Improved algorithm for the direct dynamics variational multi-configurational Gaussian method

Georgia Christopoulou, Antonia Freibert and Graham A. Worth

COLLECTIONS

View Online   Export Citation   CrossMark

ARTICLES YOU MAY BE INTERESTED IN

Improved algorithm for the direct dynamics variational multi-configurational Gaussian method

View Online   Export Citation   CrossMark

Georgia Christopoulou, [iD] Antonia Freibert, [iD] and Graham A. Worth[a) [iD]

AFFILIATIONS

Department of Chemistry, University College London, London WC1H 0AJ, United Kingdom

Note: This paper is part of the JCP Special Topic on Quantum Dynamics with *Ab Initio* Potentials.
[a)]Author to whom correspondence should be addressed: g.a.worth@ucl.ac.uk

ABSTRACT

The Direct Dynamics variational Multi-Configurational Gaussian (DD-vMCG) method provides a fully quantum mechanical solution to the time-dependent Schrödinger equation for the time evolution of nuclei with potential surfaces calculated *on-the-fly* using a quantum chemistry program. Initial studies have shown its potential for flexible and accurate simulations of non-adiabatic excited-state molecular dynamics. In this paper, we present developments to the DD-vMCG algorithm that improve both its accuracy and efficiency. First, a new, efficient parallel algorithm to control the DD-vMCG database of quantum chemistry points is presented along with improvements to the Shepard interpolation scheme. Second, the use of symmetry in describing the potential surfaces is introduced along with a new phase convention in the propagation diabatization. Benchmark calculations on the allene radical cation including all degrees of freedom then show that the new scheme is able to produce a consistent non-adiabatic coupling vector field. This new DD-vMCG version thus opens the route for effectively and accurately treating complex chemical systems using quantum dynamics simulations.

I. INTRODUCTION

Quantum molecular dynamics simulations are essential to understand a number of different phenomena, especially those occurring on the ultra-fast, femtosecond time scale, where simulations are necessary to interpret experimental data. This includes fundamental dynamical processes in chemistry, such as photodissociation[1–4] or proton transfer.[5–8] These simulations involve solving the time-dependent Schrödinger equation (TDSE) to obtain the molecular time evolution.

As solving the Schrödinger equation for complex systems is computationally difficult, the Born–Oppenheimer approximation[9] is employed to simplify the problem. The Born–Oppenheimer approximation separates the electronic and nuclear motions, and the nuclei move on potential energy surfaces (PESs) provided by the electrons.[10–14] In this way, an electronic Schrödinger equation[15] is solved for a set of fixed nuclear arrangements, yielding the potential energy surfaces for the nuclear motion in a specific electronic

state and obtaining the solution of the molecular problem within the so-called *adiabatic approximation*. In general, for an N atom molecular system, the potential energy surfaces are a function of the $3N$ nuclear Cartesian coordinates, which is completely described by $3N-6$ linearly independent internal coordinates ($3N-5$ for a linear molecule), as the Hamiltonian is invariant to rotation and translation of the entire system.

A wide range of methods have been developed to solve the TDSE. However, in standard propagation approaches, where the initially formed nuclear wavepacket and the Hamiltonian are represented by a time-independent product grid basis, the calculations become a computationally heavy process and impossible with more than a few (typically 3–4) degrees of freedom. To overcome this obstacle, the Multiconfiguration Time-Dependent Hartree (MCTDH) method[16,17] was introduced, where the wavefunction is represented by employing a basis set of time-dependent single particle functions, which can either be one dimensional or multidimensional. Figure 1(a) shows a schematic representation of this
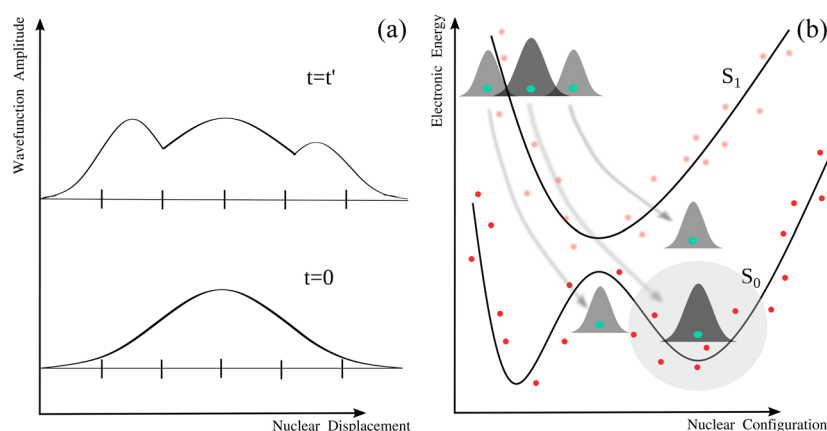
FIG. 1. Schematic representation of (a) grid-based and (b) DD-vMCG methods for excited-state dynamics.

concept. Multi-dimensional, non-adiabatic systems, where the surfaces are strongly coupled, can be treated using this method.[18–22] MCTDH is a very powerful method; however, like other grid-based methods, the computation and fitting of the potential energy surfaces is required prior to any calculation being performed. Hence, treating larger chemical systems is feasible only by introducing new approximations that remove the restrictions of the grid-based methods.

To this end, the variational Multi-Configurational Gaussian (vMCG)[23–25] method and its DD-vMCG version[24–27] were developed. This is referred to as a Gaussian wavepacket method as the time-dependent basis functions of MCTDH are replaced by parameterized Gaussian functions.[23,24,28–31] Direct dynamics is the branch of molecular dynamics simulations that solves the time-dependent Schrödinger equation by allowing for the calculation of potential energy surfaces *on-the-fly*.[27] This enables the analysis of the influence of the quantum effects on reactivity without the time-consuming need to pre-compute potential energy surfaces.

One of the major advantages of this method is the straightforward extension to larger systems that undergo long-range dynamics, and now, the only limitation is the need to calculate potential energies with a quantum chemistry (QC) method of choice. Direct dynamics simulations of photo-excited molecules are becoming the method of choice, and a number of methods and codes have been developed for this, as described in detail in a recent review of Crespo-Otero and Barbatti.[32] These include surface hopping,[33,34] *ab initio* multiple spawning (AIMS),[35,36] and multi-configurational Ehrenfest (MCE).[37] DD-vMCG has the potential advantage over all these methods that it converges faster due to its variational basis set and includes all terms and couplings in the Hamiltonian.

In a recent paper, DD-vMCG calculations were reported on the dynamics of formamide including eight electronic states.[38] This not only highlighted the potential of the method but also showed its limitations. Some limitations such as kinks in the potentials are due to the electronic structure calculations. Other limitations such as lack of the correct symmetry in the potentials were due to the implementation. Other problems, not detailed in this paper but discussed at the accompanying Faraday Discussion meeting, were the poor energy conservation and high computational cost: the simulations of the six-atom system required a number of months of CPU time. In this paper, we address these issues and present algorithmic advances in

two of the main algorithms used in DD-vMCG simulations, which save time and improve both stability and accuracy. We also show for the first time that the method provides a continuous and coherent description of the non-adiabatic coupling vector field.

The first advance is a novel algorithm that drastically reduces the time required for a simulation. To save the effort of many expensive quantum chemistry calculations during a DD-vMCG propagation, the calculated energies are stored in a database at selected configurations along the trajectories followed by the Gaussian wavepackets (GWPs). This is known as the quantum chemistry (QC) database. At each time step, the potential surfaces surrounding each GWP are then provided by a modified Shepard interpolation using the stored data points.[39]

Figure 1(b) illustrates schematically the concept employed by DD-vMCG during excited-state dynamics where the green dots represent the center of the GWP, the red ones represent the database points, and the black lines represent the fit to the database points. A key challenge of this approach is the time needed to continually reread, sort, and analyze the database, which makes the calculation of a large system very expensive. The gray circle around the GWP of the ground state shows that, at each time, only a small number of closest points are needed and not the whole database, which constitutes the underlying concept of the development work presented here to minimize the effort. The implementation of the Shepard interpolation used is also investigated and improved.

The second key ingredient in the study of non-adiabatic photochemical systems using the DD-vMCG method is the requirement of a diabatization scheme, which allows for the *on-the-fly* diabatization of multiple electronic states. For this propagation, the diabatization scheme previously introduced has the right properties.[38,40] Here, we improve its usage by using point group symmetry when setting up the QC database. We also use analysis of a simple model of non-adiabatic coupled surfaces to choose a new phase convention in the adiabatic–diabatic transformation matrices and then go on to show that this diabatization scheme is indeed able to provide a consistent, global vector field for the non-adiabatic coupling vector. This means that this scheme can, indeed, provide truly (pseudo-)diabatic potentials.

