

Smooth Things Come in Threes: A Diabatic Surrogate Model for Conical Intersection Optimization

Ignacio Fdez. Galván^{ID*,†} and Roland Lindh^{ID*,†,‡}

[†]*Department of Chemistry – BMC, Uppsala University, P. O. Box 576, SE-75123 Uppsala, Sweden*

[‡]*Uppsala Center for Computational Chemistry (UC₃), Uppsala University, PO Box 576, SE-751 23 Uppsala. Sweden*

E-mail: Ignacio.Fernandez@kemi.uu.se; roland.lindh@kemi.uu.se

Abstract

The optimization of conical intersection structures is complicated by the non-differentiability of the adiabatic potential energy surfaces. In this work, we build a pseudo-diabatic surrogate model, based on Gaussian process regression, formed by three smooth and differentiable surfaces that can adequately reproduce the adiabatic surfaces. Using this model with the restricted variance optimization method results in a notable decrease of the overall computational effort required to obtain minimum energy crossing points.

1 Introduction

Nonadiabatic processes, in which more than one Born–Oppenheimer potential energy surface (PES) affect the nuclear motion, are involved in many photophysical and photochemical phenomena, such as vision,¹ chemi- and bioluminescence,² DNA photostability,³ photosynthesis,⁴ etc. The theoretical study of such processes has been greatly developed during the past couple of decades, and involves mostly the use of quantum or mixed quantum–classical molecular dynamics simulations.^{5,6}

One of the characteristics of nonadiabatic processes is the degeneracy or near-degeneracy between adiabatic electronic states. A particular salient feature is the existence of conical intersections (CI), which have also been the object of a multitude of recent works.^{7–11} The geometries or structures where CIs occur are not isolated, but form a continuous subspace of geometries, and the most relevant regions of this subspace will be those most frequently traversed during the dynamics. It is common, however, to carry out static studies, as a complement or instead of dynamics simulations, where only the energies with lowest geometries are identified.^{12–15} These geometries are known as minimum energy CIs, or more generally as minimum energy crossing points (MECP),

The location of significant points in PESs is a fundamental task in computational studies. Over the years, a conventional paradigm for geometry optimization has emerged as

robust and efficient and is the most commonly used.¹⁶ This is based on a second-order Taylor expansion of the PES, a step size restriction, approximate Hessian and Hessian update methods. A prime example of such “conventional” methods is the restricted-step rational function optimization (RS-RFO) in redundant internal coordinates.¹⁷ The second-order expansion has some limitations, in particular it cannot accurately represent the parent surface beyond a local region around the expansion point, and this has pushed us to propose and develop an alternative optimization scheme based on a more flexible surrogate model. The new method, which we have called restricted variance optimization (RVO),^{18–20} relies on a surrogate model generated with a Gaussian process regression (GPR) variant also known as gradient-enhanced Kriging (GEK).^{21–23} The most relevant differences with respect to similar methods proposed by other authors^{24–26} is that RVO uses the empirical knowledge encoded in the approximate Hessian model function (HMF)²⁷ to define the so-called characteristic lengths of the model in internal coordinates, and that it uses the predicted uncertainty of the model to restrict the displacement during the iterations.

For the specific case of MECP optimization, there have been a number of proposed methods, generally using projection techniques or penalty functions to ensure that the energies of two crossing states are degenerate and simultaneously minimize their value.^{28–34} We have previously used the projected constrained optimization method (PCO)^{31,35,36} to successfully optimize MECPs, by including a constraint involving the energy difference. However, adapting this method to RVO is not straightforward. Firstly, although purely geometrical constraints have been implemented,¹⁹ including the energy difference would require a surrogate model that can represent accurately the energy difference itself and its gradient. Secondly, the very nature of a CI means that the PESs involved in the crossing are not differentiable at the crossing points, and this poses challenges for a surrogate model that relies on differentiability such as GEK. Lastly, for an efficient location of CIs, knowledge of the nonadiabatic coupling vector, or a sufficiently good approximation, is very valuable, and it would be desirable to include in the surrogate model as well.

In this work, we extend the RVO method to allow optimization of MECPs, either between states of the same spacial and spin symmetry (CIs) or different symmetries (e.g. singlet–triplet crossings). To this end, we build a pseudo-diabatic surrogate model from the data (energies, gradients and couplings) of the previous iterations. The model consists of three separate smooth and differentiable surfaces (two in the case of different-symmetry crossings) that when combined can reproduce the energies, gradients and couplings of the parent method, and thus can be used in combination with the constrained RVO.¹⁹ In section 2, the methodological details relevant for this work are detailed, in particular the construction of a surrogate model consistent with the presence of CIs. The performance of this method was tested in a set of MECP optimizations, for which the computational details are given in section 3, and the corresponding results are discussed in section 4. Finally, we summarize the work in section 5.

2 Theory and methods

This section includes a summary on the local description of conical intersections, followed by details on how to switch between diabatic and adiabatic representations, and how a smoothly interpolating surrogate model is constructed, to finish with a short overview of the optimization method.

2.1 Conical intersections

Conical intersections are features of most molecular systems, where two adiabatic electronic PES are exactly degenerate (although not all degeneracies correspond to CIs). They were once considered an academic curiosity, but are nowadays known to be ubiquitous in molecular systems and with very significant consequences for their photophysical and photochemical behavior. CIs have been extensively studied and described before,^{7-11,37-41} and here only the most relevant aspects for the rest of the article will be given.

In the absence of spin-orbit coupling, the degeneracy at a CI is lifted linearly with the displacement when the geometry of the system is distorted in one of two independent directions (or any combination thereof), while it is maintained for any other orthogonal displacement. Thus, the set of geometries where the two surfaces touch, the *intersection space*, is a subspace of $K - 2$ dimensions, where K is the dimensionality of the PES. At each point of the intersection space, the 2-dimensional subspace that breaks the degeneracy is known as the *branching plane*. The branching plane can be defined as the subspace spanned by two (generally nonorthogonal) vectors, the gradient difference, \mathbf{g} , and the nonadiabatic coupling (NAC), \mathbf{h} . For completeness, it is also useful to define the average gradient vector, \mathbf{s} . These can be obtained as:

$$\mathbf{s} = \frac{1}{2}(\nabla E^A + \nabla E^B) \quad (1)$$

$$\mathbf{g} = \frac{1}{2}(\nabla E^A - \nabla E^B) \quad (2)$$

$$\mathbf{h} = (E^A - E^B) \langle \Psi^A | \nabla \Psi^B \rangle \quad , \quad (3)$$

where E^A, E^B and Ψ^A, Ψ^B are the energies and wave functions of the two degenerate states. In eq. (3), the factor $E^A - E^B$ (which is identically zero in the intersection space) effectively cancels an equal denominator in $\langle \Psi^A | \nabla \Psi^B \rangle$, such that a nonzero \mathbf{h} is obtained.^{36,37} The degeneracy between the two states means that the two wave functions are not uniquely defined, as any unitary transformation of them is an equally good possibility. This also means that \mathbf{g} and \mathbf{h} are not uniquely defined for structures in the intersection space, but the branching plane – the subspace spanned by them – is. Indeed, a “rotation” of the two wavefunctions by an angle χ results in a corresponding rotation of the \mathbf{g} and \mathbf{h} vectors by an angle 2χ :

$$(\Psi^A \quad \Psi^B) \leftarrow (\Psi^A \quad \Psi^B) \begin{pmatrix} \cos \chi & -\sin \chi \\ \sin \chi & \cos \chi \end{pmatrix} \quad (4)$$

$$\begin{pmatrix} \vdots & \vdots \\ \mathbf{g} & \mathbf{h} \\ \vdots & \vdots \end{pmatrix} \leftarrow \begin{pmatrix} \vdots & \vdots \\ \mathbf{g} & \mathbf{h} \\ \vdots & \vdots \end{pmatrix} \begin{pmatrix} \cos 2\chi & -\sin 2\chi \\ \sin 2\chi & \cos 2\chi \end{pmatrix} \quad , \quad (5)$$

where the vertical dots simply indicate that the vectors are arranged as columns. A pair of orthonormal vectors, $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$, unique up to permutations and sign flips (except in highly symmetrical cases), to span the branching plane,³⁶ and these are the coordinates used for the branching plane in the figures of this article.

The adiabatic PESs around a CI, represented in the branching plane, have the familiar double-cone shape, with the two surfaces touching at the intersection point and diverging as the structure moves away from it (fig. 1). This shape makes the surfaces not differentiable at the intersection, which is problematic for optimization and dynamics methods,

typically based on PES gradients. The location of CI structures, or in general of crossing points between adiabatic surfaces, is usually done by including some kind of constraint or penalty that enforces a zero energy difference between the surfaces.^{28–30,32–34} We make use of the PCO method,^{31,35,36} that allows general arbitrary constraints and requires, at each geometry, the adiabatic energies E^A and E^B , and the \mathbf{s} , \mathbf{g} , and \mathbf{h} vectors.

