Predicting molecular photochemistry using machine-learning-enhanced quantum dynamics simulations

Gareth W. Richings and Scott Habershon*

Department of Chemistry, University of Warwick, Coventry, UK, CV4 7AL

E-mail: *S.Habershon@warwick.ac.uk
Conspectus

The processes which occur after molecules absorb light underpin an enormous range of fundamental technologies and applications, including photocatalysis to enable new chemical transformations, sunscreens to protect against the harmful effects of UV over-exposure, efficient photovoltaics for energy generation from sunlight, and fluorescent probes to image the intricate details of complex biomolecular structures. Reflecting this broad range of applications, an enormously versatile set of experiments are now regularly used to interrogate light-driven chemical dynamics, ranging from the typical ultrafast transient absorption spectroscopy used in many university laboratories to the inspiring central facilities around the world, such as the next-generation of X-ray free-electron lasers.

Computer simulations of light-driven molecular and materials dynamics are an essential route to analyse the enormous amount of transient electronic and structural data produced by these experimental sources. However, to date, direct simulation of molecular photochemistry remains a frontier challenge in computational chemical science, simultaneously demanding accurate treatment of molecular electronic structure, nuclear dynamics and the impact of non-adiabatic couplings.

To address these important challenges, and to enable new computational methods which can be integrated with state-of-the-art experimental capabilities, the last few years have seen a burst of activity in the development of “direct” quantum dynamics methods, merging machine-learning of potential energy surfaces (PESs) and non-adiabatic couplings with accurate quantum propagation schemes such as the multi-configuration time-dependent Hartree (MCTDH) method. The result of this approach is a new generation of “direct” quantum dynamics tools in which PESs are generated in tandem with wavefunction propagation, enabling accurate “on-the-fly” simulations of molecular photochemistry. These simulations offer an alternative route towards gaining quantum dynamics insights, circumventing the challenge of generating \textit{ab initio} electronic structure data for PES fitting by instead only demanding expensive energy evaluations as and when they are needed.

In this Account, we describe the chronological evolution of our own contributions to this field, focusing on describing the algorithmic developments that enable direct MCTDH simulations for complex molecular systems moving on multiple coupled electronic states. Specifically, we highlight active learning strategies for generating PESs during grid-based quantum chemical dynamics simulations, and we discuss the development and impact of novel diabatization schemes to enable direct grid-based simulations of photochemical dynamics; these developments are highlighted in a series of benchmark molecular simulations of systems containing multiple nuclear degrees-of-freedom moving on multiple coupled electronic states. We hope that the ongoing developments reported here represent a major step forward in tools for modelling excited-state chemistry such as photodissociation, proton and electron transfer, and ultrafast energy dissipation in complex molecular systems.
Tensor decomposition

New diabatization

Machine learning

Tensor decomposition

On-the-fly quantum dynamics

3
Key references


> Describes the initial integration of machine-learned potential energy surfaces, diabatization, and grid-based quantum dynamics.


> Describes how machine-learned potentials can be integrated with MCTDH quantum dynamics simulation in an active learning scheme.


> Describes improvements in efficiency available in direct, grid-based quantum dynamics due to improved kernel functions and potential energy decomposition schemes.


> Describes extension of potential energy surface decomposition scheme to three-body terms, and demonstrates use of alternative diabatization strategies in direct quantum dynamics.
Introduction

The absorption of light by molecules, and the subsequent photochemical processes which occur, underpins an ever-expanding range of applications in key technological and biological domains. Photovoltaics,\textsuperscript{5,6} sunscreens,\textsuperscript{7,8} photothermal converters,\textsuperscript{9,10} sensors,\textsuperscript{11,12} imaging,\textsuperscript{13,14} and photocatalysts\textsuperscript{15–18} are just some of the many examples where detailed knowledge of the light-molecule interactions is crucial to the chemist’s goal of molecular design for targeted functionality. Concurrently with the demands of practical applications, new experiments, such as high-intensity synchrotron radiation and X-ray free-electron lasers (XFELs),\textsuperscript{19,20} as well as the widespread deployment of methods such as ultrafast transient absorption spectroscopy, require the development of new tools for extracting chemical insights from complex data.

Against this backdrop of new applications and experiments, computational chemistry has sought to keep pace by developing improved simulation tools to provide insight into light-driven molecular dynamics. The ultimate “computational microscope” for predicting non-adiabatic chemical dynamics\textsuperscript{21–23} is direct simulation of the time-dependent Schrödinger equation (TDSE), given by:

\[ \frac{\partial \psi(R, r; t)}{\partial t} = -\frac{i}{\hbar} \hat{H} \psi(R, r; t), \]  

where \( \psi(R, r; t) \) is the time-dependent wavefunction describing the coupled nuclear (\( R \)) and electronic (\( r \)) degrees-of-freedom and \( \hat{H} \) is the full nuclear-electronic Hamiltonian operator. Most commonly, one invokes the well-known Born-Oppenheimer approximation in such simulations, with a nuclear wavepacket moving on, and transitioning between (in the case of non-adiabatic systems), different Born-Oppenheimer potential energy surfaces (PESs). After preparation of the initial wavefunction, typically mimicking the excitation of a molecular configuration on the electronic ground state to an excited electronic state (Franck-Condon approximation\textsuperscript{24,25}), wavefunction propagation with Eq. (1) enables direct access to time-
dependent properties, connecting simulation and experiment. For example, quantum dynamics simulations enable prediction of transient electronic and vibrational spectroscopies,\textsuperscript{26–31} determination of lifetimes for excited electronic states,\textsuperscript{26–29} provide estimates of product branching ratios,\textsuperscript{32} and give insight into photochemical energy transfer processes.