In Sec. I, the theory of the DD-vMCG method will be presented along with a detailed description of the QC database and interpolation approaches and also the diabatization scheme used.

Subsequently, an efficient parallel algorithm for dealing with large QC databases will be introduced accompanied with further developments in the interpolation. An improvement of the diabatization scheme based on an analytical treatment of the properties of the non-adiabatic coupling vectors and of the behavior of the adiabatic-to-diabatic transformation matrix that leads to a more accurate and efficient program for treating various chemical systems will also be described. Methodological updates to the DD-vMCG implementation are outlined along with an application to the allene cation ($C_3H_4^+$) in the Jahn–Teller split $\tilde{X}$ manifold in each case. Benchmark calculations are also presented to show the performance of the parallel algorithm. Finally, a discussion along with the main conclusions and some future work will be provided.

All the methods described here are included in the powerful and flexible Quantics package.[41,42] Together, this new implementation makes a huge step forward in making DD-vMCG a practical, competitive, and accurate approach for the simulation of non-adiabatic molecular dynamics.

II. THEORY AND METHODS

A. Variational multi-configuration Gaussian (vMCG) method

A brief overview of the key concepts of the vMCG method will be presented here. A detailed description of the method, including fundamental theories and approaches governing vMCG dynamics, can be found in various scientific publications.[23,24,40] As described in the Introduction, the DD-vMCG method aims to solve the time-dependent Schrödinger equation for a wavefunction Ψ, which depends on both nuclear coordinates, x, and time, t,

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

where $\hat{H}$ is the Hamiltonian operator, which can be written in the usual way as

$$\hat{H} = \hat{T}_N + \hat{H}_{el}, \tag{2}$$

where $\hat{T}_N$ is the nuclear kinetic energy operator and $\hat{H}_{el}$ is the clamped nucleus electronic Hamiltonian of quantum chemistry, which is a function of the nuclear coordinates.

DD-vMCG uses what is known as the *single-set* vMCG ansatz. This means that a single set of Gaussian wavepackets is used to describe the wavefunction on all states. This is more efficient than using separate sets of functions for each state as the quantum chemistry calculations need to be performed at each GWP center, so reducing the number of centers is more important than reducing the overall number of configurations. Thus, the following *ansatz* is employed:

$$\Psi(x,t) = \sum_{s=1}^{N_s} \sum_{j=1}^{N} A_{js}(t) g_j(x,t)|s\rangle, \tag{3}$$

where N_s is the number of states and N is the number of time-dependent Gaussian wavepackets in which the nuclear wavefunction is expanded. The vectors $|s\rangle$ denote the electronic states.

Each GWP employed in the above *ansatz* is a multi-dimensional parameterized function, with x being the set of coordinates, and has the following form:

$$g_j(x,t) = \exp(x^T \cdot \varsigma_j \cdot x + \xi_j \cdot x + \eta_j), \tag{4}$$

where x^T is the transpose vector of the coordinates and the time-dependent, complex parameters inside the Gaussian function are described by a square matrix, ς, a vector, ξ, and a scalar η.

Applying the Dirac–Frenkel variational principle to Eq. (3),

$$\langle \delta\Psi | H - i\hbar \frac{\partial}{\partial t} | \Psi \rangle = 0, \tag{5}$$

outputs two sets of equations of motion (EOMs), one for the set of parameters of GWPs and one for the wavefunction expansion coefficients. The nuclear dynamics using an appropriate representation of the potential energy surfaces (PESs) can then be followed by solving these EOMs employing standard numerical integrators and appropriate initial conditions.

The single-set vMCG equations of motion for the expansion coefficients can be written as follows:

$$i\dot{A}_{js} = \sum_t \sum_{kl} S_{jk}^{-1}(H_{ks,lt} - i\tau_{kl})A_{lt}, \tag{6}$$

where indices j, k, l follow the GWPs and s, t represent the electronic states. The Hamiltonian operator matrix, H, is given by

$$H_{is,jt} = \langle sg_i | \hat{H} | tg_j \rangle \tag{7}$$

$$= \langle g_i | \hat{T}_N + H_{el}^{(st)} | g_j \rangle, \tag{8}$$

where $H_{el}^{(st)}$ is the electronic Hamiltonian matrix for states s, t, i.e., the potentials and non-adiabatic couplings, which are calculated by quantum chemistry and, then, using the procedure described below, are provided in a diabatic representation. The overlap matrix, S, is expressed in the Gaussian function basis set as follows:

$$S_{ij} = \langle g_i | g_j \rangle, \tag{9}$$

and the differential overlap matrix, τ, is given by

$$\tau_{ij} = \langle g_i | \frac{\partial}{\partial t} g_j \rangle. \tag{10}$$

For reasons of stability and efficiency, the GWP width matrix ς is kept fixed and only the GWP parameter vector ξ is propagated variationally, i.e., "frozen" Gaussians are used. The scalar parameter is then constructed to make the GWPs phaseless, and the phase is carried by the expansion coefficients. The equations of motion for the GWP parameters can then be written in a vector notation as

$$i\dot{\xi} = C^{-1}Y. \tag{11}$$

Employing the usual nomenclature,[24] C and Y have the following forms:

$$C_{i\alpha,j\beta} = \rho_{ij}\left(S_{ij}^{(\alpha\beta)} - \left[S^{(\alpha 0)}S^{-1}S^{(0\beta)}\right]_{ij}\right), \tag{12}$$

$$Y_{i\alpha} = \sum_j \rho_{ij}\left(H_{ij}^{(\alpha 0)} - \left[S^{(\alpha 0)}S^{-1}H\right]_{ij}\right), \tag{13}$$

with

$$\rho_{ij} = \sum_s A_{is}^* A_{js} \tag{14}$$

being the density matrix, and the superscripts α, β in the matrix elements indicate the derivatives of the Gaussian functions regarding the parameters of the Gaussian form [Eq. (4)].

The C matrix and Y vector look quite complicated, involving the Gaussian function overlaps and Hamiltonian matrix elements in the Gaussian basis set, respectively. It can, however, be shown that the centers of the GWPs follow trajectories that are variationally coupled and not just classical. More information and definitions regarding the EOM employed in the vMCG method can be found in Refs. 23 and 24.

B. DD-vMCG: The QC database

As described in the Introduction, the DD-vMCG method is one of the most promising applications of Gaussian wavepacket dynamics in terms of accuracy and flexibility. In order to solve the EOMs, the matrix elements of the Hamiltonian must be evaluated. As outlined in Eq. (8), the Hamiltonian contains the kinetic and potential energy terms. Assuming that the dynamics run in rectilinear coordinates (e.g., Cartesian or normal modes), the matrix elements of the kinetic energy operator have a fairly simple analytic form.[23]

Using a Taylor series to second-order around the geometry at the center of a GWP, x_0, a potential energy surface can be expanded as follows:

$$V(x) = V(x_0) + g(x_0)^T.(x - x_0) + \frac{1}{2}(x - x_0)^T.H(x_0).(x - x_0), \quad (15)$$

where $V(x_0)$ denotes the energy, $g(x_0)$ denotes the gradient, and $H(x_0)$ denotes the Hessian of the potential energy surface with respect to changes in geometry. This expansion of the PES is called the local harmonic approximation (LHA), and it is also used to calculate the matrix elements in the equations of motion [Eqs. (6) and (11)].

In rectilinear coordinates, using the LHA, all the integrals can be obtained analytically using information from standard quantum chemistry (QC) calculations. However, it is not desirable to run an expensive QC calculation at each step in the propagation. Thus, QC calculations are only run when the center of a GWP has moved significantly away from a previous point. For the integration steps between the QC calculations, Shepard interpolation is used to provide the LHA based on the set of prior QC calculations. The data from the previous calculations must therefore be saved in a database, the QC database. It is important to note that at each point in configuration space and for all the electronic states involved in the dynamics, information regarding the electronic structure must be known.

The distance criteria for a significant geometric change can be defined in many ways. For example, it can be the Euclidean norm between the new structure and any database point in Cartesian coordinates. However, in practice, it is found that the maximum displacement of an atom in the new structure compared to the DB points is a more useful criterion as it distinguishes structures that are locally mobile. The distance that must be exceeded is referred to as the *dbmin* parameter.

vMCG calculations have to be run in the diabatic picture as the discontinuities in adiabatic potential surfaces cause problems for the LHA. Thus, the quantum chemistry results need to be transformed to a diabatic basis. How this is done is explained in Sec. II C, but it means that the derivative coupling must also be calculated and stored in the QC database in addition to the energy, gradient, and Hessian. To save the effort of calculating the Hessians at each point, Hessian updating is used,[23] which means that Hessians must only be calculated at the first point to be added. The updating is done in the diabatic picture.

