Dynamical Systems
Vol. 00, No. 00, Month 200x, 1–37

RESEARCH ARTICLE

Quantum Dynamics Simulations using Gaussian Wavepackets: The vMCG Method

G. W. Richings, I. Polyak, K. E. Spinlove and G. A. Worth*
School of Chemistry, University of Birmingham, Birmingham, B15 2TT, U.K.

and

I. Burghardt
University of Frankfurt, Germany
and

B. L. Lasorne
University of Montpellier, France
(*Received 00 Month 200x; in final form 00 Month 200x*)

Gaussian wavepacket (GWP) methods are an attractive way to solve the time-dependent Schrödinger equation (TDSE). They have an underlying trajectory picture that has a natural connection to semi-classical mechanics, allowing a simple pictorial interpretation of an evolving wavepacket. They also have better scaling with system size compared to conventional grid-based techniques. Here we review the *variational multi-configurational Gaussian* (vMCG) method. This is a variational solution to the TDSE, with explicit coupling between the Gaussian basis functions, resulting in a favourable convergence on the exact solution. The implementation of the method and its performance will be discussed with examples from non-adiabatic photo-excited dynamics and tunneling to show that it can correctly describe both of these strongly quantum mechanical processes. Particular emphasis is given to the implementation of the *direct dynamics* variant, DD-vMCG, where the potential surfaces are calculated on-the-fly via an interface to quantum chemistry programs.

Keywords: Quantum dynamics simulations; Gaussian Wavepackets; Theoretical chemistry; Direct quantum dynamics.

1. Introduction

Solving the time-dependent Schrödinger equation (TDSE) for molecular nuclei using time-dependent Gaussian functions as a basis set, can be traced back to the work of Heller [1, 2]. Known as Gaussian Wavepackets (GWPs), these basis functions provide a link between the delocalised description of quantum mechanics and the localised, trajectory-based, description of classical mechanics. As a result they promise a potential way of breaking the exponential scaling of quantum dynamics simulations with system size. They also provide a way to run quantum direct dynamics, calculating the potential energy surface on-the-fly, bypassing a big bottleneck to practical simulations while still including all quantum effects.

The time-dependent Schrödinger equation

$$i\hbar\frac{\partial\Psi(\mathbf{x},t)}{\partial t} = \hat{H}\Psi(\mathbf{x},t) \tag{1}$$

describes the time-evolution of the system wavefunction, Ψ, a function of the nuclear coordinates, $\mathbf{x}$, driven forward in time by the Hamiltonian operator, $\hat{H}$. Over

ISSN: 1468-9367 print/ISSN 1468-9375 online
© 200x Taylor & Francis
DOI: 10.1080/1468936YYxxxxxxxx
http://www.informaworld.com

the last few decades algorithms have been developed to solve the TDSE accurately [3–6]. These methods can be termed *grid-based* as the wavefunction and Hamiltonian are effectively represented on a grid of points, usually in coordinate space. Using these methods it is possible, at least in principle, to study reactivity in complete detail.

Initial applications of grid-based methods were to calculate state-to-state cross-sections for reactive collisions to help interpret molecular beam experiments [7–9]. Following the development of time-resolved experiments, exemplified by *femto-chemistry* pump-probe laser spectroscopy [10], more recent applications focus on time-resolved phenomena, such as intra-molecular vibrational relaxation pathways, surface crossing timescales in photo-activated processes, and the laser control of reactivity. Grid-based scattering calculations are also being used to great effect in studying ultracold chemistry, where the quantum nature of molecules and atoms dominates their behaviour. Refs. [11, 12] are typical examples. Refs. [13–15] and references therein give a good overview of the general field and techniques.

In place of the discrete basis functions of a grid representation, GWP-based methods use a basis set of, usually time-dependant, Gaussian functions. Initially, Heller used a single GWP, which in one dimension has the form [1]

$$g(x,t) = \exp \frac{1}{\hbar}(-a(x-q)^2 + ip(x-q) + i\gamma)\,, \tag{2}$$

to approximate the evolving wavepacket. It was shown that in a solution of the TDSE the GWP centre coordinate and momentum, (q,p), follow the classical trajectory, the width, α, spreads with time, and the phase, γ, carries the quantum information along the trajectory. Extending this to a multi-dimensional case, when the width becomes a matrix with off-diagonal elements carrying the correlation between modes, is straightforward [16]. It is a standard result of quantum mechanics to show that in a harmonic potential the time-evolving GWP is the exact solution, and it is assumed to be a reasonable solution in other cases when the potential is changing slowly compared to the width. This approximation has been used in a number of studies, but gives good results only for simple systems and very short timescales.

The single "thawed" GWP ansatz, with time-dependent width, α, is simply not flexible enough to follow a real molecular system. To improve this, Heller introduced the idea of describing the wavepacket by a superposition of *frozen* GWPs with constant widths [2]

$$\Psi(\mathbf{x},t) = \sum_j g_j(\mathbf{x},t) \tag{3}$$

each following a different classical trajectory. Much work has been done, especially by Heller and co-workers, to make the frozen Gaussian superposition a useful, general method by connection to semi-classical approaches [17–19]. Metiu and co-workers put the method on a more solid theoretical basis by deriving equations of motion for the Gaussian parameters using a variational "minimum error method" [20–22].

Despite the promise of a simple scheme, however, GWP propagation was found to be numerically very unstable. A resurgence of interest, continuing to the present day, was driven by the development of the *spawning* algorithm of Martínez, Ben-Nun and Levine [23, 24]. This uses an expansion of frozen GWPs with redundant

coefficients

$$\Psi(\mathbf{x},t) = \sum_j A_j(t) g_j(\mathbf{x},t)\,. \tag{4}$$

The GWPs still follow classical trajectories, but the coefficients evolve so as to variationally solve the TDSE in the GWP basis.

$$i\dot{A}_j = \sum_l \left(\langle g_j|H|g_l\rangle - i\langle g_j|\dot{g}_l\rangle\right) A_l \tag{5}$$

The redundant coefficients are very important for the stability of the method as the GWPs can be kept as normalised functions.

The name "spawning" comes from the use of a clever algorithm to expand the basis set in regions of non-adiabatic coupling where the wavefunction bifurcates, allowing an efficient description of curve-crossing processes. This spawning method has pioneered direct quantum dynamics by use of quantum chemistry methods to calculate potential energy surfaces on-the-fly [25, 26]. In this form, known as *ab initio multiple spawning* (AIMS), the spawning method has become a powerful tool for studying photo-excited reactions. Examples are the simulations of the time-resolved photo-electron spectroscopy of uracil [27], of the photo-excited dynamics of ethylene including Rydberg states [28], and of the photoactived dynamics of the chromophore of the PYP enzyme [29].

The *coupled-coherent states* (CCS) method of Shalashilin and co-workers [30] also uses sets of GWPs to solve the TDSE. It differs from the spawning approach in that the GWPs follow trajectories given by the Gaussian averaged potential rather than the classical force at the centre coordinate. It has been extended to multi-state problems as the *multi-configurational Ehrenfest* (MCE) method [31] which in turn has been set up as a direct dynamics method [32] and used to study photochemistry, such as the photodissociation of pyrrole [33]. MCE and spawning have also been combined, taking the best features of each, to produce the *multiple cloning* method [34].

In recent years a number of other elegant methods have been introduced to provide stable solutions of the GWP equations of motion. For example, Batista and co-workers introduced the matching-pursuit method [35] which re-expands the GWP basis periodically in a new basis set and has been applied to both tunneling [36] and non-adiabatic effects [37]. Habershon introduced an algorithm based on the matching-pursuit methods, removing redundant GWPs to prevent linear dependencies [38]. The Basis Expansion Leaping algorithm of Koch and Frankcombe [39] also uses re-expansion combined with judicious use of keeping GWPs fixed in time to cure problems due to overcompleteness.

The subject of this review article is a method that does not use GWPs that follow classical trajectories. Instead, it starts from a full variational solution to the TDSE with a time-dependent basis set, the multi-configurational time-dependent Hartree method (MCTDH). Originally devised as an efficient grid-based solution [40] the method is based on the wavefunction ansatz

$$\Psi(\mathbf{x},t) = \sum_{j_1,\ldots j_d} A_{j_1,\ldots j_d}(t)\varphi^{(1)}_{j_1}(Q_1,t)\ldots\varphi^{(d)}_{j_d}(Q_d,t) \tag{6}$$

where $\varphi^{(\kappa)}(Q_\kappa)$ are multi-dimensional basis functions, known as *single-particle functions* (SPFs). The MCTDH method provides equations of motion for the ex-

pansion coefficients and SPFs [6]. The MCTDH method has proved able to provide accurate quantum dynamics simulations for a wide range of systems. These include non-adiabatic state-population transfer treated with the vibronic coupling model [41] with up to 24 degrees of freedom, as well as more realistic potentials such as the 15D model used to study the strongly anharmonic vibrations in the Zundel cation [42].

The full power of the MCTDH method for solving the TDSE is seen in the multi-layer variant, ML-MCTDH [43]. Here, the multi-dimensional SPFs are in turn expanded in an MCTDH form, leading to layers of coefficients. The resulting method is able to treat hundreds of degrees of freedom and has been used by Wang, Thoss and co-workers to study large systems such as semi-conductors [44] or condensed phase proton transfer [45]. A full overview of the MCTDH method is given in a recent book [14].

To remove the restrictions of the grid, Burghardt *et al* introduced the G-MCTDH method in 1999 [46]. In this, some of the single-particle functions are replaced by Gaussian functions (although any suitable parameterised functions could be used) to give the ansatz,

$$\Psi(\mathbf{x},t) = \sum_{j_1,\dots j_f} A_{j_1,\dots j_f} \varphi^{(1)}_{j_1}(x_1,t)\dots\varphi^{(d)}_{j_d}(x_d,t) g^{(d+1)}_{j_{d+1}}(x_{d+1},t)\dots g^{(f)}_{j_f}(x_f,t)\,, \quad (7)$$

resulting in coupled equations of motion for the expansions coefficients, SPFs and Gaussian parameters. The G-MCTDH method was aimed at treating system-bath problems, where the system is highly quantum mechanical and the bath is a set of oscillators, hence well described by GWPs. Its utility has been demonstrated for high-dimensional system-bath problems [47] and vibronic coupling models [48]. The method is related to the *local coherent state approximation* (LCSA) approach of Martinazzo *et al* [49].

If no grid-based SPFs are included in the G-MCTDH method one naturally arrives at a GWP method. In the following sections we shall discuss the properties of this method, known as the *variational multi-configurational Gaussian* (vMCG) method. As a result of the full variational development, the vMCG GWPs do not follow classical trajectories but, as will be shown, the method does contain trajectory-based GWPs as a limit.

The implementation of the vMCG method will be discussed, showing how numerical problems are dealt with, and demonstrating the fast convergence on both non-adiabatic and tunneling problems. The implementation of the direct dynamics version (DD-vMCG) will also be presented. A description of the vMCG and DD-vMCG methods is given in a number of the original papers as their development proceeded [46, 48, 50–54]. Here we summarize the general algorithmic details of the vMCG method and emphasize the principles of its current implementation. We then describe the recent developments in the way direct dynamics is performed, building up a database of energies, gradients and Hessians.

2. Theory and Implementation

2.1. *Equations of motion*

Consistent with the MCTDH philosophy, the vMCG equations of motion (EOM) are obtained by applying the Dirac-Frenkel variational principle [55, 56],

$$\langle \delta\Psi | \, H - i\frac{\partial}{\partial t} \, | \Psi \rangle = 0, \tag{8}$$

guaranteeing variationally optimal evolution of the approximate nuclear wavefunction. It is taken to be written as a sum of time-dependent basis functions, multiplied by the time-dependent coefficients,

$$\Psi(\mathbf{x}, t) = \sum_{j=1}^{n} A_j(t) g_j(\mathbf{x}, t), \tag{9}$$

where the basis functions have the form of multidimensional Gaussian functions, with all degrees of freedom combined together. In matrix notation these have the following form:

$$g_j(\mathbf{x}, t) = \exp(\mathbf{x}^T \cdot \boldsymbol{\varsigma}_j \cdot \mathbf{x} + \boldsymbol{\xi}_j \cdot \mathbf{x} + \eta_j). \tag{10}$$

Note that the Dirac-Frenkel principle has been written in units in which $\hbar = 1$ for simplicity. The complex, generally time-dependent parameters of the function,

$$\boldsymbol{\Lambda}_j = \{\boldsymbol{\varsigma}_j, \boldsymbol{\xi}_j, \eta_j\}, \tag{11}$$

are represented by a square matrix (with the dimension equal to the number of degrees of freedom), a vector and a scalar respectively. Based on the choice of the matrix $\boldsymbol{\varsigma}_j$, we define three types of Gaussian functions. Following Heller we call these: thawed, with $\boldsymbol{\varsigma}_j$ containing both diagonal and off-diagonal elements, therefore allowing for coupling between different modes; separable, where $\boldsymbol{\varsigma}_j$ is diagonal; and frozen, where the diagonal elements are kept fixed during the wavepacket propagation. The latter Gaussian type is usually used for quantum and semi-classical dynamics simulations. Although thawed Gaussians have been used successfully in system-bath problems [47], in anharmonic potentials they lead to an unstable propagation [57]. Conventionally we also use frozen Gaussians, however, we have recently found strong evidence for the importance of carefully choosing the width parameter $\boldsymbol{\varsigma}_j$. Developing stable dynamics using separable or semi-separable Gaussians, e.g. where the width parameter will be evolving according to the local PES shape [58] may be the solution. Rewriting Eq. (10) for a separable case in a linear notation gives,

$$g_j(\mathbf{x}, t) = \exp(\sum_{\kappa} \varsigma_{j\kappa} x_\kappa^2 + \xi_{j\kappa} x_\kappa + \eta_j), \tag{12}$$

with κ running over the system degrees of freedom. Although, this could also be written as a product of one-dimensional Gaussian functions, the single function emphasizes the relationship of the parameters which are coupled in the dynamics.

The relationships:

$$\hbar\varsigma_{j\kappa} = -a_{j\kappa} \tag{13a}$$

$$\hbar\xi_{j\kappa} = 2a_{j\kappa}q_{j\kappa} + ip_{j\kappa} \tag{13b}$$

$$\hbar\eta_j = \sum_\kappa (-a_{j\kappa}q_{j\kappa}^2 - ip_{j\kappa}q_{j\kappa}) + i\gamma_j \tag{13c}$$

transform Eq. 12 into the more intuitive Heller form of Eq. (2),

$$g_j(\mathbf{x},t) = \exp(\frac{1}{\hbar}\sum_\kappa -a_{j\kappa}(x_\kappa - q_{j\kappa})^2 + ip_{j\kappa}(x_\kappa - q_{j\kappa}) + i\gamma_j). \tag{14}$$

Therefore, parameters $\boldsymbol{\varsigma}_j$ represent the width of the Gaussian functions, $\boldsymbol{\xi}_j$ represent linear parameters - momentum and coordinate of the centre of the function - and η_j is a sum of the remaining scalar parameters, including the phase, γ_j. It is this last parameter which, by various treatments, leads to a potentially stable propagation (discussed below).

Applying the Dirac-Frenkel variational principle (Eq. (8)) to the wavefunction (Eq. (9)), two EOM are obtained - one for the time-dependent coefficients, $A_j(t)$, and one for the parameters of the Gaussian functions, $\boldsymbol{\Lambda_j}$. This is fairly straightforward algebra starting from variations either to a coefficient or one of the Gaussian parameters

$$\delta\Psi = \delta A_j g_j \tag{15}$$

$$\delta\Psi = \delta\lambda_{j\alpha}A_j\frac{\partial g_j}{\partial\lambda_{j\alpha}} \tag{16}$$

where $\lambda_{j\alpha}$ is one of the parameters of the jth Gaussian, g_j.