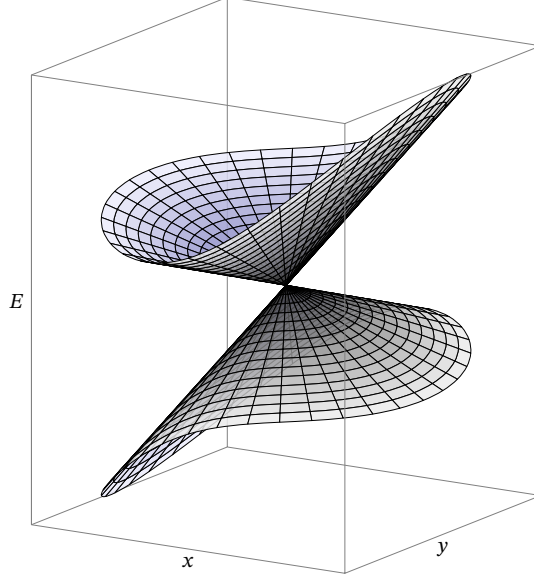


Figure 1: Representation of the adiabatic PESs around a conical intersection, in the branching plane, spanned by the x and y coordinates, with the energy along the vertical axis.

2.2 Diabatization

The lack of differentiability of the PESs can be avoided by switching to a diabatic representation of the electronic states, instead of adiabatic. Such transformation, known as diabatization, is commonly used in dynamics simulations, and there are many techniques to achieve it which are described and overviewed elsewhere.⁴² A strictly diabatic basis, in which the so-called nuclear-momentum coupling vanishes, does not in general exist,⁴³ so in practice one resorts to a quasi-diabatic basis, where the couplings are reduced to a negligible or acceptable size.⁴²

In our case, our only goal is to obtain continuous, differentiable functions that can accurately reproduce the adiabatic surfaces around a CI. For this, we consider a simple linear two-state model, in which the elements of the Hamiltonian matrix are linear functions of the nuclear coordinates \mathbf{q} :^{36,44,45}

$$\mathbf{H}(\mathbf{q}) = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \quad (6)$$

$$H_{11} = \alpha(\mathbf{q}) = \alpha_0 + \mathbf{k}_\alpha^\top \mathbf{q} \quad (7)$$

$$H_{22} = \beta(\mathbf{q}) = \beta_0 + \mathbf{k}_\beta^\top \mathbf{q} \quad (8)$$

$$H_{12} = H_{21} = \gamma(\mathbf{q}) = \gamma_0 + \mathbf{k}_\gamma^\top \mathbf{q} \quad (9)$$

Diagonalization of \mathbf{H} yields the adiabatic energies E^A and E^B as eigenvalues. This can be compactly expressed in terms of the average (τ) and half-difference (δ) energies, and inter-

preting δ and γ as the two components of a vector, with modulus λ and argument ω :

$$\tau(\mathbf{q}) = \frac{1}{2} (\alpha(\mathbf{q}) + \beta(\mathbf{q})) \quad (10)$$

$$\delta(\mathbf{q}) = \frac{1}{2} (\alpha(\mathbf{q}) - \beta(\mathbf{q})) \quad (11)$$

$$\lambda(\mathbf{q}) = \sqrt{\delta(\mathbf{q})^2 + \gamma(\mathbf{q})^2} \quad (12)$$

$$\omega(\mathbf{q}) = \text{atan2}(\gamma(\mathbf{q}), \delta(\mathbf{q})) \quad , \quad (13)$$

where the function $\text{atan2}(y, x)$ is similar to $\arctan(y/x)$, but returns an angle in the correct quadrant according to the signs of the two arguments. Then E^A and E^B are given by:

$$E^A(\mathbf{q}) = \tau(\mathbf{q}) + \lambda(\mathbf{q}) \quad (14)$$

$$E^B(\mathbf{q}) = \tau(\mathbf{q}) - \lambda(\mathbf{q}) \quad . \quad (15)$$

The vectors \mathbf{s} , \mathbf{g} and \mathbf{h} are obtained from \mathbf{k}_α , \mathbf{k}_β and \mathbf{k}_γ :

$$\mathbf{s}(\mathbf{q}) = \mathbf{k}_\tau = \frac{1}{2} (\mathbf{k}_\alpha + \mathbf{k}_\beta) \quad (16)$$

$$\mathbf{k}_\delta = \frac{1}{2} (\mathbf{k}_\alpha - \mathbf{k}_\beta) \quad (17)$$

$$\begin{pmatrix} \vdots & \vdots \\ \mathbf{g}(\mathbf{q}) & \mathbf{h}(\mathbf{q}) \\ \vdots & \vdots \end{pmatrix} = \begin{pmatrix} \vdots & \vdots \\ \mathbf{k}_\delta & \mathbf{k}_\gamma \\ \vdots & \vdots \end{pmatrix} \begin{pmatrix} \cos \omega(\mathbf{q}) & -\sin \omega(\mathbf{q}) \\ \sin \omega(\mathbf{q}) & \cos \omega(\mathbf{q}) \end{pmatrix} \quad . \quad (18)$$

We are interested in the reverse process (diabatization), i.e., obtaining the diabatic properties (α, β, γ) from the adiabatic ones $(E^A, E^B, \mathbf{s}, \mathbf{g}, \mathbf{h})$. This would be trivial if we knew the angle ω , but as it turns out, it cannot be deduced from the adiabatic data alone. In fact, the diabatization is not well defined because different sets of linear α, β, γ can lead to the same adiabatic PESs. In principle, any one of those sets is equally valid, but when performing this process at different \mathbf{q} , we would like to always obtain the same solution. The possible solutions correspond to the different values of the angle ω , so in order to obtain a consistent solution, we must choose ω appropriately.

Let us choose an arbitrary reference structure \mathbf{q}^{ref} , and define $\omega(\mathbf{q}^{\text{ref}}) = 0$. This gives

$$\mathbf{k}_\delta = \mathbf{g}(\mathbf{q}^{\text{ref}}) \quad (19)$$

$$\mathbf{k}_\gamma = \mathbf{h}(\mathbf{q}^{\text{ref}}) \quad . \quad (20)$$

For any other \mathbf{q} , the angle ω can be obtained from

$$\begin{pmatrix} \vdots & \vdots \\ \mathbf{k}_\delta & \mathbf{k}_\gamma \\ \vdots & \vdots \end{pmatrix}^+ \begin{pmatrix} \vdots & \vdots \\ \mathbf{g}(\mathbf{q}) & \mathbf{h}(\mathbf{q}) \\ \vdots & \vdots \end{pmatrix} = \begin{pmatrix} \cos \omega(\mathbf{q}) & -\sin \omega(\mathbf{q}) \\ \sin \omega(\mathbf{q}) & \cos \omega(\mathbf{q}) \end{pmatrix} \quad , \quad (21)$$

where \mathbf{A}^+ denotes the Moore–Penrose inverse of \mathbf{A} . With this ω , the same linear functions for α, β and γ will be obtained from any \mathbf{q} . Different choices of \mathbf{q}^{ref} will result in different diabatic functions, but any of them reproduces the adiabatic data.

To extend the linear model to more general functions, we start by replacing \mathbf{k}_x with

$\nabla x(\mathbf{q})$ ($x \in \{\alpha, \beta, \gamma, \delta, \tau\}$). We note that the diabatic-to-adiabatic transformation above still holds, and most of the diabaticization would work too, only the selection of ω needs to be modified, because the left-hand side in eq. (21) is now unlikely to produce an orthogonal 2×2 matrix from which ω can be extracted.

