such an approach does have a downside; as explained in Section II C 1, and relevant here too, the diabatization is path-dependent meaning that the ADT matrix at $\mathbf{q}$ may vary depending on the series of steps taken to reach it from $\mathbf{q}_0$, which is an artefact of the presence of non-removable couplings. However, in both methods the sampling of geometries is not stochastic so the results obtained during the simulations are repeatable.

A major issue, which is caused by the path dependence of the propagation and Procrustes diabatization schemes, is that the diabatic states are not unique at any given geometry, $\mathbf{q}$; they depend on the path taken to reach that point from $\mathbf{q}_0$. This can cause problems, for example, if a part of the path starts at some intermediate point $\mathbf{q}_a$ and then, by some series of steps, returns to $\mathbf{q}_a$; in this case the diabatic states calculated at the first and second visits to $\mathbf{q}_a$ will most likely not be identical. For both path-dependent approaches described here, this problem is avoided in that there are (generally) multiple paths radiating out from $\mathbf{q}_0$, points being visited in order of increasing distance from $\mathbf{q}_0$; it is thus not possible for these problematic closed loops to form. In a similar way, two paths, both starting at $\mathbf{q}_0$ could in principle follow different routes, moving away from $\mathbf{q}_0$ but then reaching the same point, $\mathbf{q}$. Again, the diabats calculated at $\mathbf{q}$ would differ, depending on which of the two routes had been taken to get there. In our methods this problem is avoided because each sampled geometry is visited once and only once, so can only be reached by a single path. Additionally, the use of the KRR variance measure, Eq. (11), ensures that the same geometry cannot be included in the database of energies more than once (it may be sampled repeatedly, but after the first acceptance into the database, all subsequent samplings will be rejected). In fact, the use of the variance measure prevents the inclusion of very nearby points in the database so, even in the case of two paths circling a conical intersection and approaching one another they will never get too close to one another and, in any case, the phases of the adiabatic wavefunctions are kept consistent when using the propagation diabatization by wavefunction overlaps, whilst the unitary nature of the Procrustes transformations ensures the uniformity of the phases of the diabatic states.

We also note that the KRR sampling of configuration space can account for the point group symmetry of the molecule, if required. Each of the normal modes is a member of an irreducible representation of the point group and it follows that, along any non-totally-symmetric mode, q_n (keeping all other coordinates constant), the adiabatic and diabatic energies will be symmetric *i.e.* $V_{ii}(q_n) = V_{ii}(-q_n)$ whilst the diabatic couplings will be symmetric or anti-symmetric (depending on the irreducible representation) *i.e.* $V^D_{ij}(q_n) = -V^D_{ij}(-q_n)$. To ensure the correct symmetries of the diabatic PESs and couplings: if a point $\mathbf{q}$ is sampled and accepted into the database, all symmetry equivalent points will also be accepted. This occurs unless their variances compared to one another fall below the usual variance threshold parameter, in which cases all points, including the initially accepted one, will be rejected. Crucially this ensures that points with a coordinate of 0 along q_n (except the unique $\mathbf{q}_0$) will always be rejected, meaning that, if the couplings are anti-symmetric along q_n, then it will be impossible to have a non-zero coupling there due to the imperfect diabatization. If the symmetry equivalent points are all accepted into the database, then the energies and other adiabatic quantities (NACTs and CASSCF expansion coefficients) are calculated at all points. Currently it is not possible to use the data at a single point then use symmetry operations to get the NACTs and CASSCF coefficients, whose magnitudes will be identical, but whose phases will not, although the NACTs are compared to ensure the calculated values have the same magnitudes (and if not, are averaged). Because of the symmetries of the adiabatic quantities calculated and used in the diabatizations, the diabatic states and couplings have the correct symmetry. This will be demonstrated later in the case of the butatriene cation.

Finally, we note that the Procrustes solution is an orthogonal matrix, not necessarily a rotation matrix, meaning that the arbitrary changes in sign of the states, which can occur in electronic structure programs, are automatically dealt with to keep the diabatic couplings (off-diagonal terms in $\mathbf{V}^D$) consistent.

III. RESULTS AND DISCUSSION

A. Lithium Fluoride

To test the Procrustes diabatization scheme, we first look at the classic case of LiF. A common feature of alkali metal halides is that, at their equilibrium geometry, the electronic ground state is ionic in character whilst there is an excited state of covalent character. At the dissociation limit the ordering of these states is reversed, the inference being that, at some point along the dissociation coordinate, there is a point at which the two states cross. The first test of our new diabatization scheme is whether the crossing of states is successfully dealt with; the avoided crossing of the adiabats being transformed

into two smoothly crossing diabats. Furthermore, if we restrict ourselves to these two states only, then the ADT matrix is of the form

$$\mathbf{A}(\mathbf{q}) = \begin{pmatrix} \cos\theta(\mathbf{q}) & \sin\theta(\mathbf{q}) \\ -\sin\theta(\mathbf{q}) & \cos\theta(\mathbf{q}) \end{pmatrix} \tag{24}$$

We constrain LiF to the z-axis, with the coordinate of the F being greater than that of the Li (*i.e.* $z_{\mathrm{F}} > z_{\mathrm{Li}}$), then transform to a coordinate system consisting of the centre-of-mass and the bond length, r. It has been shown that, for a two-state system, the rotation angle is related to the non-adiabatic coupling as[64,77,118]

$$\frac{\partial\theta}{\partial r} = \mathbf{F}_{21} = \langle\psi_2|\frac{\partial}{\partial r}\psi_1\rangle. \tag{25}$$

This can be related to the Cartesian components of the NACT by use of the chain rule such that

$$\langle\psi_i|\frac{\partial}{\partial r}\psi_j\rangle = \frac{1}{M}\langle\psi_i|m_{\mathrm{Li}}\frac{\partial}{\partial z_{\mathrm{F}}} - m_{\mathrm{F}}\frac{\partial}{\partial z_{\mathrm{Li}}}|\psi_j\rangle \tag{26}$$

where m_{Li} is the atomic mass of Li, m_{F} is the atomic mass of F and $M = m_{\mathrm{Li}} + m_{\mathrm{F}}$. Hence, by diabatizing the states and taking the derivatives of the rotation angle with respect to the bond length we can directly compare the results to the Cartesian NACTs given by an electronic structure package, in our case Molpro.

This comparison was performed by Robertson *et al*[77] when they first introduced the projection diabatization method; it was found to perform very well, correctly reproducing the peak in the coupling around the avoided crossing between the two states, as well as the smooth crossing of the states. As such we will not try to reproduce this result, but we will show the difference between the results produced by the propagation and Procrustes diabatization schemes.

The ground state geometry of LiF was optimized using the SA-CASSCF method in Molpro,[104,113,114] with the lowest two A_1 states equally weighted (Molpro employs Abelian point group symmetry, so LiF is in the C_{2v} group rather than the correct $C_{\infty v}$). The active space was 8 electrons in 12 orbitals (5 in the A_1 irreducible representation, 1 in A_2, and 3 each in B_1 and B_2, with the two closed orbitals in A_1) and a 6-311+G(3df) basis set was employed. The stretching mode frequency was calculated at the non-state-averaged CASSCF level and found to be 893 cm^{-1}, allowing the quantum dynamics to be run in a coordinate system of mass-frequency scaled normal modes.

Subsequently two DD-SM calculations were run in a development version of the *Quantics* package[119], using the propagation and Procrustes diabatization scheme respectively. In both cases a sine-DVR basis was used with 1201 members in the domain [-100,6]. The initial wavepacket was a Gaussian function centred at the origin (corresponding to the Franck-Condon point) on the excited state with width $1/\sqrt{2}$ and 0 initial momentum. The dynamics was then run with SA-CASSCF energies (and NACTs if needed) being calculated on-the-fly at points selected by Sobol sampling within three widths of the centre of the wavepackets on each state, every 1 fs (10 geometries *per* state). Electronic structure calculations were only performed if the KRR variance (Eq. (11)) was greater than 10^{-3} at each selected geometry whilst the width-parameter in the Gaussian kernel (Eq. (5)) was fixed at $\alpha = 0.05$. The dynamics was run using the default short iterative Lanczos integrator for 200 fs and a complex absorbing potential was placed at a (unitless) coordinate of -95 to absorb the dissociating wavepacket before reaching the end of the DVR grid. In addition to the dynamics calculations, a series of SA-CASSCF calculations was performed with Molpro at bond lengths between 1.0 and 9.0Å in order to get the NACTs at each point to allow comparison with those extracted from Eq. (25).

