PAPER

Curve crossing in a manifold of coupled electronic states: direct quantum dynamics simulations of formamide†

K. Eryn Spinlove, [ID] *[abd] Gareth W. Richings, [ID] [c] Michael A. Robb [ID] [d] and Graham A. Worth [ID] *[a]

Quantum dynamics simulations are an important tool to evaluate molecular behaviour including the, often key, quantum nature of the system. In this paper we present an algorithm that is able to simulate the time evolution of a molecule after photo-excitation into a manifold of states. The direct dynamics variational multi-configurational Gaussian (DD-vMCG) method circumvents the computational bottleneck problems of traditional grid-based methods by computing the potential energy functions on-the-fly, *i.e.* only where required. Unlike other commonly used direct dynamics methods, DD-vMCG is fully quantum mechanical. Here, the method is combined with a novel on-the-fly diabatisation scheme to simulate the short-time dynamics of the key molecule formamide and its acid analogue formimidic acid. This is a challenging test system due to the nature and large number of excited states, and eight coupled states are included in the calculations. It is shown that the method is able to provide unbiased information on the product channels open after excitation at different energies and demonstrates the potential to be a practical scheme, limited mainly by the quality of the quantum chemistry used to describe the excited states.

1 Introduction

Photochemistry is an area of chemistry in which quantum effects play a key role in how molecules behave. After absorbing a photon, a molecule is promoted into an excited electronic state and subsequently evolves on the excited state potential energy surface. In many cases, this state is one of a whole manifold of electronic states close in energy that are coupled by the nuclear motion, termed vibronic coupling. This coupling can lead to a breakdown of the Born–Oppenheimer

[a]Dept. of Chemistry, University College London, 20 Gordon St., London, UK. E-mail: g.a.worth@ucl.ac.uk

[b]School of Chemistry, University of Birmingham, Birmingham, B15 2TT, UK

[c]Dept. of Chemistry, University of Warwick, Gibbet Hill, Coventry, CV4 7AL, UK

[d]Dept. of Chemistry, Imperial College London, London, SW7 2AZ, UK

approximation and it is not valid to separate the electronic and nuclear motion. To correctly simulate the time evolution of a molecule after photo-excitation it is therefore necessary to solve the time-dependent Schrödinger equation (TDSE) for the nuclear wavepacket created by the photon absorption.

A variety of dynamics methods have been developed with a view of solving the TDSE. Generally these methods fall into two categories, termed semi-classical if the nuclei are treated using classical equations of motion which are coupled to the electronic motion, or quantum if the nuclei are treated quantum mechanically. Quantum molecular dynamics methods attempt to directly solve the full TDSE for both nuclei and electrons. Using a grid-based representation, standard approaches provide a pictorial representation of the evolving wavepacket and information can be extracted from the wavepacket over the course of a calculation. However, these exact methods suffer from exponential scaling with an increase in the number of degrees of freedom, and powerful numerical methods are required to describe the time evolution of the system. The multi-configuration time-dependent Hartree (MCTDH) method,[1–3] and the derivative multi-layer MCTDH method,[4] are probably the most powerful algorithms at present for quantum dynamics able to solve the TDSE for many tens of degrees of freedom.

A significant drawback of grid-based methods is the requirement for global potential energy surfaces that must be computed and fit to analytic functions before any calculations are performed. This presents a major bottleneck for polyatomic systems due to the huge number of quantum-chemistry calculations and complex fitting procedure required, although recent work by one of us is beginning to alleviate this problem.[5–7] Direct dynamics that compute the potential on-the-fly are thus receiving significant interest at present as they circumvent this problem by only calculating the potential in regions of space visited by the system. Most direct dynamics methods, however, use a semi-classical description of the evolving system, with surface hopping methods being the most popular approach. In these, an ensemble, or swarm, of trajectories are propagated on the adiabatic surfaces of a system, which are supplied by an external electronic structure method. In regions dominated by non-adiabatic effects, hopping can occur between surfaces. A variety of approaches has been developed with differing methods by which the trajectories are described and propagated, the manner by which the hopping occurs, and the effect a successful hop has on the ensemble.[8–11]

A novel semi-classical method is exact factorization.[12,13] Using a wavefunction ansatz in which the nuclear and electronic parts are factorised, similar to the Ehrenfest method, it has explicit time-dependent electronic motion. In addition, the electronic motion is explicitly coupled to the nuclear coordinate. The method has only been applied to one-dimensional systems but could prove to be a promising alternative method for direct dynamics.

Semi-classical dynamics methods have much better scaling than quantum dynamics methods, and provide a suitable framework for direct dynamics calculations, but are unfortunately not guaranteed to converge on the correct result due to the missing nuclear correlation. The underlying classical structure also means that they cannot correctly describe tunnelling through a barrier. A different approach is *ab initio* multiple spawning (AIMS).[14–16] In this, the nuclear wavepacket is described by a superposition of Gaussian basis functions, known as Gaussian wavepackets (GWPs). The expansion coefficients solve the full TDSE

while the GWPs follow classical trajectories making them suitable for direct dynamics. In principle AIMS can reach the full quantum solution, but the use of classical trajectories leads to slow convergence. AIMS has been used successfully in many applications.

A relatively recent development that can be used in direct-dynamics simulations and provides a full quantum treatment of the nuclei is the variational multi-configurational Gaussian method (DD-vMCG).[17–19] Similar to multiple spawning, it uses a superposition of GWPs as a basis set for the nuclei. The equations of motion for both GWPs and expansion coefficients are, however, variational and the GWPs follow coupled "quantum" trajectories rather than classical trajectories. Initial test calculations show that the method converges quickly on the full solution, but suffers from numerical instabilities. It also has the drawback that a diabatic representation must be used for the wavefunction and it is not clear how to diabatise a set of potentials on-the-fly.

Recently, a propagation diabatisation method was introduced.[20] This uses the relationship between adiabatic and diabatic wavefunctions to provide global diabatic surfaces from the quantum chemistry data. The relationship, however, is only exact for a complete set of states and its practical behaviour must be evaluated. Formamide (FAM), also known as methanamide, provides a suitable test system. Over the years since its first published synthesis in 1863,[21] it has been an important, but until relatively recently, often disregarded compound.

FAM has a number of industrial uses, *e.g.* as a precursor to hydrogen cyanide and as a formylating agent. It is also the smallest, most stable compound containing the four most abundant elements in the universe. FAM itself has been generally detected in the interstellar medium,[22] in comets[23,24] including Hale-Bopp, on icy grains around the protostellar object W33A,[25] and from sources SgrA and SgrB2.[26,27] In recent years, support has been growing in the apperception of FAM as a key abiotic precursor to the synthesis of pyrimidines and nucleobases in prebiotic Earth, supporting the RNA world theory. A comprehensive review, "Formamide and the origin of life" by Saladino *et al.*,[28] details the diverse conditions under which FAM has been used in the synthesis of nucleic acid bases and nucleosides, the phosphorylation of nucleosides, as a catalyst for the oligomerization and polymerization of nucleotides, and the synthesis of pre-metabolic components. As the Hadean Earth was not shielded from solar radiation, it is reasonable to consider not only the effects of temperature, catalysis and concentration but also the photo-activity of FAM in its viability as a precursor to prebiotic life.

Despite the plethora of information regarding the behaviour, uses and function of FAM, to date there have been a limited number of computational studies, even though it was the subject of early molecular orbital (MO) calculations.[29] Most of the recent work focuses on the ground state tautomerism and the importance of solvation for the proton transfer process[30,31] and the thermal decomposition pathways,[32,33] as well as the nature of the excited states and possible processes.[34–38] As of yet, however, there are no published full quantum dynamics studies on the behaviour of FAM or its acid analogue formimidic acid (FIM) after photo-excitation, despite the potential importance of light in prebiotic chemistry. A direct dynamics surface hopping study of protonated formamide has been carried out.[39] In this paper we apply the DD-vMCG method to the excited state dynamics of formamide and formamidic acid to demonstrate the utility of the method as well as to shed some light on the excited state dynamics of these key molecules.

2 Background theory: DD-vMCG

The variational multi-configurational Gaussian (vMCG)[40–43] ansatz has the form

$$\Psi(\mathbf{x}, t) = \sum_{j=1}^{n} A_j(t) g_j(\mathbf{x}, t) \tag{1}$$

where the basis functions and expansion coefficients of the system are both time-dependent. In vMCG calculations, frozen GWPs are used in which the width matrix is kept diagonal with fixed values during the wavepacket propagation. The multidimensional GWPs have the form

$$g_j(\mathbf{x}, t) = \exp\left(\sum_{\kappa} \varsigma_{j\kappa} x_\kappa^2 + \xi_{j\kappa} x_\kappa + \eta_j \right) \tag{2}$$

where κ runs over the degrees of freedom of the system.

It then follows that by defining the relationships

$$\left. \begin{aligned} \varsigma_{j\kappa} &= -a_{j\kappa} \\ \xi_{j\kappa} &= 2a_{j\kappa} q_{j\kappa} + ip_{j\kappa} \\ \eta_j &= \sum_{\kappa} \left(-a_{j\kappa} q_{j\kappa}^2 - ip_{j\kappa} q_{j\kappa} \right) + i\gamma_j \end{aligned} \right\} \tag{3}$$

Eqn (2) can be written in the more intuitive Heller form:[44,45]

$$g_j(\mathbf{x}, t) = \exp\left(\sum_{\kappa} -a_{j\kappa} \left(x_\kappa - q_{j\kappa} \right)^2 + ip_{j\kappa} \left(x_\kappa - q_{j\kappa} \right) + i\gamma_j \right) \tag{4}$$

In this form it can be seen that the quadratic parameters, ς, represent the widths of the Gaussian functions, the linear parameters, ξ, represent the momentum, p, and centre coordinate, q, of the functions, and the scalar parameters, η, represent the remaining parameters, including the phase, γ, of the functions.