The resulting EOM for the time-dependent coefficient has the following form:

$$i\dot{A}_j = \sum_{lm}[S_{jl}]^{-1}(H_{lm} - i\tau_{lm})A_m, \tag{17}$$

where $\mathbf{S}$ is an overlap matrix with the elements,

$$S_{jl} = \langle g_j | g_l \rangle , \tag{18}$$

$\mathbf{H}$ is a Hamiltonian matrix with the elements,

$$H_{jl} = \langle g_j | \hat{H} | g_l \rangle , \tag{19}$$

and $\boldsymbol{\tau}$ is an overlap time-derivative matrix, the diagonal of which is zero as a necessary constraint to keep the Gaussians normalised during the propagation. Using the chain rule,

$$\tau_{jl} = \langle g_j | \dot{g}_l \rangle = \sum_\alpha \langle g_j | \frac{\partial g_l}{\partial\lambda_{l\alpha}}\dot{\lambda}_{l\alpha} \rangle , \tag{20}$$

$\boldsymbol{\tau}$ can be explicitely written as a function of the time derivative of the Gaussian parameters, that are from now on considered to be part of a single vector $\boldsymbol{\Lambda}_j$ with indices α (with $\max(\alpha) = N_\kappa + 1$ for the frozen-width, $\max(\alpha) = 2N_\kappa + 1$ for

separable, and $\max(\alpha) = N_\kappa(N_\kappa + 1) + 1$ for thawed Gaussians, where N_κ is the number of degrees of freedom).

The EOM for the Gaussian parameters has a compact matrix form [46]:

$$i\dot{\boldsymbol{\Lambda}} = [\boldsymbol{C}]^{-1}\boldsymbol{Y} \quad .. \tag{21}$$

The EOM for the Gaussian parameters, however, becomes much more complicated when the matrix $\boldsymbol{C}$ and vector $\boldsymbol{Y}$ are expanded:

$$C_{j\alpha,l\beta} = \rho_{jl}(S_{jl}^{(\alpha\beta)} - [\boldsymbol{S}^{(\alpha 0)}\boldsymbol{S}^{-1}\boldsymbol{S}^{(0\beta)}]_{jl}) \tag{22a}$$

$$Y_{j\alpha} = \sum_l \rho_{jl}(H_{jl}^{(\alpha 0)} - [\boldsymbol{S}^{(\alpha 0)}\boldsymbol{S}^{-1}\boldsymbol{H}]_{jl}), \tag{22b}$$

where ρ_{jl} is an element of the density matrix:

$$\rho_{jl} = A_j^* A_l \tag{23}$$

and additional definitions are:

$$S_{jl}^{(\alpha\beta)} = \langle \frac{\partial g_j}{\partial \lambda_{j\alpha}} | \frac{\partial g_l}{\partial \lambda_{l\beta}} \rangle \tag{24a}$$

$$S_{jl}^{(\alpha 0)} = \langle \frac{\partial g_j}{\partial \lambda_{j\alpha}} | g_l \rangle \tag{24b}$$

$$H_{jl}^{(\alpha 0)} = \langle \frac{\partial g_j}{\partial \lambda_{j\alpha}} | H | g_l \rangle \,. \tag{24c}$$

The definitions in Eqs. (24a) and (24b) are particularly important as they, depending on the parameters α and β, correspond to the Gaussian moments of different order $M_{jl}^{(\kappa)} = \langle g_j | x_\kappa | g_l \rangle$, $M_{jl}^{(\kappa\mu)} = \langle g_j | x_\kappa x_\mu | g_l \rangle$ etc. We are particularly interested in the matrix $\boldsymbol{S}^{(0\alpha)}$, as will become clear shortly. The order up to which the $\boldsymbol{S}^{(0\alpha)}$ can represent the Gaussian moments depends on the type of the Gaussian basis function in use. For a frozen Gaussian only zero-order and linear moments are available. For a separable Gaussian all zero-order, linear and diagonal second-order moments are covered. For a thawed Gaussian all moments up to second-order can be represented.

2.2. *"CX" formalism*

The importance of the representability of the Gaussian moments by the matrix $\boldsymbol{S}^{(0\alpha)}$ becomes clear if one realizes that the separable (zeroth-order) part of the Hamiltonian can be written as a power series expansion in terms of Gaussian moments:

$$H_{jl} = M_{jl}^{(0)} X_l^{(0)} + \sum_\kappa M_{jl}^{(\kappa)} X_l^{(\kappa)} + \sum_{\kappa\mu} M_{jl}^{(\kappa\mu)} X_l^{(\kappa\mu)} + ... \tag{25}$$

Using a local harmonic approximation (LHA), the potential energy is expanded with respect to the time-dependent centre coordinate, $\boldsymbol{q}_j(t)$, of a given multi-

dimensional Gaussian function up to the second order, we can write,

$$V_j(\boldsymbol{r}) = V_{j0} + \sum_\kappa V'_{j,\kappa}(r_\kappa - q_{j\kappa}) + \frac{1}{2}\sum_{\kappa\mu} V''_{j,\kappa\mu}(r_\kappa - q_{j\kappa})(r_\mu - q_{j\mu}), \tag{26}$$

where the derivatives of the potential are all taken at the centre of $g_j(\mathbf{r}, t)$, i.e. at $\mathbf{q}_j(t)$. When in addition the kinetic energy operator is taken to have a separable form $\hat{T} = \sum_j \frac{1}{2m_j}\frac{\partial^2}{\partial r_j^2}$, the X coefficients in Eq. (25) are as follows

$$X_j^{(0)} = \sum_\kappa (\frac{\varsigma_{j\kappa}}{m_\kappa} + \frac{\xi_{j\kappa}^2}{2m_\kappa}) + V_{j0} - \sum_\kappa V'_{j,\kappa} q_{j\kappa} + \frac{1}{2}\sum_{\kappa\mu} V''_{j,\kappa\mu} q_{j\kappa} q_{j\mu} \tag{27a}$$

$$X_j^{(\kappa)} = -\frac{2}{m_\kappa}\varsigma_{j\kappa}\xi_{j\kappa} + V'_{j,\kappa} - \sum_\mu V''_{j,\kappa\mu} q_{j\mu} \tag{27b}$$

$$X_j^{(\kappa\mu)} = -\frac{2}{m_\kappa}\varsigma_{j\kappa}\varsigma_{j\mu} + \frac{1}{2}V''_{j,\kappa\mu} \tag{27c}$$

It is useful now to separate the $\boldsymbol{Y}$-vector into the two parts:

$$\boldsymbol{Y} = \boldsymbol{Y}_0 + \boldsymbol{Y}_R \quad , \tag{28}$$

with the first term, $\boldsymbol{Y}_0$, depending on the part of a Hamiltonian that can be written in terms of $\boldsymbol{S}^{(0\alpha)}$, and the second, "residual term" $\boldsymbol{Y}_R$, comprising the remaining part of the Hamiltonian. This includes higher order terms of the separable part, as well as correlation terms of the Hamiltonian. With that in mind, comparing Eqs. (22a) and (22b), Eq. (22b) can be rewritten in the following way:

$$Y_{j\alpha} = \sum_{l\beta} C_{j\alpha,l\beta} X_l^{(\beta)} + Y_{R,j\alpha}. \tag{29}$$

This leads to a simplification of the EOM for the Gaussian parameters [46, 48]:

$$i\dot{\boldsymbol{\Lambda}} = \boldsymbol{X} + \boldsymbol{C}^{-1}\boldsymbol{Y}_R. \tag{30}$$

The CX formalism has advantages. Firstly, it improves the stability of the propagation as removing part of Hamiltonian from the $\boldsymbol{C}^{-1}\boldsymbol{Y}$-term decreases the possible numerical error due to the $\boldsymbol{C}$-matrix inversion. Secondly, as described in the next section, it allows a division of the parameter EOMs into "classical" and "non-classical" parts. A more rigorous discussion of the classical limit of vMCG is given in Ref. [59].

2.3. *Separation of classical and quantum terms*

We adopt the following strategy in vMCG: only the terms corresponding to classical propagation of the Gaussian parameters are put into the $\boldsymbol{X}$ term, while all of the other, "quantum" contributions, are kept in $\boldsymbol{Y}_R$. Such a separation must be done differently for each of the Gaussian types (thawed, separable and frozen). It is common knowledge that the thawed Gaussians in a harmonic potential move classically, so that in our notation $i\dot{\boldsymbol{\Lambda}} = \boldsymbol{X}$ is the complete solution (see also [20, 46, 48]). We will, however, describe the frozen-width case here which is, surprisingly, more com-

plicated in terms of the separation of the classical and non-classical contributions [60].

In the case of frozen-width Gaussian basis functions only the κ linear parameters are varied with time, and so the zeroth-order Hamiltonian can be written as:

$$H_{jl} = S_{jl}^{(0)} X_l^{(0)} + \sum_{\kappa} S_{jl}^{(0\kappa)} X_l^{(\kappa)} + \sum_{\kappa\mu} M_{jl}^{(\kappa\mu)} X_l^{(\kappa\mu)} + ..., \tag{31}$$

(please note the difference to Eq. (25)) with $X_j^{(0)}$ and $X_j^{(\kappa)}$ given in Eq. (27a) and Eq. (27b) respectively. It is straightforward now to separate the Hamiltonian into the $\boldsymbol{Y}_0$ and $\boldsymbol{Y}_R$ terms. However the zeroth-order part of the the linear parameter EOM are of the form

$$i\dot{\xi}_{j\kappa}^{(0)} = -i2\varsigma_{j\kappa}\dot{q}_{j\kappa} - \dot{p}_{j\kappa} = X_j^{(\kappa)} \tag{32}$$

$$= -\frac{2}{m_\kappa}\varsigma_{j\kappa}\xi_{j\kappa} + V'_{j,\kappa} - \sum_{\mu} V''_{j,\kappa\mu} q_{j\mu}, \tag{33}$$

while the classical EOM for $\dot{q}_{j\kappa}$ and $\dot{p}_{j\kappa}$ have the following form:

$$\dot{q}_{j\kappa} = \frac{p_{j\kappa}}{m_\kappa} \tag{34a}$$

$$\dot{p}_{j\kappa} = -V'_{j,\kappa} \tag{34b}$$

In order to obtain $i\dot{\xi}_{j\kappa}^{(0)} = -i2\varsigma_{j\kappa}\frac{p_{j\kappa}}{m_\kappa} + V'_{j,\kappa}$, we must therefore put the terms $-\frac{2}{m_\kappa}\varsigma_{j\kappa}(\xi_{j\kappa} - ip_{j\kappa})$ and $-\sum_\mu V''_{j,\kappa\mu} q_{j\mu}$ into $\boldsymbol{Y}_R$.

Instead of propagating $\xi_{j\kappa}^{(0)} = -2\varsigma_{j\kappa}q_{j\kappa} + ip_{j\kappa}$, we propagate $\bar{\xi}_{j\kappa}^{(0)} = q_{j\kappa} + ip_{j\kappa}$, and when necessary (e.g. for calculating τ_{jl}) reconstruct $\xi_{j\kappa}^{(0)}$, using the constant width parameter, $\varsigma_{j\kappa}$. The remaining part of the $\dot{\xi}_{j\kappa}$ is obtained from $\boldsymbol{Y}_R$:

$$i\dot{\xi}_{R,j\kappa} = \sum_{l\mu} C^{-1}_{j\kappa l\mu} Y_{R,l\mu} \tag{35}$$

So overall we have:

$$i\dot{\bar{\xi}}_{j\kappa} = V'_{j,\kappa} - i\frac{p_{j\kappa}}{m_\kappa} + \mathrm{Re}(i\dot{\xi}_{R,j\kappa}) - \frac{1}{2\varsigma_{j\kappa}} i\mathrm{Im}(i\dot{\xi}_{R,j\kappa}), \tag{36}$$

Ignoring the last two terms means that each GWP will follow a classical trajectory. This connects the variational GWPs of vMCG with the classical basis functions of spawning or CCS. A detailed comparison showing the relationship between vMCG and CSS is given in [61]. It should also be mentioned that for coherent states in the harmonic well of the appropriate width these last two terms cancel.

2.4. *Phase propagation*

The variational derivation of the EOMs do not define the time-evolution of the scalar parameters, η_j. This is due to the redundancy in the wavefunction ansatz as a result of the expansion coefficients. As noted in Sec. 2.1 and explained in [46], if τ_{jj} is zero or imaginary, the Gaussian basis functions remain normalised. For the

frozen-width Gaussians we get the following form:

$$\tau_{jj} = \langle g_j | \dot{g}_j \rangle = \sum_{\alpha} \langle g_j | \frac{\partial g_j}{\partial \lambda_{j\alpha}} \dot{\lambda}_{j\alpha} \rangle = \sum_{\kappa} q_{j\kappa} \dot{\xi}_{j\kappa} + \dot{\eta}_j, \tag{37}$$

with the time-derivative of the scalar parameter, η_j, (Eq. (13c)) for a frozen-width Gaussian function being,

$$\dot{\eta}_j = \sum_{\kappa} (-2a_{j\kappa} q_{j\kappa} \dot{q}_{j\kappa} - i\dot{p}_{j\kappa} q_{j\kappa} - i p_{j\kappa} \dot{q}_{j\kappa}) + i\dot{\gamma}_j. \tag{38}$$

From Eq. (37), if the choice $\tau_{jj} = 0$ is made, then $\dot{\eta}_j$ is constrained to have the following form:

$$\dot{\eta}_j = \sum_{\kappa} (-2a_{j\kappa} q_{j\kappa} \dot{q}_{j\kappa} - i\dot{p}_{j\kappa} q_{j\kappa}), \tag{39}$$

(using Eq. (13b)) which, according to Eq. (38), implies $i\dot{\gamma}_j = \sum_{\kappa} -ip_{j\kappa}\dot{q}_{j\kappa}$. One may, however, wish to have $i\dot{\gamma}_j = 0$, which implies:

$$\dot{\eta}_j = \sum_{\kappa} (-2a_{j\kappa} q_{j\kappa} \dot{q}_{j\kappa} - i\dot{p}_{j\kappa} q_{j\kappa} - i p_{j\kappa} \dot{q}_{j\kappa}), \tag{40}$$

and $\tau_{jj} = \sum_{\kappa} -ip_{j\kappa}\dot{q}_{j\kappa}$, which is still alright, as an imaginary τ_{jj} also keeps the Gaussian function normalised, as noted above.

In the current implementation of vMCG we have chosen the latter formalism, as it enables one to omit propagation of the scalar parameter η_j since it can be simply recalculated at every step with Eq. (13c), thus simplifying the integration of the EOMs.

2.5. *Singularities and the C-matrix inversion*

The EOM for the Gaussian parameters (Eqs. (21), (22a) and (22b)) involves the inverse of the overlap, $\boldsymbol{S}$, and $\boldsymbol{C}$ matrices. Inversion of each of these matrices may become a serious numerical problem if they become close to singular. In the case of the $\boldsymbol{S}$-matrix, this happens if the Gaussian functions overlap strongly. This problem has been raised already by Metiu and coworkers [62, 63]. The behaviour of the $\boldsymbol{C}$-matrix is more complex and harder to visualise. However, if Eq. (22a) is rewritten in terms of the derivatives of Gaussian functions with respect to their parameters (derivative functions):

$$C_{j\alpha,l\beta} = \rho_{jl} \left\langle \frac{\partial g_j}{\partial \lambda_{j\alpha}} \middle| 1 - \sum_{rs} |g_r\rangle \, S^{-1}_{rs} \, \langle g_s| \, \middle| \frac{\partial g_l}{\partial \lambda_{l\beta}} \right\rangle, \tag{41}$$

it becomes evident that the $\boldsymbol{C}$-matrix represents projection of the derivative functions out of the GWP parameter space, multiplied by the density matrix, ρ_{jl}.