In Fig. 1 we present the diabatic energy surfaces produced by each method: in Fig. 1(a) from the Procrustes diabatization and in 1(b) from the propagation method (both transformed to bond length coordinates). Clearly both methods perform well, producing diabatic curves which cross smoothly as required at the same bond length. The slight downturns in the upper diabats at very short bond-lengths are due to lack of KRR sampling in that region of configurational space.

To assess the absolute, rather than relative, success of the diabatization methods we must compare the NACTs extracted from the derivatives of the ADT matrix rotation angles to those given directly by Molpro using Eq. (25). These derivatives were calculated numerically by diagonalising the KRR-fitted diabatic energies at the centre of each DVR gridpoint to get the ADT matrices, then using finite differences to get the derivatives. Plots comparing the both sets of NACTs, around the point of intersection, are presented in Figs. 2(a) for the Procrustes diabatization and 2(b) for the propagation diabatization.

It is clear from both of these plots that the NACTs calculated from the ADT rotation angles, represented by the solid, red lines, closely follow those calculated directly using Molpro (the dashed, blue lines in both plots). The propagation diabatization results almost exactly match the Molpro results, but this is to be expected as the rotation angle are determined by integration of the NACTs calculated during the dynamics. The NACTs are not exactly reproduced due to the errors inherent in the integration used in the propagation diabatization scheme.

The Procrustes diabatization-derived couplings also agree very well with Molpro calculated values in this critical region of configuration space, indicating that the method is correctly representing the system in an absolute sense. There are differences between the exact and fitted couplings due to the presence of non-removable couplings which results in an imperfect, quasi-diabatization. In the Procrustes method, the diabatization relies on the overlaps between states at different geometries, however, the existence of other states, which are not included in the calculation means that the two

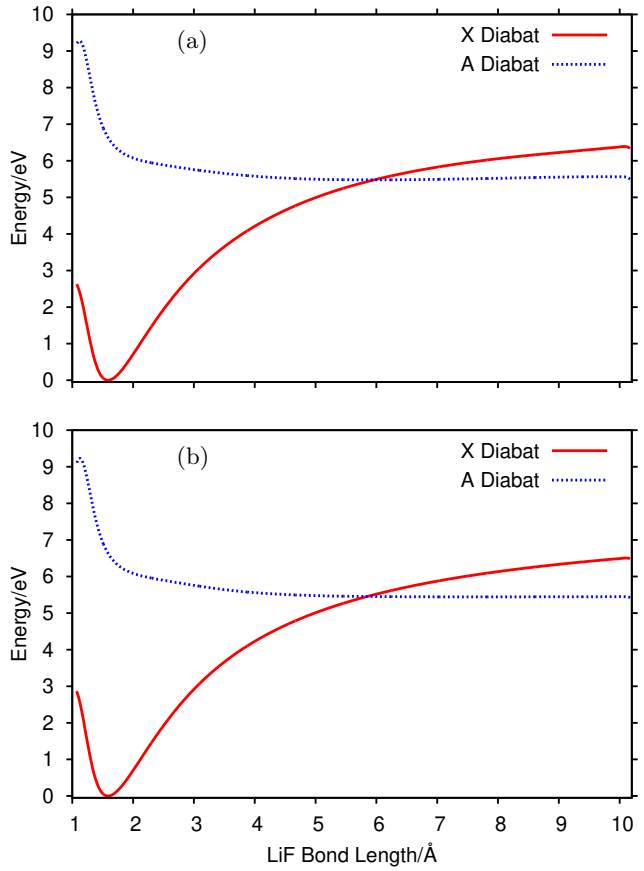

FIG. 1. Diabatic PESs calculated for LiF generated during on-the-fly dynamics using (a) Procrustes diabatization and (b) propagation diabatization.

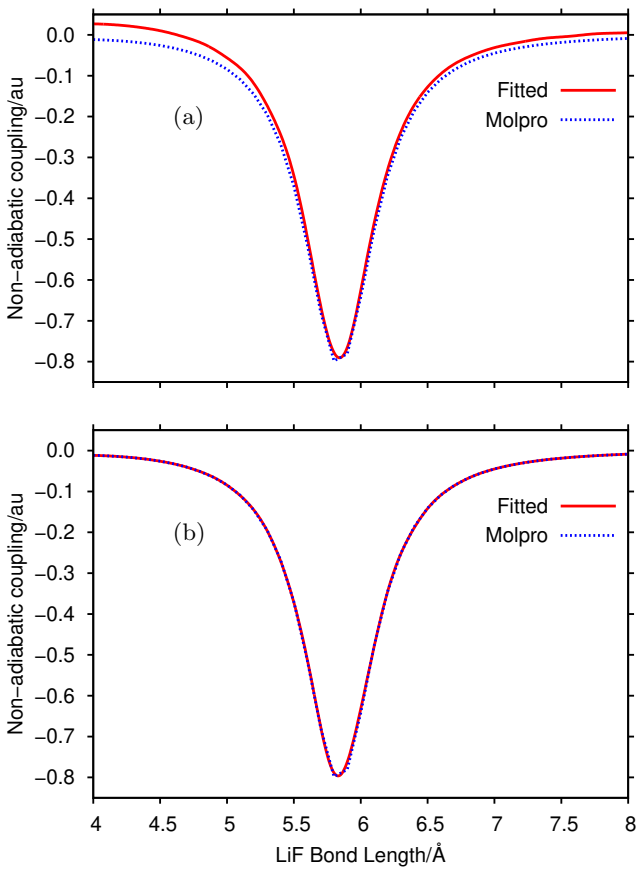

FIG. 2. Non-adiabatic coupling terms calculated for LiF from rotation angles obtained from the adiabatic-diabatic transformation matrices calculated during on-the-fly dynamics using (a) Procrustes diabatization and (b) propagation diabatization. In both plots the solid, red lines are the fitted terms and the dashed, blue lines are the coupling terms calculated directly using Molpro.

included states do not span the same space at the different geometries (in practice the norms of the columns and rows of the overlap matrices are not 1). The close agreement of the fitted and exact couplings indicates that the vast majority of the coupling is recovered, as expected.

We have demonstrated the successful Procrustes diabatization of the LiF potential energy curves (PECs) for the two lowest lying A_1 states. However, in the SI (Section IV), we include the results of some additional calculations which demonstrate the robustness of the Procrustes method when applied to LiF. In particular, we investigate the impact of the α-parameter in the KRR kernel (which allows comparison to the robustness of the propagation diabatization method) and the CI cut-off (*i.e* which overlaps of the Slater determinants are included in the total state overlaps based on their CI coefficients). It is shown there that the Procrustes method is more robust to changes in α than the propagation scheme (which fails to reproduce the surface crossings if α is set too small), and is also surprisingly resilient when terms in the state overlaps are ignored for reasons of increasing computational speed.