As with the MCTDH method, the Dirac–Frenkel variational principle is applied to the wavefunction ansatz to obtain equations of motion for the expansion coefficients

$$i\dot{A}_j = \sum_{lm} \left[\mathbf{S}_{jl} \right]^{-1} \left(\mathbf{H}_{lm} - i\tau_{lm} \right) A_m \tag{5}$$

where the overlap, $\mathbf{S}$, and Hamiltonian, $\mathbf{H}$, matrices are

$$\mathbf{S}_{jl} = \langle g_j | g_l \rangle \text{ and } \mathbf{H}_{jl} = \langle g_j | \hat{\mathbf{H}} | g_l \rangle \tag{6}$$

By collecting the GWP time-dependent parameters in a vector,

$$\Lambda_j = \{\xi_j, \eta_j\} \tag{7}$$

the equations of motion for the GWP parameters can be written in a compact matrix form as

$$i\dot{\Lambda} = [\mathbf{C}]^{-1} \mathbf{Y} \tag{8}$$

where $\mathbf{C}$ and $\mathbf{Y}$ are

$$\mathbf{C}_{j\alpha,l\beta} = \rho_{jl}(\mathbf{S}_{jl}^{(\alpha\beta)} - [\mathbf{S}^{(\alpha0)}\mathbf{S}^{-1}\mathbf{S}^{(0\beta)}]_{jl}) \tag{9}$$

$$\mathbf{Y}_{j\alpha} = \sum_l \rho_{jl}\left(\mathbf{H}_{jl}^{(\alpha0)} - [\mathbf{S}^{(\alpha0)}\mathbf{S}^{-1}\mathbf{H}]_{jl}\right) \tag{10}$$

Here, ρ_{jl} is an element of the density matrix, and the superscripts on $\mathbf{S}$ and $\mathbf{H}$ denote derivatives of the overlap and Hamiltonian matrices with respect to the GWP parameters:

$$\left.\begin{aligned}
\rho_{jl} &= A_j^* A_l \\[4pt]
\mathbf{S}_{jl}^{(\alpha\beta)} &= \left\langle \frac{\partial g_j}{\partial\lambda_{j\alpha}}\middle|\frac{\partial g_l}{\partial\lambda_{l\beta}}\right\rangle \\[4pt]
\mathbf{S}_{jl}^{(\alpha0)} &= \left\langle \frac{\partial g_j}{\partial\lambda_{j\alpha}}\middle| g_l\right\rangle \\[4pt]
\mathbf{H}_{jl}^{(\alpha0)} &= \left\langle \frac{\partial g_j}{\partial\lambda_{j\alpha}}\middle|\mathbf{H}|g_l\right\rangle
\end{aligned}\right\} \tag{11}$$

Furthermore, τ can explicitly be written as a function of the time derivative of the Gaussian parameters using the chain rule:

$$\tau_{jl} = \langle g_j|\dot{g}_l\rangle = \sum_\alpha\left\langle g_j\middle|\frac{\partial g_l}{\partial\lambda_{l\alpha}}\dot{\lambda}_{l\alpha}\right\rangle \tag{12}$$

In order to maintain the normalisation of the Gaussians during the propagation, the diagonal of the overlap time-derivative matrix, τ, can be set to zero. This defines the evolution of the otherwise undefined scalar parameters.

To evaluate the matrix elements of the potential function, the local harmonic approximation (LHA) is used, *i.e.* the potential energy is expanded to second order around each GWP time-dependent centre coordinate, $\mathbf{q}_j(t)$:

$$V_j(\mathbf{x}) = V_{j0} + \sum_\kappa V'_{j,\kappa}(x_\kappa - q_{j\kappa}) + \frac{1}{2}\sum_{\kappa\mu}V''_{j,\kappa\mu}(x_\kappa - q_{j\kappa})(x_\mu - q_{j\mu}) \tag{13}$$

In addition, if rectilinear coordinates are used, the kinetic energy operator can be taken to have a separable form, *i.e.* $\hat{T} = \sum_j -\frac{1}{2m_j}\frac{\partial^2}{\partial x_j^2}$, then the separable part of the Hamiltonian can be expanded as a power series in terms of the Gaussian moments $M_{jl}^\kappa = \langle g_j|x_\kappa|g_l\rangle$ and $M_{jl}^{\kappa\mu} = \langle g_j|x_\kappa x_\mu|g_l\rangle$ as

$$\mathbf{H}_{jl} = M_{jl}^{(0)}\mathbf{X}_l^{(0)} + \sum_\kappa M_{jl}^{(\kappa)}\mathbf{X}_l^{(\kappa)} + \sum_{\kappa\mu}M_{jl}^{(\kappa\mu)}\mathbf{X}_l^{(\kappa\mu)} + \dots \tag{14}$$

with coefficients

$$\mathbf{X}_j^{(0)} = \sum_\kappa\left(\frac{\varsigma_{j\kappa}}{m_\kappa} + \frac{\xi_{j\kappa}^2}{2m_\kappa}\right) + V_{j0} - \sum_\kappa V'_{j,\kappa}q_{j\kappa} + \frac{1}{2}\sum_{\kappa\mu}V''_{j,\kappa\mu}q_{j\kappa}q_{j\mu} \tag{15}$$

$$\mathbf{X}_j^{(\kappa)} = -\frac{2}{m_\kappa}\varsigma_{j\kappa}\xi_{j\kappa} + V'_{j,\kappa} - \sum_\mu V''_{j,\kappa\mu}q_{j\mu} \tag{16}$$

$$\mathbf{X}_j^{(\kappa\mu)} = -\frac{2}{m_\kappa}\varsigma_{j\kappa}\varsigma_{j\mu} + \frac{1}{2}V''_{j,\kappa\mu} \tag{17}$$

The **Y**-vector can also be separated into two parts:

$$\mathbf{Y} = \mathbf{Y}_0 + \mathbf{Y}_R \tag{18}$$

The part of the Hamiltonian that can be expressed in terms of $\mathbf{S}^{(0\alpha)}$ is represented by $\mathbf{Y}_0$, with the remaining part of the Hamiltonian expressed by $\mathbf{Y}_R$, known as the residual term. The residual term includes both of the correlation terms, as well as the higher order terms of the separable part of the Hamiltonian. Using the relationship between the overlap matrix elements, $\mathbf{S}^{(\alpha\beta)}$, and the Gaussian moments, M, and comparing eqn (9) and (10), eqn (18) can be rewritten as

$$\mathbf{Y}_{j\alpha} = \sum_{l\beta}\mathbf{C}_{j\alpha,l\beta}\mathbf{X}_l^{(\beta)} + \mathbf{Y}_{R,j\alpha} \tag{19}$$

Consequently, the equation of motion for the Gaussian parameters can be further simplified to

$$i\dot{\Lambda} = \mathbf{X} + \mathbf{C}^{-1}\mathbf{Y}_R \tag{20}$$

The significance of this so-called "CX formalism" is realised when the practical implications are considered. As the **C**-matrix is inverted during propagation, the removal of part of the Hamiltonian from the $\mathbf{C}^{-1}\mathbf{Y}$-term results in a reduction of possible numerical errors and hence increases the stability of the propagation. Another desirable feature is that the parameter equations of motion can be divided into classical and non-classical parts.

The separation of classical and non-classical parts is achieved in the vMCG method by putting only the classical Gaussian propagation terms into **X** with the remaining (quantum) contributions kept in the $\mathbf{Y}_R$ term.

When frozen GWPs are used, only the scalar and linear parameters are varied with time. Consequently, the Hamiltonian can be written in terms of eqn (15) and (16) as

$$\mathbf{H}_{jl} = \mathbf{S}_{jl}^{(0)}\mathbf{X}_l^{(0)} + \sum_\kappa \mathbf{S}_{jl}^{(0\kappa)}\mathbf{X}_l^{(\kappa)} + \sum_{\kappa\mu} M_{jl}^{(\kappa\mu)}\mathbf{X}_l^{(\kappa\mu)} + \dots \tag{21}$$

In this form, the linear parameter of the equations of motion eqn (20) can be written

$$i\dot{\xi}_{j\kappa} = -i2\varsigma_{j\kappa}\dot{q}_{j\kappa} - \dot{p}_{j\kappa} \tag{22}$$

$$= \mathbf{X}_j^{(\kappa)} + \sum_{l\mu}\mathbf{C}_{j\kappa l\mu}^{-1}\mathbf{Y}_{R,l\mu} \tag{23}$$

$$= -\frac{2}{m_\kappa}\varsigma_{j\kappa}\xi_{j\kappa} + V'_{j,\kappa} - \sum_\mu V''_{j,\kappa\mu}q_{j\mu} + \sum_{l\mu}\mathbf{C}_{j\kappa l\mu}^{-1}\mathbf{Y}_{R,l\mu} \tag{24}$$

where the second-order moments are put into $\mathbf{Y_R}$. If the terms $\dfrac{2}{m_\kappa}\varsigma_{j\kappa}(\xi_{j\kappa} - ip_{j\kappa})$ and $\displaystyle\sum_\mu V''_{j,\kappa\mu}q_{j\mu}$ are added to $\mathbf{Y_R}$, defining

$$i\dot{\xi}_{\mathrm{R},j\kappa} = \sum_{l\mu}\mathbf{C}_{j\kappa l\mu}{}^{-1}\mathbf{Y}_{\mathrm{R},l\mu} \tag{25}$$

results in an equation of motion for the linear parameters

$$i\dot{\bar{\xi}}_{j\kappa} = V'_{j,\kappa} - i\frac{p_{j\kappa}}{m_\kappa} + \Re\left(i\dot{\xi}_{\mathrm{R},j\kappa}\right) - \frac{1}{2\varsigma_{j\kappa}}i\Im\left(i\dot{\xi}_{\mathrm{R},j\kappa}\right) \tag{26}$$

where $\bar{\xi}_{j\kappa} = q_{j\kappa} + ip_{j\kappa}$ is propagated (as opposed to $\xi_{j\kappa} = -2\varsigma_{j\kappa}q_{j\kappa} + ip_{j\kappa}$). If the last two terms are ignored, the GWPs will follow classical trajectories. It should also be stated that for coherent states in an appropriate-width harmonic well, the last two terms of this expression cancel.