As before, we can select an arbitrary structure as \mathbf{q}^{ref} and define the (constant) \mathbf{k}_δ and \mathbf{k}_γ with eqs. (19) and (20). We realize that in the linear model the \mathbf{g} and \mathbf{h} vectors always span the same plane. In a more general case, the $\{\mathbf{g}, \mathbf{h}\}$ plane changes with \mathbf{q} . So, for any other \mathbf{q} , we first transform \mathbf{g} and \mathbf{h} such that they lie in the same plane as \mathbf{k}_δ and \mathbf{k}_γ . This transformation is based on the singular value decomposition (svd) of the inner product matrix between the two subspaces, and is the “minimal” rotation that achieves it:

$$G(\mathbf{q}) = \text{orth} \begin{pmatrix} \vdots & \vdots \\ \mathbf{g}(\mathbf{q}) & \mathbf{h}(\mathbf{q}) \\ \vdots & \vdots \end{pmatrix} \quad (22)$$

$$K = \text{orth} \begin{pmatrix} \vdots & \vdots \\ \mathbf{k}_\delta & \mathbf{k}_\gamma \\ \vdots & \vdots \end{pmatrix} \quad (23)$$

$$G(\mathbf{q})^\top K = U(\mathbf{q}) \text{diag}(\phi_1(\mathbf{q}), \phi_2(\mathbf{q})) V(\mathbf{q})^\top \quad (\text{svd of } G(\mathbf{q})^\top K) \quad (24)$$

$$\begin{pmatrix} \vdots & \vdots \\ \mathbf{g}'(\mathbf{q}) & \mathbf{h}'(\mathbf{q}) \\ \vdots & \vdots \end{pmatrix} = (KV(\mathbf{q})) (G(\mathbf{q})U(\mathbf{q}))^\top \begin{pmatrix} \vdots & \vdots \\ \mathbf{g}(\mathbf{q}) & \mathbf{h}(\mathbf{q}) \\ \vdots & \vdots \end{pmatrix}, \quad (25)$$

where $\text{orth}(\mathbf{A})$ indicates an ortonormalization of the columns of \mathbf{A} by any method, and $\text{diag}(x, y)$ a diagonal matrix with diagonal elements $\{x, y\}$. Even though \mathbf{g}' and \mathbf{h}' are now coplanar with \mathbf{k}_δ and \mathbf{k}_γ , they will probably not correspond to a unitary rotation of the latter, but we can assign a “best fit” value for ω :

$$\begin{pmatrix} c_{11}(\mathbf{q}) & c_{12}(\mathbf{q}) \\ c_{21}(\mathbf{q}) & c_{22}(\mathbf{q}) \end{pmatrix} = \begin{pmatrix} \vdots & \vdots \\ \mathbf{k}_\delta & \mathbf{k}_\gamma \\ \vdots & \vdots \end{pmatrix}^+ \begin{pmatrix} \vdots & \vdots \\ \mathbf{g}'(\mathbf{q}) & \mathbf{h}'(\mathbf{q}) \\ \vdots & \vdots \end{pmatrix} \quad (26)$$

$$\omega_g(\mathbf{q}) = \text{atan2}(c_{21}(\mathbf{q}), c_{11}(\mathbf{q})) \quad \omega_h(\mathbf{q}) = \text{atan2}(-c_{12}(\mathbf{q}), c_{22}(\mathbf{q})) \quad (27)$$

$$\omega(\mathbf{q}) = \text{avg}(\omega_g(\mathbf{q}), \omega_h(\mathbf{q})) \quad , \quad (28)$$

where $\text{avg}(x, y)$ is the circular mean of two angles, i.e., the angle equidistant and closer to both arguments.

Once a value of ω is defined, the diabaticization proceeds as before, that is,

$$\tau(\mathbf{q}) = \frac{1}{2} (E^A(\mathbf{q}) + E^B(\mathbf{q})) \quad (29)$$

$$\lambda(\mathbf{q}) = \frac{1}{2} (E^A(\mathbf{q}) - E^B(\mathbf{q})) \quad (30)$$

$$\delta(\mathbf{q}) = \lambda(\mathbf{q}) \cos \omega(\mathbf{q}) \quad (31)$$

$$\gamma(\mathbf{q}) = \lambda(\mathbf{q}) \sin \omega(\mathbf{q}) \quad (32)$$

$$\alpha(\mathbf{q}) = \tau(\mathbf{q}) + \delta(\mathbf{q}) \quad (33)$$

$$\beta(\mathbf{q}) = \tau(\mathbf{q}) - \delta(\mathbf{q}) \quad (34)$$

$$\begin{pmatrix} \vdots & \vdots \\ \nabla \delta(\mathbf{q}) & \nabla \gamma(\mathbf{q}) \\ \vdots & \vdots \end{pmatrix} = \begin{pmatrix} \vdots & \vdots \\ \mathbf{g}(\mathbf{q}) & \mathbf{h}(\mathbf{q}) \\ \vdots & \vdots \end{pmatrix} \begin{pmatrix} \cos \omega(\mathbf{q}) & \sin \omega(\mathbf{q}) \\ -\sin \omega(\mathbf{q}) & \cos \omega(\mathbf{q}) \end{pmatrix} \quad (35)$$

$$\nabla \alpha(\mathbf{q}) = \mathbf{s}(\mathbf{q}) + \nabla \delta(\mathbf{q}) \quad (36)$$

$$\nabla \beta(\mathbf{q}) = \mathbf{s}(\mathbf{q}) - \nabla \delta(\mathbf{q}) \quad (37)$$

Note that \mathbf{g}' , \mathbf{h}' , \mathbf{k}_δ and \mathbf{k}_γ are only used to define ω . A last detail is that the \mathbf{h} vector obtained from electronic structure calculations may change sign in an uncontrolled manner, due to the arbitrary phase of the wave functions. To account for this, if the angles ω_g and ω_h in eq. (27) differ by more than $\pi/2$, \mathbf{h} is replaced by $-\mathbf{h}$.

This achieves a pseudo-diabatization (we make no assumption on the size of the couplings between the corresponding states) that we expect to be smooth and consistent at least in the vicinity of \mathbf{q}^{ref} , even if it contains a CI seam. Farther from \mathbf{q}^{ref} , especially when the transformation in eq. (25) is large (the product $\phi_1 \phi_2$ is small) and when ω_g and ω_h differ significantly, this may not be the case. Moreover, this procedure considers only two surfaces and does not incorporate possible crossings with other surfaces. It is also worth noting that the nature of the wave functions or orbitals involved is never examined, only their energies and gradients/couplings are used.

2.3 Surrogate model

The RVO method is based on a GPR or GEK surrogate model²¹⁻²³ for the PESs.¹⁸ This surrogate model is built from a set of data (sample) points, and exactly reproduces the energies and gradients at the data points – it is an exact interpolator – within the specified tolerance, which is usually set close to machine precision. In particular, the model can be expressed as:

$$E^*(\mathbf{q}) = \mu(\mathbf{q}) + \mathbf{w}^\top \mathbf{v}(\mathbf{q}) \quad , \quad (38)$$

where μ is the “trend-function” or baseline, \mathbf{w} is a vector of weights, to be optimized when building the model, and \mathbf{v} is a vector of kernel functions and their derivatives. The length of the vectors is the number of independent data used to build the model, i.e., the number of data points (n) multiplied by the dimensionality of the PESs (m) plus one (all the gradient components and the energy for each data point), $n(m+1)$. The kernel function is in general given as $f(\mathbf{q}, \mathbf{q}')$, and the elements of the vector \mathbf{v} at a given \mathbf{q} are the values of $f(\mathbf{q}, \mathbf{q}_i)$ and $(\nabla f(\mathbf{q}, \mathbf{q}_i))_k$ for each data point \mathbf{q}_i and dimension k , ordered in some convenient way. The (constant) vector \mathbf{w} is obtained by ensuring that the model reproduces the input data, i.e.,

that all $E^*(\mathbf{q}_i)$ and $\nabla E^*(\mathbf{q}_i)$ match the energies and gradients, respectively, at the data points. This is accomplished by solving the equation

$$\mathbf{M}\mathbf{w} = \mathbf{y} \quad , \quad (39)$$

where \mathbf{y} is a vector that collects the energies (as $E - \mu$) and gradients at the data points, in the same order as the \mathbf{v} vector, and \mathbf{M} is the covariance matrix, containing the covariance between the data points, $f(\mathbf{q}_i, \mathbf{q}_j)$, as well as their first and second derivatives.

In our case, we use a constant value for the baseline μ , and a Matérn-5/2 covariance function^{46,47} as a kernel function:

$$f(\mathbf{q}, \mathbf{q}') = \left(\frac{5}{3}d(\mathbf{q}, \mathbf{q}')^2 + \sqrt{5}d(\mathbf{q}, \mathbf{q}') + 1 \right) \exp \left(-\sqrt{5}d(\mathbf{q}, \mathbf{q}') \right) \quad (40)$$

$$d(\mathbf{q}, \mathbf{q}') = \sqrt{\sum_{k=1}^m \left(\frac{q_k - q'_k}{l_k} \right)^2} \quad . \quad (41)$$

Here, d measures the distance between \mathbf{q} and \mathbf{q}' , with each dimension scaled by its l -value or “characteristic length”. The l values are chosen such that the model built with a single data point reproduces the approximate Hessian matrix given by the HMF²⁷ at that point.¹⁸

Thus, the process to build the pseudo-diabatic surrogate model can be summarized as the following:

1. Start with a set of structures, \mathbf{q}_i , and the associated E^A , E^B , \mathbf{s} , \mathbf{g} , \mathbf{h} for each structure.
2. Select the latest structure as reference, $\mathbf{q}^{\text{ref}} = \mathbf{q}_n$ and set $\mathbf{k}_\delta = \mathbf{g}(\mathbf{q}^{\text{ref}})$ and $\mathbf{k}_\gamma = \mathbf{h}(\mathbf{q}^{\text{ref}})$
3. For each structure, obtain the transformed vectors \mathbf{g}' and \mathbf{h}' , eqs. (22) to (25), and the angle ω with eqs. (26) to (28). Possibly flip the direction of \mathbf{h} .
4. For each structure, obtain α , β , γ , $\nabla\alpha$, $\nabla\beta$, $\nabla\gamma$, using eqs. (29) to (37).
5. Build three independent GEK surfaces, eq. (38), from $\{\alpha(\mathbf{q}_i), \nabla\alpha(\mathbf{q}_i)\}$, $\{\beta(\mathbf{q}_i), \nabla\beta(\mathbf{q}_i)\}$, and $\{\gamma(\mathbf{q}_i), \nabla\gamma(\mathbf{q}_i)\}$.