The challenge in using Eq. (1) lies in the associated computational effort. Wavefunction-based propagation schemes can suffer from rapid increase in computational demands as dimensionality increases; in the grid-based schemes explored below, this leads to the well-known “curse of dimensionality”. Although simulation schemes employing pruned and/or adaptive basis sets have been explored,\textsuperscript{33–39} such schemes are often developed on a case-by-case basis, rather than as a general solution to poor computational scaling. Similarly, deploying alternative basis sets to expand the time-dependent wavefunction (such as Gaussian wavepackets [GWPs]) coupled to approximate time-evolution strategies (such as Newtonian dynamics) can enable simulations of quantum chemical dynamics in systems containing many nuclear and electronic degrees-of-freedom; however, the inherent approximations made mean that some quantum phenomena (\textit{e.g.} tunnelling and zero-point energy conservation) remain difficult to treat accurately.\textsuperscript{40–48}

The current “gold standard” in propagating Eq. (1) is the multi-configuration time-dependent Hartree (MCTDH) scheme.\textsuperscript{22,49,50} MCTDH expands the time-evolving wavefunction as a tensor-product of single-particle functions (SPFs), each describing a small number (typically 1-4) of nuclear degrees-of-freedom. Application of the time-dependent, Dirac-Frenkel variational principle\textsuperscript{51,52} to the MCTDH wavefunction \textit{ansatz} then yields equations-of-motion for the individual SPFs and the associated expansion coefficients, with the variational nature of this strategy ensuring convergence on the numerically-exact results as the number of SPFs is increased. Ultimately, the computational effort for MCTDH exhibits the same exponential scaling as the “standard” direct-product grid-based method for solving Eq. (1);\textsuperscript{22,49,50} however, the time-dependent nature of the SPFs lowers the pre-factor associated with this scaling component, enabling MCTDH simulations of systems contain-
ing tens of degrees-of-freedom (compared to simulations of 5-6 nuclear degrees-of-freedom for “standard” methods). Furthermore, the development of the multilayer MCTDH scheme (ML-MCTDH),\textsuperscript{53–55} which exploits the underlying hierarchically-correlated nature of typical molecular PESs, has enabled the expansion of MCTDH to systems containing hundreds of nuclear coordinates. Despite this progress, MCTDH simulations remain extremely demanding with regards to computational effort, predominantly as result of the requirements associated with evaluation of the Hamiltonian matrix elements in the SPF basis set.

A second challenge in regards to computational expense is the global nature of the quantum equations-of-motion describing wavefunction evolution (Eq. (1)), which formally demand knowledge of the PES across the global configuration space of the system; in the case of non-adiabatic simulations, one similarly requires knowledge of the PES of every accessible electronic excited-state, and the couplings between them. The PESs describing nuclear motion are typically obtained from \textit{ab initio} electronic structure calculations using the Born-Oppenheimer approximation; as such, generation of an accurate global PES can be prohibitively expensive for all but the smallest systems. Strategies to address this challenge often employ an intermediate PES fitting step, where a large set of \textit{ab initio} calculations for different molecular configurations are performed and subsequently used to fit an \textit{analytical} global PES.\textsuperscript{30,56–64} While successful, an important drawback of this approach is the requirement for a PES pre-fitting step, which can itself prove non-trivial. In contrast, the focus of this article is on so-called “direct dynamics” simulations, in which PES information is only generated as-and-when required during propagation, thereby circumventing the necessity to pre-fit a PES.

Our recent work, and the focus of this Account, has developed a different approach to MCTDH-type simulations which enables a “direct-dynamics” simulation strategy. Combining regression, machine-learning (ML) and tensor decomposition schemes, along with novel diabatization strategies, we have shown how one can develop “on-the-fly” MCTDH simulations which do not require pre-fitting of a global PES. Below, we describe the origins of this
approach, before describing the extensive improvements and refinements which now enable “direct” MCTDH simulations of increasingly-complex molecular systems moving in mani-foils of coupled electronic states. We conclude by highlighting remaining computational and theoretical challenges for direct dynamics simulations using MCTDH, and offer suggestions of how these might be addressed.

Theory developments

In this Section, we begin by briefly describing the tools underpinning our direct simulation strategy, namely MCTDH, diabatization and kernel ridge regression (KRR). Subsequently, we explain how KRR can be used in “on-the-fly” MCTDH simulations, before showing how tensor decomposition and better kernel functions enable direct MCTDH simulations of increasingly complex molecular systems. An overview of these developments is given in Fig. 1.

Throughout the discussion below, our focus will be on highlighting methodological developments; simultaneously, we will give a number of representative simulations which highlight new capabilities enabled by new simulation tools. For simplicity of presentation, we will assume below that a set of normal-mode coordinates, $q$, (derived from an ab initio Hessian evaluation at a reference configuration, typically a stationary point on the PES) will be used as the active nuclear space for wavefunction propagation; extension to alternative coordinate schemes is noted later.