2.1 The direct-dynamics variational multi-configurational Gaussian method

Using the LHA, eqn (13), the Hamiltonian matrix elements required to integrate the equations of motion (eqn (5) and (8)) need only the energies, gradients and Hessians at the centres of the Gaussians at each time step. These values are easily evaluated on-the-fly *via* an interface with an external quantum chemistry software package. It is, however, undesirable to carry out these time-consuming electronic structure calculations for every point reached by the GWPs. The DD-vMCG method as implemented in the Quantics package[46] bypasses this issue by creating a database of electronic energies and other calculated information and, when possible, uses this information to construct the PESs.[47] The idea of using a database to store the information has been used before in both classical trajectory[48,49] and quantum trajectory[50] methods.

Instead of calling the external electronic structure program at each time step, the database is consulted. The Euclidean norm of the difference vector between all points of the new molecular geometry and all existing geometries in the database is then calculated as a method by which the distance between two structures can be measured. If the lowest value of the calculated norm is higher than the predefined value then the electronic structure program is called in order to calculate the energy, gradient and Hessian of the PES, the LHA constructed, and the information is stored in the database. Additional information such as the dipole moments, derivative couplings and MO coefficients may also be stored.

If the lowest value of the norm is lower than the predefined value, a modified Shepard interpolation[50] is carried out in order to obtain the energy, gradient and Hessian for the LHA. This Shepard weighted interpolation has the form

$$V(\mathbf{x}) = \sum_i \omega_i(\mathbf{x})T_i(\mathbf{x}) \tag{27}$$

where T_i is the ith database entry centred Taylor series expansion and ω_i is defined as

$$\omega_i(\mathbf{x}) = \frac{v_i(\mathbf{x})}{\displaystyle\sum_j v_j(\mathbf{x})} \tag{28}$$

where

$$v_i(\mathbf{x}) = \frac{1}{|\mathbf{x} - \mathbf{x}_i|^{2p}} \tag{29}$$

Following earlier work, $p = 2$ was used.[43] This set of equations (eqn (27)–(29)) is used to gain the interpolated energy, gradient and Hessian. The choice of the parameter to trigger a new calculation appears to be system-dependent,[43] though running an exploratory set of calculations varying this value will yield informative results. It should also be noted, from a practical perspective, that if an electronic structure calculation fails to complete, this interpolation is also used.

Although the use of the database in this way can significantly reduce the amount of real time required for a full propagation, there remains the issue of the evaluation of the Hessian matrix, which is a computationally intensive process. Various methods for the approximation of the Hessian exist requiring only a reference Hessian, and the current gradient information. Here, the Powell Hessian update algorithm[51] is used, where the updated Hessian is calculated using the following equation

$$\mathbf{H}_{\mathrm{new}} = \mathbf{H}_{\mathrm{old}} + \frac{1}{\delta \cdot \delta}(\varepsilon \otimes \delta + \delta \otimes \varepsilon) - \frac{\varepsilon \cdot \delta}{(\delta \cdot \delta)^2}\mathbf{H}_{\mathrm{old}} \cdot \delta \otimes \delta \cdot \mathbf{H}_{\mathrm{old}} \tag{30}$$

where $\mathbf{H}$, with the appropriate label, is the Hessian, δ is the vector of the position difference, and ε is the vector of the gradient difference between the "old" and "new" geometries.

At the first point in a DD-vMCG calculation, when it is performed using an empty database, an electronic structure calculation is carried out in order to obtain the energy, gradient and full Hessian, and this data is used as the reference. In the case where a populated database is used, the first entry in the database is used as the reference. The propagation continues and when the calculation reaches a point where a new electronic structure calculation is required, as opposed to an extrapolated point, only the energy and gradient at this new point are calculated. The DD-vMCG program then calculates the Euclidean norm distance between the reference and the new point, as well as calculating the distances between the reference and all other points in the database. The points in the database are then divided into two subsets. The first, "internal" subset of points comprises points which are closer to the reference than the new point. The second, "external" subset of points comprises the points which are further away from the reference than the new point. The gradient at the new point and the gradient and Hessian of each of the points in the internal subset in turn is used to obtain a set of Powell-updated Hessians for the new point, $\mathbf{H}_i$. As the distances between the new point and the internal points, d_i, have been calculated, the Hessian at the new point is given by the weighted sum

$$\mathbf{H}_{\mathrm{new}} = \frac{\displaystyle\sum_{i \in \mathrm{internal}} d_i^{-4}\mathbf{H}_{\mathrm{old}}^{(i)}}{\displaystyle\sum_{i \in \mathrm{internal}} d_i^{-4}} \tag{31}$$

The new Hessian is then added to the database and each of the Hessians in the external subset are updated in a similar manner, including the new Hessian as part of the internal subset. The performance of this Hessian update procedure has been tested, the details of which can be found in ref. 43.

A final important feature to note in the DD-vMCG method is the form of the potential energy surface, specifically the representation of the states in the calculation. Although the external electronic structure programs calculate points along adiabatic surfaces, the wavepacket is propagated along the diabatic surfaces. Given the prerequisites for the DD-vMCG method, a method by which the diabatic surfaces can be calculated is required which can even-handedly account for on-the-fly calculated surfaces with multiple states and an unknown number of state crossings. The method by which this is carried out is known as propagation diabatisation.

A unitary transformation matrix, $\mathbf{S}$, can be defined as between the adiabatic, ψ, and diabatic, φ, electronic functions.

$$\mathbf{S}_{ji} = \langle \psi_j | \varphi_i \rangle \tag{32}$$

If the adiabatic and diabatic states form complete sets, the gradient of these matrix elements can be written

$$\nabla \mathbf{S}_{ji} = \sum_{k=1}^{N} \langle \nabla \psi_j | \psi_k \rangle \langle \psi_k | \varphi_i \rangle + \sum_{k=1}^{N} \langle \psi_j | \varphi_k \rangle \langle \varphi_k | \nabla \varphi_i \rangle \tag{33}$$

the final term of which, in a strictly diabatic representation, is

$$\langle \varphi_k | \nabla \varphi_i \rangle = 0, \ \forall \, i,k \tag{34}$$

Using the definition of the non-adiabatic coupling vector,

$$\mathbf{F}_{ij} = \langle \psi_i | \nabla \psi_j \rangle = -\langle \nabla \psi_i | \psi_j \rangle = \frac{\langle \psi_i | \nabla \hat{\mathbf{H}} | \psi_j \rangle}{V_j - V_i} \tag{35}$$

where V_i and V_j are the adiabatic energies of states i and j, the adiabatic states are orthonormal and the matrix derivative can be written

$$\nabla \mathbf{S}_{ij} = -\sum_{k=1}^{N} \mathbf{F}_{ik} \mathbf{S}_{kj} \tag{36}$$

This relationship, first derived by Baer,[52] is strictly correct only when a complete basis set is used. It completely defines a diabatic basis once an initial $\mathbf{S}$ matrix is defined at a specific point, referred to as selecting the gauge for the transformation. For example, the adiabatic and diabatic representations can be set to be equal at the Franck–Condon point.

The key equation to propagation diabatisation is the matrix form of this equation

$$\nabla \mathbf{S} = -\mathbf{FS} \tag{37}$$

where the initial $\mathbf{S}$ matrix is taken to be at a point where the adiabatic and diabatic surfaces are the same, *i.e.* $\mathbf{S}$ is a unit matrix. It is the solutions to this differential

equation that define the scheme of the diabatisation. Propagating over a short time step, the formal solution to this equation is

$$\mathbf{S}(\mathbf{x} - \Delta\mathbf{x}) = \exp\left(-\int_{\mathbf{x}}^{\mathbf{x}+\Delta\mathbf{x}} \mathbf{F} \cdot d\mathbf{x}\right)\mathbf{S}(\mathbf{x}) \tag{38}$$

However, in this form it can be seen that in order to solve this equation an exponential of a matrix expression must be taken. It can also be seen that the integral of a function with an unknown analytic form must be taken. The second of these issues can be overcome using a simple numerical integration along the straight line between $\mathbf{x}$ and $\mathbf{x} + \Delta\mathbf{x}$ using the trapezium rule, resulting in a matrix of scalar parameters. The first issue is not as straightforward to solve.