Thus the $\mathbf{C}$ can become singular for two distinct reasons. The first is when the density matrix, ρ, is singular. This problem is well known from MCTDH and can be handled with a "regularisation". The second reason, however, is when the basis set is complete. In this limit, the parameter derivatives cannot be projected outside the space spanned by the Gaussian functions and $\mathbf{C}$ becomes zero. This pathological property means that the vMCG EOMs become very unstable as the

basis approaches completeness because the motion is irrelevant. Furthermore, for this reason regularisation is inappropriate when $\mathbf{C}$ is singular. When the singularity is due to the density matrix, regularisation perturbs the evolution of unpopulated functions, which is not significant. Here though it may be randomly perturbing the evolution of populated functions, introducing errors.

Standard ways to deal with singular matrices without regularisation are based on matrix decomposition techniques such as LU or Cholesky which can take advantage of efficient parallelisation (e.g. using ScaLapack libraries). Unfortunately, these techniques have failed to cope with the $\mathbf{C}$ matrix in tests.

To cure the pathological problem when approaching completeness, an algorithm termed *dynamic coupling* has been introduced which simply ignores the coupling for GWPs when it is not required [58]. At time t the time derivatives of the coefficients are calculated and used to estimate the coefficients at the end of a large step, $t+\delta t$. This is used to estimate the density matrix at $t+\delta t$ which is needed as functions unpopulated at t may become populated during δt. Using this approximate density matrix the $\mathbf{C}$ matrix is built using the GWPs at t and diagonalised. The number of eigenvalues, n_{thr}, under a threshold are counted ($\varepsilon = 10^{-6}$ has been found to be a suitable value). The functions that contribute most to these n_{thr} eigenvectors are then "decoupled" and follow classical trajectories over the long step. The reduced $\mathbf{C}$ matrix in the space of important GWPs is then inverted (with regularisation, which may be needed due to small populations) at every integration step.

Typically, a basis set is overcomplete at the start of a propagation as the initial wavepacket is a Gaussian function surrounded by basis functions. The completeness decreases as the wavepacket evolves, spreads out and develops structure. The dynamic coupling algorithm allows the space to adapt appropriately.

The standard way of regularisation, although having proved to provide stable propagation (at least for frozen-width Gaussians), is time-consuming, as diagonalisation of the $\boldsymbol{C}$-matrix is required at every integration step and becomes a computational bottleneck for systems with many degrees of freedom. We have, however, recently implemented an approach, known as Tikhonov regularisation [64], which performs well in providing a regularised inverse. In this approach the regularised inverse of matrix $\boldsymbol{A}$ is given by:

$$\tilde{\boldsymbol{A}}^{-1} = (\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Gamma}^T\boldsymbol{\Gamma})^{-1}\boldsymbol{A}^T, \tag{42}$$

where $\boldsymbol{\Gamma}$ is a unit matrix multiplied by the factor that defines the level of regularisation, and inversion of $(\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Gamma}^T\boldsymbol{\Gamma})$ may be obtained by means of a singular value decomposition (SVD) or, for example, LU decomposition, which scales better than matrix diagonalisation and SVD. With $\boldsymbol{\Gamma} = 0$, Eq. (42) reduces to an unregularised inversion of matrix $\boldsymbol{A}$ assuming $(\boldsymbol{A}^T\boldsymbol{A})^{-1}$ exists. We found that the regularisation factor of 10^{-12} gives the same results as the conventional regularisation, while lower values can lead either to an increase of integration step sizes (making propagation faster) or to making propagation less stable. Consequently its value should be chosen individually for each system under consideration.

Alternatively, after projecting out the unwanted parameters using the dynamic coupling algorithm, the $\boldsymbol{C}$-matrix rarely becomes near-to-singular and a simple matrix decomposition (e.g. LU decomposition or Cholesky factorization) can be used to get the inverse, without using any kind of regularisation. In the case of salicylaldimine, this proved to be efficient [58].

A second numerical problem is that the overlap $\mathbf{S}$ matrix can also become singular. Again there can be two different causes. The first is that two GWPs get very close in phase space and their overlap approaches 1 (linear dependence). In addi-

tion, a dense basis set can lead to an overlap matrix with small eigenvalues even when there is no specific clash. This again happens when the basis set approaches completeness. The approach we use to deal with both cases is similar to that used for the **C** matrix above. The ***S***-matrix is diagonalised at every step (which is not an expensive procedure due to the relatively small matrix dimensions) and if $n_{\rm thr}$ eigenvalues are below a given threshold (usually 10^{-6}), $n_{\rm thr}$ Gaussians are removed from the propagation as they are surplus to requirement, and the wavepacket re-expressed in the remaining functions. In this case they are permanently removed and ignored. Removing a function results in a discontinuity in the propagation, which can lead to instabilities hence this should only be done when essential. For this reason the linear dependency check and removal is only done at the start of a large step and the **S** matrix is simply regularised over the large step.

2.6. *Multi-layer vMCG*

The **C**-matrix inversion is the bottleneck in the vMCG method. The matrix has dimension $(n \times f)^2$ where n is the number of GWPs and f the number of degrees of freedom so inverting by initial diagonalisation, as required by the regularisation quickly becomes hard as this scales as $(n \times f)^3$. Even using efficient direct inversion, such as the Tikhonov method mentioned above, is hard for large matrices as these too scale as $(n \times f)^3$.

A simple way to improve on the effort required for ***C***-matrix inversion is to apply the multi-layer (ML) philosophy to the vMCG method. In this, as in ML-MCTDH, the wavefunction is represented, not by the sum of the full-dimensional Gaussian functions, but is expanded in two or more layers of functions, from which only the lowest-level functions are represented by low-dimensional Gaussians. The theory for the two-layer approach was recently developed [65].

The idea is straightforward. The wavefunction is first expanded in MCTDH form in orthonormal SPFs (Eq. (6))

$$\Psi(\mathbf{x},t) = \sum_{j_1=1}^{n_1} \dots \sum_{j_f=1}^{n_p} A_{j_1 \dots j_p}(t) \prod_{\kappa=1}^{p} \varphi_{j_\kappa}^{(\kappa)}(Q_\kappa,t) \tag{43}$$

and each SPF is then expanded in a set of frozen GWPs:

$$\mid \varphi_j\rangle = \sum_{a=1}^{m} \mid g_a\rangle S_{ab}^{-1} \langle g_b \mid \varphi_j\rangle \tag{44}$$

$$= \sum_{a=1}^{m} \mid g_a\rangle D_{aj} \tag{45}$$

The GWPs are thus a time-dependent basis for the representation of the SPFs and the **D** matrix elements are hence included in the variational solution of the TDSE. The EOMs for the A_J are those of MCTDH, while those for the GWP parameters, $\lambda_{a\alpha}$, are as in standard vMCG. The **D** matrix EOM is

$$i\dot{D}_{ci} = \sum_{lja} S_{ca}^{-1} \rho_{ij}^{-1} \langle g_a|(1-\hat{P})\langle H\rangle_{jl}|\varphi_l\rangle + \sum_m f_{mi} D_{cm} - \sum_{ab} S_{ca}^{-1} \tau_{ab} D_{bi} \tag{46}$$

with $P = \sum_r |\varphi_r\rangle\langle\varphi_r|$ the projector onto the SPF space. where $f_{ij} = i\langle\varphi_i|\dot{\varphi}_j\rangle$ is a Hermitian constraint matrix that keeps the SPFs orthonormal. Different choices

for this matrix are possible that affect the numerics, but do not affect the final result [6].

This approach reduces all matrices to be inverted to low-dimensionality and, while yet to be tested fully, should make it possible to treat hundreds of degrees of freedom, while currently only systems of about 20-30 degrees of freedom can be realistically treated.

3. Example Calculations

3.1. *Non-adiabatic Dynamics*

As a first example of the performance of the vMCG method we will take a classic model of non-adiabatic photophysics; the four-dimensional pyrazine S_2/S_1 absorption. This model describes the initial photo-excited dynamics of pyrazine and is able to reproduce the main features of the absorption spectrum [66]. As part of an extended model, it was used in the first demonstration of the power of the MCTDH method in providing a full 24-dimensional quantum dynamics simulation [67].

The Hamiltonian is a linear vibronic coupling model, in which the electronic states are diabatic by construction. Coordinates are mass-frequency weighted normal modes and vibronic coupling is provided only by a single, linear term due to symmetry. It has the form:

$$\begin{aligned}\mathbf{H} = &\sum_i \frac{\omega_i}{2}\left(-\frac{\partial^2}{\partial Q_i^2} + Q_i^2\right)\mathbf{1} + \begin{pmatrix} -\Delta & 0 \\ 0 & \Delta \end{pmatrix} \\ &+ \sum_{i=v_1,v_{6a},v_{9a}} \begin{pmatrix} \kappa_i^{(1)} & 0 \\ 0 & \kappa_i^{(2)} \end{pmatrix} Q_i + \sum_{i=v_{10a}} \begin{pmatrix} 0 & \lambda_i \\ \lambda_i & 0 \end{pmatrix} Q_i\end{aligned} \tag{47}$$

Parameters were provided by fitting to quantum chemistry calculations with subsequent adjustments to fit the experiments [67].

In vMCG, as in MCTDH, there are two alternative formulations that can be used to treat non-adiabatic problems which involve multiple electronic states. In the *multi-set* formalism, a different set of basis functions is used for each electronic state [48],

$$\Psi(\mathbf{x},t) = \sum_{s=1}^{n_s}\sum_{j_s=1}^{n_{j_s}} A_{j_s}^{(s)}(t) g_{j_s}^{(s)}(\mathbf{x},t) \tag{48}$$

where the superscript labels which of the n_s states with which the basis functions are associated. There can be a different number of functions for each state. The EOMs for the expansion coefficients also change slightly to couple the GWPs in the different states

$$i\dot{A}_j^{(s)} = \sum_{lm}[S_{jl}^{(s)}]^{-1}[(H_{lm}^{(ss)} - i\tau_{lm})A_m^{(s)} + \sum_{s'\neq s} H_{lm}^{(ss')}A_m^{(s')}] \tag{49}$$

where the superscripts denote the states.

In contrast, in the *single-set* formalism the electronic states are included as an extra degree of freedom described by a finite basis labeling the states and only a

single set of basis functions is used:

$$\Psi(\mathbf{x},t) = \sum_{s=1}^{n_s} \sum_{j=1}^{n_j} A_{j,s}(t) g_j(\mathbf{x},t)|s\rangle \tag{50}$$

In general it is found that the multi-set formalism is preferred as the basis functions are able to adapt better to the different states so the number of expansion coefficients is minimised. However, it does require more basis functions, which for direct dynamics may be a crucial expense as more quantum chemistry calculations may be required.

To demonstrate the convergence of the vMCG method, simulations were performed using the single-set formalism with different numbers of four-dimensional GWPs. In Fig. 1 (b) - (d) and (f) - (h) the results using vMCG with 20, 40 and 60 GWPs are shown against the fully converged grid-based MCTDH ones. In Fig. 1 (b) - (d) the population of the S_2 state is shown, and in (f) - (h) the absolute value of the autocorrelation function is plotted. The convergence on the exact result is clear, with the greatest error being pushed out to longer times as more functions are added. Note that the autocorrelation function goes to 240fs while the state populations only to 120fs. This is due to using the time doubling allowed by the time symmetry of the problem [6]. Similar results were obtained using the multi-set formalism [48], with convergence being reached with 40 GWPs on each state - 80 GWPs in total.

These calculations were performed using the dynamic coupling algorithm with a cutoff of 10^{-6}. The number of functions that are coupled is shown in Fig. 2 for the vMCG calculations with 60 and 40 GWPs. It can be seen that the number grows quite smoothly.

Ignoring the coupling totally, the GWPs follow classical trajectories. As we are using a single-set formalism, this is similar to an Ehrenfest approach with the GWPs moving on a potential averaged by the state populations.

The classical GWPs converge much more slowly. The results with 140 functions are shown in Fig. 1 (a) and (e). While the correct structure is seen, with peaks in both the populations and autocorrelation function in the correct places, it is seen that the recurrences are too strong and while an improvement is seen compared to smaller basis sets, clearly many more functions will be needed for a converged result.

A final word must be said about computer resources. For this simple model the MCTDH method is very efficient and needs just 8 s CPU time on a linux workstation. The vMCG calculations with 20, 40 and 60 GWPs need much longer with 219, 1,328 and 4,817 s, respectively. The classical GWPs need less time than the coupled vMCG ones, with 60 GWPs requiring only 406 s. With 140 classical GWPs, the time becomes much longer and 2418s is needed. These timings seem to suggest that vMCG is very uncompetitive. However, it is not meant to compete with MCTDH, but be a more general and flexible method able to run direct dynamics. The comparison will also be different for larger systems as the scaling is quite different. Compared to the classical GWPs there is a clear trade off between accuracy and time - vMCG is much more expensive, but converges much faster and so is potentially cheaper.

3.2. *Proton Transfer Dynamics*

For the second set of calculations, the system of interest is the salicylaldimine (SA) molecule, which can undergo proton transfer between a hydroxy and primary

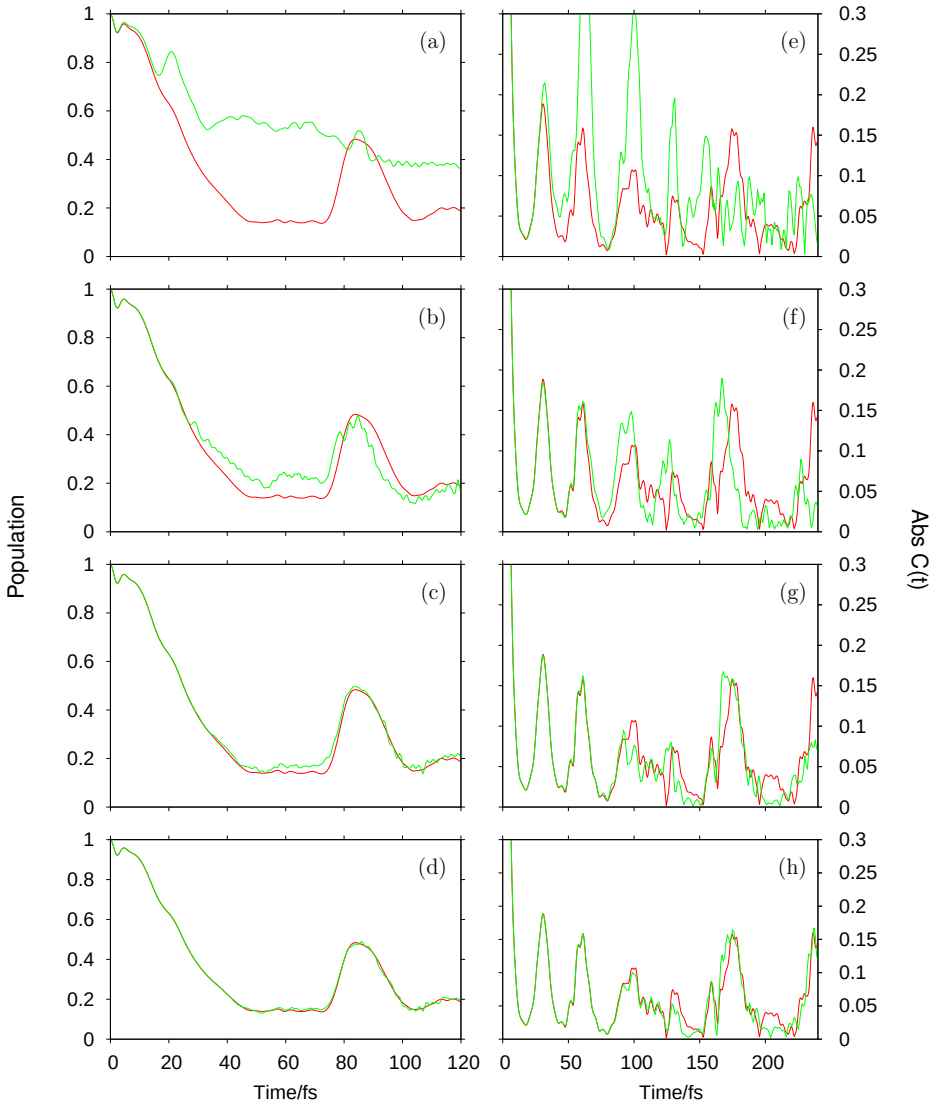

Figure 1. The population of the S_2 state (left panel) and the absolute value of the autocorrelation function (right panel) for different simulations of the photo-excited dynamics of pyrazine in a 4-mode model. In all figures the full MCTDH result is in red. (a) and (e) 140 uncoupled (classical) GWPs. (b) and (f) 20 vMCG GWPs. (c) and (g) 40 vMCG GWPs. (d) and (h) 60 vMCG GWPs. The single-set formalism was used.

