

Marching along Ridges. Efficient Location of Energy-Minimized Conical Intersections of Two States Using Extrapolatable Functions

David R. Yarkony

Department of Chemistry, Johns Hopkins University, Baltimore, Maryland 21218

Received: November 11, 2003; In Final Form: December 22, 2003

A conical intersection is a singular point in nuclear coordinate space. As a result of this singularity the parameters used to search for energy-minimized conical intersections, energy gradients, energy difference gradients, and coupling vectors vary irregularly along the search path. This irregular variation precludes the efficient use of extrapolation procedures to speed convergence. This impediment may be overcome by the use of extrapolatable functions, that is, functions that vary smoothly along the search path. From a topographical perspective this approach amounts to walking along a path parallel to the ridge of conical intersections. In this work an algorithm based on these functions is introduced and its performance discussed.

I. Introduction

In the past decade conical intersections of two states of the same symmetry have emerged as fundamental to the description of electronically nonadiabatic processes.^{1–3} Algorithms have been developed, exploiting the power of analytic gradient techniques,^{4,5} to locate these points.^{6–9} Points of conical intersection are not isolated but are continuously connected forming seams.¹⁰ The seams are $N^{\text{int}}-2$ -dimensional subspaces in the N^{int} -dimensional space of internal nuclear coordinates. For general polyatomic molecules $N^{\text{int}} - 2$ is a large number and it is desirable to determine sections of these seams for which the energy has been minimized. Although it is relatively straightforward to locate points of conical intersection, energy optimization is more costly to achieve. We attribute this difficulty to the erratic behavior of the algorithmic parameters along the search path, which precludes extrapolation. This erratic behavior might seem intrinsic to the problem because it is a consequence of the singular character of the conical intersection, the object of the search. Fortunately, this need not be the case. We have recently introduced extrapolatable functions,¹¹ a set of functions that are well-behaved along the search path. In this work we report a simple, no-additional cost, algorithm based on these functions and Hessian updating¹² and show that it can significantly improve convergence to an energy-minimized point of conical intersection.

In section II the basic search algorithm and the idea of extrapolatable functions is reviewed. In section III an algorithm based on these functions is described. The algorithm achieves quasi-second-order convergence without the need to precompute the relevant Hessians in their entirety using divided difference techniques. This will make efficient convergence possible in larger molecules. The performance of the algorithm is illustrated using numerical examples. Section IV summarizes and concludes.

II. Extrapolatable Functions and Conical Intersection:

In this section an algorithm for locating energy-minimized conical intersections is presented so that we may describe the incorporation of extrapolatable functions, functions that vary smoothly along a search path, into that algorithm.

A. Search Algorithm. Let \mathbf{Q} denote the N^{int} internal nuclear coordinates. We seek a point of conical intersection of states I and J in the subspace of nuclear coordinate space defined by

N^{con} equations

$$K^j(\mathbf{Q}) = 0 \quad j = 1, \dots, N^{\text{con}} \quad (1a)$$

A point of conical intersection satisfies

$$C^1(\mathbf{Q}) \equiv \Delta E_{\text{JI}}(\mathbf{Q}) = E_J(\mathbf{Q}) - E_I(\mathbf{Q}) = 0 \quad (1b)$$

together with the requirement

$$C^2(\mathbf{Q}) = \mathbf{c}^I(\mathbf{Q})^{\text{f}} \mathbf{H}^{\text{CSF}}(\mathbf{Q}) \mathbf{c}^J(\mathbf{Q})^{\text{f}} = 0 \quad (1c)$$

where \mathbf{H}^{CSF} is the electronic Hamiltonian in the configuration state function¹³ (ψ) basis and the \mathbf{c}^I satisfy the configuration interaction problem

$$[\mathbf{H}^{\text{CSF}}(\mathbf{Q}) - \mathbf{I}E_I(\mathbf{Q})]\mathbf{c}^I(\mathbf{Q}) = 0 \quad (2)$$

Finally, as explained in refs 3 and 11, the “f” superscript indicates that $\mathbf{c}^I(\mathbf{Q})$ is held fixed when coordinate derivatives are taken.