The equation for the propagation, eqn (38), does not guarantee that a unitary matrix $\mathbf{S}(\mathbf{x})$ will return a unitary matrix for $\mathbf{S}(\mathbf{x} - \Delta\mathbf{x})$. Following the Esry and Sadeghpour[53] method, which uses the Cayley–Hamilton form of the propagator, the unitarity of the transformation matrix can be maintained, resulting in

$$\exp\left(\frac{1}{2}\int \mathbf{F} \cdot d\mathbf{x}\right)\mathbf{S}(\mathbf{x} + \Delta\mathbf{x}) = \exp\left(-\frac{1}{2}\int \mathbf{F} \cdot d\mathbf{x}\right)\mathbf{S}(\mathbf{x}) \tag{39}$$

which is a rearranged form of eqn (38). By utilising a Taylor series expansion of the exponentials, and the resultant matrix on the left-hand side is inverted, the final transformation matrix in terms of $\mathbf{S}(\mathbf{x} - \Delta\mathbf{x})$ is obtained. Other checks need to be made, for example to account for the change in order of the surfaces when an intersection seam is crossed. These are dealt with by extrapolating the diabatic surfaces from one point to the next. Full details of the method by which this implemented in the DD-vMCG software can be found in ref. 20.

As mentioned previously, this method is correct only for a complete basis set of states. However, tests to date indicate that this propagation diabatisation scheme provides smooth, diabatic potential energy surfaces.[20]

3 Methods

3.1 Electronic structure

In a DD-vMCG calculation, external quantum chemistry programs are needed in order to calculate the energy, gradient and Hessian required for the propagation. The use of a large basis set and a high level of theory would enable higher accuracy in the results; however, these computationally intensive calculations would present a severe bottleneck to the propagations over even relatively short timescales in terms of the memory and both the computational and real time for taken for the calculation. An important consideration, particularly relevant in quantum chemistry calculations involving a large number of excited states, is the stability of the quantum chemistry calculations to the movement of atoms away from equilibrium geometries. A final consideration is that the DD-vMCG method is relatively new and computational studies have, so far, been carried out testing the ability of the code to handle large numbers of degrees of freedom, but thus far no studies have been carried out examining the ability of the code to handle large numbers of excited states. Consequently a balance between accuracy, stability, computational expense, and time expense in the calculations carried out by an

external quantum chemistry software must be considered from the outset of an investigation.

Due to the lack of computational investigations into the excited states of FAM and FIM, it was unknown as to how many excited states would be needed in the quantum chemistry calculations. Consequently, an overall emphasis on efficiency in the quantum chemistry calculations was made, whilst maintaining the intention of gaining as accurate a description of the excited states of the molecules as possible within a reasonable time constraint.

In order to be able to successfully carry out a direct dynamics calculation involving multiple states, it is necessary for the non-adiabatic couplings between all of the states to be calculated. This is particularly important when the system of interest has a manifold of excited states in close proximity between which multiple crossings and intersections may occur. Consequently, the logical choice when considering the excited state dynamics of a system from a chemical perspective is the use the complete active space self-consistent field (CASSCF) method. All electronic structure calculations in the following sections used the Molpro 2015 program[54] which has a very efficient CASSCF procedure[55] and is able to provide all of the derivative couplings between states as required for the direct dynamics.

The use of the CASSCF is, however, by no means a "black box" method. It is known that a poor selection of active space size, occupancy and orbitals can lead to poorly converged calculations and instabilities. This problem is often exacerbated with the inclusion of excited states leading to even further instabilities in the calculation. It is these issues that present a major difficulty in the use of the DD-vMCG method. In addition, the inclusion of a nitrogen atom in the system is known to increase the difficulty in the accurate selection of the orbitals within the active space. As a result, as opposed to attempting to enforce a particular selection of orbitals on a specified number of excited states, a new analytical procedure for the systematic identification of the orbitals required alongside the number of excited states was developed in order to minimise the difficulties in this CASSCF selection process.

Before any active space orbitals and number of excited states can be chosen, the chemical features of the system in question must be considered. It was decided that the initial selection of orbitals should have the capacity to describe a potential proton transfer between the two isomers. It is known that π–π^* transitions are optically bright, so in addition to the description of a proton transfer, a sufficient description of the π-bonding network would be required. In addition, due to the system size, it was also decided that, in the interest of efficiency, a relatively large active space could be chosen so as to maximise the number of configuration interactions to be calculated with the caveat that a smaller basis set would be required in order to minimise the time-scaling issue.

Consequently, the orbital selection for FAM should include the description of the N–H(1) σ-bond and σ^*-bond, the lone-pair acceptor on the oxygen and a description of the π-bonding network containing at least one π-bonding molecular orbital and one π^*-bonding molecular orbital. The selection of orbitals to be included in the active apace for FIM should hence contain the complementary O–H(1) σ and σ^* molecular orbitals, as well as the lone-pair acceptor on the nitrogen and the relevant π- and π^*-bonding molecular orbitals. It was also decided that the active space for the FAM and the FIM should be the same size. Finally,

while the choice of the 6-311++G** basis set would yield more accurate results, in the interest of efficiency and given the choice of a large active space and the as-of-yet unknown number of excited states required for the calculation, a 6-31G* basis set would be used as a compromise. After extensive testing with different CAS spaces and including different numbers of excited states, it was decided to use 10 electrons in 8 orbitals, and 8 states with equal weights were to be calculated.

The results from the studies using a state-averaged CAS(10,8) with different numbers of states are shown in the ESI.† As a result of these studies it can be seen that not only is the character of the active space important to state-averaged calculations, but also the number of states included in the calculation. This representation of the results of the state-averaged CAS calculations serves both as a tool for the selection of the active space size and the number of excited states, and as a demonstration that the analysis of the numerical and pictographic results of CAS calculations are intrinsically linked. Fig. 1 shows the pictures of the molecular orbitals of FAM and FIM in the active space from the final choice, which included 8 singlet states. The accompanying labels are used below in the state-averaged CAS characterisations.

3.2 Direct dynamics

The dynamics calculations were propagated using the ground state normal modes of the molecules with the widths of the Gaussian wavepackets defined by the frequencies. The initial distribution of the Gaussians comprised a single fully populated Gaussian at the Franck–Condon point around which the rest of the
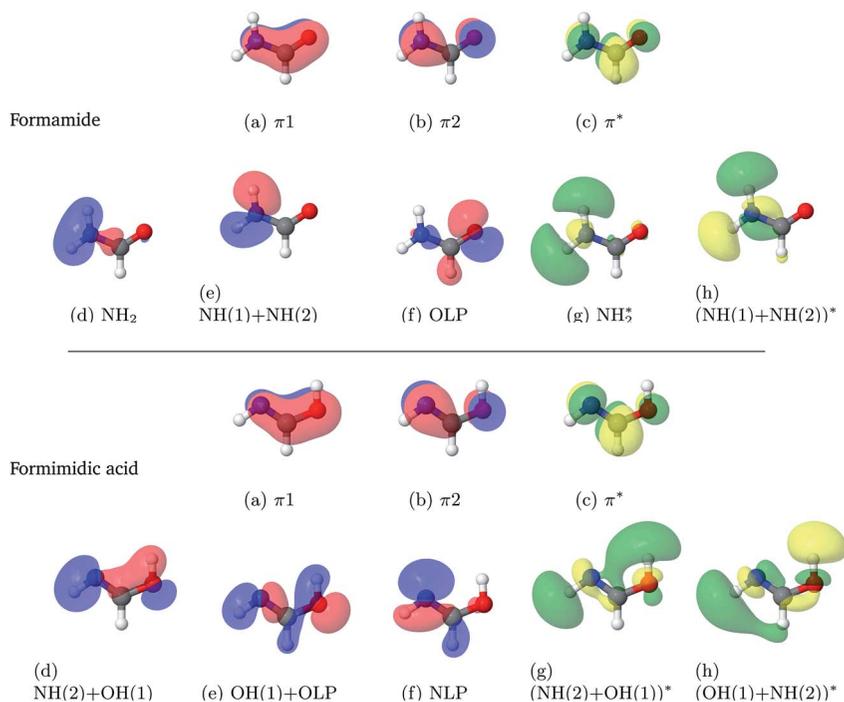

Fig. 1 Molecular orbitals of formamide (top) and formimidic acid (bottom) from SA8-CAS(10,8)/6-31G* calculations.

zero populated Gaussians are distributed, with an overlap of 0.5. In order to ensure the stability of the initial electronic structure calculation, the Gaussians were distributed in momentum space, as opposed to configuration space. The use of the single fully populated Gaussian at this point ensures that the initial wavepacket is the ground state vibrational wavefunction. As there were 8 states included in the electronic structure calculation, a preliminary calculation was made with a vertical excitation to S_1 using 8 Gaussian basis functions, with a propagation time of 150 fs.

These results allowed the selection and placement of complex absorbing potentials (CAPs). These are negative, imaginary potentials which are used to absorb a wavepacket. In grid based calculations, they are placed at the ends of grids to ensure that the wavepacket is not reflected off the end of the grid, resulting in the decoherence. In a direct dynamics calculation, if a rapid dissociation occurs, the dissociating atom or fragment gains momentum as it gets further away from the main molecule. This has a direct impact on the integrator in that rapidly decreasing time step sizes must be taken in order to gain a valid description of the system as a whole. Additional issues arise in that these rapidly changing geometries result in a larger number of points requiring electronic structure calculations. At these widely spaced geometries, the electronic structure calculations will take longer to run, and may fail. Consequently, the use of CAPs in the DD-vMCG method essentially provides a cut-off point to dissociative motion along normal mode coordinates.

The CAPs are defined as

$$iW = \eta\Theta(k(x - x_0))^n \tag{40}$$

Here, Θ is a Heaviside step function, while $[x_0, \eta, n, k]$ defines the order, n, and strength, η, of the CAP positioned at x_0 along a normal mode coordinate, where $k = \pm 1$ indicates whether it is in the positive or negative direction. When the Gaussian reaches the CAP the motion continues classically until the population of that Gaussian is zero, after which it stops. Values of $x_0 = \pm 12$, $\eta = 1.0$, $n = 3$ were used along each mode which put the CAPs far enough along the dissociative coordinates so as to not perturb the vibrational motion.