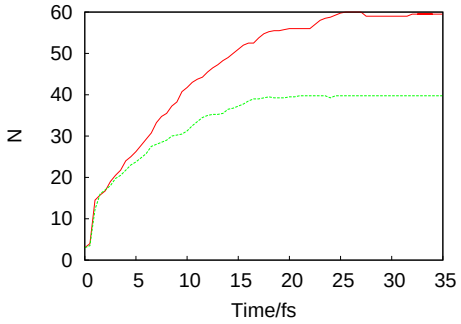

Figure 2. The number of GWP functions, N, as a function of time, that are coupled in a simulation of a 4-mode model of pyrazine with 60 GWPs (red) and 40 GWPs (green).

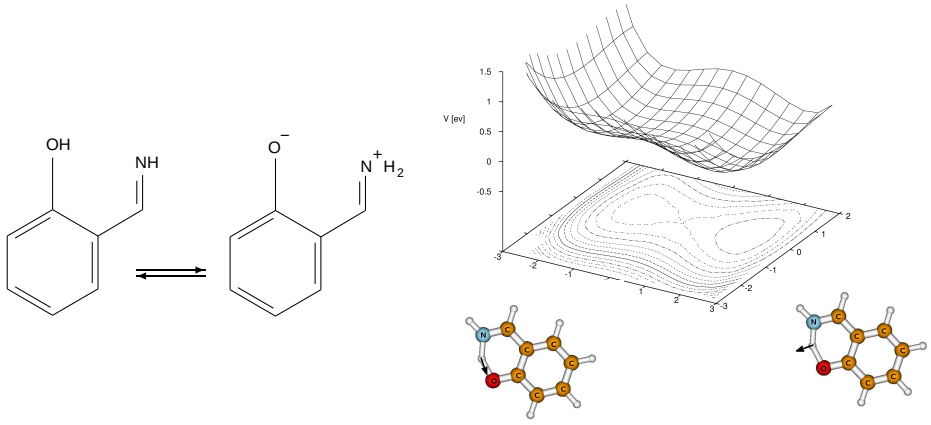

Figure 3. Representation of the proton transfer isomerisation of salicylaldimine with the two-dimensional potential surface of the main vibrations v_1 and v_{36}.

imine groups (Fig. 3). This is a convenient system to study intramolecular proton transfer due to its medium size and the planar structure of the phenol ring imposing certain constraints on the modes directly involved into the reaction. Due to the fundamental interest in weak hydrogen-bond dynamics, the intramolecular hydrogen-bonding of SA has previously been the subject of purely theoretical studies [68]. The excited-state proton-transfer reactions of SA substituted derivatives such as salicylidene methylamine or salicylideneaniline, are being actively investigated both theoretically[69–71] and experimentally[72–76]. These representatives of the family of Schiff bases are of particular interest due to their photochromism and potential of applying them as molecular photo switches in devices such as rewritable molecular memories and nanoelectronics[77]. As our purpose is to assess accuracy and efficiency of the direct quantum dynamics method and not to obtain quantitative results, potential energy surfaces have been constructed by fitting to quantum chemistry calculations at the computationally cheap restricted Hartree-Fock (RHF) level with the 3-21G* basis using the Gaussian 03 package [78]. The level of theory was chosen to allow comparison to direct dynamics calculations (see below Sec. 4.2) and provides a suitable double well potential. The surface was a simple polynomial up to 4th order and the in-plane mass-frequency weighted normal modes calculated at the transition state optimised at the RHF/3-21G* level.

$$V(\boldsymbol{Q}) = V(\boldsymbol{Q}_0) + \sum_{\alpha}^{N} \kappa_\alpha Q_\alpha + \frac{1}{2}\sum_{\alpha\beta}^{N} \gamma_{\alpha\beta} Q_\alpha Q_\beta + \frac{1}{6}\sum_{\alpha\beta}^{N} \epsilon_{\alpha\beta} Q_\alpha Q_\beta^2 \qquad (51)$$
$$+\frac{1}{24}\sum_{\alpha\beta}^{N} \iota_{\alpha\beta} Q_\alpha^2 Q_\beta^2,$$

where $\kappa_\alpha, \gamma_{\alpha\beta}, \epsilon_{\alpha\beta}$ and $\iota_{\alpha\beta}$ are the expansion coefficients serving as fitting parameters. Not all modes are coupled. For details see Ref. [58].

For this example study, a two-dimensional model is used with the v_1 and v_{36} normal modes as coordinates. The first, v_1, is the imaginary frequency mode, with frequency $\omega_1 = 1516.9674\ i$,cm^{-1}, representing motion of the proton between the oxygen and nitrogen atoms. The second, v_{36}, represents motion of the proton within the plane of the molecule perpendicular to the line joining the oxygen and nitrogen atoms, with frequency ω_{36} =2174.6706 cm^{-1}, The initial coordinates for the centre of the wavepacket were $v_1 = 0.9634$ and $v_{36} = 0.1373$ which represents a stretching

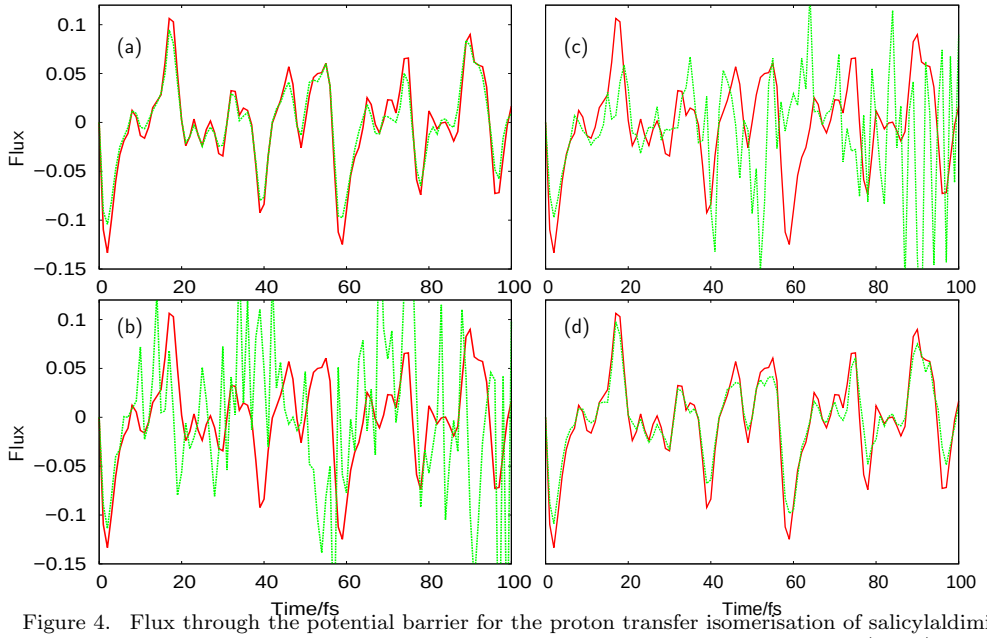

Figure 4. Flux through the potential barrier for the proton transfer isomerisation of salicylaldimine in a 2D model calculated at different levels of approximation for the nuclear wavepacket (green) against the full quantum dynamics result (red). (a) vMCG with 16 GWPs using 4th order integrals. (b) 64 uncoupled (classical) GWPs using adjusted widths. (c) vMCG with 32 GWPs using LHA. (d) vMCG with 32 GWPs using LHA and adjusted widths.

of the O–H bond. All other coordinates were kept fixed at the transition state geometry. The initial energy for this configuration is below the barrier height, thus making tunneling important. This is demonstrated by running a single classical trajectory with these initial conditions along which no crossing of the barrier is seen.

The flux through the barrier can be defined by the expectation value of the operator,

$$\hat{F} = [\Theta_1, \hat{T}_1]\,, \tag{52}$$

the commutator of a Heaviside step function, Θ_1, placed at the barrier along the proton transfer mode v_1, and the kinetic energy operator for this mode, $\hat{T}_1$. The details for calculating wavepacket flux are described in Ref. [6]. Essentially the method provides a measure of the proportion of the wavepacket passing through a dividing surface placed perpendicular to the mode in question. This flux is shown as a function of time for different wavepacket calculations in Fig. 4.

In the first set of calculations vMCG was run with different basis set sizes and integrals calculated to 4th order meaning that the result should converge on the full solution to the TDSE. The width of the initial wavepacket (and all GWP basis functions) was taken from a relaxation calculation to find the ground-state vibrational eigenfunction in the harmonic approximation. For all modes the width was $\frac{1}{\sqrt{2}}$, except for v_1 for which the width 0.57 was used. The result using 16 GWPs is shown in Fig. 4(a) compared to the results from a standard grid-based wavepacket propagation. The flux is small and highly structured, making it a hard property to reproduce, but it can be seen that the vMCG result agrees excellently with the grid-based quantum dynamics.

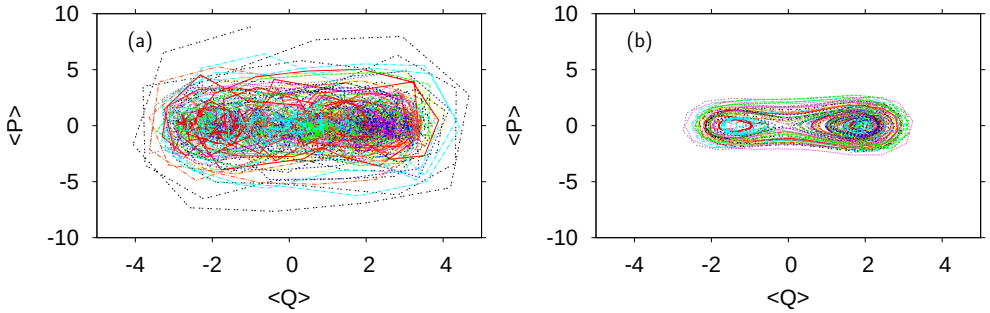

Figure 5. Trajectories followed by the centres of a set of 32 GWPs simulating the proton transfer isomerisation of salicylaldimine in a 2D model plotted in the "phase space" of mode v_1. (a) vMCG. (b) Classical GWPs.

In usual GWP calculations, and in the DD-vMCG, a local harmonic approximation (LHA) is used in which integrals are only made to 2nd order. The resultant flux when using an LHA is shown in 4(c). The difference from the exact result is purely due to the GWPs being too wide for the double-well potential to be described locally by the LHA, and thus introducing an error in the integrals. In 4(d), narrower GWPs with a width of 0.25 for v_1 and 0.65 for v_{36} have been used, and the correct result is obtained. Again, only 32 GWPs were required for a converged result. The question of how to choose the optimum width is still open. Too narrow and many functions are needed, with simulations becoming very noisy as the coupling matrix becomes singular. Too wide and errors occur in the integrals.

In Fig. 4(b) the result for the adjusted widths with no coupling between the basis functions is shown with 64 GWPs. Without coupling the GWPs follow classical trajectories. This is still a full quantum dynamics calculation as the expansion coefficients contain all of the phase information. Indeed the flux, particularly at the start, follows the full result. More functions would be required for convergence.

The reason for the marked difference between the vMCG and classical GWP behaviour can be seen in Fig. 5. This shows the trajectories of the centre of the GWPs in the "phase space" of the proton transfer mode v_1. Plotted is the value of the position, $\langle Q \rangle$, and momentum, $\langle P \rangle$. In Fig. 5(a) the trajectories from 32 GWPs calculated using vMCG and the LHA are shown, while (b) shows the trajectories from 32 uncoupled GWPs that follow classical trajectories. Both calculations started with the same initial conditions.

The classical GWP trajectories show regular periodic motion. The potential along this coordinate is a double well. Some trajectories remain in their starting well, while some have enough energy that they orbit both wells. As basis functions start in both wells, tunneling can be described by the amplitudes of the functions. The vMCG GWP trajectories, however, do not show the same simple structure. The GWPs clearly move in a less regular way and cover a much larger region of phase space. This is a general feature of vMCG GWPs and is the reason for the fast convergence - the functions move to cover the space as required by the TDSE rather than guided by classical mechanics.

4. Direct Dynamics

Having described the methods of using variational Gaussian functions as a basis for nuclear quantum dynamics and given examples of calculations using GWPs on fitted PES, we now turn to perhaps the most promising application of GWP

dynamics, the direct-dynamics version, DD-vMCG[53]. In order to integrate EOMs for the time-dependent coefficients and Gaussian parameters (Eqs. (17) and (21)), the matrix elements of the Hamiltonian need to be evaluated. If the local harmonic approximation (Eq. (26)) is adopted, and dynamics run in the normal mode or Cartesian coordinates, in which the kinetic energy operator has a simple analytic form, all one needs are energies, gradients and Hessians at the centres of Gaussian basis functions at every step. It is straightforward to evaluate these values on-the-fly via an interface with an external quantum chemical software.

In addition to avoiding the pre-calculation of a fitted PES, DD-vMCG also offers an advantage in the form of the kinetic energy operator. When using pre-fitted surfaces, it is necessary to make a careful choice of the particular coordinate system to use for the dynamics in order to avoid spurious correlations between motions and to allow a relatively simple form for the potential. However, this can lead to complicated expressions for the kinetic energy operator. DD-vMCG, on the other hand, has no need to generate a simple form for the PES, therefore coordinates can be used in which the form of the kinetic energy operator is easily evaluated. As such, DD-vMCG studies have been performed using the Cartesian coordinates of atoms [52, 79], Jacobi coordinates [52] and normal modes [51, 54, 80–84]. The major problem with using Cartesian coordinates is that the overall translational and rotational motion of the molecule are not separated from the vibrational motion which is generally of interest in nuclear quantum dynamics[52]. Jacobi coordinates ameliorate this situation by separating out the translations, but representing the initial wavepacket as a superposition of GWPs is not straightforward [79]. Normal mode coordinates go further by approximately removing the complete rotational motion. In the example calculations presented in subsequent sections, normal modes will be used for this reason.

4.1. *The Potential Energy Surface Database (PES-DB)*

As electronic structure calculations are very time-consuming, the re-calculation of energies, let alone gradients and Hessians at every point reached by all of the GWPs is undesirable. As implemented in the MCTDH package, the DD-vMCG algorithm builds up the PES on which the dynamics occurs through the use of the local harmonic approximation (LHA) and the creation of a database of electronic energies and other information [79]. The idea of such a database is intuitive and is based on the GROW philosophy, developed by Collins and co-workers to generate potential surfaces automatically using classical trajectories to sample the configuration space [85, 86]. This procedure has been adapted by Frankcombe to use vMCG functions to sample configurations using "quantum trajectories" [87].

As described above, the nuclear wavepacket is constructed as a linear combination of multi-dimensional GWPs. Each of these is centred at a particular point in configuration space, i.e. at specific molecular geometry. If, as the dynamics proceeds, a GWP reaches a geometry, $\mathbf{x}_0$, which fulfills the conditions required to calculate a new energy (to be discussed further in Sec. 4.4), then the program calls an electronic structure program in order to calculate the electronic energy at that geometry, as well as the gradient and Hessian of the PES. This information is stored in the database. Depending on the problem, further information such as the dipole moments and derivative couplings may also be stored. If a Complete Active Space SCF (CASSCF or CAS) method is used for the PES evaluation, the molecular orbital (MO) coefficients are also stored so that the CAS space can be constructed at any point.