The desired point, $\mathbf{Q}^{\text{x,II}}$, is to be the lowest energy point of conical intersection in the subspace defined in eq 1a. $\mathbf{Q}^{\text{x,II}}$ is an extremum of the Lagrangian L^{II} ,

$$\frac{\partial}{\partial \xi_k} L^{\text{II}}(\mathbf{Q}^{\text{x,II}}) = 0 \quad \text{for } \xi = \mathbf{Q}, \lambda, \zeta \quad (3)$$

where L^{II} is constructed from $P(\mathbf{Q})$ (described below) and the constraints equations, eqs 1a–1c

$$L^{\text{II}}(\mathbf{Q}, \lambda, \zeta) = P(\mathbf{Q}) + \sum_{i=1}^2 \lambda_i C^i(\mathbf{Q}) + \sum_{i=1}^{N^{\text{con}}} \zeta_i K^i(\mathbf{Q}) \quad (4)$$

With the Lagrange multipliers λ , ζ , and \mathbf{Q} as independent variables, eq 3, using eq 4, becomes

$$\nabla L^{\text{II}}(\mathbf{Q}) \equiv \nabla P(\mathbf{Q}) + \lambda_1 \mathbf{g}^{\text{II}}(\mathbf{Q}) + \lambda_2 \mathbf{h}^{\text{II}}(\mathbf{Q}) + \sum_{\alpha=1}^{N^{\text{con}}} \zeta_{\alpha} \mathbf{k}^{\alpha}(\mathbf{Q}) = 0 \quad (5a)$$

$$\Delta E^{\text{II}}(\mathbf{Q}) = 0 \quad \text{and} \quad C^2(\mathbf{Q}) = 0 \quad (5b,c)$$

$$K^i(\mathbf{Q}) = 0 \quad (5d)$$

where

$$\mathbf{k}^i(\mathbf{Q}) = \nabla K^i(\mathbf{Q}) \quad (6a)$$

$$\mathbf{h}^{kl}(\mathbf{Q}) = \mathbf{c}^{k^\dagger}(\mathbf{Q}^{x,IJ})(\nabla \mathbf{H}(\mathbf{Q}))\mathbf{c}^l(\mathbf{Q}^{x,IJ}) \approx \mathbf{c}^{k^\dagger}(\mathbf{Q})^f(\nabla \mathbf{H}(\mathbf{Q}))\mathbf{c}^l(\mathbf{Q})^f \quad (6b)$$

$$2\mathbf{g}^{IJ} = \mathbf{h}^{II} - \mathbf{h}^{JJ} \quad (6c)$$

the approximation in eq 6b improves as the seam is approached, and for later use

$$2\mathbf{s}^{IJ} = \mathbf{h}^{II} + \mathbf{h}^{JJ} \quad (6d)$$

Here \mathbf{h}^{IJ} is approximately parallel to the derivative coupling vector,¹⁴ and $2\mathbf{g}^{IJ}$ is the gradient of the energy difference. These vectors define the branching¹⁵ or $g-h$ ¹⁶ space, the space in which the conical topography is evidenced. The orthogonal complement of the branching space is the seam space. Equation 5a has the usual interpretation¹⁷ of a constrained equilibrium. The “real force” ∇P , does not vanish but is instead balanced by the “pseudoforces” due to the constraints, $2\mathbf{g}^{IJ}$, \mathbf{h}^{IJ} , and \mathbf{k}^i . It is the Lagrange multipliers that prescribe the proper balance between the real and pseudoforces.

The solution to eqs 5a–5d is obtained by iteratively solving the system of Newton–Raphson equations

$$\begin{pmatrix} \nabla \nabla L^{IJ} & 2\mathbf{g}^{IJ} & \mathbf{h}^{IJ} & \mathbf{k} & \delta \mathbf{Q} \\ 2\mathbf{g}^{IJ\dagger} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \delta \lambda_1 \\ \mathbf{h}^{IJ\dagger} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \delta \lambda_2 \\ \mathbf{k}^\dagger & \mathbf{0} & \mathbf{0} & \mathbf{0} & \delta \zeta \end{pmatrix} = \begin{pmatrix} \nabla L^{IJ} \\ \Delta E_{IJ} \\ 0 \\ \mathbf{K} \end{pmatrix} \quad (7a)$$

$$(7b)$$

$$(7c)$$

$$(7d)$$

until $\mathbf{Q}^{x,IJ}$ is found such that $\text{NRN}(\mathbf{Q}^{x,IJ}) = \|\nabla L^{IJ}, \Delta E_{IJ}, 0, \mathbf{K}\| \sim 0$.