Finally, is is worth reiterating that the Procrustes diabatization scheme does not require NACTs, but produces diabatic surfaces comparable to those produced using the propagation diabatization scheme, and correctly reproduces the NACTs around the critical curve crossing point. However, we should note that the total walltime (including all Molpro calculations) for the dynamics using the propagation dynamics was 46 minutes (on a standard desktop machine) and for the Procrustes diabatization was 92 minutes; twice as long. Nearly all of this difference is due to the extra time taken in overlapping the CASSCF wavefunctions for Procrustes diabatization. The average CPU time for calls to the overlap subroutine were 44 seconds for the Procrustes method and just under 6 seconds for the propagation scheme. The reason for this discrepancy is the difference in the number of significant configurations in the CASSCF wavefunctions needed in each method. In both calculations configurations with a maximum coefficient (for both states) of 10^{-8} or less were ignored, so the conditions were the

same for both. However, for the propagation diabatization method the number of configurations with coefficient greater than this cut-off stays at similar levels at all sampled geometries, whilst the number increased significantly for geometries further from the Franck-Condon point for the Procrustes scheme. The reason for this difference lies in the orbitals used to expand the CI vectors in each case; for the propagation scheme the natural orbitals for the relevant geometry, which produces an optimally compact CI expansion, are always used. For the Procrustes method, the orbitals are the diabatic orbitals which are kept in maximal alignment with the natural orbitals at the Franck-Condon point. As the sampled geometries get further from this reference point, the diabatic orbitals become less and less like the natural orbitals at the current point, leading to longer and longer CI expansions which take much more effort to overlap. This is a potential downside of the Procrustes method presented here; the saving in not calculating NACTs possibly being eaten up (and more) by the effort in calculating overlaps. The balance between the two factors will be system dependent; smaller active spaces and/or larger numbers of states should favour the Procrustes scheme as the effort in overlapping the wavefunctions is reduced in the former case whilst the number of NACTs needed increases in the latter (although the effort in overlapping is increased, the major contributor to the effort is the number of configurations and hence the number of overlap determinants which need to be evaluated). In the SI, we present results for the Procrustes scheme where the CI coefficient cut-off is varied to compare the saving in time and the robustness of the results.

B. Butatriene Cation

To further test the Procrustes diabatization scheme, we turn to another well studied system, the butatriene cation, whose conveniently located conical intersection close to the Franck-Condon point makes it an ideal test case for new diabatization methods. Initially we look at the dynamics for a 2-dimensional model before looking in more detail at potential energy slices along single degrees-of-freedom.

1. Dynamics in 2 Dimensions

To study the dynamics of the butatriene cation in two dimensions, three DD-SM calculations were performed, the only difference between them being the diabatization scheme used: Procrustes, propagation, or projection.

The geometry of the neutral butatriene molecule was optimized at the CASSCF(6,6)/3-21G level using Molpro, and normal-mode frequencies and coordinates were determined. DD-SM calculations were then performed in two dimensions, using as a coordinate system the mass-frequency scaled $5A_u$ (torsion, labelled q_5, with frequency $768\,\mathrm{cm}^{-1}$) and $14A_g$ (symmetric carbon stretching, labelled q_{14}, with frequency $2195\,\mathrm{cm}^{-1}$) normal modes. The ground-state and the first three excited states of the cation were included in the DD-SM calculations. Sine DVR bases were used along each normal-mode: 101 grid-points in the range $[-10, 10]$ along q_5, and 81 functions in the range $[-8, 8]$ along q_{14}. The initial wavepacket was a Gaussian product function with width $1/\sqrt{2}$ in each degree-of-freedom, centred at the Franck-Condon point (origin of the coordinate system) on the first excited diabatic state. Time evolution of the wavepacket was followed for 100 fs in total. The PESs were generated by Sobol sampling 20 geometries within three widths (calculated as $\left[\langle q^2\rangle - \langle q\rangle^2\right]^{1/2}$, where the expectation values for each degree-of-freedom are calculated from the current wavefunction) of the centre of the wavepacket on each state every 1 fs. The SA(4)-CASSCF(5,6)/3-21G energies were calculated at sampled points only if the KRR variance at each point exceeded 10^{-3}; symmetry equivalent points were additionally included by supplementing each accepted geometry with a further configuration in which the q_5 coordinate is inverted. For the projection diabatization calculation, as described in section II C 2, a total of 20 states were used in the CI-only calculations using the diabatic orbitals. After diabatization, the PESs were fitted using KRR with Gaussian functions with a width parameter $\alpha = 0.05$.

Fig. 3 shows the diabatic PESs produced by the Procrustes diabatization, and can be compared to those presented in the supporting information of Ref. 41 (using the projection diabatization) and in Ref. 98 (using the original implementation of the propagation diabatization with gradient information). The main features to note here are the smooth crossing of the surfaces, particularly around the conical intersection between the D_0 and D_1 (which correspond to the $\tilde{X}$ and $\tilde{A}$ diabats at the Franck-Condon point respectively) at the coordinate $q_{14} \approx 2$ in a planar geometry ($q_5 = 0$) as well as between the $\tilde{B}$ and $\tilde{C}$ diabats (which correspond to D_2 and D_3 at the Franck-Condon point). The diabatization method is therefore capable of producing smooth, quasi-diabatic potentials which cross at points of degeneracy in a multi-dimensional system.

Fig. 4 shows the diabatic coupling between the $\tilde{X}$ and $\tilde{A}$ surfaces as a function of molecular geometry. As explained previously,[41,96] by symmetry this function should be odd in the q_5 coordinate and this is indeed what we find here. As we move away from the planar geometry, the coupling exhibits a dependence on the symmetric stretching coordinate, q_{14}; this coupling is identically zero wherever $q_5 = 0$, but as the D_{2h} symmetry has been broken when $q_5 \neq 0$, a coordinate-dependence is now allowed. This smooth coupling surface, displaying appropriate symmetry, demonstrates the physically-consistent behaviour of the Procrustes diabatization method.

Turning now to the actual wavepacket dynamics, Fig. 5 shows the time-dependent population of the $\tilde{A}$ diabat for calculations performed using the three diabatization

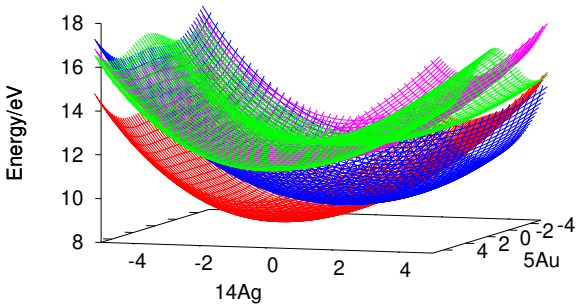

FIG. 3. KRR fit of the diabatic potential energy surfaces of the ground and first three excited states of the butatriene cation as a function of the $5A_u$ torsion and $14A_g$ symmetric-stretching mass-frequency scaled normal modes, as calculated on-the-fly using the Procrustes diabatization during the DD-SM propagation described in the text.

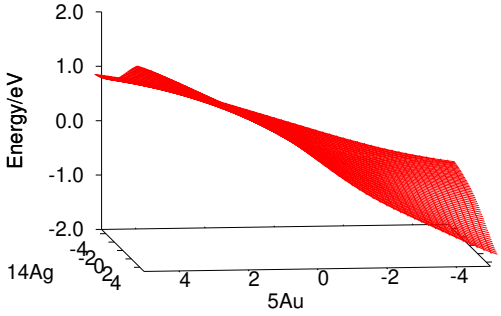

FIG. 4. KRR fit of the diabatic coupling between the $\tilde{X}$ and $\tilde{A}$ diabatic states of the butatriene cation as a function of the $5A_u$ torsion and $14A_g$ symmetric stretching mass-frequency scaled normal modes, as calculated on-the-fly using the Procrustes diabatization during the DD-SM propagation described in the text.

schemes. We note that the exact population dynamics is different for all three cases, but this is to be expected as the diabatic surfaces are different, however, the dynamics is qualitatively the same for all three: the initial, rapid de-population of the $\tilde{A}$ state as the wavepacket encounters the conical intersection, followed by partial recurrences of the population as the wavepacket rocks back and forth past the state crossing. The Procrustes diabatization thus produces diabatic PESs which not only display the correct desired physical characteristics, but which also allow the wavepacket to behave as expected.

The population dynamics qualitatively matches that seen when using pre-fitted vibronic coupling PESs[43,66,95], but we note that in those cases, the q_5 mode was a proper torsion mode, rather than the $5A_u$ normal mode used here. The direct dynamics approaches used here also automatically include the anharmonicity of the PESs, which was not included in the fitted surfaces. As a result of these two facts we would not expect quantitative agreement between the dynamics given when using on-the-fly fitting of the PESs when compared to that on the pre-fitted surfaces.