With this data, the PES can be expanded in a Taylor series to second-order

around that geometry:

$$\mathbf{V}(\mathbf{x}) = \mathbf{V}(\mathbf{x}_0) + \mathbf{g}(\mathbf{x}_0) \,.\, (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2} (\mathbf{x} - \mathbf{x}_0) \,.\mathbf{H}(\mathbf{x}_0) \,.\, (\mathbf{x} - \mathbf{x}_0) \tag{53}$$

where $\mathbf{g}(\mathbf{x}_0)$ and $\mathbf{H}(\mathbf{x}_0)$ are the gradient and Hessian of the adiabatic PES with respect to changes in nuclear geometry, both evaluated at $\mathbf{x}_0$. This expansion of the PES is the basis of the LHA which is then used to calculate the necessary matrix elements in the EOMs. The calculation of up to second-derivatives of the PES, however, has a further use.

Where no new database point is required we use modified Shepard interpolation [87] to get the energies, gradients and Hessians. To do so, the Euclidean norm of the difference vector of all atomic coordinates is calculated for the new point and all current database points to give a measure of distance between the structures. The LHA data for the new point is then provided by the Shepard weighted interpolation formula:

$$V(\boldsymbol{q}) = \sum_i \omega_i(\boldsymbol{q}) T_i(\boldsymbol{q}), \tag{54}$$

where

$$\omega_i(\boldsymbol{q}) = \frac{\nu_i(\boldsymbol{q})}{\sum_j \nu_j(\boldsymbol{q})}, \tag{55}$$

and

$$\nu_i(\boldsymbol{q}) = \frac{1}{|\boldsymbol{q} - \boldsymbol{q}_i|^{2p}}, \tag{56}$$

with $p = 2$ showing the best results. T_i is the Taylor series expansion of the energy centred at the i^{th} database entry. As the energy, gradients and Hessian are stored for every entry, the expansion is truncated at second order so that the local harmonic approximation is used once again. The same formula as in Eq. (54) is used for gradients and Hessians, although the Taylor series have to be truncated at the first order for gradients, and at the zeroth order for Hessians. The second-order Taylor series for the PES is then formed, according to equation (53) for each database point, $\mathbf{x}_i$, and the energy evaluated at the new geometry. *Ab-initio* calculation of the Hessians might easily become a bottleneck for direct dynamics, and so it is possible to apply a Hessian update algorithm, such as was done by Frankcombe (see below) [88].

4.2. *Examples*

Having set out the basic methodology involved in carrying out a DD-vMCG calculation, it seems appropriate to provide a few simple examples of calculations using this method to demonstrate the current state-of-the-art. The example systems have been chosen so as to illustrate features of practical calculations which need to be considered when trying to achieve accurate results in a timely fashion.

For the first set of calculations, the system of interest is the proton transfer in the salicylaldimine molecule used in Sec. 3.2. In order to carry out direct-dynamics we need to select an appropriate electronic structure method with which to calculate the PES. In the case of SA, the same level of theory is used to construct the surfaces

above: restricted Hartree-Fock (RHF) with the 3-21G* basis, as implemented in the Gaussian 03 package[78].

Again, we restrict ourselves to a two-dimensional model, using as coordinates the unitless mass-frequency weighted normal modes calculated at the transition state (optimised using RHF/3-21G* in Gaussian 03) between the two forms seen in Fig. 3. The modes chosen are again modes v_1 and v_{36}, the most important for the proton transfer, with v_1 representing motion of the proton between the oxygen and nitrogen atoms and v_{36}, which represents motion of the proton, within the plane of the molecule, perpendicular to the line joining the oxygen and nitrogen atoms. All propagations in this section start with the wavepacket centred at coordinates $v_1 = 0.9634$ and $v_{36} = 0.1373$ (in terms of the mass-frequency weighted coordinates), with a width in each direction 0.5530 and 0.7188 respectively. The individual GWPs have widths of 0.25 and $\frac{1}{\sqrt{2}}$ along the two modes respectively and are spaced at 0.649 intervals. The other coordinates are kept fixed at the enol minimum, providing a slightly different system to the analytic surface propagations which kept the frozen coordinates at their values for the transition state.

In order to assess the quality of the calculations we need to select a property for use as a comparison between calculations. As in Sec. 3.2, the property of interest is the flux of the wavepacket over the barrier between the two PES minima representing the structures in Fig. 3.

4.3. *Convergence with Number of GWPs*

The first set of results presented are for the flux of the wavepacket over the proton transfer barrier, as a function of the increasing number of GWPs used to represent the wavepacket. All calculations were started with an empty database of PES points. The PES was then constructed on-the-fly by evaluating the energy, gradient and Hessian of the molecule whenever the geometry of the centre of the GWP differed by more than 0.1 Bohr from any previously calculated database point (as measured by taking the Euclidean norm of the difference vector between all atomic coordinates). In Fig. 6(a) we show the flux over a total propagation period of 100 fs with an increasing number of GWPs from 8 to 64.

From this plot it is clear that the flux expectation value converges with an increasing number of GWPs. Indeed the qualitative shape of the plot for the largest number of GWPs is reproduced reasonably well by a propagation using only 16 GWPs. We also note that the plot with just 8 GWPs, whilst accurate over the first 10 fs or so, becomes increasingly inaccurate after that point. For example, the pair of peaks at around 20 fs, observed for larger numbers of GWPs, is just a single maximum when using just 8 GWPs. From the peak at around 30 fs to that at 70 fs, the maxima and minima occur earlier in the propagation when using only 8 GWPs, but they are broadly correct. However, after 80 more major errors have begun to impact the dynamics. With 16 or 32 GWPs, the differences are essentially quantitative with the peaks and troughs in the same places.

As described in Sec. 2.5, when using large numbers of GWPs, the problem of linear dependencies within the basis can occur as the GWP basis has no orthogonality condition attached. When this happens, the offending function is removed from the basis set as the propagation proceeds. This was the case when using 32 or 64 GWPs here, and in the case of the 64 GWP propagation this seems to have introduced unwanted fluctuations in the flux after 60 fs.

The results show that for vMCG a larger number of GWPs are required to accurately represent longer time scale dynamics. The reason for this is common to all wavepacket dynamics methods: as the wavepacket proceeds away from its

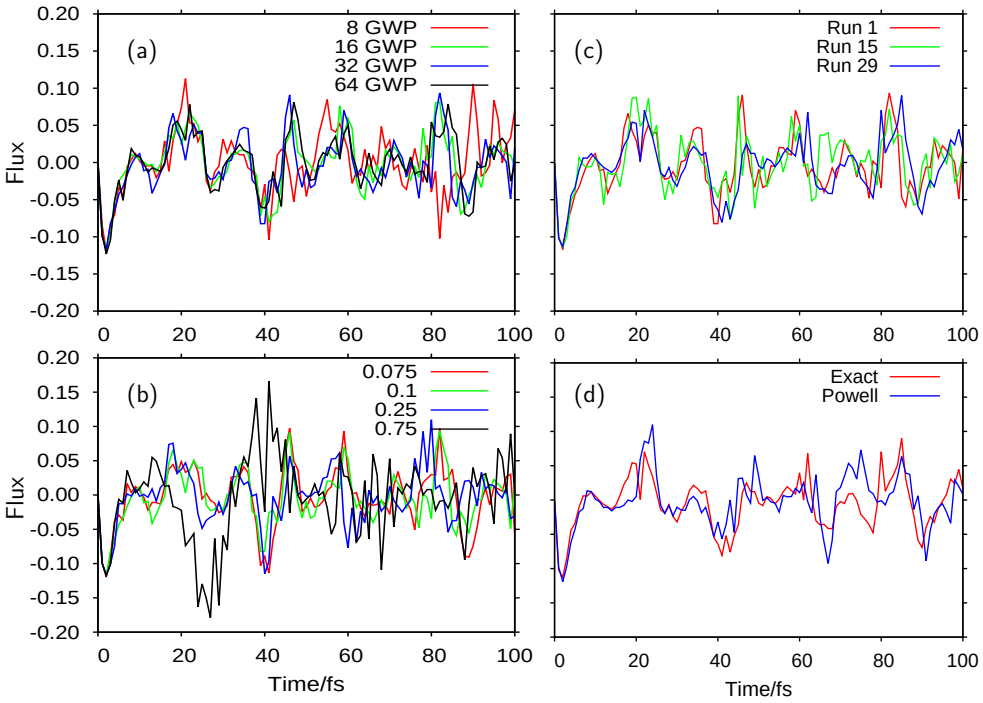

Figure 6. Wavepacket flux along the ν_1-mode in the proton transfer of salicylaldimine (a) as a function of increasing number of GWPs. (b) as a function of the geometry difference parameter, dbmin. (c) as a function of increasing database size. Plots for the initial run, fifteenth run and the final, twenty eighth run (at which point no further database points were added and the flux was invariant) are shown. (d) when using Hessian matrices calculated using Gaussian 03 and when using the Powell update outlined in the text.

relatively localised starting point it can spread, or even bifurcate, as time goes on, and so a larger number of basis functions is required to represent the dynamics.

There is, of course, a downside to using a larger number of GWPs, which is the increased computational effort required. For DD-vMCG, increasing the number of basis functions increases the amount of time needed to propagate the wavepacket simply due to the rise in the number of parameters. For example, with 8 GWPs, the CPU time for the propagation of the wavepacket was just 221 s and for 32 GWPs, it was 7929 s. However, with DD there is also the overhead of the electronic structure calculations being performed, during the propagation, to build up the PES. For SA, each G03 frequency calculation used about 72 s of CPU time, and for the 8 GWP calculation 132 database points were calculated, whereas for the 32 GWP calculation, 785 database points were calculated. In other words, nearly 36 times more effort is required to propagate a wavepacket constructed from 32 basis functions than one made up of 8 functions, but we also need about 6 times the effort to create the PES. In general, the effort in generating the PES will be significantly greater than that used in propagating the wavepacket. However, the additional effort is minimal compared to the work needed to generate points and fit functions to get an analytic PES.

We will return to the issue of the time taken in performing the electronic structure calculations in the next section.

4.4. *Convergence with dbmin*

With the large computational effort required to generate the PES, the second parameter we look at when deciding how to best generate the dynamics of the molecular system is how often to evaluate new electronic structure results. This

can be controlled by altering a parameter called *dbmin* in the MCTDH package.

As the GWPs move with the wavepacket, they explore regions of the configuration space of the molecule. If the molecular geometry represented by the centre of the GWP has not been considered before, we can either perform a weighted expansion of previously calculated energies (using gradients and Hessian in the LHA) to get the potential at that point, or we can simply calculate a new value by performing an electronic structure calculation. Which method to use is determined by dbmin. A measure of the difference between the new geometry and those already in the database is gained by taking the Euclidean norm of the difference vector between all atomic coordinates. If the minimum value of this norm when comparing the new point to all database geometries is less than dbmin, then an expansion of the potential is performed, but if greater then a new point is calculated.

To assess the effect of various dbmin values, calculations have been performed on SA using the same initial conditions as in the previous section, but varying dbmin between 0.075 Bohr and 0.75 Bohr. As a good balance between speed and accuracy, all calculations were performed using 32 GWPs. The flux expectation values as a function of dbmin are presented in Fig. 6(b).

Taking the plot using the lowest value of dbmin to be the reference, as it is produced using the most tightly spaced *ab initio* points, we can compare the convergence of the flux as the dbmin value approaches this ideal. The first point to note is that the plot for a dbmin of 0.75 Bohr, whilst following other plots quite well for the first 10 fs or so of the propagation, diverges soon after. In particular the minimum around 25 fs is significantly underestimated and at later times there is a strong oscillatory behaviour not seen in the other plots. A value of dbmin as large as 0.75 Bohr clearly does not give an accurate description of the PES in this case.

Reducing dbmin to 0.25 Bohr produces a marked improvement in the results, the spurious feature at 25 and 40 fs being replaced with the expected minima, both with essentially the correct magnitudes. The minimum at 25 fs is perhaps a little early in comparison to the reference but is broadly correct. Greater inaccuracies, however, arise at later propagation times. At 60 fs, we see a maximum followed by a minimum which occurs earlier in the propagation than in the reference case. This asynchronicity follows through to the peak at about 80 fs and the minimum at 90 fs. It thus appears that reducing the dbmin value allows the calculation to maintain accuracy for longer, but that eventually the accumulation of errors in the PES, and consequently the dynamics, leads to qualitative inaccuracies.

This problem is repeated when reducing dbmin further to 0.1 Bohr. This plot follows the reference closely for almost the entirety of the propagation: the previous mentioned features occur at the correct times, with the correct magnitudes. Perhaps the only significant inaccuracy is the slightly over deep minimum at 90 fs. Details of the flux, however, are still quite different to the reference, indicating the sensitivity of this property to inaccuracies in the dynamics.

As mentioned above, the reason for the inaccuracy in the propagations using higher values of dbmin is that as time progresses, differences in the PES become more significant and the wavepackets diverge more and more. So, for dbmin=0.075 Bohr, far more database points are calculated (1262) than for dbmin=0.75 Bohr where only 122 points are calculated. In effect, the dynamics for dbmin=0.75 Bohr proceeds on a surface which is 10 times more coarsely defined than for the lowest value of dbmin. This means that far more of the points in configuration space sampled by the GWPs relied on an extrapolated PES than when using the lower value. Also, the extrapolated values of the potential would be calculated using database points, on average, further away from the point in question, leading to

greater inaccuracy as the harmonic assumption behind the extrapolation becomes less valid. As such, the PES for the larger dbmin is much more approximate than that for the lower dbmin, so we would expect, and indeed get, different dynamics. On a more harmonic PES it may be the case that larger values of dbmin are less injurious on the quality of the results than in this case.

This parameter is likely to be quite system dependent and in practice, should be set quite high in test calculations and then reduced in subsequent tests in order to reach convergence of the property of interest. Calculation of new database points is very expensive (here approximately 72 s in each case), so one should not set dbmin too low. In this case a dbmin of 0.1 Bohr was found to be satisfactory. We will proceed on this basis (as we did in the previous section) and use dbmin=0.1 Bohr for all subsequent calculations.

4.5. *Convergence of the Database*

So far the calculations presented here have used DD-vMCG starting from an empty database, i.e. with nothing known about the PES. However, as the database is updated during the progress of the dynamics, it follows that wavepacket motion will change as more points are added. Consider a single GWP which finds its way to the same point in configuration space at two different times in the dynamics, with identical momentum. If this point is near enough to a database point that the potential can be extrapolated, and if new database points have been added in the period between the first and second visits to the point, then the extrapolated potential (and likewise the gradient etc.) will be different, and as such the subsequent motion of the GWP will differ.

Thus to get a true representation of the dynamics, propagations should be carried out using the same initial conditions but using the database created from a previous run. During this run, potential extrapolations are used where possible and new database points are added as needed. The process is repeated until no new database points are added and the dynamics becomes invariant as measured by some property (in our case the flux expectation value). This re-using and adding to a database is a central feature of the DD-vMCG implementation. Electronic structure calculations are only done if required, and the database from a small calculation can be used for a larger calculation, again saving effort.

To demonstrate the database convergence, we have taken the same SA model as above with a 32 GWP basis, and setting dbmin=0.1 Bohr. The first run began with an empty database and generated 785 points as noted above in Sec. 4.3. This database was then used as a starting point for the next calculation, with more points being added as appropriate. We repeated until the flux expectation value in subsequent calculations became invariant. The results are presented in Fig. 6(c) for three of the runs: the initial, the fifteenth and the twenty eighth, at which point convergence was achieved (a twenty ninth run confirmed this).