B. Rotational Invariance and Its Consequences. For any point \mathbf{Q} on the seam, the degenerate wave functions for states I and J are defined only up to a rotation among themselves by θ . As a result¹⁸ \mathbf{g}^{IJ} and \mathbf{h}^{IJ} are defined up to a rotation by 2θ . The \mathbf{g}^{IJ} and \mathbf{h}^{IJ} for the rotated states, $\mathbf{g}^{IJ,(\theta)}$ and $\mathbf{h}^{IJ,(\theta)}$, are related to the original or nascent $\mathbf{g}^{IJ,(0)} = \mathbf{g}^{IJ}$ and $\mathbf{h}^{IJ,(0)} = \mathbf{h}^{IJ}$ by

$$\begin{pmatrix} \mathbf{g}^{IJ,(\theta)} \\ \mathbf{h}^{IJ,(\theta)} \end{pmatrix} = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{pmatrix} \begin{pmatrix} \mathbf{g}^{IJ} \\ \mathbf{h}^{IJ} \end{pmatrix} \quad (8)$$

Note that in general $\|\mathbf{g}^{IJ,(\theta)}\| \neq \|\mathbf{g}^{IJ}\|$ and $\|\mathbf{h}^{IJ,(\theta)}\| \neq \|\mathbf{h}^{IJ}\|$ because \mathbf{g}^{IJ} and \mathbf{h}^{IJ} are not normalized. The flexibility in the definition of these vectors has several consequences. The values of the Lagrange multipliers are not unique. If θ in eq 8 is increased by θ^0 then λ_j ($j = 1, 2$) must then be modified to preserve eq 5a.

More significant in the present context is the ability to choose \mathbf{h}^{IJ} and \mathbf{g}^{IJ} orthogonal.¹⁸ In particular, $\mathbf{g}^{IJ,(\theta)} \equiv \bar{\mathbf{g}}^{IJ}$ and $\mathbf{h}^{IJ,(\theta)} \equiv \bar{\mathbf{h}}^{IJ}$ will be orthogonal for

$$\tan 4\theta = \frac{2\mathbf{h}^{IJ} \cdot \mathbf{g}^{IJ}}{(\mathbf{h}^{IJ} \cdot \mathbf{h}^{IJ}) - (\mathbf{g}^{IJ} \cdot \mathbf{g}^{IJ})} \quad (9)$$

The $\bar{\mathbf{h}}^{IJ}$ and $\bar{\mathbf{g}}^{IJ}$ obtained from eq 9 are unique up to transpositions and sign changes¹⁸ and all the $\mathbf{h}^{IJ,(\theta)}$ and $\mathbf{g}^{IJ,(\theta)}$ associated with a particular point of conical intersection yield the single orthogonal pair $\bar{\mathbf{h}}^{IJ}$ and $\bar{\mathbf{g}}^{IJ}$. Thus to each $\mathbf{Q}^{x,IJ}$ there is associated a unique pair $\bar{\mathbf{h}}^{IJ}$ and $\bar{\mathbf{g}}^{IJ}$, a consequence of which is illustrated in Figure 1. \mathbf{g}^{IJ} is shown at two points, $\mathbf{Q}^{(i)}$ and $\mathbf{Q}^{(i+1)}$ on the search path near but not on the seam. $\mathbf{g}^{IJ}(\mathbf{Q}^{(i)})$ and $\mathbf{g}^{IJ}(\mathbf{Q}^{(i+1)})$ differ markedly because $\mathbf{Q}^{(i)}$ and $\mathbf{Q}^{(i+1)}$, whose separation is small, fall on opposite sides of the cone so that the gradients point in opposite directions. This is a consequence of the square

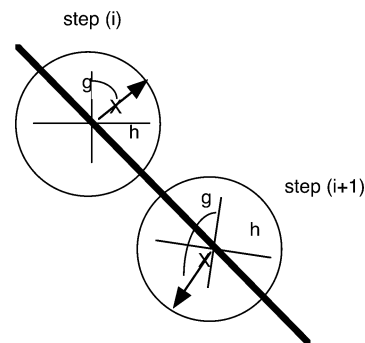


Figure 1. Schematic representation of slowly varying $\bar{\mathbf{g}}^{IJ}$ and $\bar{\mathbf{h}}^{IJ}$, denoted $\bar{\mathbf{g}}$ and $\bar{\mathbf{h}}$, and an erratically varying gradient, ∇E_I or \mathbf{g}^{IJ} , denoted by \rightarrow for two points on the search path.

root singularity in the energy (see below). Thus the $\mathbf{g}^{IJ}(\mathbf{Q})$ behave erratically along the path. However, because distinctly different \mathbf{g}^{IJ} and \mathbf{h}^{IJ} yield the same $\bar{\mathbf{g}}^{IJ}$, $\bar{\mathbf{h}}^{IJ}$ (it is not surprising and it is shown by example in ref 1), $\bar{\mathbf{g}}^{IJ}$ and $\bar{\mathbf{h}}^{IJ}$ are similar at $\mathbf{Q}^{(i)}$ and $\mathbf{Q}^{(i+1)}$; that is, they are slowly varying functions of the search path.