The DD-vMCG method builds a database of geometries, energies and other information, such as molecular orbitals, along the course of a propagation. Multiple runs of a calculation, using the database constructed from previous calculation(s), will result in fewer electronic structure calculations required, as well as the opportunity for further regions of configuration space to be explored. Consequently, the databases from this first set of calculations were used for subsequent calculations.

A second set of calculations was propagated starting on the S_1, S_2 and S_3 states and with 8, 24 and 48 GWPs, each with a propagation time of 150 fs. The databases constructed from these propagations were then merged into one database which was then used for the final propagations on the three states, with 48 GWPs. The final databases for both formamide and formimidic acid at the end contained approximately 6500 structures. It is the results of this final set of calculations that are used for the following analysis.

Faraday Discussions

4 Results

4.1 Electronic structure

The first question to be answered is what are the ground state structures of FAM and FIM? Optimised structures at the CCSD/6-311+G* level of theory were first obtained for the two structures and the transition state for proton transfer from the N to the O. It was found, in agreement with previous work,[37] that the minimum structure of FAM has a non-planar structure with the NH_2 group having a pyramidalisation angle of 25.4°. The minimum energy of the planar structure with C_s symmetry was therefore also obtained. The energies of these structures relative to the FAM non-planar minimum are listed in Table 1. It can be seen that FIM is approximately 0.5 eV less stable than FAM, and there is an appreciable barrier to proton transfer on the ground state surface of approximately 2.5 eV. The barrier to NH_2 inversion in FIM is, however, tiny, of the order of 0.01 eV. The wells are thus too shallow to contain vibrational states and the molecule will appear to be planar for spectroscopic processes. The planar structure was thus taken as the starting structure for the dynamics. Table 1 also includes labels distinguishing the 3 hydrogen atoms that are used in the following analysis.

In Table 2 the normal modes of formamide and formimidic acid are listed, calculated at the CCSD/6-311+G* and the SA8-CAS(10,8)/6-31G* levels of theory. The relevant frequencies and characterisations are also shown. All frequencies were calculated on structures optimised at the relevant level of theory. The normal mode vectors are shown graphically in the ESI.†

For formamide, in the first row of Table 2, it can be seen that the normal mode, ν_1, is an imaginary frequency. This is the NH_2 inversion mode that connects the global C_1 minimum as discussed above. It can be seen that the frequencies calculated at the two levels of theory are close to each other with the largest difference for ν_{12} where the CASSCF frequency is higher by approximately 250 cm^{-1}. As the structure has C_s symmetry, there exist two categories of normal mode representing in-plane and out-of-plane motion, where only three modes, ν_1, ν_3 and ν_4, represent this out-of-plane motion. It should be noted that despite the fact that the motion of H(2) and H(3) appear to have their "own" vibrational modes, the motion of H(1) is always coupled to the motion of other atoms.

The normal mode frequencies of formimidic acid are seen to be all in good agreement between the 2 methods. As with FAM, there exist two categories of

Table 1 Energies of key formamide (FAM) and formimidic acid (FIM) geometries calculated at the CCSD/6-311+G* level of theory. The energy of the FAM C_1 optimised geometry is in hartrees. All other reported energies are relative to this value and are in eV
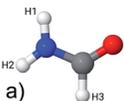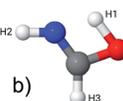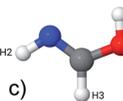
FAM (C_1)	FAM (C_s)	TS	FIM
	0.00778	2.4416	0.6185
−169.493998	a)	b)	c)

Table 2 The normal mode frequencies of formamide and formimidic acid calculated at different levels of theory at the optimised C_s structures. The numbering of the hydrogen atoms is as in Table 1(a) and (c). "ip" means in-plane and "oop" out-of-plane

Label	Formamide Frequency (cm⁻¹)			Formimidic acid Frequency (cm⁻¹)		
	CCSD/6-311+G*	CAS(10,8)/SA8 6-31G*	Character	CCSD/6-311+G*	CAS(10,8)/SA8 6-31G*	Character
ν_1 (A″)	350.35i	320.66i	NH$_2$ oop wag	596.58	630.43	OH oop wag
ν_2 (A′)	576.24	598.06	H1 ip rocking	611.55	635.19	H2 + H1 ip wag
ν_3 (A″)	619.77	601.74	NH$_2$ oop torsion	846.36	911.91	H2 + H3 oop in-phase
ν_4 (A″)	1045.95	1014.98	H3 oop bend	1066.07	1062.53	H3 + H2 oop out-of-phase
ν_5 (A′)	1079.99	1108.68	H2 ip rocking	1097.61	1150.16	NH ip bend
ν_6 (A′)	1293.11	1356.26	H3 H1 ip out-of-phase	1232.28	1284.95	H1 + H2 + H3 ip wag
ν_7 (A′)	1443.86	1397.43	H3 ip bend	1404.31	1447.58	H1 + H2 in-phase ip wagging
ν_8 (A′)	1680.88	1803.58	NH$_2$ ip bend	1437.46	1515.33	H3 ip bend
ν_9 (A′)	1827.88	2031.85	CO stretch	1750.26	1854.47	CH3 ip rocking
ν_{10} (A′)	3008.20	2640.00	CH stretch	3155.30	3126.07	NH stretch
ν_{11} (A′)	3636.12	3855.27	NH$_2$ sym stretch	3554.58	3379.82	OH stretch
ν_{12} (A′)	3772.01	4029.16	NH$_2$ anti-sym stretch	3776.70	3449.48	CH + OH stretch

normal mode representing in-plane and out-of-plane motion where only three modes, ν_1, ν_3 and ν_4, represent this out-of-plane motion.

Excitation energies, oscillator strengths and characterisations from the SA8-CAS(10,8)/6-31G* calculations of FAM and FIM are summarised in Table 3. The orbitals are shown in Fig. 1. It can be seen that the S_6 state, characterised as a π–π^* transition, is the brightest state for both FAM and FIM, while the S_3 and S_5 states have smaller but significant oscillator strengths. In FAM, the S_6 and S_3 states are both a mixture of π–π^* and Olp–NH_2^* transitions, while S_5 is predominantly Olp–NH_2^*, where Olp is the oxygen lone pair. In FIM, the S_6 and S_3 states are both characterised as being a mixture of π–π^* and Nlp–$(NH(2) + OH(1))^*$ character, where Nlp is the nitrogen lone pair, while S_5 is predominantly Nlp–$(NH(2) + OH(1)^*)$. The S_1 (Nlp–π^*) state in FIM also has a small, but significant, oscillator strength. It can also be seen that the oscillator strengths of the S_3 and S_5 states of FIM are greatly increased in comparison to the same states in FAM. This comparative increase in the oscillator strength in the S_3 state is likely due to the dominant π–π^* transition in the FIM, which is only a secondary contribution to the equivalent state in FAM.

Table 4 compares the results from the SA8-CAS(10,8)/6-31G* calculations on formamide with literature values. Experiments have determined five main bands in the UV absorption spectrum, which have been assigned to three valence ($n\pi^*$ and $\pi\pi^*$) and two Rydberg states. The presence of Rydberg states is well-known in small nitrogen-containing molecules, and a basis set with diffuse functions is required to treat them correctly. The calculations of Hirst et al.,[34] Serrano-Andrés and Fülscher[35] and Antol et al.[37] all include diffuse functions and obtain a large number of Rydberg states. For example, Serrano-Andrés and Fülscher, using a huge 15-orbital active space and large atomic natural orbital (ANO) basis set specifically designed for this, obtained 22 states in the 5 eV range between the $n\pi^*$

Table 3 Formamide energies, oscillator strength and coefficients of the main configurations (values > 0.15) from a SA8-CAS(10,8)/6-31G* calculation

	Formamide			Formimidic acid		
	Energy (eV)	Oscillator strength (au)	Main configurations	Energy (eV)	Oscillator strength (au)	Main configurations
S_1	5.607	0.0008	0.69 (Olp–Pi*)	6.935	0.0105	0.687 (Nlp–Pi*)
S_2	8.015	0.0004	0.66 (Pi–NH + NH)	8.935	0.0009	0.56 (Pi–(OH + NH)*) + 0.35 (Pi–(NH + OH)*)
S_3	8.159	0.0225	0.54 (Olp–NH + NH) + 0.34 (Pi–Pi*)	9.117	0.1325	0.48 (Pi–Pi*) + 0.37 (Nlp–(OH + NH)*)
S_4	9.118	0.0000	0.66 (Pi–NH + NH)	10.079	0.0045	0.56 (Pi–(NH + OH)*) + 0.36 (Pi–(OH + NH)*)
S_5	10.033	0.0710	0.63 (Olp–NH + NH)	10.137	0.1723	0.48 (Nlp–(NH + OH)*) + 0.47 (Nlp–(OH + NH)*)
S_6	10.574	0.7258	0.44 (Pi–Pi*) + 0.37 (Olp–NH + NH)	11.267	0.4634	0.41 (Nlp–(NH + OH)*) + 0.39 (Pi–Pi*) + 0.30 (Nlp–(OH + NH)*)
S_7	11.450	0.0013	0.55 (Pi–NH + NH) + 0.36 (Pi–Pi*)	11.831	0.0000	0.676 (OH + Olp–Pi*)

Table 4   The main vertical excitation energies in eV, with oscillator strengths in brackets, from various methods. Rydberg states from the various calculations have been selected and assigned on the basis of energy and oscillator strength

Band	Expt.[56,57]	This work CAS(10,8) 6-31G*	Hirst[34] MRCI/CAS(6,13) 6-31+G**	Serrano-Andrés[35] CASPT2(6,15) ANO	Antol[37] MRCISD aug-cc-pVDZ
W ^{1}A″ (nπ*)	5.82 (0.002)	5.61 (0.001)	5.86 (0.001)	5.61 (0.001)	5.78 (0.001)
R1 ^{1}A″	6.35	8.01 (0.001)	6.14 (0.022)	6.52 (0.024)	6.77 (0.022)
V1 ^{1}A′ ($\pi\pi$*)	7.36 (0.37)	8.16 (0.022)	7.40 (0.156)	7.41 (0.371)	7.71 (0.338)
R2 ^{1}A′	7.72	10.03 (0.071)	7.50 (0.041)	7.72 (0.101)	7.84 (0.017)
Q ^{1}A′ ($\pi\pi$*)	9.23	10.57 (0.726)	7.94 (0.149)	10.50 (0.131)	

and higher $\pi\pi$* states. These Rydberg states are not included in our calculations as the basis set was chosen for efficiency in the direct dynamics. It can be seen that the CAS space chosen, however, places the valence states with reasonable accuracy compared to the other calculations and experiment. It would also be expected that as the molecule moves away from the high symmetry starting point that the Rydberg states will mix strongly with the valence states and the description of these will be adequate for our purposes of testing the DD-vMCG method with a realistic, but challenging system.