It is clear from this that there is no qualitative change in the dynamics for the different databases, but there are significant numerical differences, particularly at later times. In terms of the database size, there were 785 entries after the first run and 2066 at convergence. In other words nearly 1300 extra database points were calculated on top of the initial set. So, with regards to the extra time needed to get the database converged, the extra electronic structure calculations required about 92,232 s. The principle additional effort, though, was in the actual propagation part of the calculations. The CPU time for each of the later calculations was significantly greater than for the first run (at 7929 s) with a maximum of 70,634 s for the eleventh run. The minimum time was in fact for the first run, being more than twice as fast as

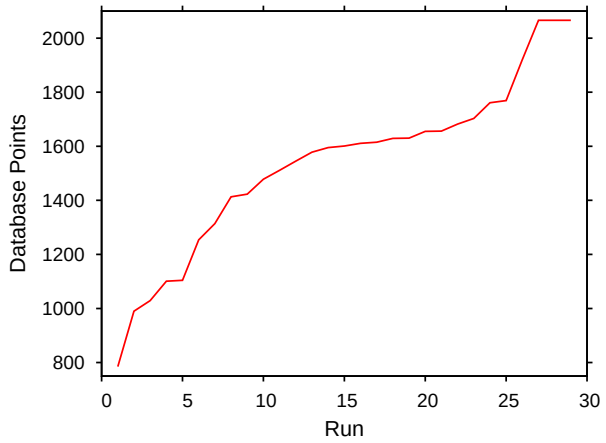

Figure 7. Total number of points in the direct dynamics database as a function of the number of propagations performed, using the database of the previous run as a starting point. Two-mode model of salicylaldimine using a basis of 32 GWPs and a dbmin of 0.1 Bohr.

the next quickest run. The reason for the significantly increased timings is twofold. First, it is found that inexplicably the changes in the PES lead to shorter integration steps, leading to longer propagations. Second, the larger database requires more time to sort through and read, which is presently done at each step.

Overall, another 806,922 s were required to run the propagations, meaning that overall (including electronic structure calculations) an extra 899,154 s were needed to converge the database, after the initial 64,449 s. This is a significant additional cost in time, but for a numerically accurate result the effort is necessary. It is, however, difficult to comment on how much 'effort' the different parts of the propagation take relative to each other. There are three sources of effort: the scaling with system size for the EOMs, the number of integration steps, and the quantum chemistry calculations. Each system will perform differently in these three areas. Here we are using a very cheap quantum chemistry method (RHF/3-21G*), which took only 70 s. The tunneling seems to force small integration steps and thus the time taken for the dynamics dominates.

In Fig. 7 we show the total number of points in the database at the completion of each of the runs. For approximately the first twenty runs, the database appears to be growing fairly smoothly with only a single extra point added during the 21^{st} calculation. It may thus appear to be reasonable to assume that the calculations had converged at this point and stop performing further calculations. However, this would have been an error as we can see that another 151 points in the 26^{th} propagation and a further 146 during the 27^{th}, beyond which no more points were added. This illustrates the point that it is important to continue to run calculations until both no more database points are added and no further change in the expectation value is seen between consecutive runs before convergence can be claimed. This is all the more important for a very sensitive expectation value such as a flux, where large numerical differences between the plots were seen for consecutive runs, even when only a few extra database points were added.

It should be noted that the database is only converged for propagations of 100 fs or less. If we wanted to propagate for a longer time using the final database as a starting point, it would, in general, be necessary to add extra points as different regions of configuration space are explored. This in turn leads to the need to run further calculations to converge that database.

4.6. *Hessian Update vs. Full Hessian*

As noted above, it is necessary to include the potential energy, gradient vector and Hessian matrix in the database so as to be able to use the LHA for the calculation of extrapolated energies, gradients and Hessians, and also for the calculation of matrix elements. However, the evaluation of the molecular Hessian is a very computationally intensive process, even when methods with analytic derivatives are available as for RHF, and even more so for electronic structure methods where numerical differentiation has to be used. As illustrated in Sec. 4.3, the effort in performing a DD-vMCG calculation would be much reduced if we could avoid the explicit evaluation of the Hessian at every database point.

Fortunately there are a wide range of methods for the approximation of Hessian matrices using only gradient information and a reference Hessian. In the MCTDH program a version of the Powell update algorithm[88] has been implemented, which avoids any problems if the Hessian becomes singular. The update is

$$\mathbf{H}_{\mathrm{New}} = \mathbf{H}_{\mathrm{Old}} + \frac{1}{\boldsymbol{\delta}.\boldsymbol{\delta}} \left(\boldsymbol{\epsilon} \otimes \boldsymbol{\delta} + \boldsymbol{\delta} \otimes \boldsymbol{\epsilon} \right) - \frac{\boldsymbol{\epsilon}.\boldsymbol{\delta}}{\left(\boldsymbol{\delta}.\boldsymbol{\delta}\right)^2} \mathbf{H}_{\mathrm{Old}}.\boldsymbol{\delta} \otimes \boldsymbol{\delta}.\mathbf{H}_{\mathrm{Old}} \tag{57}$$

where $\mathbf{H}$ is the Hessian, labeled appropriately, $\boldsymbol{\delta}$ is the position difference vector and $\boldsymbol{\epsilon}$ is the gradient difference vector between the two geometries in question.

Our implementation proceeds as follows when starting with an empty database: The electronic structure calculation for the first GWP is run including the evaluation of the Hessian to give the so-called reference point (if starting with a non-empty database, the first entry is taken as the reference). As the propagation proceeds, extrapolated points are treated as before, but when a new database point is needed, only the energy and gradient are calculated. The program then calculates the distance between the new point and the reference point (as detailed above when dealing with the dbmin parameter). It also calculates the distances between all points in the database and the reference point. The space of database points is then divided into two parts; an internal subset of points which are closer to the reference point than the new point and an external subset which are further away.

Using the gradient at the new point and the gradient and Hessian at one of the points in the internal subset, a Powell update for the Hessian at the new point is calculated, $\mathbf{H}_i$. This is repeated for all points in the internal subset. Subsequently, having calculated the distance between the new point and the internal points, d_i, the Hessian at the new point is given by the following weighted sum:

$$\mathbf{H}_{\mathrm{New}} = \frac{\sum_{i \in \mathrm{Internal}} d_i^{-4} \mathbf{H}_{\mathrm{Old}}^{(i)}}{\sum_{i \in \mathrm{Internal}} d_i^{-4}} \,. \tag{58}$$

As a new Hessian has been added to the database close to the reference point, it is then necessary to update the Hessians at all of the points in the external subset by a similar weighted Powell update (this time including the Hessian at the new point). This ensures that at every step, the Hessian at each point has been generated by extrapolation from all points closer to the reference than itself.

To test the effectiveness of using an approximate Hessian at each database point, we repeated the calculations on salicylaldimine using 32 GWPs and a dbmin of 0.1 Bohr using the Powell update as outlined. The calculations were then repeated until a converged database was achieved. The results for the flux expectation value are presented in Fig. 6(d).

We can see from this plot, that the use of the Hessian update gives a result which is in very good agreement with that obtained when calculating the exact

Hessian at each database point. The results are quantitatively different, but qualitatively correct with nearly all peaks and troughs in the right positions. The main difference is the feature between 70 and 80 fs where a small trough in the full Hessian results has been replaced by a peak in the results using the Hessian update. However, subsequent to this the plots follow one another much more closely until completion of the propagation. Having shown that good results can be obtained using an approximate Hessian, we need to show that the loss of numerical accuracy is compensated for by a gain in calculation speed.

Each electronic structure calculation when calculating the Hessian took 72 s of time. In contrast, when only calculating the gradients each electronic structure calculation only takes about 13 s, a saving of a factor of 5.5 in this particular molecular case. At convergence, the database contained 2066 entries for the exact Hessian and 2263 for the Powell Hessian. Even with the slightly greater number of electronic structure calculations required, the time saving from that portion of the calculations was by a factor of more than five. However, this saving was negated by the longer time that the actual propagations took with the Hessian update method. With the full Hessian, 814851 s of CPU time were required for all of the propagations, but with the Powell updated Hessian 1325401 s were required. This is an increase in time of more than 60 % due to short integrator time steps in the early parts of the propagation, the larger database and the Hessian updating itself.

We also applied the method to a different two-mode model of SA, this time using mode 1 (the motion of the proton between the oxygen and nitrogen atoms) and mode 18 (a ring-breathing type mode). This system was a much easier one on which to achieve database convergence as mode 18 is fairly harmonic, and only eight and seven calculations were required to reach convergence for the full and Powell Hessian methods respectively. With regards to the actual timings, the full Hessian method took a total of 261576 s to reach convergence whilst the Powell Hessian method needed 127706 s (including all electronic structure evaluations). As such, less than half the CPU time was required to complete the propagations when using the updated Hessian.

It would thus seem that the approximate Hessian can provide significant savings in time dependent on the system studied. It is likely to be crucial, though, when using high-level quantum chemistry methods for which the calculation of the Hessian is very time consuming.

4.7. *Multi-State System: Butatriene*

So far the main factors which need to be considered when using DD-vMCG to propagate a wavepacket on a single PES have been considered. However, the MCTDH package also contains the necessary code to allow one to propagate wavepackets on multiple surfaces.

Several prior studies have been carried out on multi-state PES using the DD-vMCG method [54, 80–84, 89, 90] after the initial study by Worth *et al* on the butatriene cation[51]. The current implementation focuses on the use of the CASSCF method as the appropriate electronic structure method to calculate both ground and excited state PES on-the-fly, but there is no reason why any other excited-state electronic structure method could not also be used. The main requirement is a method that treats all states in a balanced way across a range of geometries and can provide non-adiabatic couplings.

Running DD-vMCG on multiple states proceeds in much the same way as on a single PES as far as the generation and use of the database of points is concerned. One addition that has to be made when using CASSCF regards the definition of the

active space used in the calculations. As is well known, the orbitals constituting the active space do not have to be neighbours on the energy scale of the Hartree-Fock MOs, so if this is not to be the case, we must keep track of the orbitals included in the active space. The DD-vMCG program goes about this by saving the MO coefficients in the database in the order they appear in the CASSCF expansion. When a new database point is required, the nearest point in the database is located and its MO coefficients read from the database. These can then be copied into the Gaussian input file and the Gaussian program will use them in that order, giving the correct active space. It should also be noted that state-averaged CASSCF can also be used if required, the correct option being set as appropriate.

As described in Sec. 3.1 for standard vMCG, there are two ways of running DD-vMCG on multiple states, the so-called single-set and multi-set methods. The former relies on using a single set of GWPs for all states, in other words the GWPs on each state have the same positions on those states, whilst in the latter, each PES has a set of GWPs which are allowed to move independently of those on the other states. The single-set method seems more appropriate for direct dynamics calculations as it requires fewer functions overall, and therefore fewer evaluations of the potential surfaces.

When performing dynamics on multiple states care must be taken at geometries where two or more electronic states become very close in energy. This is particularly true around conical intersections where the adiabatic PES calculated by electronic structure calculations become non-differentiable and the non-adiabatic coupling diverges. To circumvent this problem, the dynamics is better carried out on quasi-diabatic states, which are unitary transformations of the adiabatic states to produce smoothly varying PES at these intersections.

Currently implemented in the MCTDH package is a version of the regularisation diabatisation method of Köppel [51, 80, 91–93], which we will briefly describe, following Worth *et al* [51], for completeness. The adiabatic potential energy matrix, $\mathbf{V}$, is related by a similarity transform (defined by the unitary matrix, $\mathbf{U}$) to the diabatic potential matrix, $\mathbf{W}$, i.e.

$$\mathbf{W} = \mathbf{U}^{\dagger}\mathbf{V}\mathbf{U}\,. \tag{59}$$

For a two state system, this can be split into two parts using the sum and difference of the adiabatic energies i.e. defining $\Sigma = 1/2\,(V_{11} + V_{22})$ and $\Delta = 1/2\,(V_{22} - V_{11})$, we have that

$$\mathbf{W} = \Sigma\mathbf{1} + \Delta\mathbf{U}^{\dagger}\,.\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\mathbf{U} \tag{60}$$

Using the invariance of the trace of a matrix under unitary transformation, we get $\Sigma = 1/2\,(W_{11} + W_{22})$, and so

$$\mathbf{W} = \Sigma\mathbf{1} + \Delta\begin{pmatrix} -\Delta_W & W_{12} \\ W_{12} & \Delta_W \end{pmatrix}/\Delta \tag{61}$$

where Δ_W is the diabatic analogue of Δ and W_{12} is the diabatic coupling matrix element. As Δ approaches zero close to a conical intersection, the second term in Eq. (61) becomes singular. By assuming that the adiabatic and diabatic states are equal at the intersection (which is valid as the transformation matrix, $\mathbf{U}$, can be multiplied by an arbitrary matrix which is constant over configuration space), the

matrix in equation (61) can be expanded, to first-order, around the intersection as

$$\mathbf{W}^{(1)} = \begin{pmatrix} -\boldsymbol{\delta} & \boldsymbol{\lambda} \\ \boldsymbol{\lambda} & \boldsymbol{\delta} \end{pmatrix} . \mathbf{Q} \tag{62}$$

where $\mathbf{Q}$ is the displacement vector from the intersection, $\boldsymbol{\delta}$ is the gradient difference vector of the two adiabatic PES and $\boldsymbol{\lambda}$ is the derivative coupling vector, both evaluated at the intersection. Finally, by approximating the difference in the adiabatic energies, Δ, by $\Delta^{(1)} = \sqrt{|\boldsymbol{\delta}|^2 + |\boldsymbol{\lambda}|^2}$ we get the expression for the regularised diabatic states

$$\mathbf{W} = \Sigma \mathbf{1} + \Delta \frac{\mathbf{W}^{(1)}}{\Delta^{(1)}} . \tag{63}$$

As the denominator is positive, we thus avoid the problem of the singularity that was seen in Eq. (61), and the dynamics can proceed on smooth surfaces. By construction, the eigenvalues of the regularised diabatic surfaces reproduce the adiabatic surfaces. The linear approximation to the diabatisation angle, however, introduces an error into the non-adiabatic couplings which are only exact at the reference conical intersection.

The drawback of this method is that it relies on locating a conical intersection prior to performing the DD-vMCG calculation. In effect information about the potential dynamics must be known before carrying out the propagations. This rather goes against the spirit of direct dynamics, so much current work is ongoing to develop methods to avoid this necessity and to calculate diabatic states on-the-fly with no prior knowledge of the PES.

In order to demonstrate the DD-vMCG method on multiple states, we look at a two-mode, two-state model of the butatriene cation. The modes chosen are v_5, the A_u torsion mode, and v_{14}, the totally symmetric carbon-carbon stretching mode. The normal modes and associated frequencies ($767.6\,\mathrm{cm}^{-1}$ and $2196.2\,\mathrm{cm}^{-1}$ respectively) were calculated at the neutral, ground state minimum energy geometry using the CAS(6,6)/3-21G* method in Gaussian 03. The conical intersection at a planar geometry between the two cation states was then located using the state-averaged CAS(5,6) method (with equal weights to each) and the same basis set for use as the reference point for the diabatisation.

DD-vMCG dynamics were then run in the single-set formalism, with the wavepacket starting on the excited cation state at the neutral molecule Franck-Condon point, and allowed to run for 50 fs. The populations of the wavepacket on each state were recorded and plots for the population of the excited state with increasing numbers of GWPs are shown in Fig. 8.

It can be seen from the plot that the same qualitative pattern is given by all calculations - a rapid de-population of the excited state as the conical intersection is encountered, followed by a partial re-population and then repeat. There are, however, quantitative differences. The calculation with only 5 GWPs gives results which are far from smooth, with jagged jumps in population. This is augmented by addition of more GWPs such that the population changes more smoothly although there is some roughness even with 49 GWPs. It can be seen from the trough at about 15 fs, the 13 GWP calculation generally underestimates the population of the excited state when compared to the 25 and 49 GWP calculations. The calculations with the largest numbers of GWPs follow each other quite well and are converging towards one another as would be hoped.