Thus, each time a new point is calculated, the geometry and all the related information, both adiabatic raw data and data in the diabatic picture, are added to the QC database. DD-vMCG can be run without using a QC database, like other direct dynamics methods, but it has various advantages. One of the obvious advantages when using a QC database is the smaller number of QC calculations needed, which is important when an expensive method is employed to calculate the electronic structure quantities. Even if a cheap method is used, the QC database offers the advantage that it can be reused after the completion of the propagation or if the calculation terminates unexpectedly, one does not have to recalculate the points previously computed. However, at each propagation step, for every GWP at each integration step, the program needs to reread, sort, and analyze this database. Thus, as the dynamics progress and especially in the case of large and complex chemical systems where higher numbers of GWPs need to be used, running direct dynamics in this way can become a computationally quite heavy process even when QC calculations are not being performed.

It should be noted that the QC database is not reliant on the use of an LHA and provides data that can be used in more general potential fitting.

C. Diabatization

Within the DD-vMCG method, a general *on-the-fly* diabatization scheme is provided by the so-called propagation diabatization.[40,43] This method is based on the propagation of the adiabatic-to-diabatic transformation matrix along the paths followed by the GWPs and is, in principle, able to handle an arbitrary number of states with an unknown number of surface crossings.

The basic idea of the propagation diabatization method is to use the relationship[44]

$$\nabla S = -FS \quad (16)$$

between the transformation matrix, S, and the derivative coupling matrix, F, to evaluate S at each point. Integration of Eq. (16) along a path between two molecular geometries x and $x + \Delta x$ describes the propagation of the adiabatic-to-diabatic transformation matrix over some short step Δx with starting point x. The formal solution is given by

$$S(x + \Delta x) = \exp\left(-\int_x^{x+\Delta x} F(x')dx'\right)S(x) \quad (17)$$

and defines the diabatization scheme in essence. As a straightforward numerical integration generally does not guarantee to yield a unitary matrix, Eq. (17) can be rearranged as

$$\exp\left(\frac{1}{2}\int F(x')dx'\right)S(x + \Delta x) = \exp\left(-\frac{1}{2}\int F(x')dx'\right)S(x) \quad (18)$$

according to the method of Esry and Sadeghpour.[45] Then, the exponentials are calculated using a Taylor expansion, and by inversion of the resultant expression for the matrix $\exp\left(\frac{1}{2}\int F(x')dx'\right)$ on the left-hand side, the final transformation matrix, $S(x + \Delta x)$, may be obtained.

Equation (16) is exact if a complete (infinite) set of electronic states is considered.[44] However, it is obviously impossible to handle an infinite number of states in practical calculations, and the electronic basis must be truncated. If the electronic states included in the calculations form a sub-basis where the extended curl, divergence, and quantization conditions hold, as extensively discussed in Ref. 46, the couplings to the excluded states can be safely ignored and Eq. (16) is approximately fulfilled. If these conditions are not met, then integrating Eq. (16) is path dependent and the transformation matrix S may not be consistently defined.

To correctly obtain the diabatic states, it is also important to get consistent global phases for the coupling elements. For this, it is instructive to consider the topography of the potential surfaces around a conical intersection. For this, a two-state diabatic potential matrix, W, may be expanded around the intersection as

$$W = \begin{pmatrix} \kappa_1 x & \lambda y \\ \lambda y & \kappa_2 x \end{pmatrix} + \left(\frac{\omega_x}{2} x^2 + \frac{\omega_y}{2} y^2 \right) \mathbf{1}, \qquad (19)$$

where $\mathbf{1}$ is the unit matrix, x and y are the mass–frequency scaled normal mode coordinates with vibrational frequencies ω_x and ω_y, respectively, and κ_1 and κ_2 and λ are the gradients and non-adiabatic coupling defined in the diabatic electronic basis by

$$\kappa_i = \langle \psi_i | \nabla \hat{H}_{el} | \psi_i \rangle \big|_0 \quad \text{for} \quad i = 1, 2, \qquad (20)$$

$$\lambda = \langle \psi_1 | \nabla \hat{H}_{el} | \psi_2 \rangle \big|_0, \qquad (21)$$

where ψ_i are the diabatic electronic wavefunctions at the intersection point. For later use, the expression of the diabatic potential matrix can be rewritten as

$$W = \Sigma \mathbf{1} + \begin{pmatrix} -\delta x & \lambda y \\ \lambda y & \delta x \end{pmatrix}, \qquad (22)$$

where $\Sigma = \frac{1}{2}(\omega_x x^2 + \omega_y y^2 + \kappa_1 x + \kappa_2 x)$ and $\delta = \frac{1}{2}(\kappa_2 - \kappa_1)$.

By definition, the general form of the derivative coupling matrix is given by

$$F = \begin{pmatrix} 0 & F_{12} \\ -F_{12} & 0 \end{pmatrix}, \qquad (23)$$

and it contains only one non-vanishing coupling vector F_{12}. Noting that the transformation matrix, S, is unitary and can be defined by a rotation angle, θ,

$$S = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \qquad (24)$$

and using the rotation angle that diagonalizes the diabatic potential matrix [Eq. (22)], it directly follows from Eq. (16) that the derivative coupling vector is given by

$$F_{12} = \frac{1}{2((\delta x)^2 + (\lambda y)^2)^{\frac{1}{2}}} \begin{pmatrix} \lambda y \\ -\delta x \end{pmatrix}. \qquad (25)$$

Hence, the derivative coupling vectors form a vector field where the vectors circulate around the conical intersection. This analytical model of the derivative coupling vectors satisfies the curl condition, which for the two-dimensional case may be simplified to

$$\frac{\partial}{\partial y} F_{12}^{(x)} - \frac{\partial}{\partial x} F_{12}^{(y)} = 0. \qquad (26)$$

Moreover, substitution of the derivative coupling matrix F and the adiabatic-to-diabatic transformation matrix S into Eq. (16) yields the following differential equation for the transformation angle:

$$\nabla\theta = -F_{12}. \qquad (27)$$

If the transformation angle is set to 0 at the starting point R_0, the solution of Eq. (27) is given by

$$\theta = -\int_{R_0}^{R} F_{12} dR, \qquad (28)$$

where the integral has to be taken along a given path. If the path is closed, the integral mentioned above has to be an integer multiple of π, leading to the following quantization condition:

$$-\oint F_{12} dR = n\pi, \qquad (29)$$

where n is an integer if the loop encloses a conical intersection and 0 otherwise.

From an examination of this model, the following procedure is followed to obtain the diabatic states for any number of states using propagation diabatization. For the first point in the QC database, the adiabatic and diabatic surfaces are taken to be equivalent. This sets the global gauge and effectively defines the diabatic states. For subsequent points, integration of the transformation matrix to the new point is made from an old point, taken as the closest point in the QC database. Using the diabatic surfaces at the old point, predicted diabatic surfaces at the new point are also obtained, which give a predicted new transformation matrix and predicted adiabatic data. This allows the following procedure to be run:

1. The order of states in the predicted diabatic model and the diabatic model at the old point is compared. If they have swapped the order, the conical intersection seam is crossed while taking the step and integration of the transformation matrix cannot account for this. The predicted transformation matrix is thus used to rotate the adiabatic data at the new point to the diabatic representation at the new point.

2. The overlap between the predicted and calculated derivative coupling is found, and if the absolute value of the overlap is small, taken as less than $1/\sqrt{2}$, the vector field is not continuous, probably due to an intruder state. Again, the predicted transformation matrix is used with the calculated adiabatic energies to provide the diabatic surfaces at the new point. This effectively makes one diabatic surface change character in a controlled way. If the overlap is good, but negative, the sign on the calculated derivative coupling is corrected and the transformation matrix propagated from the old point to the new using the altered vector.

3. The energy gap of the adiabatic states at the new point is checked. If it is less than a threshold apart (0.05 eV), this is classed as a region of degeneracy where the quantum chemistry is not to be trusted. The predicted diabatic surfaces are stored in the QC database, along with the adiabatic raw data.

4. If the QC calculation has failed (e.g., CAS has failed to converge), the predicted diabatic surfaces are stored along with the predicted adiabatic surfaces.

5. If none of these conditions are met, the points are in a continuous part of the diabatic space and the propagation diabatization

is used to propagate the transformation matrix from the old to new point, to obtain the new diabatic potentials from the adiabatic data.