4.2   Quantum dynamics simulations

The diabatic state populations as a function of time for formamide and for-mimidic acid after vertical excitation to the S_1, S_2 and S_3 states for each of the 8, 24 and 48 GWP propagations are shown in the ESI.† These demonstrate that the results are not changing significantly with increasing basis set, but the results become smoother. It can be seen, most significantly in the S_1 plots, that the increase from 8 to 24 GWPs resulted in a more smooth representation of the decay and redistributions of the density into the different states, while the improvement from the 24 to 48 GWP is less significantly pronounced. This implies that the calculations are reaching convergence.

In Fig. 2 the diabatic state populations from the 48 GWP propagations are shown, with a comparison to the total density. The diabatic state populations give information as to how the total wavepacket has been distributed, along the period of propagation, into the various states included in the calculation. The amount of the population absorbed by the CAPs cannot directly be separated from the population transfer in the analysis of this set of graphs. However, by also plotting the total density with the diabatic state populations, the features of population transfer can more easily be resolved from the population that is being absorbed by the CAP. It can be seen that in both molecules, the timescale for the decay from the S_1 state is significantly longer than that from the higher lying states, and that FAM has a longer lifetime than FIM after all excitations.

In FAM, approximately 20% of the population is still in the S_1 state after 70 fs, while the population decay from the higher lying states is complete by around 30 fs. It can be seen that in the S_1 state, density begins to flow into the CAPs after around 20 fs, while in the S_2 state this begins to occur after 10 fs and in the S_3 state this occurs after around 8 fs, indicating that the initial drop in the population at
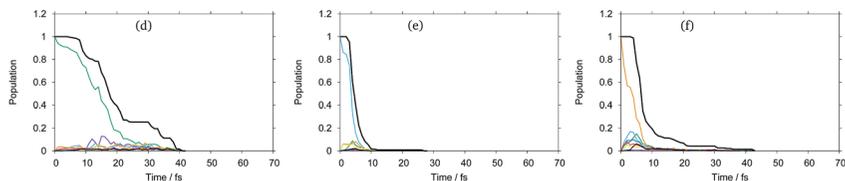
Formamide
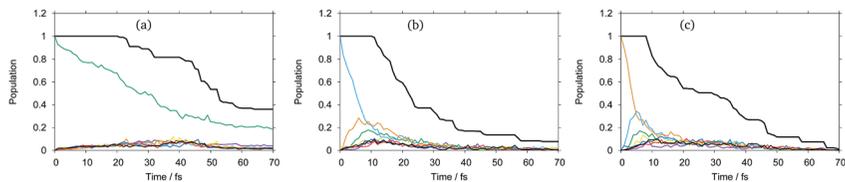
Formimidic Acid

Fig. 2 Diabatic state populations from DD-vMCG simulations of formamide (top) and formamidic acid (bottom) starting with a vertical excitation to various states: (a and d) S_1, (b and e) S_2 and (c and f) S_3. Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red; S_7: black. Total density (norm²): thick black.

the beginning of the propagation is due to population transfer. In the S_1 state, as mentioned previously, the rate of population transfer is significantly slower than in the other excited states. It can also be seen that the population transfer is relatively evenly distributed across all of the excited states in the calculation.

In the S_2 state it can be seen that after approximately 6 fs about 25% of the population has been transferred to the S_3 (orange) state, with around 18% of the population being transferred to the S_1 (green) state by 9 fs. A similar fast population transfer is also seen in the S_3 plot, where about 32% of the population is transferred to the S_2 (light blue) state with the next most significant population transfer going into the S_1 state. Population transfer also takes place from the S_3 to S_4 (yellow) states.

In FIM after excitation to S_1, the most significant diabatic state population transfer is into the S_0 state at about 15 fs, and the population transfer is complete by around 30 fs. The transfer is, however, minor. It can be seen instead that the decrease in total density follows the trend in the decreasing population of the S_1 state, suggesting that the S_1 state is displaying dissociative behaviour. In the S_2 state, the decrease in the total density occurs rapidly between 4 and 10 fs, with relatively little population transfer occurring to S_1 and S_3. In the S_3 state, it can be seen that there is a fast population transfer, mostly to the S_2 state, but also to the S_5 and S_6 states and a small amount to the S_4 state, and the total density does not decrease over this short timescale. It appears that in this S_3 calculation, when the S_1 and S_7 states become populated, the total density proceeds to decrease rapidly, supporting the suggestion that the S_1 and perhaps the S_7 states are dissociative.

As a result of this analysis it can be seen that while dissociative behaviour can be seen in all sets of results, it is not possible from this analysis to state conclusively that particular states exhibit dissociative behaviours, or to know what the products are. A useful description of the distribution of the products of a calculation can be obtained by assigning a weight to a particular GWP. An estimate can then be obtained of the density going into the channel described by the trajectory at the centre of a GWP. The density is given by

$$\langle \Psi | \Psi \rangle = \sum_{ij} A_i^* \langle g_i | g_j \rangle A_j \tag{41}$$

$$= 2\Re \sum_i |A_i|^2 + \sum_{i<j} A_i^* \mathbf{S}_{ij} A_j \tag{42}$$

When the overlaps between the functions are divided evenly, the gross Gaussian population,[58] or GGP, is hence defined where

$$\mathbf{GGP}_i = |A_i|^2 + \Re \sum_{j \neq i} A_i^* \mathbf{S}_{ij} A_j \tag{43}$$

$$= \Re \sum_j A_i^* \mathbf{S}_{ij} A_j \tag{44}$$

In order to use these GGPs, it is first necessary to perform a visual inspection of the geometries defined by the trajectory of the centre of the GWPs. The fraction of the density, defined by the GGP, along each of the product channels was then categorised, allowing analysis of the product distribution. The results of this analysis of the trajectories of the GWPs for formamide and formimidic acid can be seen in Fig. 3. It should be noted that the >1 population in the first few femtoseconds of the propagation is likely due to numerical instabilities in the method.

In Fig. 3, representing the product distribution of FAM, upon cursory inspection it can be seen that while the product distribution in the S_2 and S_3 states is dominated by one behaviour, the number of product channels, and hence the product distribution in the S_1 state, is more evenly distributed between the
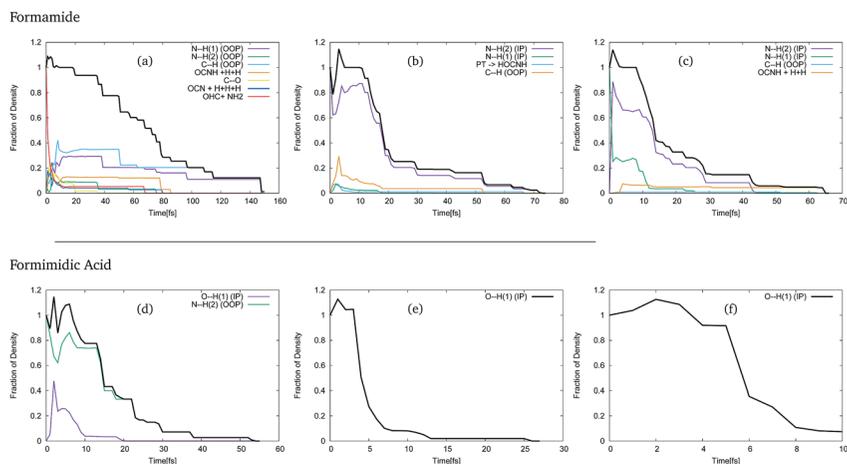

Fig. 3 The fraction of density going into different product channels from DD-vMCG simulations of formamide (top) and formimidic acid (bottom) with SA8-CAS(10,8)/6-31G* potential surfaces starting in different states: (a and d) S_1, (b and e) S_2 and (c and f) S_3. Each line, or series, represents the different products defined either by the bond that breaks or by the products formed. As in the characterisation of the vibrational frequencies, IP and OOP signify whether the dissociation occurs in- or out-of-plane.

possible pathways. In the S_1 state, it can be seen that dissociation in the first 70 to 80 fs of the propagation occurs along a multitude of pathways, the principal of which represent the N–H(1) and C–H out-of-plane motions, which correspond to the ν_1, ν_2 and ν_3 vibrational modes of the system. It is in the movement in these channels that the density continues after 80 fs, until all of the Gaussians have been absorbed by the caps. In both the S_2 and the S_3 states, the principal product channel is defined by the N–H(2) bond stretching. In the S_2 state it can be seen that a very small amount of proton transfer occurs. In this process, the H(1) dissociates and the H(3) subsequently transfers to the oxygen. Although the proton transfer involving the H(1) atom was expected, this mechanism of proton transfer was somewhat unexpected. In the S_3 state it can be seen that a secondary product channel to the nuclear motion in the first 12 fs is observed along the N–H(1) stretching mode. It should be noted that while the C–H out-of-plane bond-breaking is observed in all three of these states, NH_2 and NH_1 bond-breaking is along in-plane motion for S_2 and S_3, but out-of-plane motion for S_1.