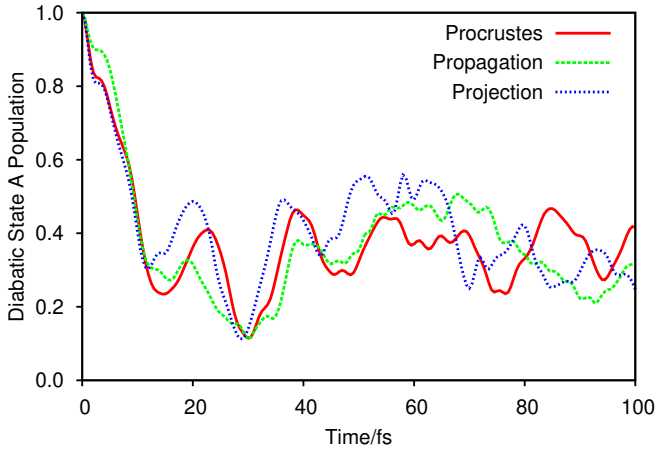

FIG. 5. Populations of the $\tilde{A}$ diabatic state of the butatriene cation during 4-state dynamics on-the-fly: the solid, red line is the result due to using the Procrustes diabatization to generate the PESs; the green, hashed, line is the result when using the propagation diabatization; the blue, dashed line results when using the projection diabatization.

Even though the dynamic behaviour on the diabatic PESs is different depending on the diabatization scheme used, we might expect the dynamics to be similar when viewed within the adiabatic picture. However, although it is possible to calculate adiabatic populations within the *Quantics* package (the diabatic PESs are diagonalised at each DVR gridpoint, allowing the transformation of the wavepacket at those points to the adiabatic representation), this comparison would itself be misleading. In particular, the initial wavepacket is placed on the first excited diabatic PES, but the diabatic surfaces are slightly different for each diabatization method. As a result, the initial adiabatic wavepackets in this analysis would each be different; in other words, the initial conditions of the adiabatic dynamics obtained for each diabatization scheme are different, so the comparison between the population dynamics is not particularly meaningful. Ideally, we could prepare the initial wavepacket on an adiabatic state, before transforming it to whichever diabatic representation was being used before allowing the dynamics to proceed. To achieve this, we have implemented a method of transforming the initial wavepacket between the adiabatic and diabatic representations when diabatic PESs

are already available.

To carry out a calculation using an adiabatic, initial wavefunction one must first run a calculation in the normal, diabatic representation as a way of sampling configuration space and hence building an approximation to the global, diabatic PES. A subsequent dynamics calculation can then be performed using this pre-constructed PES to which no additional database points are added during the course of the wavepacket propagation. The initial wavefunction for this second calculation is defined as usual, however, before the dynamics begins, this adiabatic wavefunction is transformed to the diabatic representation. To do this, each DVR gridpoint is taken in turn and the diabatic energies and couplings are evaluated using the pre-fitted KRR functions before being diagonalized to give the adiabatic energies. The eigenvectors calculated at the same time form the ADT matrix at the DVR location which can then be applied to the adiabatic wavefunction at that point, giving the initial wavefunction in the diabatic representation, which is now in general non-zero on each state. This initial diabatic-representation wavefunction can then be used to perform the dynamics simulation as normal.

A difficulty with this procedure is the fact that the eigenvectors at each DVR gridpoint have arbitrary phases, meaning that the initial, transformed wavefunction can randomly change sign between adjacent gridpoints. To avoid this, the ADT matrix at the DVR gridpoint closest to the centre of the adiabatic wavepacket is taken as the reference point and, moving outwards from that location, the phases of the vectors at each DVR point are compared to those at the nearest gridpoint which is also closer to the reference point. The sign of each eigenvector is changed if the maximum overlap with those at the nearby gridpoint is less than zero. In this way a consistent phase propagates outwards across configuration space.

To test the population dynamics of the butatriene cation using an adiabatic, initial wavepacket, calculations were performed with the databases created by initial dynamics runs using the three diabatization methods. For the Procrustes and propagation diabatization calculations, the databases created by the calculations presented above were used. For the projection diabatization calculation, a new wavepacket propagation was run to generate a database of energies; the initial conditions were the same as described above except that only four states (rather than 20) were included in the second, CASSCF calculation (using unoptimized, diabatic orbitals), meaning that the adiabatic-diabatic transformation was only in the same space of states as both the Procrustes and propagation diabatization calculations. Using more than four states would not have allowed a useful comparison because the set of adiabatic states available would have been different in this case. The second set of calculations, using the previously-created databases of energies, were carried out using the same initial conditions as described above for the initial calculations; in particular, the initial wavefunction which was taken to be an adiabatic wavefunction on the D_1 state before being transformed to the diabatic representation.

In Fig. 6, we present the population of the adiabatic D_1 states as calculated during the calculations using the databases from the Procrustes and projection diabatizations. We note that the results from using the propagation diabatization scheme are not included as the agreement with the other two methods was found to be poor. The reason for this is the inaccuracy of the created diabatic states, which leads to a poor fit of the adiabatic states at the locations of the DVR gridpoints. Further calculations were attempted using the propagation diabatization, but with different KRR fitting parameters. It was found that the results were quite variable when changing these parameters, so it is difficult to be confident about what is the "correct" result, hence its omission here. This uncertainty about the reliability of the propagation diabatization scheme in the context of KRR fitting was mentioned in Section II C 1 and will be further illustrated in the next Section.

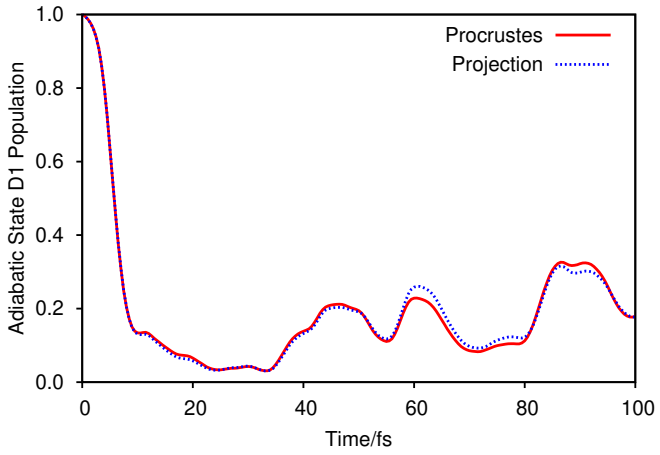

FIG. 6. Populations of the D_1 adiabatic state of the butatriene cation during four-state dynamics on-the-fly: the solid, red line is the result due to using the Procrustes diabatization to generate the PESs; the blue, dashed line results when using the projection diabatization.

Returning to the results presented in Fig. 6, we see excellent agreement between the Procrustes and projection diabatization results, as would be expected. The slight variation between the results reflects the difference in which wavefunction is used as a reference when diabatizing the energies at a particular geometry; for the projection method the comparison is always directly with the point, $\mathbf{q}_0$, whilst for the Procrustes method the comparison is to the nearest geometry which is also nearer to $\mathbf{q}_0$. The fits of the PESs are thus slightly different, which is reflected at the DVR gridpoints where the wavefunction is defined, leading to different dynamics.

2. Potential Energy Curves in One-Dimension: $5A_u$ Mode

In light of the difficulty in comparing population dynamics, to more clearly examine the differences between the diabatic PESs for the three diabatization schemes we examine PES cuts along the two normal modes of the butatriene cation.

In Fig. 7 we show the diabatic PECs of the first four states along mode $5A_u$, calculated using the three diabatization schemes in *Quantics*, with the other normal-mode coordinate being fixed at zero. The same level of electronic structure theory was used as in the two-dimensional case, but this time the potential energy was simply calculated at the location of the 51 sine DVR gridpoints spread evenly in $[-6, 6]$ in order to avoid any possible artefacts from the KRR fitting. The only need for KRR in the calculations was in the fitting of the adiabatic potentials used in integrating the NACTs when using the propagation diabatization scheme (Fig. 7(b)), where a width parameter of $\alpha = 10.0$ was used to account for the close proximity of the gridpoints.