In all cases, before the final adiabatic-to-diabatic transformation is made, the phase of the eigenvectors that forms the transformation is changed to ensure that the diagonal elements are positive, and the transformation matrix represents a proper rotation. As a result, the rotation angle in a two-state case will be between $\frac{\pi}{2}$ and $-\frac{\pi}{2}$.

D. The use of symmetry

If a molecule has symmetry, the potential energy surfaces are also symmetric with respect to the various symmetry operations of the relevant point group. To ensure this symmetry, when adding a point to the QC database, all the unique symmetry replicas of the structure, i.e., those generated by the operations of the point group, should also be added to the database. In DD-vMCG, this has been introduced as follows.

First, using a template structure with the correct symmetry, a mapping is set up so that after each symmetry operation, the atoms are returned to their original position. This permutation mapping ensures that the symmetry replicas relevant to the region of configuration space being explored are provided. This template structure is often the molecular geometry at the center of the initial wavepacket for a simulation as the ground-state equilibrium structure (the Franck–Condon point) is usually the highest possible symmetry.

Once the mapping is known, each time a point is to be added, the replicas generated by the symmetry operation and permutation are checked for uniqueness. If the structures are closer to the original structure than the database distance criteria (dbmin), they are averaged to provide a new structure and a new set of replicas via operation and permutation are produced. Any replicas now closer to the original structure than dbmin are ignored as they should be identical.

The quantum chemistry calculation is then performed on the original (or averaged) structure to obtain the energy, gradient, and derivative coupling. The energy is taken as identical for each unique replica, and the gradient and derivative coupling vectors are provided by appropriate symmetry operations. The diabatization procedure is then performed for each unique structure, along with the Hessian updating, to provide the symmetrized diabatic potentials.

E. Modified Shepard interpolation

In cases where there is no need for a new point to be calculated, a modified Shepard interpolation[39] can be employed to obtain the energies, gradients, and Hessian matrices. The first step is to calculate the Euclidean norm of the difference vector of all atomic coordinates between the new point and each point in the QC database. The potential energy can then be calculated by employing a weighted average of the Taylor series over the N_d data points,

$$V(\boldsymbol{x}) = \sum_{i=1}^{N_d} w_i(\boldsymbol{x}) T_i(\boldsymbol{x}), \tag{30}$$

where T_i is the Taylor expansion and w_i is the weight function that weights the contribution of the Taylor expansion at point i and has the following form:

$$w_i(\boldsymbol{x}) = \frac{v_i(\boldsymbol{x})}{\sum_{j=1}^{N_d} v_j(\boldsymbol{x})}. \tag{31}$$

In its simplest form, data points that are not close to $\boldsymbol{x}$ are assigned smaller weights compared to the ones that are very close by employing the primitive weight function v_i,

$$v_i(\boldsymbol{x}) = \frac{1}{|\boldsymbol{x} - \boldsymbol{x}_i|^{2p}}, \tag{32}$$

with

$$2p > 3N - 3, \tag{33}$$

where N denotes the number of atoms and i denotes the location inside the database. If the exponent p is sufficiently large, then, in the limit $N_{d\to\infty}$, Eq. (30) converges to the exact potential.

To make this procedure more efficient, a maximum weight can be used as a criterion to exclude from the PES calculation the data points a large distance away from the new geometry. Together with this selection procedure, the aforementioned method allows for a remarkably accurate calculation of the PES for small chemical systems ($N < 4$).

In a series of papers,[39,47–50] the different forms of the weight function have been investigated. Thompson *et al.* concluded that a better performance is obtained using the following expression for the weight function:[51]

$$v_i(\boldsymbol{x}) = \left\{ \left[\frac{\|\boldsymbol{x} - \boldsymbol{x}_i\|}{rad_i} \right]^{2q} + \left[\frac{\|\boldsymbol{x} - \boldsymbol{x}_i\|}{rad_i} \right]^{2p} \right\}^{-1}, \tag{34}$$

with

$$2q \geq 2 \quad \text{and} \quad 2p \geq 3N - 3, \tag{35}$$

where p and q are positive integers and $p \gg q$ ensures that the first term of this equation is sum-dominant when $\|\boldsymbol{x} - \boldsymbol{x}_i\| < rad_i$, whereas the second term is sum-dominant when $\|\boldsymbol{x} - \boldsymbol{x}_i\| > rad_i$.

The values of $q = 2$ and $p = 12$ were found to be reasonable values. The rad_i distance is a type of *confidence radius* supplied by a close point in the database, j,

$$rad_i = \|\boldsymbol{x}_j - \boldsymbol{x}_i\|, \tag{36}$$

and the selection of the point j depends on the number of nearest points to be included and the molecule under investigation.

F. Allene radical cation

The test system to be used to demonstrate the new implementations in the DD-vMCG code is the $\tilde{X}$ state of the allene radical cation. This is a molecule that has been previously studied using a vibronic model Hamiltonian,[52] and the nature of the coupling between the components of the degenerate state is known, making it an ideal target for the present study as a comparison of the results can support the accuracy of the method. It is also of a sufficient size to be a challenge for direct dynamics simulations and has the added difficulty of having a conical intersection directly at the Frank–Condon point, which can cause problems for simulations by introducing a strong non-adiabatic coupling from the start.

The equilibrium structure of neutral allene, shown in Fig. 2, belongs to the point group D_{2d}. It has 15 normal vibrational modes, which are classified by their irreducible representations,

$$\Gamma = 3A_1 + B_1 + 3B_2 + 4E. \qquad (37)$$

The electronic ground and first excited states of the allene cation, denoted by $\tilde{X}^2E$, are doubly degenerate at the Franck–Condon point. The allene cation provides a representative example of the $E \otimes \beta$ Jahn–Teller effect, where the symmetry of the state is lowered by coupling to pairs of modes with B_1 and B_2 symmetry. Cederbaum et al.[53] showed in their theoretical treatment of the Jahn–Teller effect in the allene cation that the only modes that strongly couple the ionic doubly degenerate ground state are the $v_5(B_1)$ torsional and $v_{11}(B_2)$ antisymmetric $C - C$ stretching modes. The conical intersection formed along these two modes takes place exactly at the Franck–Condon point. In the area around that point, the next energetically higher electronic states are well separated. Thus, the investigation of the potential surfaces of the allene cation can focus on a system involving only two modes and two states. This can be related to the analytic model discussed above with the $v_5(B_1)$ and $v_{11}(B_2)$ vibrations of the ionic allene system the x and y coordinates of the diabatic potential matrix.

G. DD-vMCG simulations

DD-vMCG nuclear dynamics were performed on the ground and first excited states of the ionic allene molecule including all degrees of freedom as mass–frequency scaled normal modes. The electronic structure calculations on the allene cation were performed using MOLPRO 2015.1[54] at the CASSCF(3,4)/6-31G* level of theory. The normal modes were obtained from a CASSCF(4,4)/6-31G* calculation at the neutral ground-state geometry.

The GWPs used for the basis functions have a width $1/\sqrt{2}$ along all normal coordinates. In the mass–frequency scaled coordinate system used, this is the width of the neutral ground-state vibrational eigenfunction in the harmonic approximation. To form the desired initial wavepacket, one of these functions is placed at the Franck–Condon point with a momentum of 0 and a coefficient of 1.0 for the configuration including the second state. All other GWPs are then displaced in phase space and given coefficients of 0.0, i.e., the initial wavepacket is an exact representation of the neutral ground-state eigenfunction placed vertically into the ion manifold,

$$\Psi(\boldsymbol{x}, t = 0) = 1.0 \times g_j(\boldsymbol{x} = 0, 0)|2\rangle. \qquad (38)$$
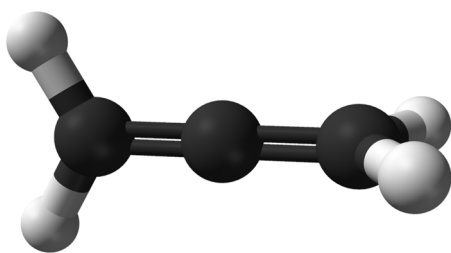

FIG. 2. The equilibrium ground-state structure of allene.

One of the key features of the vMCG method compared to other trajectory-based methods is that results are insensitive to the initial placing of initially undefined GWPs. Poor choices result in the integrator having to work harder at the start and may result in instability, but the variational nature of the trajectories means that the GWPs will always provide the optimal basis set that spans the same space. For DD-vMCG, it has been found that the best choice is to have all GWPs start at the same coordinate and to be displaced in momentum space, chosen by randomly stepping along the normal modes. This ensures that a single quantum chemistry calculation provides the surfaces for all functions in the first few femtoseconds.