Quantum dynamics with MCTDH

We begin by outlining the standard method (SM) and MCTDH methods for modelling quantum molecular dynamics in adiabatic and non-adiabatic systems. Together, we describe these as grid-based methods because the wavefunctions are (ultimately) expanded using basis sets of time-independent, grid-like discrete variable representation (DVR) functions along
Figure 1: Overview of developments in direct MCTDH quantum dynamics simulations. Several developments are highlighted in this article, including (a) active learning strategies to generate PESs in tandem with wavefunction, (b) efficient tensor decompositions of PES functions, and (c) novel diabatization schemes.

For a system with $N_s$ electronic states, the SM ansatz on state $s$ is a linear combination of all elements of the tensor product of the DVR basis, $X_J$, with time-dependent coefficients, $C_J^{(s)}$. Using the Dirac-Frenkel variational principle (DFVP), a set of coupled equations-of-motion for the coefficients can then be derived:

$$i\hbar \dot{C}_J^{(s)} = \sum_{u=1}^{N_s} \sum_L \langle X_J|\hat{H}^{(su)}|X_L\rangle C_L^{(u)}.$$  \hspace{1cm} (2)

The method described above provides the numerically-exact solution of the TDSE for the given initial conditions but, because of the exponential increase in computational effort with increasing system size, we must use MCTDH and its smaller time-dependent basis of SPFs
to reduce computational demands.

In MCTDH, each SPF, \( \varphi_{j\kappa}^{(s)}(Q_{\kappa}, t) \), is expanded in the DVRs of a small number (1 ≤ \( n_{\kappa} \leq \sim 4 \)) of degrees-of-freedom, i.e. \( Q_{\kappa} = (q_{k_{1}}, \cdot \cdot \cdot , q_{k_{n_{\kappa}}}) \), and a tensor product of all SPFs, \( \Phi_{J}^{(s)}(Q, t) \), is used to expand the total system wavefunction. Application of the DFVP gives two sets of coupled equations-of-motions, one for the SPFs and one for the wavefunction expansion coefficients:

\[
i\hbar \dot{\varphi}_{j\kappa}^{(s)} = \left(1 - \hat{P}_{j\kappa}^{(s)}\right) \left(\rho_{j\kappa}^{(s)}\right)^{-1} \sum_{u} \langle \hat{H}^{(su)} \rangle^{(\kappa)} \varphi_{u\kappa}^{(s)}, \tag{3a}\]

\[
i\hbar \dot{A}_{J}^{(s)} = \sum_{u} \sum_{L} \langle \Phi_{J}^{(s)}|\hat{H}^{(su)}|\Phi_{L}^{(u)} \rangle A_{L}^{(u)} \tag{3b}\]

Here, \( \hat{P}_{j\kappa}^{(s)} \) projects along mode \( \kappa \) onto the SPF space, and \( \left(\rho_{j\kappa}^{(s)}\right)^{-1} \) is the inverse density matrix associated with \( \kappa \). Constructing a Hartree product of SPFs in all modes except \( \kappa \) gives a function, \( \Phi_{J}^{(s)} \), from which are built single-hole functions, \( \Psi_{j\kappa}^{(s)} = \sum_{J} A_{jJ}^{(s)} \Phi_{J\kappa}^{(s)} \), which are used to construct the elements of the mean-field matrix, \( \langle \hat{H}^{(su)} \rangle_{jl}^{(\kappa)} = \langle \Psi_{j\kappa}^{(s)}|\hat{H}^{(su)}|\Psi_{l\kappa}^{(u)} \rangle \). The standard MCTDH method described above, as well as all of the developments highlighted below, are implemented in the Quantics code;\(^{65}\) further extensive reviews of MCTDH have also been published previously.\(^{22,49,50}\)

An effective implementation of Eq. (3) hinges upon efficient evaluation of the Hamiltonian matrix elements \( \langle \Phi_{J}^{(s)}|\hat{H}^{(su)}|\Phi_{L}^{(u)} \rangle \) and \( \langle \Psi_{j\kappa}^{(s)}|\hat{H}^{(su)}|\Psi_{l\kappa}^{(u)} \rangle \). In order to efficiently evaluate these matrices in the DVR basis set, a standard requirement is that the PES operator is in a ‘sum-of-products’ form; this demands an expensive pre-fitting, in which a set of \( \text{ab initio} \) evaluations of adiabatic energies and non-adiabatic couplings are performed at a range of molecular configurations, before a ‘sum-of-products’ form (for example, in the form of a vibronic coupling Hamiltonian\(^{30,56-61}\)) is generated. This is a key roadblock to ‘direct’ MCTDH simulations; our work below is one route to addressing this challenge.
Diabatization strategies

At this point, it is worth noting that non-adiabatic MCTDH simulations require *diabatic* electronic states and couplings, but typical *ab initio* electronic structure codes only provide *adiabatic* states. As such, diabatization strategies are an important consideration which must be integrated into any new MCTDH methodology.

In our “on-the-fly” MCTDH, described in more detail below, a key design principle is to only use a limited number of local PES evaluations in order to construct the global PES description required for evolution using Eq. (3). As a result, in non-adiabatic simulations, we also only have access to non-adiabatic coupling matrix elements (NACMEs) at selected geometries, rather than across the entire accessible DVR grid; as such, the diabatization schemes used in our direct MCTDH simulations are adapted to this sparse application.