In Fig. 3, representing the product distribution of FIM, upon cursory inspection it can be seen that while the number of product channels for the S_1 state is greater than in the S_2 and S_3 states, as was the case with FAM, only two product channels are seen. In all three states the O–H(1) in-plane stretching motion, characterised as the ν_{11} mode, to the point of dissociation is seen, while in the S_1 state the N–H(2) out-of-plane motion, characterised in a combination of the ν_3 and ν_4 modes, is also seen. In the S_1 state it is the N–H(2) out-of-plane motion that is dominant, though the timescale over which the O–H(1) stretching motion is seen is comparable to the timescale in the S_2 state.

As a result of this investigation it can be seen that a number of different product channels are available in the dynamics of formamide, whereas the OH stretching motion in formimidic acid is dominant.

During the course of direct dynamics propagation, the energies are calculated and stored in the database. As a result, 1- and 2-dimensional cuts of the potential energy surfaces can be made. By looking at the adiabatic and diabatic representation of the same mode, or modes, regions displaying non-adiabatic features such as avoided crossings, or conical intersections, can be determined. From a practical perspective, these surfaces can also be used to determine if the direct dynamics calculations have been propagated for sufficient time periods, if the diabatisation scheme has been successful, and other technical faults which are displayed as discontinuities on the surfaces. Using the observations of the GWP trajectories, the surface cuts were selected along the modes which were most clearly important for the dissociation of the molecules.

In Fig. 4, the adiabatic and diabatic cuts of the potential energy surfaces of FAM along the ν_{11} and ν_{12} modes are shown, characterised as the N–H_2 symmetric and anti-symmetric stretching modes. The NH dissociations occur along a combination of the modes. The potential energy surfaces are smooth, with the diabatic states correctly crossing between states.

Along ν_{11} in Fig. 4(a) and (b), there is a dissociative channel in the negative displacement direction shown as the green, S_1 state in the adiabatic representation, which in the diabatic representation becomes the orange, previously S_3, state. However, in both the diabatic and adiabatic representations it can be seen that there exists a kink in the surfaces, approximately located at a displacement of −2. From the shape of the curves at this point it is likely that there is a higher
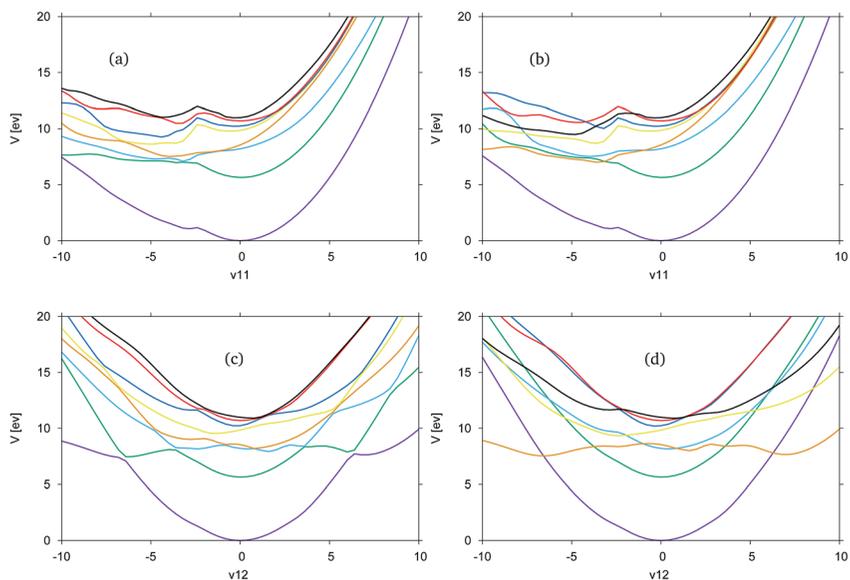
Fig. 4 Cuts through the SA8-CAS(10,8)/6-31G* potential energy surfaces of formamide from DD-vMCG simulations. (a) ν_{11} (N–H_2 symmetric stretch) adiabatic, (b) ν_{11} diabatic, (c) ν_{12} (N–H_2 antisymmetric stretch) adiabatic and (d) ν_{12} diabatic.

energy state which cuts down through most of the states, and is not represented in the CAS space used.

Along ν_{12} in Fig. 4(c) and (d) at the Franck–Condon point it can be seen that the blue, S_2, and orange, S_3, states are very close in energy and crossing between these states occurs very close to the Franck–Condon point. It is this orange state, in the diabatic picture, that leads to the NH-dissociation channel. It should be noted that these curves are not symmetric.

In Fig. 5 the adiabatic and diabatic cuts of the potential energy surface along the ν_{10} (Fig. 5(a) and (b)) and ν_{11} (Fig. 5(c) and (d)) modes of formimidic acid, characterised as the N–H_2 and the OH stretching modes, are shown. These modes represent the equivalent modes to the ν_{11} and ν_{12} of FAM. In Fig. 5(a) and (b) there appears to be a major failure in the calculation in the negative direction along the mode, at about -4. This has been identified as a failure in the electronic structure calculations. Consequently, the only comments that can be made about this representation are that the dark blue (S_5) and yellow (S_4) states are very close in energy along the positive direction of the mode, and that the crossings between these and the red (S_6) and black (S_7) states are well-resolved. It appears as if there may be a dissociative state in the negative direction along the mode though it is unclear as to whether this state originates from a state within the precalculated manifold.

In Fig. 5(c) and (d) it can be seen that the surfaces are much more smooth and they clearly show a dissociative state which is the green, S_1, state in the adiabatic representation, or the light blue, S_2, state in the diabatic representation. As the states are well-resolved in this mode it is clear that the OH dissociative pathway was open to the dynamics of the molecule. In the diabatic representation there is a good representation of the crossings between the states, although a region
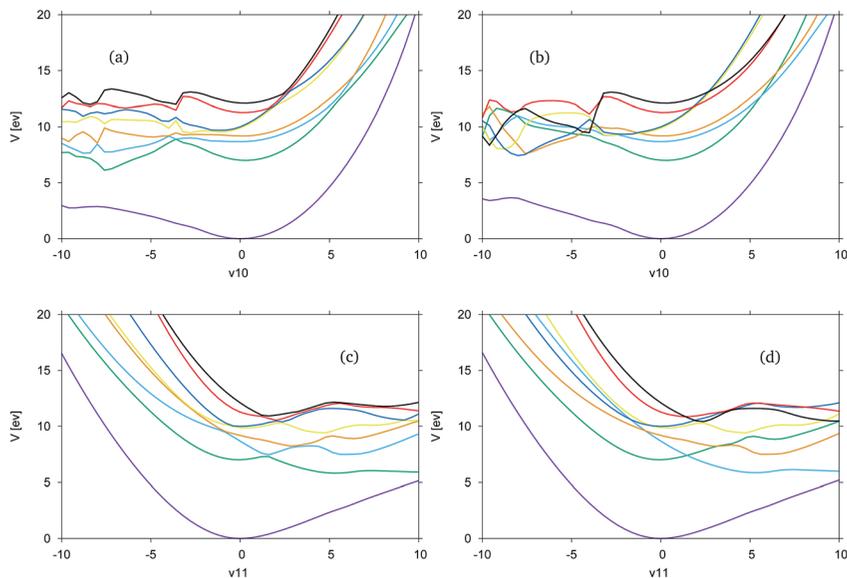
Faraday Discussions

Fig. 5 Cuts through the SA8-CAS(10,8)/6-31G* potential energy surfaces of formimidic acid from DD-vMCG simulations. (a) ν_{10} (N−H(2) stretch) adiabatic, (b) ν_{10} diabatic, (c) ν_{11} (O−H(1) stretch) adiabatic and (d) ν_{11} diabatic.

approximately located at 2, in the positive direction shows a significant convergence of the yellow, blue, red and black states, S_4 to S_7, and it is unclear if this region has been correctly resolved.

5 Conclusions

In this paper we have presented a study of the photo-dissociation dynamics of two key molecules, formamide and formimidic acid, using a new fully quantum mechanical direct dynamics algorithm. The method is able to capture the entire quantum nature of the coupled nuclear and electronic motions after photo-excitation into a manifold of states. Here, eight states were considered with the coupling between them treated using a novel propagation diabatisation scheme. It provides potential surfaces from quantum chemistry calculations on-the-fly, thus circumventing the need for pre-fitted global surfaces, and the diabatic surfaces cross smoothly where crossing points are found in the adiabatic surfaces. This should provide a significant step forward when calculating the quantum dynamics of polyatomic molecules as the provision of global diabatic surfaces is a prohibitive step in most cases.