In Fig. 7(a) we present the diabatic PECs generated when using the Procrustes diabatization, illustrating the smoothness of the curves, particularly in the crossing of the double-well diabat $\tilde{C}$ across diabats $\tilde{A}$ and $\tilde{B}$. Considering Fig. 7(b), the results from the propagation diabatization, we see that the basic shapes of the diabats are similar to those given by the Procrustes diabatization but that some of the smoothness has been lost; at coordinates $q_5 \approx \pm 3.5$ we see kinks in diabats $\tilde{X}$, $\tilde{A}$ and $\tilde{C}$. Diabats $\tilde{B}$ and $\tilde{C}$ cross as seen when using Procrustes diabatization, but there is no crossing of diabats $\tilde{A}$ and $\tilde{C}$ suggesting that they are behaving adiabatically with respect to one another, forming an avoided crossing which also affects diabat $\tilde{X}$. The presence of this avoided crossing is confirmed when considering the adiabatic plots (not shown), showing that the propagation diabatization scheme is failing to correctly locate a crossing of the PECs. The reason for these failures is that the NACTs are orthogonal to the path along q_5 meaning that the integrand in Eq. (15) is zero at all points along q_5. As a result, the ADT matrix, which is the unit matrix at the reference geometry, $q_5 = 0$, stays as the unit matrix between these states and thus cannot reorder the states as required. At the intersection there is a discontinuity in the ADT matrix which cannot be dealt with naturally using a smooth integral. In the original implementation of the propagation diabatization scheme,[66,98] which uses gradient information, this was dealt with by projecting the adiabatic energies along the integration path in order to predict the presence of a crossing, and hence to fix the ADT matrix, but this is not possible here.

In Fig. 7(c) we show the PECs formed when using the projection diabatization scheme. The major difference here from the other two results is the shape of diabat $\tilde{C}$, which is now a single well rather than the double well seen in Fig. 7(a). The other three states have the same shape as those seen in Fig. 7(a), although diabats $\tilde{X}$ and

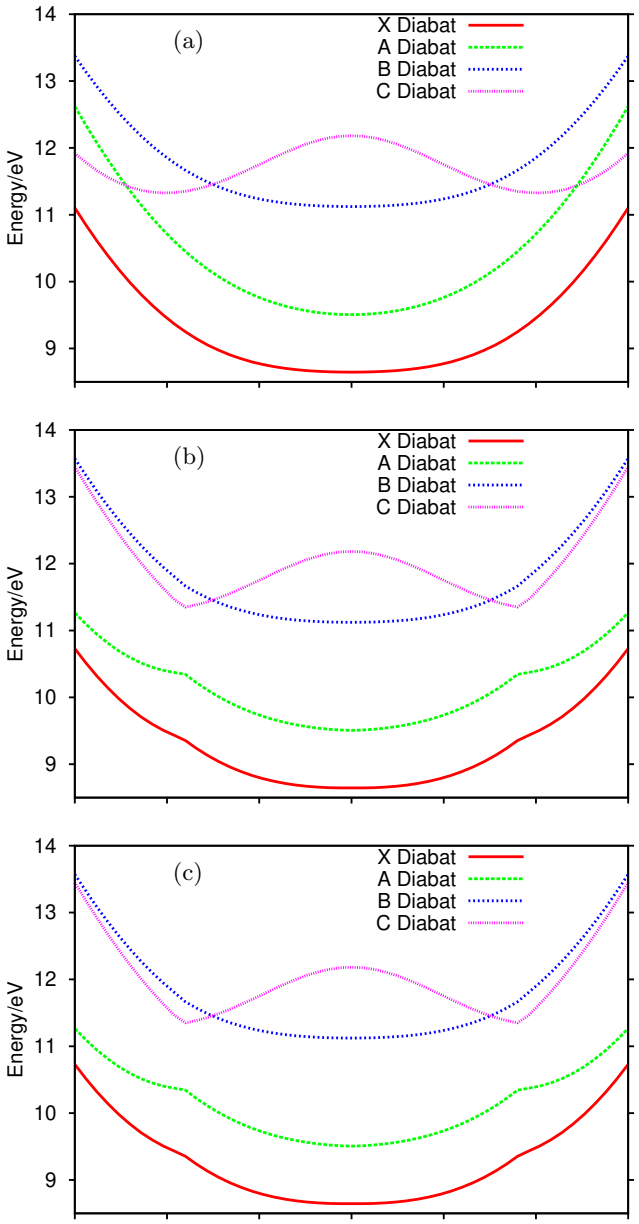

FIG. 7. Diabatic potential energy curves of the butatriene cation as a function of the mass-frequency scaled $5A_u$ normal mode: (a) Curves calculated using the Procrustes diabatization scheme; (b) energies calculated with the propagation diabatization method; (c) projection diabatization generated diabats. In all plots the red, solid line is the $\tilde{X}$ diabat, the green, dashed line is the $\tilde{A}$ diabat, the blue, hashed line is the $\tilde{B}$ diabat, and the pink, dashed line is the $\tilde{C}$ diabat.

$\tilde{A}$ attain higher energies at the extremities when using the projection method. The reason for this is the mixing in of higher lying states; a total of 20 states were included in the CI-only calculations in the projection scheme as, outlined in Section II C 2, from which components could be included in the 4 diabats. We have performed equivalent Procrustes diabatization calculations with increasing numbers of states (all significantly higher lying in this region of configuration space) and the double well character of diabat $\tilde{C}$ diminishes as the peak at $q_5 = 0$ flattens out with increasing numbers of states. This reinforces the point made in Section II C 2 that the projection diabatization scheme is able to follow diabats more closely due to this large space of states whereas the Procrustes and propagation methods are constrained to the space of (in this case) 4 states.

In conclusion, the propagation diabatization struggles in this case, failing to locate a state crossing which is dealt with smoothly by the Procrustes diabatization. The projection diabatization scheme, by dint of its including a larger number of states gives quite different diabats from the other two methods.

3. Potential Energy Curves in 1 Dimension: 14A$_g$ Mode

Having considered PECs along mode $5A_u$, we now examine the PECs of the butatriene cation as a function of q_{14} along mode $14A_g$ (with the normal-mode coordinate $5A_u$ fixed at zero). The calculation setup was the same as described in the previous section, with the same DVR and KRR width parameter (for the propagation diabatization). The results of the three calculations are shown in Fig. 8.

In Fig. 8(a) we show the PECs along mode $14A_g$ generated by the Procrustes diabatization scheme. The smooth crossing of diabats $\tilde{X}$ and $\tilde{A}$ at the well known conical intersection at $q_{14} \approx 2$ is of note as is the crossing of states $\tilde{A}$ and $\tilde{B}$ at $q_{14} \approx -4$. The slight kink in diabat $\tilde{C}$ at $q_{14} \approx 4.5$ is the result of a fifth diabat which crosses below (and becomes) $\tilde{C}$ at this point. The presence of this state was confirmed by carrying out calculations along mode $14A_g$ including more states in the adiabatic manifold. The kink in the PECs indicates that we could include more states in the calculation in order to more accurately represent the PEC in this part of configuration space. However, we note that there is no discontinuity in the PEC even though the strict diabat $\tilde{C}$ has left the space of included states, the new state being included naturally as a result of the propagation-type nature of the method, moving from geometry to geometry. As grid-based dynamics does not need the gradient of the PESs, the discontinuity in this property is not the issue it would be if a trajectory-based method were to be used.