Although discussed in detail elsewhere, Fig. 2 highlights different diabatization schemes which have been developed in our recent work on direct quantum dynamics. Our own work has followed the initial development of the so-called ”propagation” diabatization scheme by Richings and Worth\(^66,67\) (Fig. 2(a)), where the diabatization matrix \(A(q)\) (which transforms from adiabatic to diabatic states at a given configuration, \(q\)) is propagated along a series of related configurations; in the original work, this approach was used in the context of trajectory-based vMCG simulations, and we have subsequently shown how it can be integrated with active learning of the PES, as described below\(^1\). We have also tested the ”projection” diabatization\(^68\) scheme in direct-dynamics (Fig. 2(b)); this method, rather than exploiting the propagation of the diabatization matrix, instead uses a sequence of molecular orbital and expansion coefficient rotation operations designed to bring a CASSCF-type wavefunction at some geometry, \(q\), into best agreement with a CASSCF wavefunction at a reference geometry, \(q_0\), (hence diabatizing the states at \(q\))\(^4,69\). Finally, we have recently proposed the Procrustes diabatization scheme (Fig. 2(c));\(^70\) this approach, compatible with both trajectory-based and grid-based quantum dynamics, uses the idea of maintaining the electronic character of the adiabatic states in the diabatization transformation. This re-
results in a matrix similarity problem known as the orthogonal Procrustes problem, which can be readily solved to give an approximate diabatization matrix. All of these different diabatization schemes are available in the Quantsics simulation package.\textsuperscript{65}

Figure 2: Pictorial representations of diabatization schemes used or developed in our recent work, including (a) propagation diabatization, in which an adiabatic-diabatic transformation matrix, $A$, is propagated between different configurations by integrating $F$, the matrix containing the non-adiabatic coupling terms between all pairs of electronic states, $\psi_i$ and $\psi_j$, (b) projection diabatization, in which $A$ is derived for CASSCF-type wavefunctions by seeking molecular orbital and coefficient rotations to maximize overlap between diabats at different configurations, and (c) Procrustes diabatization, in which one directly seeks $A$ to minimize changes in electronic states between nearby configurations, where $I$ is the unit matrix, $S^A$ is the matrix of overlaps between the adiabatic electronic states, and hence $P$ is the overlap matrix between the diabatic states at $q_0$ and the adiabatic states at $q_1$.

**Kernel ridge regression for PES representation**

With the underlying theory of MCTDH simulations and diabatization described above, we now describe how these methods can be accelerated using ML. As noted above, efficient im-
plementation of MCTDH requires that a global PES is available, or can be generated using \textit{ab initio} electronic structure calculations at a range of molecular configurations. Furthermore, the MCTDH equations-of-motion are most efficient when the PES is represented in a ‘sum-of-products’ form, enabling efficient integral evaluation over SPFs.

The core idea of our work is to replace the pre-fitted PES required in MCTDH with a PES representation that is \textit{automatically} generated during the MCTDH wavefunction propagation. The idea of exploring a PES which is generated during a dynamics calculations has been explored previously, most notably in the GROW strategy developed by Collins and coworkers.\textsuperscript{71,72} The key difference in our work is the focus on using MCTDH as the underlying quantum propagation scheme (in contrast to trajectory-based methods), as well as our emphasis on non-adiabatic simulations. In our work, we opted to move away from the Taylor-series based representations which have been favoured in previous methods, and instead adopt PES representations which are widely-used and well-understood within the ML community. In particular, our focus has been on Gaussian process regression (GPR) and the related KRR method;\textsuperscript{73–76} KRR in particular is a regression strategy which fits the bill in terms of satisfying the demands for a direct MCTDH scheme.

Considering an adiabatic, ground-state PES $V(q)$ to begin, and given a set of $N$ reference points comprising molecular configurations, $\{q^i\}$, and corresponding PES values, $\{V(q^i)\}$, a KRR approximation to the PES can be written as

$$V(q) = \sum_{i=1}^{N} w_i k(q, q^i).$$

(4)

Here, $q^i$ is the molecular configuration of the $i$-th reference point and $q$ is a given molecular configuration at which the PES is required; in the context of MCTDH, this would typically correspond to one of the underlying DVR grid-points. The $k(q, q^i)$ are a set of $N$ \textit{kernel functions}, described in more detail below.

In the simplest interpretation of the KRR method, the weights, $\{w_i\}_{i=1}^{N}$, can be obtained
by ensuring that the KRR PES values are correct at the \( N \) reference points. This leads to an \( N \times N \) matrix equation which is solved to give the \( N \) weights:

\[
\mathbf{w} = \mathbf{Kv}. \tag{5}
\]

Here, \( \mathbf{K} \) is the covariance matrix, defined as

\[
K_{ij} = k(q^i, q^j) + \gamma^2 \delta_{ij}, \tag{6}
\]

and \( \mathbf{v} \) is a vector containing the PES values, \( V(q^i) \), evaluated at the reference configurations. The regularization parameter, \( \gamma^2 \), (typically \( 10^{-8} \)) provides numerical stability to the matrix solution.