Simulations with up to 20 GWPs were performed for 100 fs. These are not fully converged calculations as the aim here is to discuss the efficiency and reliability of the new code rather than provide accurate results on the allene cation dynamics. A measure of the quality of a calculation is given by the Gross Gaussian Populations (GGPs),[55]

$$GGP_i = Re \sum_s \sum_j A_{is}^* S_{ij} A_{js}, \qquad (39)$$

which divide the density of the wavepacket between the GWPs. In all simulations using 20 GWPs, there are at least two GWPS with a GGP of around 0.01, i.e., they carry only 1% of the wavepacket, and adding more functions will not make a major difference to the overall wavefunction shape and quantities such as state populations will be at least qualitatively correct. The state populations from the simulations with 20 GWPs and a simulation with 25 GWPs indeed show the same behavior, with similar differences seen by changing the parameters controlling the potential description to adding functions.

To circumvent numerical problems coming from excitation into a point of degeneracy, the initial reference geometry for the database was defined by a displacement of 0.4 units along the antisymmetric C–C stretching mode, where the states are close to the conical intersection but not degenerate. At this point, which is not at the center of the initial wavepacket, Hessians need to be calculated for both states. All other Hessians are provided by the Hessian updating procedure.

Following the concept of the modified Shepard interpolation presented above, in the DD-vMCG implementation, two different parameters are employed for this step, which are both user defined. The first one, n_{db}, is the number of closest points in the QC database, which is to be included for the interpolation in a local database for each GWP. Furthermore, the confidence radius is defined by the n_{conf} closest point. The n_{conf} has to be less than or equal to n_{db}, and both numbers are based on the distribution of the data points in the QC database. This development accompanied with the implementation of the modified Shepard interpolation scheme yields a more efficient way with respect to the computational time and also leads to more stable propagation with respect to the energy and norm conservation of the system.

III. RESULTS AND DISCUSSION

A. Use of local databases

The first part of this study was to improve the efficiency of a DD-vMCG simulation by optimizing the use of the QC database used to obtain the potential surfaces by Shepard interpolation. The

simple implementation of DD-vMCG within the Quantics code used to date involves reading, sorting, and analyzing the whole QC database at each propagation step. This approach results in a significant fraction of the total calculation time being spent on this procedure, which for direct dynamics simulations of large molecules, or when a QC database contains many points, presents a limitation for successfully treating complex chemical systems.

Here, we address this challenge by initially determining local databases containing the n_{db} closest points in the QC database for each GWP, selected by calculating the distances between the center of a GWP and all the reference points in the QC database. For efficiency, this procedure is selected to only be repeated at t_{large} time intervals for all the GWPs. It is important that the local databases are updated often enough to ensure that there are enough points close to the center of the moving GWP at all times, allowing an efficient interpolation that does not lead to integration errors or instabilities. The time t_{large} is thus chosen to represent when the displacement of a GWP center is significant.

The local databases are represented by employing a single-linked list array whose size matches the number of GWPs and consists of a series of individual node elements pointing to the location of each of the n_{db} closest points in the database. A linked list is a linear data structure where its elements are linked using pointers and are not stored at a contiguous location. Computationally, with this approach, the upper limit on the number of elements is not fixed and does not need to be predefined, which offers a great flexibility. In addition, it is cheaper to use linked lists instead of arrays when possible as inserting a new element in an array of elements requires a space that is created for this new element by shifting the existing ones. In terms of the Quantics code, the advantage of this approach is twofold. During the integration steps, the computational effort is remarkably reduced, and thus, the DD-vMCG method can be used for more complex processes, and a larger number of GWPs can be used to describe the wavefunction. Furthermore, the local databases are independent, so the code can be efficiently parallelized.

Following this approach, the part of the code dealing with the calculations of the energy, gradients, and Hessian matrices was parallelized. For this to be efficient, load balancing is critical. Thus, the efficient task distribution is based on the number of GWPs. The developed version using the local DB approach is designed to speed up both the serial and the parallel calculations, and a demonstration of its efficiency and capabilities is presented in this section.

To explore the speedup and the performance of the new code in comparison with the full QC database, a series of calculations have been conducted. The speedup is defined as the ratio of the serial runtime of the full QC database to the serial runtime of the local DB to perform the same direct dynamics calculation. Our aim was to create databases with an increasing number of total data points so that the comparison offers a clear picture of the speed improvement. Initially, a set of calculations with a decreasing dbmin value was conducted where the code performed only QC calculations at each step and results were stored in the database. Since the dbmin parameter, the distance criterion for adding new structures to the QC database, controls how often a QC calculation is performed, decreasing its value leads to databases with a higher number of data points. For each of these databases created from the aforementioned step, serial calculations employing both the full QC database and

local databases implementations were performed. In these calculations, the code only reads the database, calculates the PES from the database, and does not perform further QC calculations. During all these serial calculations, most of the variables such as the number of GWPs (10), the total propagation time (200 fs), and the interpolation confidence radius n_{conf}(6) have been kept constant. The number of points in the local databases, the n_{db} parameter, is also affected by dbmin, and when dbmin is reduced, more points are needed to cover the space required for the interpolation compared to a larger value of dbmin. Hence, the various dbmin values together with the suitable n_{db} parameter for each case are listed in Table I.

In Table I, the total elapsed time needed for all calculations is listed, along with the speedup defined as the ratio of the elapsed time with the full QC database to the elapsed time with the local databases. It can be seen that the time required for a calculation is greater when the number of points in the QC database increases. Moreover, using local DB is always faster compared to using the full QC database, and in the last case, with 5094 points, using the local database is 27 times faster. This confirms the idea that larger systems can be successfully treated with DD-vMVG. Another interesting result is that the total time needed for the local databases when the number of data points is similar, e.g., 215, 317 and 665 points, is almost constant in contrast with the full QC database, breaking the negative dependence between the total propagation time and the number of points in the database occurring during a full QC database calculation.

At the same time, the choice of the parameter n_{db} is potentially crucial as, in addition to the time and efficiency, it will also affect the dynamics and, hence, the accuracy of the calculation. To further understand the impact of n_{db} on the dynamics, the second example of the DD-vMCG calculation from Table I (dbmin = 0.2, 317 points) was employed as a starting point. Serial calculations where no new points are added in the database but only reading the existing ones were then performed using local databases with different values of n_{db}. Additionally, a similar calculation of only reading the current database was conducted employing the full DC database version. As shown in Fig. 3, changing the number of points does not significantly change the population over time. Identical results with those of employing the full database [Fig. 3(a)] can be achieved with local databases containing ten points, as depicted in Fig. 3(d). As explained above, the number of points inside the local database affects the interpolation and, thus, the integration of the DD calculation, which explains the behavior in Figs. 3(b) and 3(c), where the

TABLE I. The total elapsed time for serial DD-vMCG calculations with 10 GWPs using the full QC database and local databases, varying the number of points in the QC database by using different distance criteria (dbmin values). For all the calculations, the n_{conf} parameter was always equal to 6.

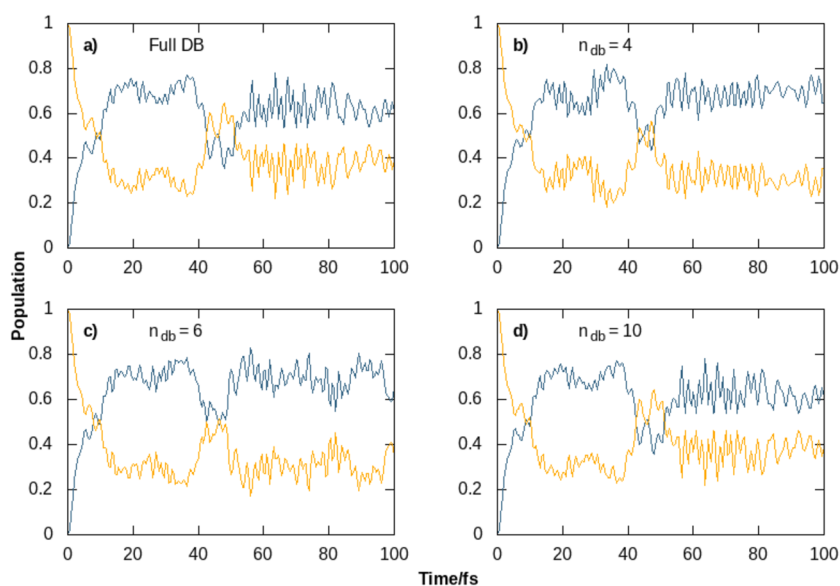| Distance criterion	0.25	0.20	0.15	0.10	0.05
n_{db}	10	10	20	20	40
	Number of data points				
	215	317	665	1292	5094
Full QC DB (s)	3033	3226	4613	18833	125374
Local DBs (s)	2766	2771	2756	3196	4723
Speedup	1.1	1.2	1.7	5.9	26.5

FIG. 3. The population of S_0 (blue line) and S_1 (orange line) states for the same DD-vMCG calculation of the allene cation employing (a) the QC database with 317 points and local databases with (b) 4 points, (c) 6 points, and (d) 10 points.

population is somewhat different compared to the full QC database calculation. Since for the allene cation a quite small n_{db} can be used, as shown in Fig. 3, altering the n_{conf} parameter did not have any impact on the results. Potentially, for molecules that require more points in the local databases and thus a larger n_{db} value, changing n_{conf} will possibly also affect the dynamics.