The adiabatic energies and gradients can be obtained from these surfaces by diagonalizing the corresponding Hamiltonian, eq. (6), and they reproduce by construction the initial data in step 1, except for the possible sign flip of \mathbf{h} .

2.4 Optimization

The RVO method, based on a GEK surrogate model, has been described in previous works.^{18–20} In short, the surrogate model is built from the electronic structure data of the previous iteration, and a stationary point is then located on the surrogate model through a number of microiterations. The progress of the microiterations is limited by the uncertainty or predicted variance of the surrogate model, which, for the case of GEK, can be computed as:

$$s^2(\mathbf{q}) = \frac{\mathbf{y}^\top \mathbf{M}^{-1} \mathbf{y}}{n} [1 - \mathbf{v}(\mathbf{q})^\top \mathbf{M}^{-1} \mathbf{v}(\mathbf{q})] \quad , \quad (42)$$

such that the 95 % confidence interval for the prediction is $E^*(\mathbf{q}) \pm 1.96\sqrt{s^2(\mathbf{q})}$. Once the stationary point is found on the surrogate model (or the maximum variance is encountered), a new electronic structure calculation is performed for that structure and that completes a macroiteration. For the next macroiteration, a new surrogate model is built, now including the data just computed.

Constraints of different types can be included in the optimization thanks to the PCO.^{19,35} This method is based on defining a unitary transformation of the coordinates \mathbf{q} that allows separating these degrees of freedom into two subspaces, one that is constrained and one that is optimized. At each microiteration, the coordinates in the constrained subspace are modified in order to fulfill the constraints, while the coordinates in the optimized subspace are modified with a general optimization method such as, for example, RS-RFO.

In ref. 19 it was noted that the implementation at the time did not support the use of nongeometrical constraints with RVO, because it needs the possibility of obtaining the value of the property being constrained during the microiterations, when no electronic structure calculations are performed. The optimization of MECPs is one of the cases that involves nongeometrical constraints,^{31,36} in particular the energy difference between two states is constrained to zero. Specifically, the optimization of CIs requires not only the energy difference between two states, but also the nonadiabatic coupling vector between them. With the pseudo-diabatization described above, all the required quantities can be obtained from the surrogate model, and the PCO can be applied as with any other constraint.

The case of crossings between states of different spin multiplicity is similar to CIs, but the process is somewhat simplified. When there is no coupling between the states, it can be assumed that γ is identically zero everywhere and therefore $E^A = \alpha$, $E^B = \beta$, only two surfaces are needed (\mathbf{h} is not used either), and no pseudo-diabatization is required as there is no singularity.

3 Computational details

The methods described above have been implemented in OpenMolcas,^{48,49} and are publicly available in its latest version. This software has been used for all the quantum chemistry calculations in this work. As in previous works,^{18,19} we set a baseline value μ for the GEK surrogate models that is $10.0 E_h$ above the maximum energy value among the data points. This is done independently for each of the energy surfaces (α , β), but for the γ surface we set $\mu = 0$. The l values obtained from the HMF are used for all the surfaces.

The optimization of MECPs has been tested for the same systems as in ref. 36 (fig. 2), and with similar settings. Optimizations were done at state-average complete active space self-consistent field (SA-CASSCF) level, the basis set was ANO-RCC with double- ζ -plus-polarization contraction,⁵⁰ and the atomic compact Cholesky decomposition (acCD)⁵¹ was employed in all calculations to treat two-electron integrals, with the default threshold of $10^{-4} E_h$. The convergence thresholds for the optimizations were the defaults in OpenMolcas (rms displacement and step size of $1.2 \cdot 10^{-3} a_0$ and $3.0 \cdot 10^{-4} E_h a_0^{-1}$, respectively; maximum components 1.5 times these values), plus a requirement for the energy difference between the crossing states to be below $10^{-5} E_h$. No spacial symmetry was enforced in any case.

For each system, at least one S_0/S_1 MECP was optimized, the number of active electrons, orbitals, and averaged states for each case is specified in table 1. The starting structures are the same as in ref. 36.

Additionally, from the same starting structures we optimized S_0/T_1 MECPs. For most of

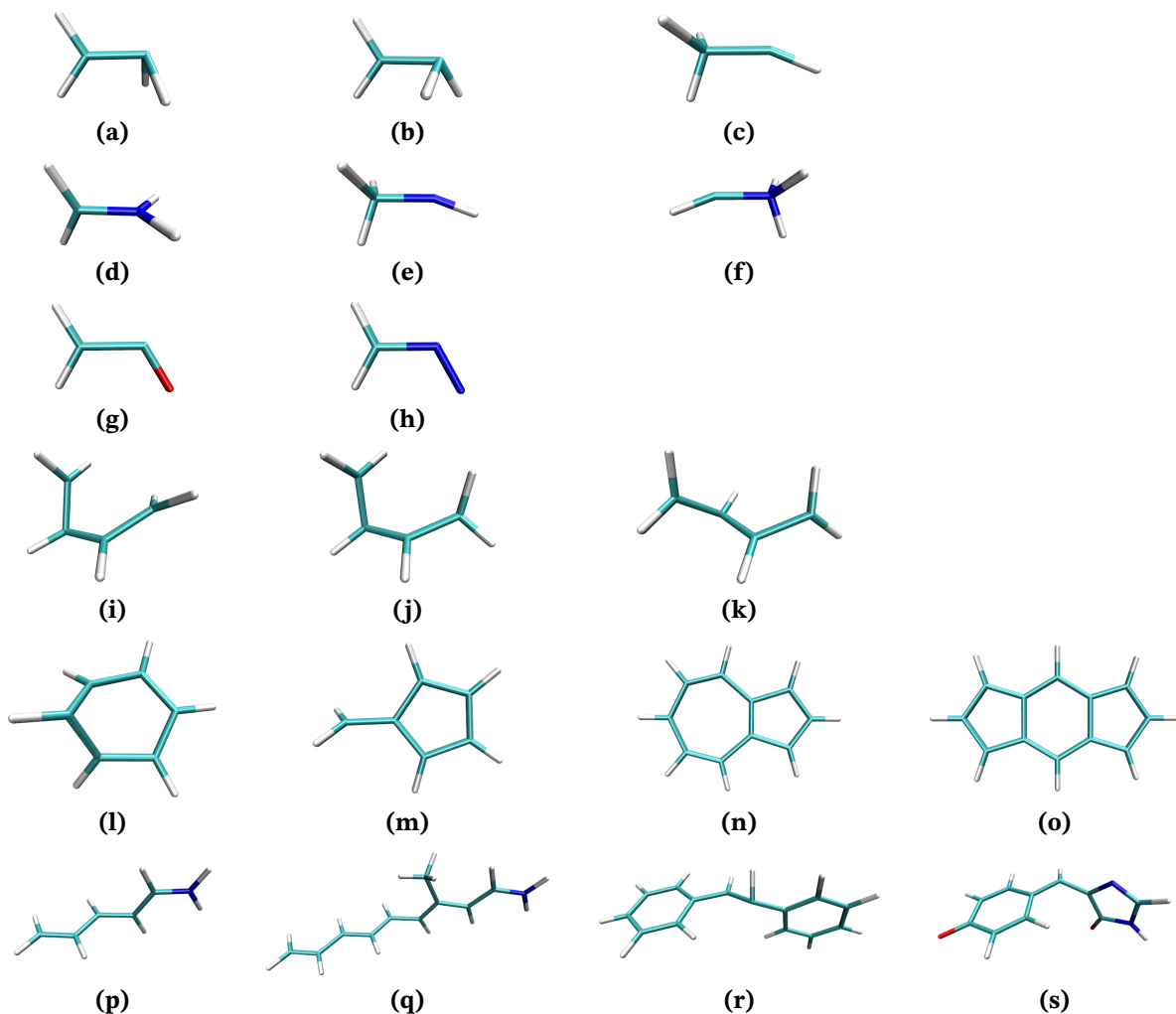


Figure 2: Structures studied in this work, shown at their optimized S_0/S_1 MECP from ref. 36.