There are a number of important aspects about KRR which makes this approach suitable as a PES generation scheme for MCTDH. First, the kernel functions, \( k(q, q^i) \), can be chosen such that Eq. (4) is in the sum-of-products form required for efficient MCTDH simulations; for example, using Gaussian product kernel functions, integrals over the SPFs and individual KRR components can be performed in a straightforward manner. Second, the determination of the weights is straightforward too; as described below, when coupled with active PES learning in MCTDH simulations, the updating of weights as the set of reference configurations changes can also be achieved using efficient matrix-inverse update schemes. Third, the choice of kernel function can be optimized to reduce the number of configurations required for accurate PES description, as described in more detail below. Finally, we can define a variance function,

\[
\sigma^2(q) = k(q, q) + \gamma^2 - k^T K^{-1} k, \tag{7}
\]

which gives a measure of the accuracy of the KRR fit at any point, \( q \), and can thus be used to guide the selection of additional reference points. Given these advantages, the rest of this Account implicitly assumes the use of KRR for PES generation in MCTDH simulations.
It is also worth noting that the KRR scheme above is directly applicable to non-adiabatic simulations too. Specifically, if a set of diabatic electronic states and inter-state couplings are available at a set of $N$ molecular configurations (for example, obtained by using one of the diabatization schemes noted in Fig. 2), then KRR representations can be generated directly for both the diagonal (state) contributions, $V_{jj}(q)$, and the couplings, $V_{jk}(q)$. This multi-state PES representation can then be used directly in non-adiabatic MCTDH simulations.

With KRR providing a PES for wavefunction propagation, the remaining aspect of a direct dynamics strategy is the generation of the reference configurations for PES construction; this is discussed in the next section, where we also give examples of KRR-based direct dynamics simulations using MCTDH.

**Direct quantum dynamics with active PES learning**

Although KRR offers a simple approach to generating a global PES representation for MCTDH simulations, an important idea that helps turn this approach into an efficient direct dynamics strategy is to use active learning to expand the KRR reference points throughout a simulation. The key idea underlying this approach is the fact that the PES on which a wavefunction evolves does not necessarily need to be accurate across all molecular configurations at all times; in other words, the choice of reference configurations can be dynamically updated to provide an accurate description of the PES only in the region in which the wavefunction adopts significant amplitude. This approach, outlined in Fig. 3, enables a direct dynamics MCTDH scheme where molecular configurations for PES construction are iteratively added in tandem with wavefunction propagation.

The initial implementation of our grid-based direct-dynamics scheme\textsuperscript{1,2} used a full kernel
Figure 3: Active learning strategy for on-the-fly quantum dynamics simulations. (a) Starting from an initial nuclear wavefunction $\Theta(q; t_0)$, reference configurations (black circles) are generated from an approximate sampling function. (b) PES evaluations at the reference configurations are used to generate a KRR PES representation. (c) Subsequent wavefunction propagation explores new regions of configuration space; new points are sampled and added to the KRR PES, leading to a PES representation which evolves in tandem with $\Theta(q; t)$.

with the following form:

$$k^{\text{Full}}(q, q^i) = \exp \left( -\alpha |q - q^i|^2 \right)$$

$$= \prod_{f=1}^{N_f} \exp \left( -\alpha \left( q_f - q^i_f \right)^2 \right)$$

$$= \prod_{f=1}^{N_f} k(q_f, q^i_f).$$

(8)
Here, $\alpha$ is an adjustable parameter determining the length-scale of the kernel in configuration space. Importantly, the representation of the KRR PES using Eqs. (4) and (8) has the required ‘sum-of-products’ form for MCTDH. Given this kernel, active-learning construction of the PES (concurrent with nuclear dynamics), proceeds as follows. First, the PES value is evaluated at the center of the initial wavepacket (as given by the expectation value of the position operator in each degree-of-freedom). At regular times (usually every 1 fs), beginning with $t = 0$ fs, we then randomly sample molecular geometries in an $N_f$-dimensional hyper-cuboid in coordinate space, centered on the middle of the wavepacket on the given state (the procedure is repeated for all states), the dimensions of which are determined by the wavefunction width along each degree-of-freedom (typically, we use six times the wavepacket width). At each selected geometry, we evaluate the variance (Eq. (7)) using the current database of reference points; if the variance is greater than some user-defined value, the KRR PES at the selected point is deemed insufficiently accurate and the PES (and non-adiabatic couplings, if required) is evaluated and added to the database of KRR reference points. After this sampling, the KRR PES weights are re-generated by solving Eq. (5), and wavefunction propagation proceeds until the next sampling time. In this way, the KRR PES is generated alongside wavefunction propagation; we also note that this approach can be iterated to generate increasingly accurate and complete KRR PESs, if required.

**Examples of initial methodology**

The capabilities of our initial combination of MCTDH, active learning, and diabatization, are highlighted in Fig. 4.

The initial test of our active learning scheme focussed on simulating the dynamics of butatriene using a 2-mode model with 2 electronic states, employing propagation diabatization to generate diabatic states (Fig. 4(a)); here, direct dynamics simulations (using SM) performed using *ab initio* electronic structure calculations (CASSCF/3-21G) gave comparable diabatic state populations when compared to previous simulations on a pre-fitted PES.\textsuperscript{1}
Over the first 20 fs of the dynamics the agreement is very good and remains qualitatively so for the remaining 80 fs; differences are to be expected because a proper torsion coordinate was used with the pre-fitted PESs, whilst we were restricted to using a torsion-like normal mode with the on-the-fly calculation.