In Fig. 8(b) we present the result of the calculation using the propagation diabatization method and immediately notice the same feature in diabat $\tilde{C}$ at $q_{14} \approx 4.5$

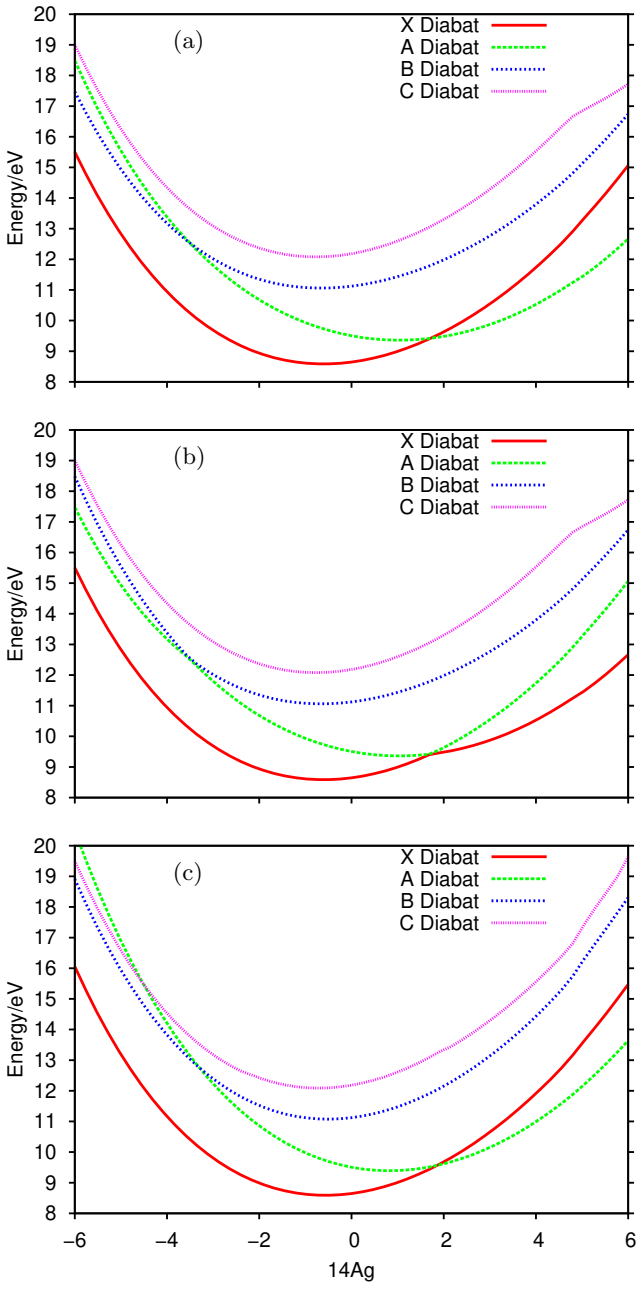

FIG. 8. Diabatic potential energy curves of the butatriene cation as a function of the mass-frequency scaled $14A_g$ normal mode: (a) Curves calculated using the Procrustes diabatization scheme; (b) energies calculated with the propagation diabatization method; (c) projection diabatization generated diabats. In all plots the red, solid line is the $\tilde{X}$ diabat, the green, dashed line is the $\tilde{A}$ diabat, the blue, hashed line is the $\tilde{B}$ diabat, and the pink, dashed line is the $\tilde{C}$ diabat.

as observed with the Procrustes method. The propagation diabatization scheme is, as noted before, also able to naturally incorporate intruder states into the manifold of diabatic states when they appear. However, the propagation diabatization scheme shows the same failures as was seen along mode $5A_u$; specifically, it fails to describe the crossing of diabats $\tilde{X}/\tilde{A}$ and of $\tilde{A}/\tilde{B}$ which were correctly seen using the Procrustes method. The reason, as was true in the case of the q_5 PECs, is the orthogonality of the NACTs to the integration path, which results in discontinuities in the ADT matrix at the points of intersection; the symmetry of the molecule is the key issue here.

In Fig. 8(c) we present the PECs along mode $14A_g$ created using the projection diabatization scheme. Clearly the method correctly describes the same state crossings as were seen in the Procrustes case, with an additional crossing between diabats $\tilde{A}$ and $\tilde{C}$ at $q_{14} \approx -4.5$. This additional crossing is another manifestation of the diabatization scheme's use of the additional states it has at its disposal to allow the following of the diabats as strictly as possible. All of the states reach significantly higher energies at large negative coordinates than in the other two cases as a result of this mixing too. At $q_{14} \approx 4.5$ we see a very slight kink in diabat $\tilde{C}$, but in the other direction to that seen with the other methods. Again the diabat is being more strictly followed when using the projection diabatization method, but this feature can prove a weakness in the method, as was noted earlier in relation to a previous study[44]. The projection diabatization method can and will follow the diabats more strictly than the other two methods, but this can be an issue when the wavepacket dynamics covers a large region of configuration space; diabats which were high in energy at the Franck-Condon point but become lower in energy elsewhere may not be included in the reference manifold in which the wavepacket must move. The Procrustes and propagation schemes are more pragmatic; at each point N diabats are formed from the N lowest energy adiabats as seen at $q_{14} \approx 4.5$ here, so the wavepacket can move along the lowest energy paths available rather than being constrained to move on the diabats defined at the reference points, hence running the risk of being diabatically trapped in a region of configuration space if there is no state low enough in energy to allow a way out.

IV. CONCLUSIONS

In this Article, we have presented a new diabatization scheme which uses the orthogonal Procrustes method to rotate a set of electronic wavefunctions at one molecular geometry to most closely match those at a nearby, reference geometry. The method has been implemented in a development version of the *Quantics* package for use within the DD-GB class of methods and its utility has been demonstrated here using the simple DD-SM approach on the classic cases of LiF and the butatriene cation. Because the potential energy operator, which is constructed by fitting functions to the diabatic energies generated by the method, is applied directly to the time-independent DVR basis, the method also works within the DD-MCTDH method presented before (the difference to the DD-SM method used here is in the manner in which the potential energy integrals are included within the equations-of-motion). Subsequent contributions from us will demonstrate the use of the Procrustes diabatization scheme within the DD-MCTDH method (including the secondary fitting used on top of the initial KRR fitting of the PESs[41,42]), but our main concern here was the diabatization method itself and the demonstration of its effectiveness in providing smooth, properly-coupled PESs.

We have also taken this opportunity to compare the Procrustes diabatization scheme with those previously implemented for use with DD-GB methods in the *Quantics* code, namely the propagation and projection diabatization approaches. We have discussed the advantages and limitations of all three methods, but overall we feel that the new method brings together the good points of the other two methods whilst avoiding some of their pitfalls. The Procrustes method has the ability to take account of new states coming down in energy, and becoming significant with respect to the dynamics, in different parts of configuration space as is done with the propagation method but not the projection approach. We avoid the need to calculate NACTs and are able to reproduce the crossing of diabats in pathological cases, as seen with the one-dimensional cuts of butatriene where symmetry meant that the propagation scheme could not cope with the discontinuity in the ADT matrix that the projection and Procrustes methods took in their stride. The path-dependence of the Procrustes diabatization method is an outstanding issue, as it is with the propagation diabatization, but the choice of paths is repeatable and avoids the closed loops which may cause issues with the non-uniqueness of the diabats, whilst precautions are taken to ensure the correct symmetry of the PESs and couplings.

A further, possible issue with the new diabatization scheme (as well as with the projection diabatization method) may also arise if very large systems are considered: when calculating the overlap of wavefunctions at successive geometries, even though the individual changes in the MOs may be small, the cumulative effect of all of the overlaps having a magnitude less than 1 is that the total wavefunction overlap tends to zero. The practical effect of this may be to make the diabatization prone to rounding error. This has not been accounted for in the current method as the systems under consideration are sufficiently small, but will be worth bearing in mind in future developments.

We also note that all three diabatization methods, as implemented, rely on the use of CI-type wavefunction methods (usually CASSCF) to allow the straightforward overlap of wavefunctions. Such methods are not the most

easily used by non-specialists, so further work will be required to extend the diabatization schemes to more black box electronic structure methods if the dynamics methods are to be taken up as useful tools more widely.

Overall, we commend the Procrustes scheme as a diabatization method which could be used, not just for on-the-fly grid-based dynamics but for those who wish to pre-fit PESs, or use trajectory-based dynamics methods. Future work will focus on applying the new method to larger systems as well as combining it with spin-orbit couplings to allow diabatic nuclear dynamics on different spin states to be carried out. We will also be looking into extending the number of coordinate systems we can use with our method; currently we are limited to using normal modes because of the automatic construction of the kinetic energy operator, but it would be desirable to be able to use, for example, valence-type coordinates to allow easier study of larger magnitude dynamics including photoisomerization or the photostability of DNA bases.

V. SUPPLEMENTARY MATERIAL

Supplementary material providing more details of the methods described herein, as well as results augmenting those presented here, is available for this paper. The method used to overlap two CASSCF wavefunctions at different geometries using Molpro is described in the first section. In the second section the propagation diabatization scheme used for DD-GB methods based on KRR fitting is described in detail as it differs in certain ways from the originally published algorithm used with DD-vMCG. The third section describes the Procrustes diabatization algorithm step-by-step when using Molpro. The fourth and final section contains extra results on the potential energy curves of LiF, particularly demonstrating the robustness of the Procrustes diabatization method to variation of user-controlled parameters: the first part deals with the effect of changing the width of the KRR kernel; the second part considers the effect of ignoring increasing numbers of configurations when overlapping CASSCF wavefunctions.