The local database approach was parallelized using OpenMP. The most accurate way to evaluate parallel performance is by running the same problem on 1 central processing unit (CPU) and on n CPUs and comparing the total elapsed time for the iteration in all cases. The aim is to explore the scalability of the parallel algorithm as a measure of its capacity to effectively utilize an increasing number of cores. Thus, the parallel speedup and efficiency of the new code are examined. The definition of Amdahl's law[56] was employed for the speedup ratio, which is determined as the ratio of the serial run-time of the best sequential algorithm for solving a problem to the time taken by the parallel algorithm to solve the same problem on p cores.

Test calculations were carried out using the newly developed algorithm for the allene system. Initially, two different databases were created by running serial calculations employing 10 and 20 GWPs. Then, parallel calculations employing different numbers of cores where the code was only reading points from the database without performing any new quantum chemistry calculation were conducted for both examples. The total propagation time, (200 fs), the interpolation confidence radius n_{conf}(6), and the total number of points in the local databases, $n_{db} = 40$, were kept constant. In Table II, the computational cost for the local database approach employing different numbers of cores is summarized.

Since the main coding development was focused on the database, separate time data accompanied with the parallel speedup and efficiency are presented for both the part of the code that deals only with the database and the total calculation. Thus, the comparison only with the parallel code that deals with the database depicts that in both examples, the speedup increases with the

number of cores, showing that the new local database approach has a great impact on the efficiency of the method. For the second example (20 GWPs), a parallel job on 20 cores was ten times faster, which reveals a very good parallel performance even though for this example the advantage of a parallel calculation is quite small. The parallel speedup seems to follow similar pattern for both cases, and a linear scaling is achieved, which is in accordance with Amdahl's law.

TABLE II. The total wall-clock time for a parallel implementation of DD-vMCG using local databases, with different numbers of cores.[a]

	Number of cores				
	1	5	10	15	20
5094 DB points—10 GWPs					
Total calculation					
Total time (s)	4723	2050	1511	1599	1623
Parallel speedup	1.0	2.3	3.2	3.0	2.9
Database code					
Total time (s)	3724	1033	587	598	604
Parallel speedup	1.0	3.6	6.3	6.2	6.2
Efficiency	1.0	0.7	0.6	0.4	0.3
9797 DB points—20 GWPs					
Total calculation					
Total time (s)	18 016	8903	5459	5004	4504
Parallel speedup	1.0	2.0	3.3	3.6	4.0
Database code					
Total time (s)	11 120	2926	1544	1463	1112
Parallel speedup	1.0	3.8	7.1	7.6	10.0
Efficiency	1.0	0.8	0.7	0.5	0.5

[a]Performed on a single node with 2 × 10 core Xeon 2.3 GHz CPUs.

The parallel efficiency is defined as the ratio of speedup to the number of cores. In this way, a good estimation of the fraction of time for which a processor is usefully utilized can be determined. A program that scales linearly has a parallel efficiency close to 1. Usually, a task-parallel program is more efficient than a data-parallel program. Parallel codes can more rarely achieve superlinear behavior as a result of an efficient cache usage per worker. The overall efficiency of the parallel program goes down as the number of the cores is increased for the same problem. This is the case for all parallel programs. The efficiency will increase when moving to more complex calculations as shown in Table II where the second test with more GWPs and more data points has a better performance.

Moving to the comparison of the total calculation, Table II shows that the speedup for the two parallel calculations of allene follows Amdahl's law, achieving the maximum speedup in the case where the number of nodes matches the number of GWPs. For the first example (10 GWPs) for both the part of the code that deals with the database and the total calculation, the efficiency drops significantly when the number of cores exceeds the number of GWPs. This might seem odd, but it is, in fact, an expected behavior as the part of the code dealing with the database is parallelized based on the number of GWPs. The communication of the extra cores is causing more delays, so it is better to keep the number of cores always less or equal to the number of GWPs.

As the total wall-clock time depicts, the largest sizes of memory and disk are required for dealing with the database when the number of data points is large, as in this example. The improvement of the parallelization in the case of the total calculation will be reduced by the performance of the communication subsystem of the hardware and the overhead of the parallel process itself, which explains why the total improvement from serial to parallel is not massive. Additionally, it must be also noted that only some bits of the total code are parallelized. In general, a code with its parallelizable component comprising 90% of total computation time can at best achieve a ten-fold speedup with lots of workers.[57] A code that is 50% parallelizable speeds up twofold with lots of workers. The results thus indicate that the speedup for the parallel calculations of allene with the local databases is useful, but other parts of the code (e.g., calculation of integrals) need to be improved.

B. Description of branching space

An accurate solution of the TDSE requires good energy and norm conservation. In the simulations run here, with either the full
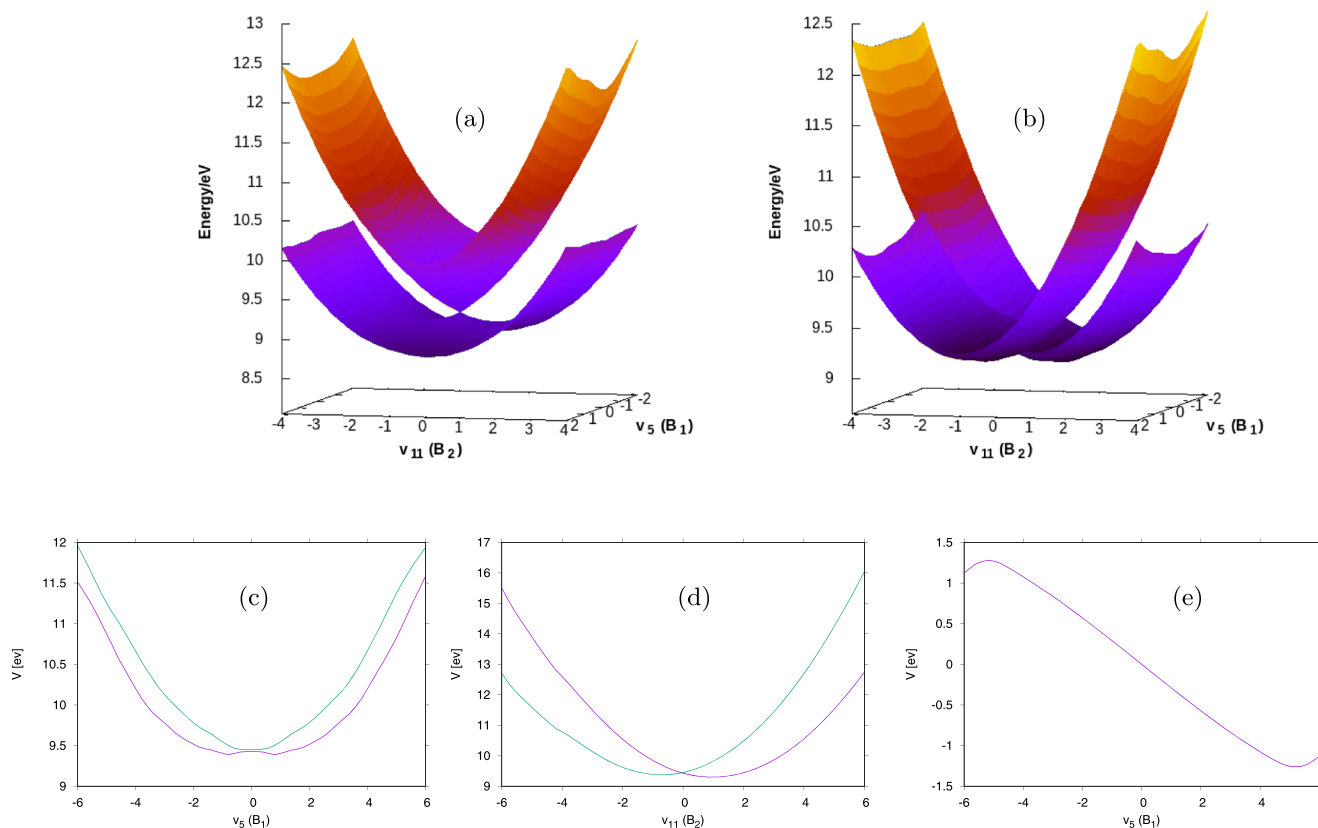