Subsequent simulations of a 3-mode model of proton tunnelling in malonaldehyde (again, using SM) further verified the accuracy of our initial approach (Fig. 4(b)), demonstrating excellent agreement between simulations performed on full PES grids without active learning. Here, 2545 total energy evaluations were used in our scheme, in contrast to 9261 required in the full-grid scheme. Increasing dimensionality further, we subsequently demonstrated 6-mode simulations of proton tunnelling in salicylaldimine; because this system was too large to model accurately with standard full-grid schemes, we instead used MCTDH. Simulations performed on both an analytical PES, and directly using Hartree-Fock energy evaluations, demonstrated the capability of this scheme in reproducing subtle features of the coupled nuclear dynamics (Fig. 4(c)), namely the flux of the wavepacket through a dividing surface, which is particularly sensitive to the form of the PES. The initial behaviour of the flux agrees well between the results using the fitted surface (dashed blue line in Fig. 4(c)) and those using KRR; this agreement is particularly good for the duration of the dynamics when using the full database of energies sampled from the fitted function (solid purple line in Fig. 4(c)(ii)) but, as expected, less so when using ab initio energies (Fig. 4(c)(i)) because the potential is no longer restricted to the functional form chosen when fitting the PES. These calculations also highlighted an important drawback, namely the large number (> 10^4) of energy evaluations required in the KRR scheme; however, this is still much smaller than the < 400 × 10^8 points in the full DVR grid. Based on these initial observations, it became clear that further improvements were required to expand the scope of the direct MCTDH simulations strategy, prompting the developments now described below.
(a) Butatriene

(b) Malonaldehyde

(c) Salicylaldimine

Figure 4: Examples of direct quantum dynamics methods using original developments. (a) Non-adiabatic dynamics in a 2-mode/2-state model of butatriene, and proton tunnelling in (b) a 3-mode model of malonaldehyde, and (c) a 6-mode model of salicylaldimine: (i) KRR PES fitted using energies sampled from the analytical PES; (ii) KRR PES fitted using \textit{ab initio} calculated energies.

**Improving simulation efficiency**

The direct MCTDH method described above enables wavefunction propagation without a pre-fitted a global PES. In the initial implementation, however, we found that simulations speeds are dramatically impacted as the number of elements in the KRR PES representation grows; in this Section, we show how our improved direct MCTDH methodology addresses this challenge.

**Additive Kernels.** Having demonstrated the fundamental accuracy of the KRR-generated PES by inference from the excellent agreement with the observables resulting from initial low-dimensional calculations (Fig. 4), it was clear that the computational effort required
to achieve similar accuracy with even moderate numbers of degrees-of-freedom would be prohibitive and, as such, improvements in efficiency were required.

As noted above, a fundamental issue in efficiency from these earlier calculations was found to be the large number of terms in the KRR PES expansion (Eq. (4)) when using the full kernel (Eq. (8)), each of which must be treated separately when evaluating the MCTDH equations-of-motion, particularly when transforming PES integrals in the DVR basis to the time-dependent SPF basis. We note that such an issue does not affect the SM, where the DVR integrals arising from each term in Eq. (4) can simply be summed after the sampling step (e.g. every 1 fs) and the results used in every step of the propagation until the next set of reference points is added to the database; however, the SM is still limited by strong exponential scaling of computational effort.

The problem with the efficiency of direct MCTDH is exacerbated by the scaling of the number of terms required in Eq. (4) as the number of degrees-of-freedom increases when using the full kernel. The full-width-at-half-maximum (FWHM) of the Gaussian function along each degree-of-freedom is $2\sqrt{\alpha \ln 2}$; this measure roughly approximates the range of influence of the Gaussian, hence the total volume, $V$, in $N_f$-dimensions is the volume of the hypersphere given by

$$V = \frac{(4\pi \alpha \ln 2)^{N_f/2}}{\Gamma(N_f/2 + 1)}. \quad (9)$$

As such, for DVR grids with a representative length of $l$ along all degrees-of-freedom, the proportion of volume of the DVR grid (volume $V'$) occupied by each kernel function is

$$P = \frac{(2\sqrt{\pi \alpha \ln 2})^{N_f}}{l^{N_f} \Gamma(N_f/2 + 1)}. \quad (10)$$

So, if $l > 2\sqrt{\alpha \ln 2}$, $P$ decreases exponentially with increasing dimension $N_f$, meaning that the number of reference points must similarly increase to maintain the accuracy of the PES representation.$^3$

To overcome this difficulty, we introduced an additive kernel$^3$ that uses a sum of Gauss-
sian functions with increasing dimension:

\[ k^{\text{Add}}(q, q^n) = \sum_{\kappa=1}^{N_f} k(q_\kappa, q^n_\kappa) + \sum_{\kappa<\lambda}^{N_f} k(q_\kappa, q^n_\kappa)k(q_\lambda, q^n_\lambda) + \cdots + k^{\text{Full}}(q, q^n). \]  

(11)

By truncating this kernel at two- or three-dimensional terms, we can expand the PES in terms of sums of low-dimensional functions which capture the important coupling between degrees-of-freedom. Such kernels maintain the symmetry and semi-positive-definiteness of \( K \), hence the approach to KRR PES fitting is identical to that described in the section, Direct quantum dynamics with active PES learning.