DD-vMCG on multiple states is thus seen to be a viable method, and certainly works well in this case. It is, however, limited by the diabatisation scheme currently

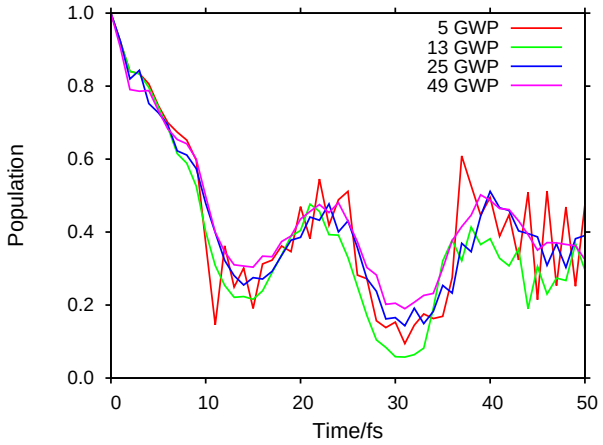

Figure 8. Population of the $\tilde{A}$-state in the butatriene cation over time as a function of increasing number of GWPs.

implemented. In addition to the need to locate any conical intersections before carrying out the dynamics, the method is limited to two states only. More general, "black-box", diabatisation schemes will need to be developed and implemented in order to fully exploit the potential of the method.

4.8. *Relaxation*

In addition to DD-vMCG propagation on one or two electronic states within the MCTDH package, it is possible to use imaginary time relaxation [6] within the DD-vMCG framework as a method of locating a minimum energy nuclear eigenfunction on a particular PES. The general idea behind the method is to take the general expression for a wavepacket

$$\psi(t) = \sum_j c_j \exp(-iE_j t)\,\phi_j \tag{64}$$

where c_j and E_j are the coefficient and energy associated with the eigenfunctions, ϕ_j, of the time-independent Hamiltonian. Making the replacement $t \to \tau = it$, and rescaling the energies so that the minimum value is 0, propagation takes place in τ. As the propagation proceeds, the higher energy components decay exponentially until a wavepacket consisting of just the lowest energy eigenfunction is left.

The implementation within DD-vMCG takes a similar form to that with the propagation in real time, with the PES created by electronic structure calculations on-the-fly. To demonstrate this method, we return to the two-dimensional model of salicylaldimine used in the first few sections of this discussion of DD-vMCG. Taking the same initial conditions as in section 4.3, we allow the wavepacket to relax for 100 fs. As an example, the expectation value of the position operator along mode ν_1 is presented in Fig. 9 for an increasing number of GWPs from 8 to 64.

As is apparent from this plot, the relaxation of the wavepacket occurs successfully for all numbers of GWPs, with all expectation values having reached their final values within 50 fs. It is also clear that the relaxations are converging towards a position expectation value of between 1.4 and 1.5 (the 64 GWP calculation reaches 1.434) with all calculations ending up in that range. It should also be noted that the position expectation value along the other mode, as well as the wavepacket widths along both modes, also converge in much the same way, meaning that the wavepacket has fully relaxed after 100 fs.

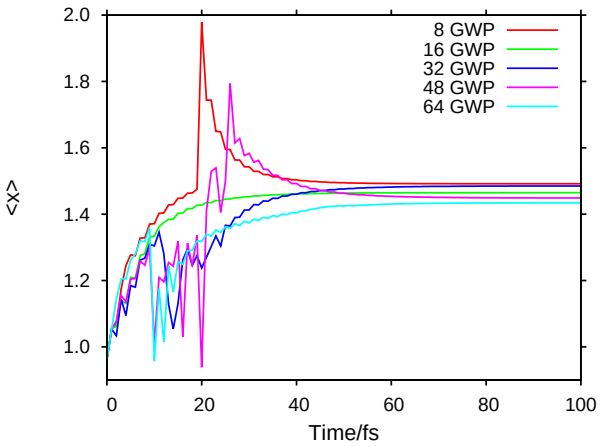

Figure 9. Position expectation value along the v_1 mode of a relaxing wavepacket constructed from an increasing number of GWPs.

With regards to computational effort for a DD-vMCG relaxation, this is much reduced when compared to a propagation over the same timescale. Taking the 32 GWP calculation as an example, we recall that a total of 64449 s of CPU time was required to propagate the wavepacket for 100 fs, generating in the process 785 database points. To relax the wavepacket for 100 fs, the actual relaxation took just 421 s, and 44 database points were created. As such, about 3589 s of CPU time was needed, a reduction in effort over the propagation by a factor of nearly 18. Clearly, a major component of this is the 18 times reduction in the number of database points. However, the actual dynamics component is reduced by a factor of almost 19. The reduction in database size is easily rationalised by considering the nature of the two processes being studied. For relaxation, the wavepacket simply moves to the lowest energy configuration, all GWPs moving down in energy, then stays put. Only a very small portion of configuration space needs to be sampled when compared to the much more unpredictable real time dynamics. The smaller database also reduces the computational effort for the dynamics itself, as much less effort has to be put into extrapolating energies for intermediate points. The major saving in the timing for the dynamics is in the integration of the equations of motion: much longer time steps are possible during relaxation, being in the range of 0.01-0.23 fs rather than 0.001-0.04 fs, as near convergence the wavepacket is stationary.

5. Properties of vMCG Method

Much of this review has been concerning the convergence properties of the vMCG and DD-vMCG methods. The reasoning behind this is that for a class of methods to be considered reliable it is necessary to be confident that, were unlimited resources available, they converge to the "correct" answer.

The dynamics generated by the MCTDH method converge towards the result given by a full expansion of the nuclear wavepacket on a grid [41, 42]. Thus if vMCG converges on the MCTDH result, vMCG will converge on the exact result. This was demonstrated in Sec. 3.1 where the dynamics of a wavepacket constructed from 60 GWPs was found to converge to the MCTDH result, both using the same PES. That using more than twice as many classical GWPs yields an inaccurate result indicates the importance of using the variationally coupled vMCG basis functions.

It was noted that the vMCG results were obtained after much longer calculation times than those obtained using MCTDH. However, vMCG, should become more efficient for larger systems as the scaling with system size is lower than that for MCTDH. It is unfortunately not possible to provide a scaling factor. The bottleneck in vMCG propagation is the time to invert the C-matrix that couples the GWPs, and this matrix has dimensions $(n_{\mathrm{gwp}} \times n_{\mathrm{dof}})^2$, where n_{gwp} and n_{dof} are the numbers of GWPs and degrees of freedom respectively. The number n_{gwp} is highly system dependent and so there is no clear scaling with system size. It should be noted though that it will be less than exponential.

The vMCG method should also provide significant time savings when applied in the context of direct dynamics. The vast majority of effort in performing quantum dynamics calculations is in the generation of the PES on which the nuclear wavepacket moves, both in terms of the number of electronic structure calculations needed and, most significantly, the time needed to fit analytic functions to those points. The ability to successfully calculate the PES on-the-fly should save a significant amount of effort overall.

The second half of this work demonstrated the convergence properties of and effort needed by the DD-vMCG method. In Secs. 4.3-4.6 we were concerned with the proton transfer dynamics of salicylaldimine, in particular the flux of the wavepacket through the barrier dividing the two isomers. These results should be compared to those in Sec. 3.2 which used the MCTDH and vMCG methods on fitted surfaces for the same system. Comparison of the plots in Figures 4 and 6 indicate that DD-vMCG generates a flux in good agreement with the other methods. Exact agreement should not be expected as the PES are formed differently, with a best fit function for vMCG and MCTDH and interpolations of ab initio data for the direct dynamics.

Sec. 4.3 demonstrated that, as with vMCG on a fitted surface, DD-vMCG converges with respect to the addition of extra GWP basis functions. A good result was obtained using 32 GWPs in just under 18 hours of CPU time (propagation and electronic structure combined). Indeed qualitatively good results could be obtained using 8 GWPs in less than $2\frac{3}{4}$ hours. This is significantly less effort than that required to calculate a PES and then run dynamics.

In Sec. 4.4 t was shown that using a more finely divided PES, in terms of the proximity of the geometries at which electronic structure calculations were performed, allowed more accurate representation of the results, at the cost of increased computational effort. In Sec. 4.5 we demonstrated that it is possible to saturate the database of points from which the PES is generated, and as such get a result which is fully converged in terms of the fact that subsequent runs, using the same initial conditions, will yield an identical result. It is, however, worth noting that the result is qualitatively correct after only one run, and that only quantitative corrections are made during the subsequent calculations. The effort to converge the database is considerable, both in terms of the extra electronic structure calculations and the longer propagation times which result.

In Sec. 4.6 it was shown that a good result can also be obtained using an approximate Hessian, with a large potential saving in time for the quantum chemistry calculations due to no longer needing to perform the expensive evaluation of the exact Hessian. It was explained further in, Sec. 4.8 that, not only does standard wavepacket propagation converge for DD-vMCG, but also the location of PES minima by the imaginary time propagation. This is a way of finding the global minimum on a potential surface.

Finally, in Sec. 4.7 it was shown that the principles outlined for SA on the ground state PES also hold in the case of the non-adiabatic dynamics of the butatriene

cation. DD-vMCG is able to reproduce the transfer of wavepacket population between states, the result converging with respect to the number of GWPs. As such, once more general diabatisation schemes have been implemented, it should be possible to accurately treat the dynamics of any reasonable molecular system on any number of electronic states.

6. Conclusion

Quantum dynamics simulations using Gaussian wavepackets are becoming an important tool in studying chemical reactivity. GWPs have properties that promise to describe quantum effects while breaking the exponential scaling with system size of grid-based methods. In recent years a variety of methods have been proposed as full solutions to the TDSE. Most of the recent development has been in algorithms that can be applied to non-adiabatic photochemistry, where the dominant quantum effects are in electronic state population transfer. Spawning and MCE are probably the main two examples. Spawning in particular is a fairly mature method proven to provide data relevant to the interpretation of time-resolved experiments. Fewer fully quantum GWP methods have been developed to treat general scattering problems. The CCS method, the single-state version of MCE, is one example.

Efficiency and convergence properties are the main differences between methods. Like many GWP based methods, Spawning, CCS and MCE all use GWPs that follow classical trajectories. These provide a reasonable basis, but it is not optimal. There is only a guarantee that the result will converge on the full quantum dynamics result with a complete basis set, but no guarantee that this can be obtained. In particular, the classical basis means that basis functions do not tunnel through barriers so not all regions of configuration space may be sampled.

In addition to this sampling problem, GWPs suffer from the inherent problem that they provide a non-orthonormal basis set. This can lead to numerical instabilities, particularly when the basis set approaches completeness. Other methods, such as matching pursuit and basis expansion leaping, aim to solve these numerical problems. But the cost is a much expanded basis set.

As with all quantum dynamics methods, a problem is how to obtain suitable potential surfaces for polyatomic systems. Direct dynamics methods are becoming increasingly popular as a way of circumventing the need to calculate surfaces *a priori*. These use quantum chemistry programs to provide the potential surfaces on-the-fly as and when required by the evolving basis functions. GWPs are ideal for direct dynamics due to their localised nature which means only local information about a surface is needed. Direct dynamics versions of spawning (AIMS) and MCE have been developed and used successfully in simulations of photochemical processes.

The vMCG method, the subject of this review, is a fully variational solution to the TDSE based on a time-dependent Gaussian Wavepacket basis. It is completely general, directly applicable to all types of problems and the variational basis leads to good convergence. Very good results have been shown above for both tunneling and non-adiabatic systems, with only a few GWPs capturing the major physics of these processes. There is still much work to be done to make the method efficient enough to study larger molecular systems, but the present results are promising.

Direct dynamics is also a trivial extension. The use of a DB allows real potential surfaces to be built up. The data so gathered is a valuable resource that can be used to generate more accurate surfaces, only collecting data where it is needed. This could also have great impact in grid-based quantum dynamics when multi-

dimensional potential surfaces are required. While one cannot afford to do a full global fit in 30 dimensions or more, it may be possible to accurately fit the relevant parts from data acquired by direct dynamics. The good convergence of vMCG really does pay dividends in a direct dynamic form, as the quantum chemistry requirements outweigh the dynamics so that saving on functions is the crucial part.

The algorithm is now at an interesting stage. The numerical problems seem to be understood and contained and work can now be done on improving the efficiency. The ML-vMCG algorithm should bring big advantages, analogous to the huge step made possible by ML-MCTDH for the grid-based method. The results shown in this review are good evidence that the vMCG method has a future in developing quantum dynamics to become a general tool.

7. Acknowledgments

We would like to acknowledge all the people who have been involved in aspects of the development and implementation of the vMCG and DD-vMCG methods. These include Charlotte Allan, Mike Bearpark, Terry Frankcombe, Kousik Giri, Werner Koch, João Malhado, David Mendive-Tapia, Mike Robb, Sarah Römer and Matthias Ruckenbauer, all of whom have worked on aspects of the theory and / or implementation. The work has been funded mostly by the EPSRC and the Leverhulme Trust. Iakov Polyak is supported by a Marie-Curie Fellowship.

References

[1] E.J. Heller, *Time-dependent approach to semiclassical dynamics*, J. Chem. Phys. 62(4) (1975), pp. 1544–1555.
[2] E.J. Heller, *Frozen Gaussians: A very simple semiclassical approximation*, J. Chem. Phys. 75 (1981), pp. 2923–2931.
[3] R. Kosloff, *Time-dependent quantum-mechanical methods for molecular dynamics.*, J. Phys. Chem. 92 (1988), pp. 2087–2100.
[4] C. Leforestier et al., *A comparison of different propagation schemes for the time dependent Schrödinger equation.*, J. Comp. Phys. 94 (1991), pp. 59–80.
[5] R. Kosloff, *Propagation Methods for Quantum Molecular Dynamics*, Annu. Rev. Phys. Chem. 45(1) (1994), pp. 145–178.
[6] M.H. Beck et al., *The multiconfiguration time-dependent Hartree method: A highly efficient algorithm for propagating wavepackets.*, Phys. Rep. 324 (2000), pp. 1–105.
[7] E. McCullough and R.E. Wyatt, *Dynamics of the Collinear H+H_2 Reaction. I. Probability Density and Flux*, J. Chem. Phys 54 (1971), pp. 3578–3591.
[8] J.V. Lill, G.A. Parker, and J.C. Light, *The discrete variablefinite basis approach to quantum scattering*, J. Chem. Phys 85 (1986), p. 900.
[9] J. Zhang *Theory and Application of Quantum Molecular Dynamics*, World Scientific, Singapore, 1999.
[10] A.H. Zewail, *Femtochemistry: Atomic-Scale Dynamics of the Chemical Bond* , J. Phys. Chem. A 104(24) (2000), pp. 5660–5694.
[11] P.S. Zuchowski and J.M. Hutson, *Low-energy collisions of NH_3 and ND_3 with ultracold Rb atoms*, Phys. Rev. A 79 (2009), p. 062708.
[12] F. Lique et al., *Non-adiabatic coupling and resonances in the F + H_2 reaction at low energies*, J. Chem. Phys. 134 (2011), p. 231101.
[13] K.C. Kulander (ed.) *Time-dependent methods for quantum dynamics*, Elsevier, Amsterdam, 1991.
[14] H.D. Meyer, F. Gatti, and G.A. Worth (eds.) *High dimensional quantum dynamics: Basic Theory, Extensions, and Applications of the MCTDH method*, VCH, Weinheim, Germany, 2008.
[15] F. Gatti *Molecular Quantum Dynamics*, Springer, Heidelberg, Germany, 2014.
[16] S.Y. Lee and E.J. Heller, *Exact time-dependent wave packet propagation: Application to the photodissociation of methyl iodide*, J. Chem. Phys. 76(6) (1982), pp. 3035–3044.
[17] M. Sepúlveda, S. Tomsovic, and E. Heller, *Semiclassical propagation: How long can it last?*, Phys. Rev. Lett. 69(3) (1992), pp. 402–405.
[18] E.J. Heller, *Cellular dynamics: A new semiclassical approach to time-dependent quantum mechanics*, J. Chem. Phys. 94 (1991), pp. 2723–2729.
[19] D. Huber and E.J. Heller, *Generalized Gaussian wave packet dynamics*, J. Chem. Phys. 87 (1987), pp. 5302–5311.
[20] S. Sawada et al., *A strategy for time dependent quantum mechanical calculations using a Gaussian wave packet representation of the wave function*, J. Chem. Phys. 83 (1985), pp. 3009–3027.