Table 1: Active space specifications for the S_0/S_1 MECP optimizations. n_e , n_o , n_s : number of electrons, orbitals and states, respectively, in the SA-CASSCF procedure.

Molecule	Structure	(n_e, n_o)	n_s
ethylene	(a), (b), (c)	(2,2)	2
methaniminium	(d), (e), (f)	(2,2)	2
ketene	(g)	(2,3)	2
diazomethane	(h)	(2,3)	2
butadiene	(i)	(4,4)	3
”	(j), (k)	(4,4)	2
benzene	(l)	(6,6)	2
fulvene	(m)	(6,6)	2
azulene	(n)	(10,10)	2
s-indacene	(o)	(12,12)	2
PSB3	(p)	(6,6)	2
Me-PSB5	(q)	(10,10)	2
stilbene	(r)	(2,2)	3
GFP chromophore	(s)	(2,2)	3

these calculations, the same active spaces as in table 1 was used, but with no state averaging, as both singlet and triplet states are the lowest in their multiplicity. The differences and exceptions are listed in table 2, in particular, for **(n)** and **(o)** the S_1/T_1 MECP was optimized instead, as the proximity of the S_0/S_1 crossing made the optimization unstable.

Table 2: Specific details for the S_0/T_1 MECP optimizations. By default, active spaces are the same as in table 1, with no state averaging. The notations CASSCF(n_e, n_o) and SA(n_s) are used.

Structure	Specific changes.
(g), (h)	S_0 with SA(2)
(i), (j), (k)	CASSCF(2,2)
(m)	S_0 with SA(2)
(n), (o)	S_1/T_1 MECP, S_1 with SA(2)
(r), (s)	S_0 with SA(3)

Root mean square deviations (rmsd) between molecular structures were computed with the rmsd Python package,⁵² considering possible mirrorings and atom permutations to minimize the difference.

4 Results

We show first the results for the S_0/T_1 optimizations. Table 3 compares the optimizations performed with the conventional RS-RFO method and with the RVO as newly implemented for MECPs. Apart from the number of iterations needed to reach convergence, the rmsd between the optimized structures of both methods is also given, as well as the difference between the optimized MECP energies (E^\times), where a negative sign indicates the RVO structure is more stable.

The first thing to notice is that in most cases RVO converges in fewer iterations than RS-RFO. Even in some cases where RS-RFO is efficient, RVO can still save one or two iterations, and in more difficult cases, like **(r)** and **(s)**, the savings can be more significant. In general, the differences in both geometry and energy are very small, indicating that the two methods converged to essentially the same structure. The cases where the results seem to be significantly different are **(c)**, **(f)**, **(m)** and **(s)**, and in these, not only does RVO take fewer (or as many) iterations than RS-RFO, but also achieves a lower final energy.

The optimized S_0/T_1 MECP for **(c)** is characterized by a H–C–C–H dihedral close to 180° . It is 179.3° with RVO, but 175.2° with RS-RFO. Similarly, in the case of **(f)** the H–N–C–H dihedral is 179.2° with RVO and 177.4° with RS-RFO. The differences in **(m)** and **(s)** are much smaller and not worth detailing.

It should be noted that the structures in fig. 2 are S_0/S_1 MECPs, so they do not reflect in all cases the structure of the S_0/T_1 MECP. For example, **(a)** and **(b)** converge to the same twisted structure, similar to **(d)**, while **(i)**, **(j)** and **(k)** converge to a structure with a 3-member ring, and **(m)** is almost planar.

Overall, it seems clear that at least for these systems the RVO method represents an improvement over the conventional RS-RFO. We would like to point out that although the RVO optimization is computationally more expensive than RS-RFO, this cost increase is completely negligible compared to the cost of the electronic structure calculations,

Table 3: Number of (macro)iterations to converge the S_0/T_1 MECP structures, rmsd and energy difference between the two methods ($\Delta E^\times = E_{\text{RVO}}^\times - E_{\text{RS-RFO}}^\times$).

	RS-RFO	RVO	rmsd (pm)	ΔE^\times (mE_h)
(a)	24	22	0.034	−0.0001
(b)	10	14	0.055	−0.0002
(c)	10	10	2.267	−0.0167
(d)	6	6	0.007	−0.0000
(e)	10	7	0.008	0.0003
(f)	34	10	0.989	−0.0034
(g)	8	6	0.005	0.0005
(h)	9	7	0.011	0.0000
(i)	17	13	0.019	0.0000
(j)	16	15	0.007	0.0000
(k)	14	11	0.025	−0.0001
(l)	14	12	0.019	0.0000
(m)	12	11	0.142	−0.0018
(n)*	6	6	0.003	−0.0007
(o)*	6	6	0.023	−0.0002
(p)	8	10	0.044	−0.0001
(q)	30	31	0.059	−0.0000
(r)	26	14	0.038	0.0001
(s)	35	18	0.152	−0.0003

* S_1/T_1 MECP.

and the number of iterations is therefore an accurate measure of performance, at least for systems of up to a few dozen atoms.

Having established the good behavior of RVO with two surfaces, we discuss now the results for S_0/S_1 MECP optimizations, where the surrogate model is given by the pseudo-diabatic surfaces α, β, γ . The comparison between RVO and RS-RFO is given in table 4. The difference now is more important than for the S_0/T_1 MECPs. Only for **(e)** and **(p)** does RVO take one or two more iterations (and it still converges to lower energy), while in all other cases it takes significantly fewer iterations, sometimes less than half. In terms of rmsd and energy differences, **(a)** and **(e)** stand out, while **(m)**, **(p)**, **(r)** are also slightly larger than the rest.

Table 4: Number of (macro)iterations to converge the S_0/S_1 MECP structures, rmsd and energy difference between the two methods ($\Delta E^\times = E_{\text{RVO}}^\times - E_{\text{RS-RFO}}^\times$).

	RS-RFO	RVO	rmsd (pm)	ΔE^\times (mE_h)
(a)	34	5	21.065	5.3841
(b)	14	7	0.097	-0.0011
(c)	20	13	0.031	0.0003
(d)	5	4	0.007	-0.0005
(e)	14	15	8.417	-0.9213
(f)	16	14	0.024	0.0008
(g)	10	7	0.010	-0.0005
(h)	10	7	0.004	-0.0004
(i)	19	10	0.027	-0.0002
(j)	29	11	0.017	0.0000
(k)	12	10	0.023	-0.0001
(l)	7	6	0.021	0.0003
(m)	17	13	0.349	-0.0008
(n)	9	6	0.016	0.0025
(o)	6	5	0.009	-0.0002
(p)	8	10	0.123	-0.0002
(q)	32	23	0.025	0.0001
(r)	34	17	0.119	-0.0005
(s)	17	11	0.055	0.0006

In the case of **(a)**, RVO converged to the symmetric structure shown in fig. 2, but RS-RFO found the same as **(b)** instead. As discussed in ref. 36, the symmetric structure is not a minimum, but a saddle point in the intersection space, and with RS-RFO, probably due to numerical noise, the symmetry is broken and the optimization falls to a minimum, which explains the large number of iterations and lower energy found with RS-RFO. For **(e)**, the main structural difference is the C–N–H angle, that is 159.6° with RVO and 174.0° with RS-RFO. Given the rather large energy difference between both structures, it does not look like the surface is very flat, and we assume that in this case RS-RFO got stuck at or close to a saddle point.

Although the surrogate model built for RVO is only intended to be used for optimization purposes (not, for instance, to run molecular dynamics simulations on it), it is instructive to examine how well it reproduces the “true” surfaces around a CI, and how it differs from a simpler linear model. We take as an example the optimized S_0/S_1 MECP of **(p)**. In ref. 36

the linear model was already analyzed for this system, and it was found that while it was valid for regions very close to the CI, it deviates appreciably from the computed energies as farther away, and can give qualitatively wrong predictions beyond $\sim 0.03 a_0$. We represent in fig. 3 the shape of the adiabatic surfaces obtained from the GEK surrogate model in the branching plane around the MECP, up to a distance of $0.1 a_0$, the deviation from the linear model is evidenced by the curved shape of the radial grid lines, particularly clear in the lower surface. We then compare this GEK prediction with actual SA-CASSCF single-point calculations for structures around the rim of this figure, and plot them in fig. 4. It is seen that the linear model (wrongly) predicts minima along the $\pm x$ direction; the single-point calculations, however, show that the real energies are much higher. The model obtained from the GEK surfaces follows much more closely the computed energies, although there are still some deviations. It must be emphasized that the GEK model is *not* built to reproduce these energies, but only those of the latest 10 iterations between the initial structure and the final optimized MECP. The right panel of fig. 4 shows the corresponding α , β and γ surfaces. Similar comparisons for the other systems confirm that the GEK model provides a much better approximation to the SA-CASSCF energies than a simple linear model.