VI. DATA AVAILABILITY

Data from Figs. 1-8 are available at wrap.warwick.ac.uk/132731.

VII. ACKNOWLEDGEMENTS

We thank Dr. Christopher Robertson for useful discussions had during the course of this work. The authors gratefully acknowledge funding by the Engineering and Physical Sciences Research Council (EPSRC) through the award of grant EP/S028986/1. The authors also gratefully acknowledge high-performance computing resources provided by the Scientific Computing Research Technology Platform at the University of Warwick.

[1] G. Worth, H.-D. Meyer, and L. Cederbaum, J. Chem. Phys. **105**, 4412 (1996).
[2] A. Raab, G. Worth, H.-D. Meyer, and L. Cederbaum, J. Chem. Phys. **110**, 936 (1999).
[3] T. J. Penfold and G. A. Worth, J. Chem. Phys. **131**, 064303/1 (2009).
[4] S. Neville and G. Worth, J. Chem. Phys. **140**, 034317/1 (2014).
[5] K. Saita, M. Nix, and D. Shalashilin, Phys. Chem. Chem. Phys. **15**, 16227 (2013).
[6] T. J. Penfold, E. Gindensperger, C. Daniel, and C. M. Marian, **118**, 6975 (2018).
[7] M. Fumanal, F. Plasser, S. Mai, C. Daniel, and E. Gindensperger, J. Chem. Phys. **148**, 124119 (2018).
[8] M. Fumanal, E. Gindensperger, and C. Daniel, J. Chem. Theory Comput. **13**, 1293 (2017).
[9] C. T. Middleton, K. de La Harpe, C. Su, Y. K. Law, C. E. Crespo-Hernández, and B. Kohler, Annu. Rev. Phys. Chem. **60**, 217 (2009).
[10] M. Richter, P. Marquetand, J. González-Vázquez, I. Sola, and L. González, J. Phys. Chem. Lett. **3**, 3090 (2012).
[11] H. R. Hudock, B. G. Levine, A. L. Thompson, H. Satzger, D. Townsend, N. Gador, S. Ullrich, A. Stolow, and T. J. Martínez, J. Phys. Chem. A **111**, 8500 (2007).
[12] J. C. Tully and R. K. Preston, J. Chem. Phys. **55**, 562 (1971).
[13] J. C. Tully, J. Chem. Phys. **93**, 1061 (1990).
[14] M. Richter, P. Marquetand, J. González-Vázquez, I. Sola, and L. González, J. Chem. Theory Comput. **7**, 1253 (2011).
[15] J. C. Tully, Faraday Discuss. **110**, 407 (1998).
[16] D. F. Coker, in *Computer Simulation in Chemical Physics*, edited by M. P. Allen and D. J. Tildesley (Kluwer Academic, Dordrecht, 1993) pp. 315–377.
[17] M. S. Topaler, M. D. Hack, T. C. Allison, Y.-P. Liu, S. L. Mielke, D. W. Schwenke, and D. G. Truhlar, J. Chem. Phys. **106**, 8699 (1997).
[18] B. R. Smith, M. J. Bearpark, M. A. Robb, F. Bernardi, and M. Olivucci, Chem. Phys. Lett. **242**, 27 (1995).
[19] M. J. Bearpark, F. Bernardi, M. Olivucci, M. A. Robb, and B. R. Smith, J. Am. Chem. Soc. **118**, 5254 (1996).
[20] R. Mitrić, J. Petersen, and V. Bonacić-Koutecký, Phys. Rev. A **79**, 053416/1 (2009).
[21] P. Lisinetskaya and R. Mitrić, Phys. Rev. A **83**, 033408/1 (2011).
[22] M. D. Hack, A. Jasper, Y. L. Volobuev, D. W. Schwenke, and D. G. Truhlar, J. Phys. Chem. A **103**, 6309 (1999).
[23] M. Ben-Nun, J. Quenneville, and T. J. Martínez, J. Phys. Chem. A **104**, 5161 (2000).
[24] M. Ben-Nun and T. J. Martínez, Adv. Chem. Phys. **121**, 439 (2002).
[25] T. Mori, W. Glover, M. Schuurman, and T. Martínez, J. Phys. Chem. A **116**, 2808 (2012).
[26] D. Shalashilin, J. Chem. Phys. **130**, 244101/1 (2009).
[27] K. Saita and D. Shalashilin, J. Chem. Phys. **137**, 22A506/1 (2012).
[28] M. H. Beck, A. Jäckle, G. A. Worth, and H. D. Meyer, Phys. Rep. **324**, 1 (2000).
[29] H. Wang, J. Phys. Chem. A **119**, 7951 (2015).
[30] O. Vendrell and H.-D. Meyer, J. Chem. Phys. **134**, 044135/1 (2011).
[31] H. Wang and M. Thoss, J. Chem. Phys. **119**, 1289 (2003).
[32] I. Burghardt, H.-D. Meyer, and L. S. Cederbaum, J. Chem. Phys. **111**, 2927 (1999).
[33] G. Worth and I. Burghardt, Chem. Phys. Lett. **368**, 502 (2003).
[34] K. Giri, E. Chapman, C. S. Sanz, and G. Worth, J. Chem. Phys. **135**, 044311/1 (2011).
[35] M. Persico and G. Granucci, Theo. Chem. Acc. **133**, 1526/1 (2014).