[21] S.I. Sawada and H. Methiu, *A Gaussian wavepacket method for studying time dependent quantum mechanics in a curve-crossing system: low energy motion, tunneling, and thermal dissipation*, J. Chem. Phys. 84 (1986), pp. 6293–6311.
[22] S. Sawada and H. Methiu, *A multiple trajectory theory for curve crossing problems obtained by using a Gaussian wave packet representation of the nuclear motion*, J. Chem. Phys. 84 (1986), pp. 227–238.
[23] T.J. Martínez, M. Ben-Nun, and R.D. Levine, *Multi-Electronic-State Molecular Dynamics: A Wave Function Approach with Applications*, J. Phys. Chem. 100 (1996), pp. 7884–7895.
[24] M. Ben-Nun and T.J. Martínez, *Nonadiabatic molecular dynamics: Validation of the multiple spawning method for a multidimensional problem*, J. Chem. Phys. 108 (1998), pp. 7244–7257.
[25] T.J. Martínez and R.D. Levine, *First-principles molecular dynamics on multiple electronic states: A case study of NaI*, J. Chem. Phys. 105(15) (1996), pp. 6334–6341.
[26] M. Ben-Nun and T.J. Martínez, Ab Initio *Quantum Molecular Dynamics*, Adv. Chem. Phys. 121 (2002), pp. 439–512.
[27] H.R. Hudock et al., *Ab initio molecular dynamics and time-resolved photoelectron spectroscopy of electronically excited uracil and thymine*, J. Phys. Chem. A 111(1) (2007), pp. 8500–8508.
[28] T. Mori et al., *Role of Rydberg States in the Photochemical Dynamics of Ethylene*, J. Phys. Chem. A 116 (2012), pp. 2808–2818.
[29] C. Ko et al., *Ab initio excited-state dynamics of the photoactive yellow protein chromophore*, J. Am. Chem. Soc. 125 (2003), pp. 12710–12711.
[30] D.V. Shalashilin and M.S. Child, *The phase space CCS approach to quantum and semiclassical molecular dynamics for high-dimensional systems*, Chem. Phys. 304(1-2) (2004), pp. 103–120.
[31] D.V. Shalashilin, *Quantum mechanics with the basis set guided by Ehrenfest trajectories: theory and application to spin-boson model.*, J. Chem. Phys. 130(24) (2009), pp. 244101–11.
[32] K. Saita and D.V. Shalashilin, *On-the-fly ab initio molecular dynamics with multiconfigurational Ehrenfest method.*, J. Chem. Phys. 137(22) (2012), pp. 22A506–8.
[33] K. Saita, M.G.D. Nix, and D.V. Shalashilin, *Simulation of ultrafast photodynamics of pyrrole with a multiconfigurational Ehrenfest method.*, PCCP 15(38) (2013), pp. 16227–35.
[34] D.V. Makhov et al., *Ab initio multiple cloning algorithm for quantum nonadiabatic molecular dynamics*, J. Chem. Phys. 141(5) (2014), pp. 054110–11.
[35] Y. Wu and V.S. Batista, *Matching-pursuit for simulations of quantum processes*, J. Chem. Phys. 118 (2003), pp. 6720–6724.
[36] Y. Wu and V.S. Batista, *Matching-pursuit split-operator Fourier-transform simulations of excited-state intramolecular proton transfer in 2-(2'-hydroxyphenyl)-oxazole*, J. Chem. Phys. 124 (2006), pp. 224305–8.
[37] X. Chen and V.S. Batista, *Matching-pursuit/split-operator-Fourier-transform simulations of excited-state nonadiabatic quantum dynamics in pyrazine*, J. Chem. Phys. 125(12) (2006), p. 124313.
[38] S. Habershon, *Linear dependence and energy conservation in Gaussian wavepacket basis sets*, J. Chem. Phys. 136(1) (2012), pp. 014109–8.
[39] W. Koch and T.J. Frankcombe, *Basis expansion leaping: A new method to solve the time-dependent schrödinger equation for molecular quantum dynamics*, Phys. Rev. Lett. 110 (2013), pp. 263202–5.
[40] H.D. Meyer, U. Manthe, and L.S. Cederbaum, *The Multi-Configurational Time-Dependent Hartree Approach*, Chem. Phys. Lett. 165 (1990), pp. 73–78.
[41] G.A. Worth et al., *Using the MCTDH wavepacket propagation method to describe multimode non-adiabatic dynamics*, Int. Rev. Phys. Chem. 27(3) (2008), pp. 569–606.
[42] O. Vendrell, F. Gatti, and H.D. Meyer, *Full dimensional (15D) quantum-dynamical simulation of the protonated water dimer II: Infrared spectrum and vibrational dynamics*, J. Chem. Phys. 127 (2007), pp. 184303–10.
[43] H. Wang and M. Thoss, *Multilayer formulation of the multiconfiguration time-dependent Hartree theory*, J. Chem. Phys. 119 (2003), pp. 1289–1299.
[44] I. Kondov, M. Thoss, and H. Wang, *Quantum dynamics of photoinduced electron transfer reactions in dye-semiconductor systems: Description and application to coumarin 343-TiO_2*, J. Phys. Chem. C 111 (2007), pp. 11970–11981.
[45] I.R. Craig, H. Wang, and M. Thoss, *Proton transfer reactions in model condensed-phase environments: Accurate quantum dynamics using the multilayer multiconfiguration time-dependent Hartree approach*, J. Chem. Phys. 127 (2007), pp. 144503–13.
[46] I. Burghardt, H.D. Meyer, and L.S. Cederbaum, *Approaches to the approximate treatment of complex molecular systems by the multiconfiguration time-dependent Hartree method*, J. Chem. Phys. 111(7) (1999), pp. 2927–2938.
[47] I. Burghardt, M. Nest, and G.A. Worth, *Multiconfigurational system-bath dynamics using gaussian wavepackets: Energy relaxation and decoherence induced by a finite-dimensional bath*, J. Chem. Phys. 119 (2003), pp. 5364–5378.
[48] I. Burghardt, K. Giri, and G.A. Worth, *Multimode quantum dynamics using Gaussian wavepackets: The Gaussian-based multiconfiguration time-dependent Hartree (G-MCTDH) method applied to the absorption spectrum of pyrazine.*, J. Chem. Phys. 129(17) (2008), pp. 174104–14.
[49] R. Martinazzo et al., *A local coherent-state approximation to system-bath quantum dynamics*, J. Chem. Phys. 125 (2006), pp. 194102–16.
[50] G.A. Worth and I. Burghardt, *Full quantum mechanical molecular dynamics using Gaussian wavepackets*, Chem. Phys. Lett. 368 (2003), pp. 502–508.
[51] G.A. Worth, M.A. Robb, and I. Burghardt, *A novel algorithm for non-adiabatic direct dynamics using variational Gaussian wavepackets*, Faraday Discuss. 127 (2004), pp. 307–323.
[52] B. Lasorne, M. Robb, and G. Worth, *Direct quantum dynamics using variational multi-configuration Gaussian wavepackets. Implementation details and test case.*, Phys. Chem. Chem. Phys. 9 (2007), pp. 3210–3227.
[53] G.A. Worth, M.A. Robb, and B.L. Lasorne, *Solving the time-dependent Schrödinger equation for nuclear motion in one step: direct dynamics of non-adiabatic systems*, Mol. Phys. 106(16-18) (2008),

pp. 2077–2091.
[54] D. Mendive-Tapia et al., *Towards converging non-adiabatic direct dynamics calculations using frozen-width variational Gaussian product basis functions.*, J. Chem. Phys. 548(22) (2012), pp. 22A548–10.
[55] P.A.M. Dirac, *Note on exchange phenomena in the Thomas atom*, Proc. Cambridge Philos. Soc. 26 (1930), pp. 376–385.
[56] J. Frenkel *Wave Mechanics*, Clarendon Press, Oxford, U.K., 1934.
[57] G.A. Worth and I. Burghardt Unpublished Work.
[58] I. Polyak, C. Allan, and G.A. Worth, *An accurate description of tunneling using direct quantum dynamics simulation*, J. Chem. Phys. xx (2015), p. xxx Submitted.
[59] S. Römer and I. Burghardt, *Towards a variational formulation of mixed quantum-classical molecular dynamics*, Mol. Phys. 111 (2013), pp. 3618–3624.
[60] I. Burghardt, K. Giri, and G.A. Worth, *Multimode quantum dynamics using Gaussian wavepackets: The Gaussian-based multiconfiguration time-dependent Hartree (G-MCTDH) method applied to the absorption spectrum of pyrazine*, J. Chem. Phys. 129(17) (2008).
[61] D.V. Shalashilin and I. Burghardt, *Gaussian-based techniques for quantum propagation from the time-dependent variational principle: Formulation in terms of trajectories of coupled classical and quantum variables.*, J. Chem. Phys. 129(8) (2008), pp. 084104–9.
[62] S. Sawada et al., *A strategy for time-dependent quantum-mechanical calculations using a gaussian wave packet representation of the wave-function*, J. Chem. Phys. 83(6) (1985), pp. 3009–3027.
[63] R. Heather and H. Metiu, *A numerical study of the multiple gaussian representation of time-dependent wave-functions of a morse oscillator*, J. Chem. Phys. 84(6) (1986), pp. 3250–3259.
[64] A.N. Tikhonov, *On the Stability of Inverse Problems*, Dokl. Akad. Nauk SSSR 39(5) (1943), pp. 195–198.
[65] S. Romer, M. Ruckenbauer, and I. Burghardt, *Gaussian-based multiconfiguration time-dependent Hartree: A two-layer approach. I. Theory.*, J. Chem. Phys. 138 (2013).
[66] C. Woywod et al., *Characterization of the S_1–S_2 conical intersection in pyrazine using ab initio multiconfiguration self-consistent-field and multireference configuration-interaction methods.*, J. Chem. Phys. 100 (1994), pp. 1400–1413.
[67] A. Raab et al., *Molecular Dynamics Of Pyrazine After Excitation To The S_2 Electronic State Using A Realistic 24-mode Model Hamiltonian.*, J. Chem. Phys. 110 (1999), pp. 936–946.
[68] M. Petkovic and O. Kuhn, *Multidimensional hydrogen bond dynamics in salicylaldimine: Coherent nuclear wave packet motion versus intramolecular vibrational energy redistribution*, J. Phys. Chem. A 107(41) (2003), pp. 8458–8466.
[69] J.M. Ortiz-Sanchez et al., *Theoretical Study on the Excited-State Intramolecular Proton Transfer in the Aromatic Schiff Base Salicylidene Methylamine: an Electronic Structure and Quantum Dynamical Approach*, J. Phys. Chem. A 110 (2006), pp. 4649–4656.
[70] J. Jankowska et al., *Photophysics of Schiff Bases: Theoretical Study of Salicylidene Methylamine*, ChemPhysChem. 13(18) (2012), pp. 4287–4294.
[71] J. Jankowska et al., *Excited-State Intramolecular Proton Transfer: Photoswitching in Salicylidene Methylamine Derivatives*, ChemPhysChem. 15 (2014), pp. 1643–1652.
[72] H. Lee and T. Kitagawa, *Tautomeric Equilibrium of Salicylidene Schiff-Base - UV-Visible Absorption and Raman-Spectroscopic Studies*, Bull. Chem. Soc. Jpn. 59(9) (1986), pp. 2897–2898.
[73] K. Kownacki et al., *Laser-Induced Absorption and Fluorescence Studies of Photochromic Schiff-Bases*, Chem. Phys. Lett. 227(3) (1994), pp. 270–276.
[74] A. Grabowska, K. Kownacki, and L. Kaczmarek, *Structural aspects of photochromism of the internally H-bonded Schiff bases. A minimal chromophore*, Acta. Phys. Pol. A 88(6) (1995), pp. 1081–1088.
[75] C. Okabe et al., *Ultrafast excited-state dynamics in photochromic N-salicylideneaniline studied by femtosecond time-resolved REMPI spectroscopy*, J. Chem. Phys. 121(19) (2004), pp. 9436–9442.
[76] J. Grzegorzek, A. Filarowski, and Z. Mielke, *The photoinduced isomerization and its implication in the photo-dynamical processes in two simple Schiff bases isolated in solid argon*, Phys. Chem. Chem. Phys. 13(37) (2011), pp. 16596–16605.
[77] A. Staykov et al., *Photoswitching of Conductance through Salicylidene Methylamine*, J. Phys. Chem. C 118(47) (2014), pp. 27539–27548.
[78] M.J. Frisch et al., Gaussian 03, Revision C.02; .
[79] B. Lasorne et al., *Direct quantum dynamics using variational multi-configuration Gaussian wavepackets*, Phys. Chem. Chem. Phys. 432 (2006), pp. 604–609.
[80] C.S.M. Allan et al., *A straightforward method of analysis for direct quantum dynamics: application to the photochemistry of a model cyanine.*, J. Phys. Chem. A 114(33) (2010), pp. 8713–8729.
[81] D. Asturiol et al., *Exploring the sloped-to-peaked S_2/S_1 seam of intersection of thymine with electronic structure and direct quantum dynamics calculations*, Phys. Chem. Chem. Phys. 12(19) (2010), pp. 4949–4958.
[82] D. Mendive-Tapia et al., *Controlling the Mechanism of Fulvene S_1/S_0 Decay: Switching Off the Stepwise Population Transfer*, Phys. Chem. Chem. Phys. 12 (2010), pp. 15725–15733.
[83] M. Araújo et al., *The Molecular Dissociation of Formaldehyde at Medium Photoexcitation Energies: a Quantum Chemistry and Direct Quantum Dynamics Study*, J. Chem. Phys. 131(14) (2009), pp. 144301–144308.
[84] B. Lasorne et al., *Controlling S_1/S_0 Decay and the Balance between Photochemistry and Photostability in Benzene: A Direct Quantum Dynamics Study*, J. Phys. Chem. A 112 (2008), pp. 13017–13027.
[85] J. Ischtwan and M.A. Collins, *Molecular potential energy surfaces by interpolation*, J. Chem. Phys. 100 (1994), pp. 8080–8088.
[86] M.A. Collins and D.H. Zhang, *Application of interpolated potential energy surfaces to quantum reactive scattering*, J. Chem. Phys. 111 (1999), pp. 9924–9931.
[87] T.J. Frankcombe, M.A. Collins, and G.A. Worth, *Converged quantum dynamics with modied Shepard interpolation and Gaussian wave packets*, Chem. Phys. Lett. 489(4-6) (2010), pp. 242–247.
[88] T.J. Frankcombe, *Using Hessian update formulae to construct modified Shepard interpolated potential*

energy surfaces: application to vibrating surface atoms, J. Chem. Phys. 140(11) (2014), pp. 114108–6.

[89] B. Lasorne et al., *Automatic generation of active coordinates for quantum dynamics calculations: Application to the dynamics of benzene photochemistry*, J. Chem. Phys. 128(12) (2008), pp. 124307–124310.

[90] B. Lasorne et al., *Controlling Product Selection in the Photodissociation of Formaldehyde: Direct Quantum Dynamics from the S_1 Barrier*, J. Phys. Chem. A 7 (2010), pp. 12016–12020.

[91] A. Thiel and H. Köppel, *Proposal and numerical test of a simple diabatization scheme*, J. Chem. Phys. 110(19) (1999), pp. 9371–9383.

[92] J. Gronki, S. Mahapatra, and H. Köppel, *Construction scheme for regularized diabatic states*, J. Chem. Phys. 115(6) (2001), pp. 2377–2388.

[93] H. Köppel and B. Schubert, *The concept of regularized diabatic states for a general conical intersection*, Mol. Phys. 104 (2006), pp. 1069–1079.