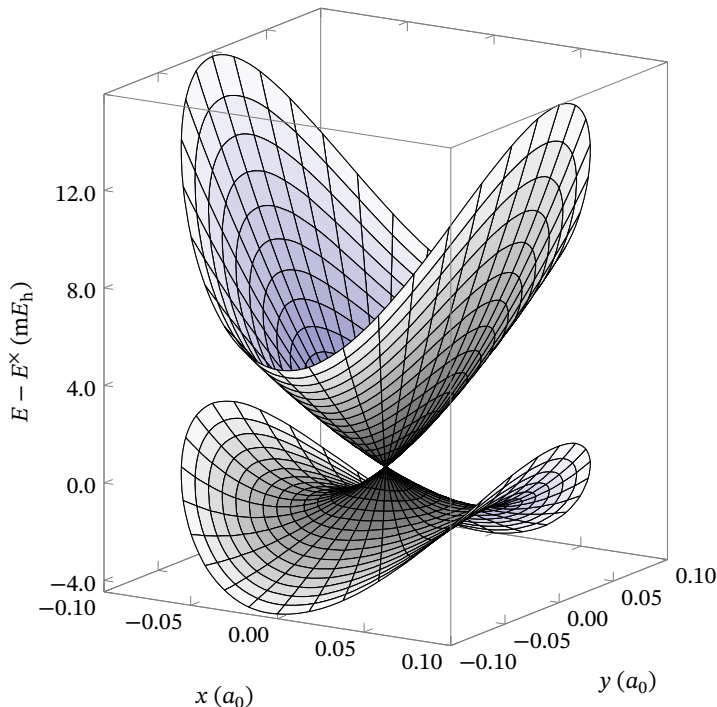


Figure 3: Representation of the adiabatic PESs obtained with the GEK surrogate model around the optimized S_0/S_1 MECP of **(p)**.

In both S_0/S_1 and S_0/T_1 MECP optimizations, it was found that the structure that needed most iterations to converge was **(q)**. This correlates to its being the most flexible system in the set, but we observe that in this case most of the iterations, for the two optimization methods, are spent in a 60° rotation of the CH_3 group, which results in a stabilization of around $1.1 \text{ kcal mol}^{-1}$. It can be expected that an overshooting procedure such as the one implemented in ref. ²⁴ could improve the performance of RVO, especially when the surrogate model is expressed in internal coordinates. However, we did not use overshooting, so this remains a possible area of improvement.

As a summary, for S_0/T_1 MECs, the use RVO reduced the total number of iterations

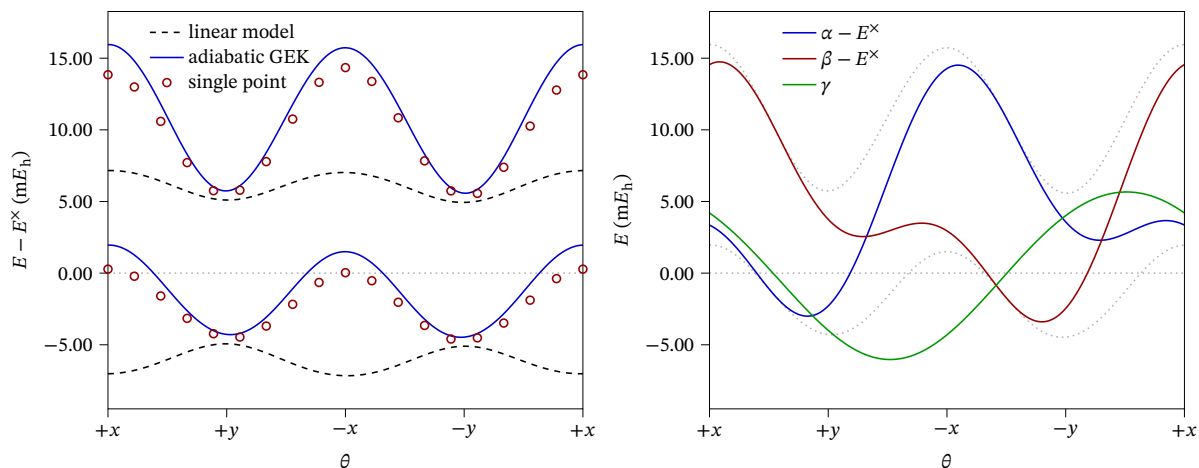


Figure 4: Plot of the PESs in a circle around the optimized S_0/S_1 MECP of **(p)** with a radius of $r = 0.01 a_0$, in the branching plane. Left: adiabatic model surfaces obtained with linear model and with the GEK surrogate model, as well as single-point calculations at selected geometries. Right: The three pseudo-diabatic surfaces from which the adiabatic GEK on the left (shown also as dotted curves) is obtained.

from 295 to 229 (a 22 % reduction), for S_0/S_1 MECPs the reduction is from 279 to 189 (32 %), excluding **(a)**, where both methods converge to clearly different structures.

5 Conclusions

We have implemented a pseudo-diabatization process that allows representing the crossing between two adiabatic PESs as a combination of three smooth, continuous pseudo-diabatic surfaces. This is used to build a surrogate model for the RVO method in order to efficiently locate MECPs. The test calculations reported here indicate that this extension to RVO achieves a noticeable reduction in the number of iterations (energy and gradient evaluations) required, both for crossings between states of different spin and for CIs. It can be noted that although most of the test MECPs in this work involve the S_0 ground state, there is nothing in the method that is specific for the ground state, so it can be straightforwardly applied to crossings between excited states.

The properties used to build the surrogate model are only those used in the conventional optimization: energies, gradients and nonadiabatic coupling. Comparison of the model with single-point energy calculations show that the adiabatic PESs around a CI are well approximated, beyond what a linear model can provide. However, it is reminded that the model is only intended to be a local approximation in the vicinity of the final optimized structure, and not as a global representation of the surfaces.

Acknowledgement

The authors acknowledge the Swedish Research Council (VR, Grant No. 2020-03182) for funding. This research was partially supported by the project *AI4Research* at Uppsala University. Some of the calculations were performed on computer resources provided by the Swedish National Infrastructure for Computing (SNIC), partially funded by the Swedish

Research Council (grant 2018-05973), at the National Supercomputer Centre in Sweden (NSC, Linköping University) and UPPMAX (Uppsala University).

Supporting Information Available

Input files and relevant output files for each of the MECF optimizations with OpenMolcas.