The additive kernel exhibits better coverage of configuration space than the full kernel. For example, closer analysis shows that the fraction of an \( N_f \)-dimensional DVR product grid (length \( l \) in each dimension, \( V = l^{N_f} \)) occupied by the one-dimensional terms in Eq. 11 is

\[ P = \frac{2\sqrt{\alpha\ln 2}}{l}. \]  

(12)

Notably, this fraction is independent of \( N_f \) so, even for increasing dimensionality, the 1-dimensional terms are non-zero in a constant proportion of space. Similarly for two-dimensional terms, we find:

\[ P = \frac{4\pi\alpha\ln 2}{l^2}, \]  

(13)

where, again, the proportion is independent of \( N_f \) (although it is obviously smaller than for one-dimensional case). The upshot is, as demonstrated in our previous work and below, that the use of additive kernels enables much greater efficiency in PES generation.

**Tensor decomposition of PES.** The additive kernel of Eq. 11 reduces the number of reference points required to represent the PES at a given accuracy (and hence reduces the number of electronic structure calculations needed), but the speed-up in MCTDH dynamics does not immediately manifest itself because the kernel contains a large number of individual terms, each of which must be treated separately. It was therefore necessary to further reduce
the number of terms in the PES expansion; this can be achieved by secondary fitting of the KRR PES using decomposition. \(^3\)

In our implementation of this idea, a secondary PES decomposition is performed in addition to the KRR fitting procedure outlined above. First, the KRR PES is evaluated at the DVR gridpoints along each degree-of-freedom (with the remaining coordinates fixed at \(q = 0\)), giving a set of one-dimensional contributions to the PES expansion and reducing the number of one-dimensional terms from \(N \times N_f\) to \(N_f\) (using the additive kernel).

A second step then accounts for couplings between each pair of degrees-of-freedom; to do so, singular value decomposition (SVD) is used. A two-dimensional grid of coordinates is created from the DVR gridpoints of a pair of degrees-of-freedom, and the KRR PES evaluated at each point. The one-dimensional PES terms, calculated previously, are then subtracted from these points to give a two-dimensional residue matrix \(V^{\kappa \lambda}\), which is subsequently subjected to SVD:

\[
V^{\kappa \lambda} = U \Sigma W^T. \tag{14}
\]

It follows that the value of the residue at some point, \((q^i_\kappa, q^j_\lambda)\), is (in order of decreasing singular values, \(\sigma_k\)), given by

\[
V^{\kappa \lambda}(q^i_\kappa, q^j_\lambda) = \sum_{k=1}^{\min(N_\kappa, N_\lambda)} \sigma_k u_{ik} w_{jk}, \tag{15}
\]

and hence the contributions to the two-dimensional term along each degree-of-freedom are

\[
V^{\kappa \lambda}_{\kappa(k)}(q^i_\kappa) = \sqrt{\sigma_k} u_{ik} \tag{16a}
\]

\[
V^{\kappa \lambda}_{\lambda(k)}(q^j_\lambda) = \sqrt{\sigma_k} w_{jk}. \tag{16b}
\]
To exclude insignificant terms we define a residual error,

\[ ||R||^2 = \sum_{k=m+1}^{\min(N_a,N_b)} \sigma_k^2, \quad (17) \]

where \( m \) is chosen so that \( ||R||^2 \) is less than a pre-chosen parameter (usually \( 10^{-3} \)). Of course, this decomposition can be extended to three-dimensional terms using tensor decomposition (and could similarly be extended further); in such cases, it is necessary to use the full kernel, or additive kernel including terms of at least three dimensions.\(^4\)

**Cholesky Updates, Sobol sequences and spin-forbidden processes.** In addition to the improvements highlighted above, it is also worth noting three further updates to our direct MCTDH scheme which help expand its scope. First, we have shown that simulation times when using active learning and KRR can be dramatically reduced by Cholesky update procedures.\(^4\) Here, instead of performing inversion of the covariance matrix, \( K \), at the end of each sampling step to solve Eqs. (5) and (7), whenever new reference configurations are added to the KRR set, one can instead use the well-known Cholesky update scheme; tests for a six-mode model of salicylaldimine demonstrated that this simple change can reduce direct MCTDH calculation times by an order-of-magnitude. As a by-product of this approach, more efficient KRR sampling is achieved because the variance calculated at each sampled point is based on all previous reference points, rather than just those added in previous sampling steps; use of an up-to-date variance can lead to fewer points being required.

Second, we have found that random sampling to generate KRR reference points can introduce an unwanted non-deterministic nature to our direct MCTDH simulations; we subsequently showed that adopting deterministic quasi-random sampling (particularly based on Sobol sequences\(^78-80\)) removes this limitation, such that simulations with the same initial conditions and sampling set-up will reproducibly yield the same results.\(^4\)

Finally, we have recently shown how our direct grid-based wavefunction simulation scheme can be extended to model systems in which spin-forbidden processes (\( i.e. \) inter-system cross-
ing) play a role; here, we have demonstrated how the influence of spin-orbit coupling can be integrated into our KRR PES generation scheme, and confirmed our results in simulations of SO₂ and thioformaldehyde.⁸¹

**Example simulations for updated methodology**

Following on from Fig. 4, Fig. 5 highlights the state-of-the-art capabilities of our direct MCTDH scheme, employing the algorithmic improvements noted above.³,⁶⁹,⁸¹,⁸² As shown in Fig. 5, the range of molecular systems which have now been studied using our MCTDH schemes has moved beyond the initial proof-of-concepts in Fig. 4. Fig. 5(a) first highlights the impact of using additive kernel functions. Here, 6-mode simulations of salicylaldimine with a full 6-D product kernel give poor agreement with the expected MCTDH results because of poor coverage of the high-dimensional PES; in contrast, the additive-kernel/SVD method reproduces the expected time-dependent proton flux to a very high degree-of-accuracy.³ Even more encouragingly, the full-dimensional kernel simulations used > 10⁴ energy evaluations to generate the KRR PES, whereas the additive-kernel/SVD used just ≃ 2,200.