A feature of the method that is also useful is that the Gaussian wavepacket basis functions are free to travel anywhere, and thus any open channels can be explored. The local nature of the basis functions allows a simple analysis of the importance of the different channels by analysing the trajectories of the GWP centres and weighting them by the Gross Gaussian Populations that are related to the density carried by each basis function. In the case of formamide and formimidic acid, it is found that both molecules dissociate in less than 100 fs, with FIM being the faster of the two. The simulations, moreover, show that the

dissociation in both cases depends on the initial excitation, *i.e.* the starting energy. In both cases, excitation to S_2 or S_3 results in very fast direct dissociation of a single hydrogen atom – the O–H bond in FIM or the N–H or C–H bonds in FAM. In contrast, excitation to S_1 leads to a longer lifetime and more open channels. In the case of FIM, the main dissociation is now the N–H bond, while in FAM, a number of fragmentations occur. Thus, in both cases, hydrogen loss appears to lend a degree of photostability to the molecule.

The study also shows the limitations of the present method. The main problem is the quantum chemistry. Despite careful initial benchmarking, the CASSCF calculations are seen to have problems in the N–H dissociation channels, resulting in non-smooth surfaces. The algorithm also needs to be improved to retain the symmetry of the surfaces. Further work is required to remedy these problems. The diabatic surfaces also need to be tested beyond visual inspection for any inconsistencies. The potential of the method is, however, clear, and should lead to a powerful tool able to simulate quantum mechanical behaviour in a straightforward way.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 H.-D. Meyer, U. Manthe and L. S. Cederbaum, *Chem. Phys. Lett.*, 1990, **165**, 73–78.
2 M. H. Beck, A. Jäckle, G. A. Worth and H.-D. Meyer, *Phys. Rep.*, 2000, **324**, 1–105.
3 H.-D. Meyer, F. Gatti and G. A. Worth, *Multidimensional Quantum Dynamics: MCTDH Theory and Applications*, Wiley VCH, 2009.
4 H. Wang and M. Thoss, *J. Chem. Phys.*, 2003, **119**, 1289–1299.
5 G. Richings and S. Habershon, *Chem. Phys. Lett.*, 2017, **683**, 228–233.
6 G. Richings and S. Habershon, *J. Chem. Theory Comput.*, 2017, **13**, 4012–4024.
7 G. Richings and S. Habershon, *J. Chem. Phys.*, 2018, **148**, 134116.
8 M. D. Hack and D. G. Truhlar, *J. Phys. Chem. A*, 2000, **104**, 7917–7926.
9 M. Richter, P. Marquetand, J. González-Vázquez, I. Sola and L. González, *J. Chem. Theory Comput.*, 2011, **7**, 1253–1258.
10 S. Mai, P. Marquetand and L. González, *Int. J. Quantum Chem.*, 2015, **115**, 1215–1231.
11 C. C. Martens, *J. Phys. Chem. Lett.*, 2016, **7**, 2610–2615.
12 A. Abedi, N. T. Maitra and E. K. U. Gross, *Phys. Rev. Lett.*, 2010, **105**, 123002.
13 S. K. Min, F. Agostini, I. Tavernelli and E. K. Gross, *J. Phys. Chem. Lett.*, 2017, **8**, 3048–3055.
14 J. Quenneville, M. Ben-Nun and T. J. Martínez, *J. Phys. Chem. A*, 2000, **104**, 5161–5175.

15 M. Ben-Nun and T. J. Martínez, *Adv. Chem. Phys.*, 2002, **121**, 439–512.

16 B. Curchod and T. Martinez, *Chem. Rev.*, 2018, **118**, 3305–3336.

17 G. A. Worth, M. A. Robb and I. Burghardt, *Faraday Discuss.*, 2004, **127**, 307–323.

18 B. Lasorne, M. A. Robb and G. A. Worth, *Phys. Chem. Chem. Phys.*, 2007, **9**, 3210–3227.

19 G. A. Worth, M. A. Robb and B. L. Lasorne, *Mol. Phys.*, 2008, **106**, 2077–2091.

20 G. W. Richings and G. A. Worth, *J. Phys. Chem. A*, 2015, **119**, 12457–12470.

21 A. W. Hofmann, *J. Am. Chem. Soc.*, 1863, **16**, 72–74.

22 P. M. Solomon, *Phys. Today*, 1973, **26**, 32–40.

23 D. Bockelée-Morvan, D. C. Lis, J. E. Wink, D. Despois, J. Crovisier, R. Bachiller, D. J. Benford, N. Biver, P. Colom, J. K. Davies, E. Gérard, B. Germain, M. Houde, D. Mehringer, R. Moreno, G. Paubert, T. G. Phillips and H. Rauer, *Astron. Astrophys.*, 2000, **353**, 1101–1114.

24 D. C. Lis, D. M. Mehringer, D. Benford, M. Gardner, T. G. Phillips, D. Bockelée-Morvan, N. Biver, P. Colom, J. Crovisier, D. Despois and H. Rauer, *Earth, Moon, Planets*, 1997, **78**, 13–20.

25 W. A. Schutte, A. C. A. Boogert, A. G. G. M. Tielens, D. C. B. Whittet, P. A. Gerakines, J. E. Chiar, P. Ehrenfreund, J. M. Greenberg, E. F. van Dishoeck and T. de Graauw, *Astron. Astrophys.*, 1999, **343**, 966–976.

26 W. H. Flygare, R. C. Benson, H. L. Tigelaar, R. H. Rubin and G. W. J. Swenson, *Molecules in the galactic environment*, John Wiley and Sons Inc, New York, 1973.

27 C. A. Gottlieb, P. Palmer, L. J. Rickard and B. Zuckerman, *Astrophys. J.*, 1973, **182**, 699–710.

28 R. Saladino, C. Crestini, S. Pino, G. Costanzo and E. Di Mauro, *Phys. Life Rev.*, 2012, **9**, 84–104.

29 M. A. Robb and I. G. Csizmadia, *Theor. Chim. Acta*, 1968, **10**, 2.

30 X. C. Wang, J. Nichols, M. Feyereisen, M. Gutowski, J. Boatz, A. D. J. Haymet and J. Simons, *J. Phys. Chem.*, 1991, **95**, 10419–10424.

31 D. Guzmán-Angel, R. Inostroza-Rivera, S. Gutiérrez-Oliva, B. Herrera and A. Toro-Labbé, *Theor. Chem. Acc.*, 2016, **135**, 37.

32 V. S. Nguyen, H. L. Abbott, M. M. Dawley, T. M. Orlando, J. Leszczynski and M. T. Nguyen, *J. Phys. Chem. A*, 2011, **115**, 841–851.

33 V. S. Nguyen, T. M. Orlando, J. Leszczynski and M. T. Nguyen, *J. Phys. Chem. A*, 2013, **117**, 2543–2555.

34 J. D. Hirst, D. M. Hirst and C. L. Brooks III, *J. Phys. Chem.*, 1996, **100**, 13487–13491.

35 L. Serrano-Andrés and M. P. Fülscher, *J. Am. Chem. Soc.*, 1996, **118**, 12190–12199.

36 N. A. Besley and J. D. Hirst, *J. Am. Chem. Soc.*, 1999, **121**, 8559–8566.

37 I. Antol, M. Vazdar, M. Barbatti and M. Eckert-Maksić, *Chem. Phys.*, 2008, **349**, 308–318.

38 N. V. Tukachev, V. A. Bataev, A. V. Abramenkov and I. A. Godunov, *Comput. Theor. Chem.*, 2016, **1080**, 23–32.

39 I. Antol, M. Barbatti, M. Eckert-Maksić and H. Lischka, *Monatsh. Chem.*, 2008, **139**, 319–328.

40 I. Burghardt, M. Nest and G. A. Worth, *J. Chem. Phys.*, 2003, **119**, 5364–5378.

41 G. A. Worth and I. Burghardt, *Chem. Phys. Lett.*, 2003, **368**, 502–508.

42 I. Burghardt, K. Giri and G. A. Worth, *J. Chem. Phys.*, 2008, **129**, 174104–174114.

43 G. W. Richings, I. Polyak, K. E. Spinlove, G. A. Worth, I. Burghardt and B. Lasorne, *Int. Rev. Phys. Chem.*, 2015, **34**, 269–308.

44 E. J. Heller, *J. Chem. Phys.*, 1975, **62**, 1544–1555.

45 E. J. Heller, *J. Chem. Phys.*, 1981, **75**, 2923–2931.

46 G. A. Worth, K. Giri, G. W. Richings, M. H. Beck, A. Jäckle and H.-D. Meyer, *Quantics package, Version 1.1*, 2015.

47 B. Lasorne, M. J. Bearpark, M. A. Robb and G. A. Worth, *Phys. Chem. Chem. Phys.*, 2006, **432**, 3210–3227.

48 J. Ischtwan and M. A. Collins, *J. Chem. Phys.*, 1994, **100**, 8080–8088.

49 M. A. Collins and D. H. Zhang, *J. Chem. Phys.*, 1999, **111**, 9924–9931.

50 T. J. Frankcombe, M. A. Collins and G. A. Worth, *Chem. Phys. Lett.*, 2010, **489**, 242–247.

51 T. J. Frankcombe, *J. Chem. Phys.*, 2014, **140**, 114106–114108.

52 M. Baer, *Chem. Phys. Lett.*, 1975, **35**, 112–118.

53 B. Esry and H. Sadeghpour, *Phys. Rev. A: At., Mol., Opt. Phys.*, 2003, **68**, 42706.

54 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. Köppl, 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, http://www.molpro.net.

55 P. J. Knowles and H.-J. Werner, *Chem. Phys. Lett.*, 1985, **115**, 259–267.

56 H. Basch, *J. Chem. Phys.*, 1968, **49**, 5007–5018.

57 J. M. Gingell, N. J. Mason, H. Zhao, I. C. Walker and M. R. F. Siggel, *Chem. Phys.*, 1997, **220**, 191–205.

58 C. S. M. Allan, B. Lasorne, G. A. Worth and M. A. Robb, *J. Phys. Chem. A*, 2010, **114**, 8713–8729.