References

- (1) Hofmann, K. P.; Lamb, T. D. Rhodopsin, light-sensor of vision. *Prog. Retinal Eye Res.* **2023**, *93*, 101116, DOI: [10.1016/j.preteyeres.2022.101116](https://doi.org/10.1016/j.preteyeres.2022.101116).
- (2) Vacher, M.; Fdez. Galván, I.; Ding, B.-W.; Schramm, S.; Berraud-Pache, R.; Naumov, P.; Ferré, N.; Liu, Y.-J.; Navizet, I.; Roca-Sanjuán, D.; Baader, W. J.; Lindh, R. Chemi- and Bioluminescence of Cyclic Peroxides. *Chem. Rev.* **2018**, *118*, 6927–6974, DOI: [10.1021/acs.chemrev.7b00649](https://doi.org/10.1021/acs.chemrev.7b00649).
- (3) Crespo-Hernández, C. E.; Cohen, B.; Hare, P. M.; Kohler, B. Ultrafast Excited-State Dynamics in Nucleic Acids. *Chem. Rev.* **2004**, *104*, 1977–2020, DOI: [10.1021/cr0206770](https://doi.org/10.1021/cr0206770).
- (4) Mirkovic, T.; Ostroumov, E. E.; Anna, J. M.; van Grondelle, R.; Govindjee; Scholes, G. D. Light Absorption and Energy Transfer in the Antenna Complexes of Photosynthetic Organisms. *Chem. Rev.* **2016**, *117*, 249–293, DOI: [10.1021/acs.chemrev.6b00002](https://doi.org/10.1021/acs.chemrev.6b00002).
- (5) Yonehara, T.; Hanasaki, K.; Takatsuka, K. Fundamental Approaches to Nonadiabaticity: Toward a Chemical Theory beyond the Born–Oppenheimer Paradigm. *Chem. Rev.* **2012**, *112*, 499–542, DOI: [10.1021/cr200096s](https://doi.org/10.1021/cr200096s).
- (6) Crespo-Otero, R.; Barbatti, M. Recent Advances and Perspectives on Nonadiabatic Mixed Quantum–Classical Dynamics. *Chem. Rev.* **2018**, *118*, 7026–7068, DOI: [10.1021/acs.chemrev.7b00577](https://doi.org/10.1021/acs.chemrev.7b00577).
- (7) Atchity, G. J.; Xantheas, S. S.; Ruedenberg, K. Potential energy surfaces near intersections. *J. Chem. Phys.* **1991**, *95*, 1862–1876, DOI: [10.1063/1.461036](https://doi.org/10.1063/1.461036).
- (8) Klessinger, M. Conical Intersections and the Mechanism of Singlet Photoreactions. *Angew. Chem. Int. Ed.* **1995**, *34*, 549–551, DOI: [10.1002/anie.199505491](https://doi.org/10.1002/anie.199505491).
- (9) Domcke, W., Yarkony, D. R., Köppel, H., Eds. *Conical Intersections. Electronic Structure, Dynamics & Spectroscopy*; Advanced Series in Physical Chemistry 15; World Scientific, 2004; DOI: [10.1142/5406](https://doi.org/10.1142/5406).
- (10) Domcke, W., Yarkony, D. R., Köppel, H., Eds. *Conical Intersections. Theory, Computation and Experiment*; Advanced Series in Physical Chemistry 17; World Scientific, 2011; DOI: [10.1142/7803](https://doi.org/10.1142/7803).
- (11) Matsika, S. Electronic Structure Methods for the Description of Nonadiabatic Effects and Conical Intersections. *Chem. Rev.* **2021**, *121*, 9407–9449, DOI: [10.1021/acs.chemrev.1c00074](https://doi.org/10.1021/acs.chemrev.1c00074).
- (12) Polyakov, I. V.; Grigorenko, B. L.; Epifanovsky, E. M.; Krylov, A. I.; Nemukhin, A. V. Potential Energy Landscape of the Electronic States of the GFP Chromophore in Different Protonation Forms: Electronic Transition Energies and Conical Intersections. *J. Chem. Theory Comput.* **2010**, *6*, 2377–2387, DOI: [10.1021/ct100227k](https://doi.org/10.1021/ct100227k).
- (13) Gozem, S.; Krylov, A. I.; Olivucci, M. Conical Intersection and Potential Energy Surface Features of a Model Retinal Chromophore: Comparison of EOM-CC and Multireference Methods. *J. Chem. Theory Comput.* **2012**, *9*, 284–292, DOI: [10.1021/ct300759z](https://doi.org/10.1021/ct300759z).

- (14) Jorner, K.; Dreos, A.; Emanuelsson, R.; El Bakouri, O.; Fdez. Galván, I.; Börjesson, K.; Feixas, F.; Lindh, R.; Zietz, B.; Moth-Poulsen, K.; Ottosson, H. Unraveling factors leading to efficient norbornadiene–quadricyclane molecular solar-thermal energy storage systems. *J. Mater. Chem. A* **2017**, *5*, 12369–12378, DOI: [10.1039/c7ta04259k](https://doi.org/10.1039/c7ta04259k).
- (15) Fdez. Galván, I.; Brakestad, A.; Vacher, M. Role of conical intersection seam topography in the chemiexcitation of 1,2-dioxetanes. *Phys. Chem. Chem. Phys.* **2022**, *24*, 1638–1653, DOI: [10.1039/d1cp05028a](https://doi.org/10.1039/d1cp05028a).
- (16) Schlegel, H. B. Geometry optimization. *WIREs Computational Molecular Science* **2011**, *1*, 790–809, DOI: [10.1002/wcms.34](https://doi.org/10.1002/wcms.34).
- (17) Bakken, V.; Helgaker, T. The efficient optimization of molecular geometries using redundant internal coordinates. *J. Chem. Phys.* **2002**, *117*, 9160–9174, DOI: [10.1063/1.1515483](https://doi.org/10.1063/1.1515483).
- (18) Raggi, G.; Fdez. Galván, I.; Ritterhoff, C. L.; Vacher, M.; Lindh, R. Restricted-Variance Molecular Geometry Optimization Based on Gradient-Enhanced Kriging. *J. Chem. Theory Comput.* **2020**, *16*, 3989–4001, DOI: [10.1021/acs.jctc.0c00257](https://doi.org/10.1021/acs.jctc.0c00257).
- (19) Fdez. Galván, I.; Raggi, G.; Lindh, R. Restricted-Variance Constrained, Reaction Path, and Transition State Molecular Optimizations Using Gradient-Enhanced Kriging. *J. Chem. Theory Comput.* **2021**, *17*, 571–582, DOI: [10.1021/acs.jctc.0c01163](https://doi.org/10.1021/acs.jctc.0c01163).
- (20) Lindh, R.; Fdez. Galván, I. Molecular structure optimizations with Gaussian process regression. In *Quantum Chemistry in the Age of Machine Learning*; Dral, P. O., Ed.; Elsevier, 2023; pp 391–428, DOI: [10.1016/b978-0-323-90049-2.00017-2](https://doi.org/10.1016/b978-0-323-90049-2.00017-2).
- (21) Liu, W.; Batill, S. Gradient-Enhanced Response Surface Approximations Using Kriging Models. 9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization. 2002; DOI: [10.2514/6.2002-5456](https://doi.org/10.2514/6.2002-5456).
- (22) Han, Z.-H.; Görtz, S.; Zimmermann, R. Improving variable-fidelity surrogate modeling via gradient-enhanced kriging and a generalized hybrid bridge function. *Aero. Sci. Technol.* **2013**, *25*, 177–189, DOI: [10.1016/j.ast.2012.01.006](https://doi.org/10.1016/j.ast.2012.01.006).
- (23) Ulaganathan, S.; Couckuyt, I.; Ferranti, F.; Laermans, E.; Dhaene, T. Performance study of multi-fidelity gradient enhanced kriging. *Struct. Multidiscip. Optim.* **2015**, *51*, 1017–1033, DOI: [10.1007/s00158-014-1192-x](https://doi.org/10.1007/s00158-014-1192-x).
- (24) Denzel, A.; Kästner, J. Gaussian process regression for geometry optimization. *J. Chem. Phys.* **2018**, *148*, 094114, DOI: [10.1063/1.5017103](https://doi.org/10.1063/1.5017103).
- (25) del Río, E. G.; Mortensen, J. J.; Jacobsen, K. W. Local Bayesian optimizer for atomic structures. *Phys. Rev. B* **2019**, *100*, DOI: [10.1103/physrevb.100.104103](https://doi.org/10.1103/physrevb.100.104103).
- (26) Meyer, R.; Hauser, A. W. Geometry optimization using Gaussian process regression in internal coordinate systems. *J. Chem. Phys.* **2020**, *152*, 084112, DOI: [10.1063/1.5144603](https://doi.org/10.1063/1.5144603).
- (27) Lindh, R.; Bernhardsson, A.; Karlström, G.; Malmqvist, P.-Å. On the use of a Hessian model function in molecular geometry optimizations. *Chem. Phys. Lett.* **1995**, *241*, 423–428, DOI: [10.1016/0009-2614\(95\)00646-l](https://doi.org/10.1016/0009-2614(95)00646-l).