[36]G. A. Worth, M. A. Robb, and B. Lasorne, Mol. Phys. **106**, 2077 (2008).
[37]G. A. Worth, M. A. Robb, and I. Burghardt, Faraday Discuss. **127**, 307 (2004).
[38]B. Lasorne, M. A. Robb, and G. A. Worth, Phys. Chem. Chem. Phys. **9**, 3210 (2007).
[39]G. Richings, I. Polyak, K. Spinlove, G. Worth, I. Burghardt, and B. Lasorne, Int. Rev. Phys. Chem. **34**, 269 (2015).
[40]G. Richings and S. Habershon, J. Chem. Theory Comput. **13**, 4012 (2017).
[41]G. W. Richings, C. Robertson, and S. Habershon, J. Chem. Theory Comput. **15**, 857 (2019).
[42]G. Richings and S. Habershon, J. Chem. Phys. **148**, 134116/1 (2018).
[43]G. Richings and S. Habershon, Chem. Phys. Lett. **683**, 228 (2017).
[44]G. W. Richings, C. Robertson, and S. Habershon, Faraday Discuss. **216**, 476 (2019).
[45]C. E. Rasmussen and C. K. Williams, *Gaussian Processes for Machine Learning* (The MIT Press, Cambridge, Massachusetts, 2006).
[46]C. Williams, in *Handbook of Brain Theory and Neural Networks*, edited by M. Arbib (The MIT Press, Cambridge, Massachusetts, 2002) pp. 466–470.
[47]A. Bartók and G. Csányi, Int. J. Quantum. Chem. **115**, 1051 (2015).
[48]J. P. Alborzpour, D. P. Tew, and S. Habershon, J. Chem. Phys. **145**, 174112/1 (2016).
[49]L. Mones, N. Bernstein, and G. Csányi, J. Chem. Theory Comput. **12**, 5100 (2016).
[50]J. Quinonero-Candela, C. E. Ramussen, and C. K. I. Williams, in *Large-Scale Kernel Machines*, edited by L. Bottou, O. Chapelle, D. DeCoste, and J. Weston (The MIT Press, Cambridge, Massachusetts, 2007) pp. 203–224.
[51]D. Duvenaud, H. Nickisch, and C. E. Rasmussen, in *Neural Information Processing Systems Conference* (2011).
[52]K. Chalupka, C. K. Williams, and I. Murray, J. Mach. Learn Res. **14**, 333 (2013).
[53]D. Hu, Y. Xie, X. Li, L. Li, and Z. Lang, J. Phys. Chem. Lett. **9**, 2725 (2018).
[54]P. O. Dral, M. Barbatti, and W. Thiel, J. Phys. Chem. Lett. **9**, 5660 (2018).
[55]W.-K. Chen, X.-Y. Liu, W.-H. Fang, P. O. Dral, and G. Cui, J. Phys. Chem. Lett. **9**, 6702 (2018).
[56]J. Westermayr, M. Gastegger, M. F. Menger, S. Mai, L. González, and P. Marquetand, Chem. Sci. **10**, 8100 (2019).
[57]T. Pacher, C. Mead, L. Cederbaum, and H. Köppel, J. Chem. Phys. **91**, 7057 (1989).
[58]C. A. Mead and D. G. Truhlar, J. Chem. Phys. **77**, 6090 (1982).
[59]M. Baer, Chem. Phys. Lett. **35**, 112 (1975).
[60]M. Baer, Chem. Phys. **15**, 49 (1976).
[61]M. Baer, J. Phys. Chem. A **104**, 3181 (2000).
[62]M. Baer, Mol. Phys. **75**, 49 (1992).
[63]A. Alijah and M. Baer, J. Phys. Chem. A **104**, 389 (2000).
[64]H. J. Werner and W. Meyer, J. Chem. Phys. **74**, 5802 (1981).
[65]B. Esry and H. Sadeghpour, Phys. Rev. A **68**, 042706/1 (2003).
[66]G. Richings and G. Worth, J. Phys. Chem. A **119**, 12457 (2015).
[67]A. Das, D. Mokhopadhyay, S. Adhikari, and M. Baer, Chem. Phys. Lett. **517**, 92 (2011).
[68]A. Das, D. Mokhopadhyay, S. Adhikari, and M. Baer, Int. J. Quantum. Chem. **112**, 2561 (2012).
[69]A. Das and D. Mokhopadhyay, Int. J. Quantum. Chem. **112**, 2767 (2012).
[70]T. Pacher, L. Cederbaum, and H. Köppel, J. Chem. Phys. **89**, 7367 (1988).
[71]R. J. Cave and J. F. Stanton, J. Chem. Phys. **140**, 214112/1 (2014).
[72]W. Domcke and C. Woywod, Chem. Phys. Lett. **216**, 362 (1993).
[73]W. Domcke, C. Woywod, and M. Stengle, Chem. Phys. Lett. **226**, 257 (1994).
[74]F. Venghaus and W. Eisfeld, J. Chem. Phys. **144**, 114110/1 (2016).
[75]C. E. Hoyer, X. Xu, D. Ma, L. Gagliardi, and D. G. Truhlar, J. Chem. Phys. **141**, 114104/1 (2014).
[76]J. Aragó and A. Troisi, J. Chem. Phys. **142**, 164107/1 (2015).
[77]C. Robertson, J. González-Vázquez, I. Corral, S. Díaz-Tendero, and C. Díaz, J. Comp. Chem. **40**, 794 (2019).
[78]I. Petsalakis, G. Theodorakopoulos, and C. Nicolaides, Chem. Phys. Lett. **185**, 359 (1991).
[79]R. Cimiraglia, J.-P. Malrieu, M. Persico, and F. Spiegelmann, J. Phys. B: At. Mol. Phys. **18**, 3073 (1985).
[80]D. Simah, B. Hartke, and H.-J. Werner, J. Chem. Phys. **111**, 4523 (1999).
[81]H. Nakamura and D. G. Truhlar, J. Chem. Phys. **115**, 10353 (2001).
[82]H. Nakamura and D. G. Truhlar, J. Chem. Phys. **117**, 5576 (2002).
[83]H. Nakamura and D. G. Truhlar, J. Chem. Phys. **118**, 6816 (2003).
[84]A. Thiel and H. Köppel, J. Chem. Phys. **110**, 9371 (1999).
[85]H. Köppel, J. Gronki, and S. Mahapatra, J. Chem. Phys. **115**, 2377 (2001).
[86]C. Evenhuis and T. J. Martínez, J. Chem. Phys. **135**, 224110/1 (2011).
[87]R. Thürwächter and P. Halvick, Chem. Phys. **221**, 33 (1997).
[88]T. Romero, A. Aguilar, and F. Gadea, J. Chem. Phys. **110**, 6219 (1999).
[89]P. H. Schönemann, Psychometrika **31**, 1 (1966).
[90]J. C. Gower, WIRES Computational Statistics **2**, 503 (2010).
[91]M. Liu, X. Chen, A. Grofe, and J. Gao, J. Phys. Chem. Lett. **9**, 6038 (2018).
[92]C. W. Bauschlicher Jr. and S. R. Langhoff, J. Chem. Phys. **89**, 4246 (1988).
[93]N. Balakrishnan, B. Esry, H. Sadeghpour, S. Cornett, and M. Cavagnero, **60**, 1407 (1999).
[94]A. Bandrauk and J.-M. Gauthier, J. Phys. Chem. **93**, 7552 (1989).
[95]C. Cattarius, G. Worth, H.-D. Meyer, and L. Cederbaum, J. Chem. Phys. **115**, 2088 (2001).
[96]G. Worth, P. Hunt, and M. Robb, J. Phys. Chem. A **107**, 621 (2003).
[97]S. Sardar, A. K. Paul, P. Mondal, B. Sarkar, and S. Adhikari, Phys. Chem. Chem. Phys. **10**, 6388 (2008).
[98]G. Richings and G. Worth, Chem. Phys. Lett. **683**, 606 (2017).
[99]P. Dirac, Proc. Cambridge Philos. Soc. **26**, 376 (1930).
[100]J. Frenkel, *Wave Mechanics* (Clarendon Press, Oxford, 1934).
[101]I. Sobol, USSR Comput. Maths. Math. Phys. **7**, 86 (1967).
[102]I. Antonov and V. Saleev, USSR Comput. Maths. Math. Phys. **19**, 252 (1979).
[103]S. Joe and F. Y. Kuo, ACM Trans. Math. Software **29**, 49 (2003).
[104]H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, W. Györffy, D. Kats, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppel, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, and M. Wang, "Molpro, version 2015.1, a package of ab initio programs," (2015), www.molpro.net.
[105]D. J. Tannor, *Introduction to Quantum Mechanics A Time-Dependent Perspective* (University Science Books, Sausalito, California, 2007).
[106]G. Granucci, M. Persico, and A. Toniolo, J. Chem. Phys. **114**, 10608 (2001).
[107]A. Mandal, S. S. Yamijala, and P. Huo, J. Chem. Theory Comput. **14**, 1828 (2018).

[108] J. S. Sandoval C., A. Mandal, and P. Huo, J. Chem. Phys. **149**, 044115/1 (2018).
[109] F. Plasser, G. Granucci, J. Pittner, M. Barbatti, M. Persico, and H. Lischka, J. Chem. Phys. **137**, 22A514/1 (2012).
[110] B. K. Kendrick, C. A. Mead, and D. G. Truhlar, Chem. Phys. **277**, 31 (2002).
[111] N. Matsunaga and D. R. Yarkony, Mol. Phys. **93**, 79 (1998).
[112] O. Godsi, C. R. Evenhuis, and M. A. Collins, J. Chem. Phys. **125**, 104105/1 (2006).
[113] H.-J. Werner and P. J. Knowles, J. Chem. Phys. **82**, 5053 (1985).
[114] P. J. Knowles and H.-J. Werner, Chem. Phys. Lett. **115**, 259 (1985).
[115] A. Troisi and G. Orlandi, J. Chem. Phys. **118**, 5356 (2003).
[116] L. Cederbaum, J. Schirmer, and H.-D. Meyer, J. Phys. A: Math. Gen. **22**, 2427 (1989).
[117] J. B. Delos, Rev. Mod. Phys. **53**, 287 (1981).
[118] F. Gatti, B. Lasorne, , H.-D. Meyer, and A. Nauts, eds., *Applications of Quantum Dynamics in Chemistry* (Springer, Gewerbestrasse 11, 6330 Cham, Switzerland, 2017).
[119] G. Worth, K. Giri, G. Richings, I. Burghardt, M. Beck, A. Jäckle, and H.-D. Meyer, "The Quantics Package, Version 1.1," Tech. Rep. (University of Birmingham, Birmingham, UK, 2015).