FIG. 4. Cuts through the potential energy surfaces of the allene cation obtained from a DD-vMCG simulation. Two-dimensional cuts through the (a) adiabatic and (b) diabatic surfaces in the space of the $v_5(B_1)$ and $v_{11}(B_2)$ normal modes. The diabatic surfaces along (c) the $v_5(B_1)$ and (d) the $v_{11}(B_2)$ normal modes, along with the diabatic coupling along $v_5(B_1)$ (e). All other coordinates have a value of 0.
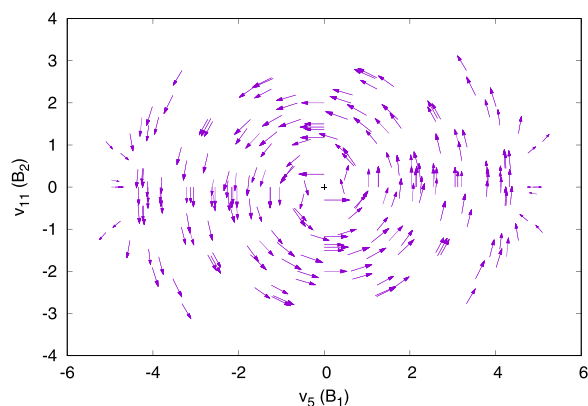
FIG. 5. Vector field formed by the derivative coupling vectors along the $v_5(B_1)$ and $v_{11}(B_2)$ normal modes.

QC database or the local databases, the energy (in eV) was conserved to three decimal places and the norm was conserved to seven decimal places over the simulations, showing that DD-vMCG propagation of the allene cation is stable. In addition, it is important that the diabatization procedure used is consistent and provides diabatic potentials that are consistent with the full solution. To show how the diabatization behaves, a full-dimensional DD-vMCG nuclear dynamics calculation was performed on allene using 10 GWPs with a dbmin value of 0.2. A local database of $n_{db} = 10$ points and an interpolation of $n_{conf} = 6$ were used as mentioned above. The dynamics were run for 100 fs with data output every 0.5 fs.

In contrast to the preceeding calculations, which focused on reading of the database, here, the symmetry of the system was used when adding points to the QC database, as described in Sec. II D. A QC database of 411 points was created, with only 70 quantum calculations being performed. Without the use of symmetry 317, calculations were made. The adiabatic and diabatic potential energy surfaces in the v_5, v_{11} branching space of the allene cation generated from the direct dynamics calculation are shown in Figs. 4(a) and 4(b), respectively. These are the modes known to be the most important in forming the conical intersection between the states.
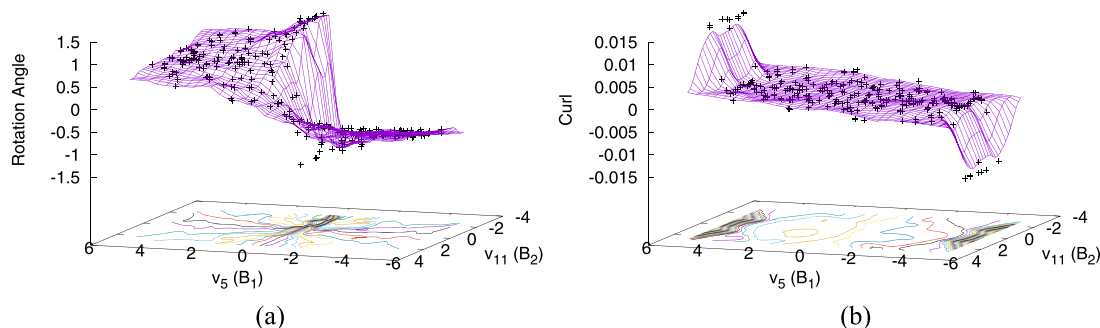
Both the adiabatic and diabatic surfaces are perfectly smooth, and the adiabatic surfaces feature a well-defined conical intersection at the Franck–Condon point.

Figures 4(c) and 4(d) show one-dimensional cuts of the diabatic surfaces along the v_5 and v_{11} modes. The diabatic coupling along v_5 is also shown in Fig. 4(e): along v_{11}, the coupling is negligible. Thus, the diabatic potentials have the form of a two-state, two-mode conical intersection in the diabatic representation of Eq. (19), with v_5 being the off-diagonal *coupling* mode and v_{11} being the on-diagonal *tuning* mode. Note that diabatic coupling along v_5 turns over at ±4.5. This is due to the interaction with a further excited state, as will be seen in the analysis below.

The potential surfaces can be analyzed in more detail to show that in addition to giving smooth surfaces, the procedure used is providing coherent, global diabatic potentials. In Fig. 5, the vector field of the derivative coupling vectors in the adiabatic picture obtained from the diabatic potential surfaces generated during the dynamics calculations on the allene cation is depicted. Note that the magnitude of the vectors was scaled up by a factor of 50 for a better visualization. As shown, the vectors are very well behaved, and the shape of the vector field clearly indicates that a conical intersection is present at the Franck–Condon geometry represented by the origin of the normal modes coordinates. Note that at values of the $v_5(B_1)$ coordinate of ±4.5, the vector field is changing direction. This is due to a higher lying state becoming degenerate with S_1 in this region.

A key quantity defining the coherence of the diabatic states is the behavior of the adiabatic–diabatic transformation angle, θ, shown in Fig. 6(a). The angle values are between $-\pi/2$ and $\pi/2$ and vary smoothly, except for the sharp jump from $-\pi/2$ to $\pi/2$ at $v_5(B_1) = 0$ for negative values of $v_{11}(B_2)$. This is a direct result from the quantization condition Eq. (29) for a two-dimensional system. This diabatization is keeping the phase of the transformation globally consistent. In addition, the curl condition [Eq. (26)] should be 0 for the diabatization to be rigorous. The non-trivial curl component of each derivative coupling vector was calculated at each point in the space of the $v_5(B_1)$ and $v_{11}(B_2)$ coordinates. From Fig. 6(b), it is shown that the curl of each vector is almost everywhere very close to 0, emphasizing that the diabatization is a good approximation to truly diabatic states in this region. Significant values of the curl are found for large values of $v_5(B_1)$, which can be

FIG. 6. Value of (a) the adiabatic-to-diabatic transformation angle θ along the $v_5(B_1)$ and $v_{11}(B_2)$ normal modes. (b) Non-trivial curl component of each derivative coupling vector in the $v_5(B_1)$ and $v_{11}(B_2)$ space. The crosses are the data points.

seen in the region where the vector field changes direction due to the change in state character.

IV. CONCLUSION

From the results mentioned above, it is clear that the DD-vMCG method is able to provide good quality diabatic surfaces for multi-dimensional non-adiabatic simulations directly from a set of quantum chemistry calculations. An efficient parallel algorithm using local databases for the potential surfaces around each basis function has been established, which leads to an improved implementation of the method in the Quantics package. In addition, analysis of the potential surfaces shows that they have the character of coherent global diabatic states.

Using the 15 dimensional, two-state allene radical cation as a test system, the efficiency and accuracy of the new algorithm were examined by performing test calculations employing local databases of different sizes and the full quantum chemistry database. Test calculations showed that including only the closest points during the propagation of each basis function leads to a significant speedup, solving a major bottleneck encountered by the DD-vMCG method when treating complex chemical systems. The speedup becomes even larger when using a parallel version of the algorithm, as well as more Gaussian wavepackets to describe the molecule under investigation. Smooth adiabatic and diabatic potential energy surfaces are produced, as well as smooth couplings between the diabatic states. Both the total energy and norm conservation are adequate, demonstrating the stability of the interpolation scheme used.