- (28) Ragazos, I. N.; Robb, M. A.; Bernardi, F.; Olivucci, M. Optimization and characterization of the lowest energy point on a conical intersection using an MC-SCF Lagrangian. *Chem. Phys. Lett.* **1992**, *197*, 217–223, DOI: [10.1016/0009-2614\(92\)85758-3](https://doi.org/10.1016/0009-2614(92)85758-3).
- (29) Manaa, M. R.; Yarkony, D. R. On the intersection of two potential energy surfaces of the same symmetry. Systematic characterization using a Lagrange multiplier constrained procedure. *J. Chem. Phys.* **1993**, *99*, 5251–5256, DOI: [10.1063/1.465993](https://doi.org/10.1063/1.465993).
- (30) Bearpark, M. J.; Robb, M. A.; Schlegel, H. B. A direct method for the location of the lowest energy point on a potential surface crossing. *Chem. Phys. Lett.* **1994**, *223*, 269–274, DOI: [10.1016/0009-2614\(94\)00433-1](https://doi.org/10.1016/0009-2614(94)00433-1).
- (31) Anglada, J. M.; Bofill, J. M. A reduced-restricted-quasi-Newton-Raphson method for locating and optimizing energy crossing points between two potential energy surfaces. **1997**, *18*, 992–1003, DOI: [10.1002/\(sici\)1096-987x\(199706\)18:8<992::aid-jcc3>3.0.co;2-l](https://doi.org/10.1002/(sici)1096-987x(199706)18:8<992::aid-jcc3>3.0.co;2-l).
- (32) Ciminelli, C.; Granucci, G.; Persico, M. The Photoisomerization Mechanism of Azobenzene: A Semiclassical Simulation of Nonadiabatic Dynamics. **2004**, *10*, 2327–2341, DOI: [10.1002/chem.200305415](https://doi.org/10.1002/chem.200305415).
- (33) Levine, B. G.; Coe, J. D.; Martínez, T. J. Optimizing Conical Intersections without Derivative Coupling Vectors: Application to Multistate Multireference Second-Order Perturbation Theory (MS-CASPT2). *J. Phys. Chem. B* **2008**, *112*, 405–413, DOI: [10.1021/jp0761618](https://doi.org/10.1021/jp0761618).
- (34) Maeda, S.; Ohno, K.; Morokuma, K. Updated Branching Plane for Finding Conical Intersections without Coupling Derivative Vectors. *J. Chem. Theory Comput.* **2010**, *6*, 1538–1545, DOI: [10.1021/ct1000268](https://doi.org/10.1021/ct1000268).
- (35) De Vico, L.; Olivucci, M.; Lindh, R. New General Tools for Constrained Geometry Optimizations. *J. Chem. Theory Comput.* **2005**, *1*, 1029–1037, DOI: [10.1021/ct0500949](https://doi.org/10.1021/ct0500949).
- (36) Fdez. Galván, I.; Delcey, M. G.; Pedersen, T. B.; Aquilante, F.; Lindh, R. Analytical State-Average Complete-Active-Space Self-Consistent Field Nonadiabatic Coupling Vectors: Implementation with Density-Fitted Two-Electron Integrals and Application to Conical Intersections. *J. Chem. Theory Comput.* **2016**, *12*, 3636–3653, DOI: [10.1021/acs.jctc.6b00384](https://doi.org/10.1021/acs.jctc.6b00384).
- (37) Yarkony, D. R. Diabolical conical intersections. *Rev. Mod. Phys.* **1996**, *68*, 985–1013, DOI: [10.1103/revmodphys.68.985](https://doi.org/10.1103/revmodphys.68.985).
- (38) Yarkony, D. R. Conical Intersections: The New Conventional Wisdom. *J. Phys. Chem. A* **2001**, *105*, 6277–6293, DOI: [10.1021/jp003731u](https://doi.org/10.1021/jp003731u).
- (39) Baer, M. *Beyond Born–Oppenheimer*; John Wiley & Sons, Inc., 2006; DOI: [10.1002/0471780081](https://doi.org/10.1002/0471780081).
- (40) Schuurman, M. S.; Yarkony, D. R. On the Characterization of Three-State Conical Intersections Using a Group Homomorphism Approach: The Two-State Degeneracy Spaces. *J. Phys. Chem. B* **2006**, *110*, 19031–19039, DOI: [10.1021/jp0607216](https://doi.org/10.1021/jp0607216).

- (41) Boeije, Y.; Olivucci, M. From a one-mode to a multi-mode understanding of conical intersection mediated ultrafast organic photochemical reactions. *Chem. Soc. Rev.* **2023**, *52*, 2643–2687, DOI: [10.1039/d2cs00719c](https://doi.org/10.1039/d2cs00719c).
- (42) Shu, Y.; Varga, Z.; Kanchanakungwankul, S.; Zhang, L.; Truhlar, D. G. Diabatic States of Molecules. *J. Phys. Chem. A* **2022**, *126*, 992–1018, DOI: [10.1021/acs.jpca.1c10583](https://doi.org/10.1021/acs.jpca.1c10583).
- (43) Mead, C. A.; Truhlar, D. G. Conditions for the definition of a strictly diabatic electronic basis for molecular systems. *J. Chem. Phys.* **1982**, *77*, 6090–6098, DOI: [10.1063/1.443853](https://doi.org/10.1063/1.443853).
- (44) Yarkony, D. R. Conical Intersections: Their Description and Consequences. In *Conical Intersections. Electronic Structure, Dynamics & Spectroscopy*; Domcke, W., Yarkony, D. R., Köppel, H., Eds.; Advanced Series in Physical Chemistry 15; World Scientific, 2004; pp 41–127, DOI: [10.1142/9789812565464_0002](https://doi.org/10.1142/9789812565464_0002).
- (45) Vértési, T.; Vibók, Á.; Halász, G. J.; Baer, M. On the peculiarities of the diabatic framework: New insight. *J. Chem. Phys.* **2004**, *120*, 2565–2574, DOI: [10.1063/1.1635352](https://doi.org/10.1063/1.1635352).
- (46) Matérn, B. *Spatial Variation*; Lecture Notes in Statistics 36; Springer New York, 1986; DOI: [10.1007/978-1-4615-7892-5](https://doi.org/10.1007/978-1-4615-7892-5).
- (47) Minasny, B.; McBratney, A. B. The Matérn function as a general model for soil variograms. *Geoderma* **2005**, *128*, 192–207, DOI: [10.1016/j.geoderma.2005.04.003](https://doi.org/10.1016/j.geoderma.2005.04.003).
- (48) Fdez. Galván, I.; Vacher, M.; Alavi, A.; Angeli, C.; Aquilante, F.; Autschbach, J.; Bao, J. J.; Bokarev, S. I.; Bogdanov, N. A.; Carlson, R. K.; Chibotaru, L. F.; Creutzberg, J.; Dattani, N.; Delcey, M. G.; Dong, S. S.; Dreuw, A.; Freitag, L.; Frutos, L. M.; Gagliardi, L.; Gendron, F.; Giussani, A.; González, L.; Grell, G.; Guo, M.; Hoyer, C. E.; Johansson, M.; Keller, S.; Knecht, S.; Kovačević, G.; Källman, E.; Li Manni, G.; Lundberg, M.; Ma, Y.; Mai, S.; Malhado, J. P.; Malmqvist, P. Å.; Marquetand, P.; Mewes, S. A.; Norell, J.; Olivucci, M.; Oppel, M.; Phung, Q. M.; Pierloot, K.; Plasser, F.; Reiher, M.; Sand, A. M.; Schapiro, I.; Sharma, P.; Stein, C. J.; Sørensen, L. K.; Truhlar, D. G.; Ugandi, M.; Ungur, L.; Valentini, A.; Vancoillie, S.; Veryazov, V.; Weser, O.; Wośowski, T. A.; Widmark, P.-O.; Wouters, S.; Zech, A.; Zobel, J. P.; Lindh, R. OpenMolcas: From Source Code to Insight. *J. Chem. Theory Comput.* **2019**, *15*, 5925–5964, DOI: [10.1021/acs.jctc.9b00532](https://doi.org/10.1021/acs.jctc.9b00532).
- (49) Aquilante, F.; Autschbach, J.; Baiardi, A.; Battaglia, S.; Borin, V. A.; Chibotaru, L. F.; Conti, I.; De Vico, L.; Delcey, M.; Fdez. Galván, I.; Ferré, N.; Freitag, L.; Garavelli, M.; Gong, X.; Knecht, S.; Larsson, E. D.; Lindh, R.; Lundberg, M.; Malmqvist, P. Å.; Nenov, A.; Norell, J.; Odelius, M.; Olivucci, M.; Pedersen, T. B.; Pedraza-González, L.; Phung, Q. M.; Pierloot, K.; Reiher, M.; Schapiro, I.; Segarra-Martí, J.; Segatta, F.; Seijo, L.; Sen, S.; Sergentu, D.-C.; Stein, C. J.; Ungur, L.; Vacher, M.; Valentini, A.; Veryazov, V. Modern quantum chemistry with [Open]Molcas. *J. Chem. Phys.* **2020**, *152*, 214117, DOI: [10.1063/5.0004835](https://doi.org/10.1063/5.0004835).
- (50) Roos, B. O.; Lindh, R.; Malmqvist, P.-Å.; Veryazov, V.; Widmark, P.-O. Main Group Atoms and Dimers Studied with a New Relativistic ANO Basis Set. *J. Phys. Chem. A* **2004**, *108*, 2851–2858, DOI: [10.1021/jp031064+](https://doi.org/10.1021/jp031064+).

- (51) Aquilante, F.; Gagliardi, L.; Pedersen, T. B.; Lindh, R. Atomic Cholesky decompositions: A route to unbiased auxiliary basis sets for density fitting approximation with tunable accuracy and efficiency. *J. Chem. Phys.* **2009**, *130*, 154107, DOI: [10.1063/1.3116784](https://doi.org/10.1063/1.3116784).
- (52) Kromann, J. C. Calculate Root-mean-square deviation (RMSD) of Two Molecules Using Rotation. <http://github.com/charnley/rmsd>, (Accessed on 2023-04-26).

TOC Graphic