An even more ambitious MCTDH simulation is shown in Fig. 5(b), illustrating the time-dependent excited-state population in a 12-mode/2-state model of pyrazine. These simulations employed SVD fitting and projection diabatization, and used a total of 2,744 direct CASSCF(10,8)/DZP ab initio calculations to build the KRR representation of the PES.³ Here, the calculated properties, including transient state populations and calculated absorption spectra, are in good qualitative agreement with previous MCTDH simulations using pre-fitted PESs. Finally, we highlight the results of Fig. 5(c), which show recent full-dimensional (6-D) non-adiabatic MCTDH simulations of thioformaldehyde, using SVD processing, Procrustes diabatization, and account of spin-orbit coupling; simulations were performed directly using CASSCF/cc-PVTZ.⁸¹ These simulations only required 303 ab initio energy evaluations; the ability to perform these direct simulations, without the requirement of PES pre-fitting, is in our opinion an extremely useful addition to the MCTDH simulation
methodology.

Figure 5: Examples of simulations performed with improved direct grid-based dynamics method; recent molecules studied include thioformaldehyde, pyrazine, salicylaldimine, ethene, and several mycosporine-like amino acid molecules possessing six-membered rings (top-left).

Conclusions

This Account has highlighted efforts to develop direct dynamics simulation schemes by merging established wavefunction propagation methods with ML and novel diabatization strategies (Fig. 1). Starting from simple KRR PESs in the “standard” grid-based approach to modelling quantum dynamics, we have highlighted how subsequent developments now enable increasingly complex molecular systems to be simulated for periods of time which enable comparison to either experiments or alternative simulations. To date, our largest direct MCTDH/SM simulations have comprised up to 12 nuclear degrees-of-freedom (with
3 states) and at most 12 electronic states (with 3 degrees-of-freedom).

Current work is focussed on pushing this size-limit even further, using a combination of methodological and computational developments. A particular point of interest is the development of direct MCTDH schemes which can employ alternative coordinate schemes, going beyond standard normal-mode coordinates, such as hyperspherical or valence coordinates. Allowing the use of other coordinate systems would make the treatment of systems involving, for example, *cis-trans* isomerisation or dissociation more practical, the use of a single coordinate to replace multiple normal modes potentially reducing the computational effort significantly. In such cases, these non-orthogonal coordinates introduce more complex kinetic energy operators. We have focussed on normal modes so far, as the kinetic energy operator has a particularly simple form; however, we anticipate that a KRR-type method could be used to ‘learn’ the underlying metric tensor transformations which are required in evaluating the associated matrix elements for wavefunction propagation. A related aspect is the choice of kernel in the KRR fitting; to date we have used Gaussian functions in both the full (Eq. (8)) and additive (Eq. (11)) kernels but KRR does not restrict us to these choices, only that the kernels should form a symmetric and positive semi-definite covariance matrix (Eq. (6)). Such kernels are not ideal for dissociation modes, for example, where the PES is formed of a steep barrier becoming a gently curved minimum and ending nearly flat as the dissociation limit is approached; in such cases, the spacing of sampled points is unaffected by the shape of the PES, but ideally we would want the density of those points to be proportional to the gradient of the PES, with the kernel width being inversely related to the gradient. The implementation of different kernels is straightforward, so a study on the appropriate choices for different modes will be a feature of our upcoming work. Another challenging aspect of MCTDH-based schemes is in judging convergence with respect to SPFs; an interesting prospect here is to employ a hierarchical scheme, in which small SPF sets are used in initial exploration of PESs, enabling initial approximation generation by KRR, followed by increasingly large calculations in which the quality of the KRR PES is
monitored to assess simulation convergence. These, and other, topics are ongoing domains of interest and development.

All of the simulation strategies highlighted in this Account are now available in the established *Quantics* simulation code, providing opportunities for further developments and application within the wider chemical dynamics community.

**Acknowledgement**

The authors gratefully acknowledge support from the Engineering and Physical Sciences Research Council (EPSRC), UK (grant EP/S028986/1), and The Leverhulme Trust (RPG-2016-055). The authors also gratefully acknowledge the Scientific Computing Research Technology Platform (SCRTP) at the University of Warwick for the provision of high-performance computing facilities.

**Author biographies**

**Gareth Richings** is a Senior Research Fellow in the Department of Chemistry at the University of Warwick, UK. His research interest is the nuclear quantum dynamics of non-adiabatic systems. He is an author of the *Quantics* package.

**Scott Habershon** is Professor of Computational and Theoretical Chemistry in the Department of Chemistry at the University of Warwick, UK. His research interests include development of simulation methods for predictive chemical dynamics, particularly photochemical modelling and reaction-mechanism discovery.
References


(60) Penfold, T. J.; Spesvytsev, R.; Kirkby, M.; Minns, R. S.; Parker, D.; Fielding, H.; Worth, G. Quantum dynamics study of the competing ultrafast intersystem crossing


35