Since a correct representation of the diabatic potential energy surfaces is crucial for nuclear dynamics, an analysis of the surfaces was made. It has been proven that point group properties should be used to both further improve the efficiency of the method and ensure the correct symmetry of the potential energy surfaces. The diabatization scheme is shown to provide a coherent vector field for the non-adiabatic coupling around the conical intersection in the allene system and to cope with the phase shift in the adiabatic–diabatic transformation rotation angle. Application of the curl condition to the vector field shows that throughout the low energy space accessible to the allene radical cation after formation, a good diabatic representation is made. The curl condition also picks up that the diabatization is less correct in certain regions of configuration space where an examination of the vector field and potential surfaces indicates that a further electronic state becomes involved.

The novelty of this development work is in making DD-vMCG a competitive, efficient, and accurate method for nonadiabatic direct quantum dynamics simulations. Systems that were not computationally feasible can now be efficiently treated with improved implementation. Thus, future work will focus on treating larger chemical systems with the developed DD-vMCG version and on further comparing the outcomes and performance with other available quantum molecular dynamics methods such as trajectory surface hoping.[58]

ACKNOWLEDGMENTS

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request. The Quantics code is available from the GitLab repository upon reasonable request from the corresponding author.

REFERENCES

[1] S. P. Neville and G. A. Worth, J. Phys. Chem. A **118**, 11975 (2014).

[2] C. W. Heaps and D. A. Mazziotti, J. Chem. Phys. **145**, 064101 (2016).

[3] M. Araújo, B. Lasorne, A. L. Magalhães, G. A. Worth, M. J. Bearpark, and M. A. Robb, J. Chem. Phys. **131**, 144301 (2009).

[4] X. Xu, J. Zheng, K. R. Yang, and D. G. Truhlar, J. Am. Chem. Soc. **136**, 16378 (2014).

[5] T. Chatterjee, M. Mandal, A. Das, K. Bhattacharyya, A. Datta, and P. K. Mandal, J. Phys. Chem. B **120**, 3503 (2016).

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

[7] X. Chen, Y. Zhao, H. Zhang, J. Xue, and X. Zheng, J. Phys. Chem. A **119**, 832 (2015).

[8] I. Polyak, C. S. M. Allan, and G. A. Worth, J. Chem. Phys. **143**, 084121 (2015).

[9] M. Born and R. Oppenheimer, Ann. Phys. **389**, 457 (1927).

[10] A. Varandas, *Conical Intersections: Electronic Structure, Dynamics and Spectroscopy* (World Scientific, 2004), pp. 205–270.

[11] J. Murrell, S. Carter, S. Farantos, P. Huxley, and A. Varandas, *Molecular Potential Energy Functions* (Wiley, Chichester, 1984).

[12] A. Varandas, *Reaction and Molecular Dynamics*, Lecture Notes in Chemistry Vol. 75 (Springer, Berlin, 2000), pp. 33–56.

[13] Z. J. Z. Hui, *Theory and Application of Quantum Molecular Dynamics* (World Scientific, Singapore, 1998).

[14] A. J. C. Varandas, Int. Rev. Phys. Chem. **19**, 199 (2000).

[15] E. Schrödinger, Phys. Rev. **28**, 1049 (1926).

[16] H.-D. Meyer, U. Manthe, and L. S. Cederbaum, Chem. Phys. Lett. **165**, 73 (1990).

[17] G. A. Worth, M. H. Beck, A. Jäckle, and H.-D. Meyer, The Heidelberg MCTDH Package, Version 8.2, H.-D. Meyer, version 8.3 (2002), Version 8.4 (2007). current Version: 8.4.17, see http://mctdh.uni-hd.de, 2019.

[18] C. A. Schroeder, E. Pluhařová, R. Seidel, W. P. Schroeder, M. Faubel, P. Slavíček, B. Winter, P. Jungwirth, and S. E. Bradforth, J. Am. Chem. Soc. **137**, 201 (2015).

[19] S. P. Neville and G. A. Worth, J. Chem. Phys. **140**, 034317 (2014).

[20] K. Giri, E. Chapman, C. S. Sanz, and G. Worth, J. Chem. Phys. **135**, 044311 (2011).

[21] N. Hill, M.S. thesis, University of Birimingham, 2016.

[22] F. Otto, J. Chem. Phys. **140**, 014106 (2014).

[23] G. W. Richings, I. Polyak, K. E. Spinlove, G. A. Worth, I. Burghardt, and B. Lasorne, Int. Rev. Phys. Chem. **34**, 269 (2015).

[24] G. A. Worth, M. A. Robb, and I. Burghardt, Faraday Discuss. **127**, 307 (2004).

[25] G. A. Worth, in *Multi-Dimensional Potential Energy Surfaces*, edited by M. Law (CCP6, Daresbury, UK, 2008), pp. 79–87.

[26] G. A. Worth, P. Hunt, and M. A. Robb, J. Phys. Chem. A **107**, 621 (2003).

[27] G. A. Worth, M. A. Robb, and B. Lasorne, Mol. Phys. **106**, 2077 (2008).

[28] I. Burghardt, M. Nest, and G. A. Worth, J. Chem. Phys. **119**, 5364 (2003).

[29] G. A. Worth and I. Burghardt, Chem. Phys. Lett. **368**, 502 (2003).

[30] I. Burghardt, K. Giri, and G. A. Worth, J. Chem. Phys. **129**, 174104 (2008).

[31] B. Lasorne, M. A. Robb, and G. A. Worth, Phys. Chem. Chem. Phys. **9**, 3210 (2007).

[32] R. Crespo-Otero and M. Barbatti, Chem. Rev. **118**, 7026 (2018).

[33] M. Barbatti, G. Granucci, M. Persico, M. Ruckenbauer, M. Vazdar, M. Eckert-Maksić, and H. Lischka, J. Photochem. Photobiol., A **190**, 228 (2007).

[34] S. Mai, P. Marquetand, and L. González, Int. J. Quantum Chem. **115**, 1215 (2015).

[35] M. Ben-Nun and T. J. Martínez, Adv. Chem. Phys. **121**, 439 (2002).

[36] B. F. E. Curchod and T. J. Martínez, Chem. Rev. **118**, 3305 (2018).

[37] K. Saita and D. V. Shalashilin, J. Chem. Phys. **137**, 22A506 (2012).

[38] K. E. Spinlove, G. W. Richings, M. A. Robb, and G. A. Worth, Faraday Discuss. **212**, 191 (2018).

[39] T. J. Frankcombe, M. A. Collins, and G. A. Worth, Chem. Phys. Lett. **489**, 242 (2010).

[40] G. W. Richings and G. A. Worth, Chem. Phys. Lett. **683**, 606 (2017).

[41] G. A. Worth, K. Giri, G. W. Richings, M. H. Beck, A. Jäckle, and H.-D. Meyer, Quantics package, version 1.1, 2015.

[42] G. A. Worth, Comput. Phys. Commun. **248**, 107040 (2020).

[43] G. W. Richings and G. A. Worth, J. Phys. Chem. A **119**, 12457 (2015).

[44] M. Baer, Chem. Phys. Lett. **35**, 112 (1975).

[45] B. Esry and H. Sadeghpour, Phys. Rev. A **68**, 042706 (2003).

[46] M. Baer and G. D. Billing, *The Role of Degenerate States in Chemistry*, Advances in Chemical Physics Vol. 124 (Wiley, Hoboken, NJ, 2002).

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

[48] O. Godsi, C. R. Evenhuis, and M. A. Collins, J. Chem. Phys. **125**, 104105 (2006).

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

[50] J. Ischtwan and M. A. Collins, J. Chem. Phys. **100**, 8080 (1994).

[51] K. C. Thompson, M. J. T. Jordan, and M. A. Collins, J. Chem. Phys. **108**, 564 (1998).

[52] K. E. Spinlove, M. Vacher, M. Bearpark, M. A. Robb, and G. A. Worth, Chem. Phys. **482**, 52 (2017).

[53] L. S. Cederbaum, W. Domcke, and H. Köppel, Chem. Phys. **33**, 319 (1978).

[54] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, W. Györffy, D. Kats, T. Korona, R. Lindh *et al.*, Molpro, version 2015.1, a package of *ab initio* programs, 2015, see http://www.molpro.net.

[55] C. S. M. Allan, B. Lasorne, G. A. Worth, and M. A. Robb, J. Phys. Chem. A **114**, 8713 (2010).

[56] G. M. Amdahl, in *AFIPS '67 (Spring) Proceedings of the April 18–20, 1967, Spring Joint Computer Conference* (Association for Computing Machinery, New York, 1967), pp. 483–485.

[57] R. E. Bryant and D. R. O'Hallaron, *Computer Systems: A Programmer's Perspective* (Pearson, India, 2016).

[58] J. C. Tully and R. K. Preston, J. Chem. Phys. **55**, 562 (1971).