C. Choice of P . In the past, because algorithms seek an energy-minimized conical intersection, we have used $P = E_I$ (or E_J) in L^{IJ} .¹⁹ However, as in the case of \mathbf{g}^{IJ} , these ∇E_M behave erratically when the search path is near the seam, again as a consequence of the square root singularity in the energy. On the other hand the trace of the energies, $E^T = (E_I + E_J)/2$ with $\nabla E^T = \mathbf{s}^{IJ}$ does not involve the square root singularity (it is invariant under a unitary transformation of states I and J) and, as was shown by example in ref 1, is an extrapolatable function.¹¹ Note that although $P = E^T$ will not produce the same search path as $P = E_I$, the end result will be the same because on the seam of conical intersection if E_I is an extremum so is E_J .

III. Locating Energy-Minimized Sections of a Conical Intersection Seam

The material in the previous section, which was the subject of ref 1, provides the foundation for developing improved implementations of the Newton–Raphson equations for locating energy-minimized conical intersections. Note that all quantities in eq 7 are exactly and efficiently evaluated using analytic gradient techniques^{4,5} with the exception of the term $\nabla \nabla L^{IJ}$. It is the approximation of this term that is the central issue in this work. In the past we have used $P(\mathbf{Q}) = E_I(\mathbf{Q})$ and approximated $\nabla \nabla L^{IJ}$ using divided differences of ∇L^{IJ} , which is evaluated analytically. To avoid problems related to the erratic behavior of quantities in eq 7a, the divided differences were constructed at a “safe distance” from the conical intersection. The use of divided differences of ∇L^{IJ} to evaluate $\nabla \nabla L^{IJ}$ becomes problematical as the number of internal coordinates grows. The construction of $\nabla \nabla L^{IJ}$ using data obtained at a safe distance from the conical intersection can also be limiting. Here we develop an approach using extrapolatable functions that avoids these limitations.

A. Incorporating Extrapolatable Functions. Our approximation for $\nabla \nabla L^{IJ}$ is based on the following analysis of eq 7. Equations 7b and 7c are

$$\Delta E_{IJ}(\mathbf{Q}) + 2\mathbf{g}^{IJ}(\mathbf{Q}) \cdot \delta \mathbf{Q} = 0 \quad (7b')$$

$$\mathbf{h}^{IJ} \cdot \delta \mathbf{Q} = 0 \quad (7c')$$

Equation 7b' is just the first-order expansion of the energy

difference using nondegenerate perturbation theory. It requires $\delta\mathbf{Q}$ to move along, \mathbf{g}^{J} , the energy difference gradient, the tuning coordinate,² to recover the degeneracy. As $\delta\mathbf{Q}$ approaches the seam, nondegenerate perturbation theory becomes untenable. One manifestation of this is a nonnegligible (first-order) contribution to ΔE_{IJ} from $\mathbf{h}^{\text{J}} \cdot \delta\mathbf{Q}$. See below. Equation 7c' prevents $\delta\mathbf{Q}$, when it is not parallel to \mathbf{g}^{J} , from picking up a first-order contribution to ΔE_{IJ} .¹¹ As a result of the inclusion of both eqs 7b and 7c the path generated by the iterative solution to eq 7 first reaches a partially energy optimized region of the seam (see section IIIB) and then ultimately achieves full energy minimization.

We can exploit this observation as follows. For a point \mathbf{Q}^x on the seam, that is $\Delta E_{\text{IJ}} = 0$, but with eq 7a not satisfied, eqs 7b and 7c become

$$\begin{pmatrix} \mathbf{g}^{\text{J}}(\mathbf{Q}) \cdot \delta\mathbf{Q} \\ \mathbf{h}^{\text{J}}(\mathbf{Q}) \cdot \delta\mathbf{Q} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (10a)$$

which when multiplied by the matrix in eq 8 gives

$$\begin{pmatrix} \mathbf{g}^{\text{J},(\theta)}(\mathbf{Q}) \cdot \delta\mathbf{Q} \\ \mathbf{h}^{\text{J},(\theta)}(\mathbf{Q}) \cdot \delta\mathbf{Q} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (10b)$$

which is a specific case of the general observation that $\delta\mathbf{Q}$ must be perpendicular to the current $g-h$ plane. Whereas a specific \mathbf{g}^{J} is required to satisfy eq 7b' away from the seam, *on the seam any $\mathbf{g}^{\text{J},(\theta)}$, $\mathbf{h}^{\text{J},(\theta)}$ pair will suffice*. This result can be used to facilitate energy minimization, which is the provenance of eq 7a

$$-\nabla L^{\text{J}} = [\nabla s^{\text{J}}(\mathbf{Q}) + \lambda_1 \nabla \mathbf{g}^{\text{J}}(\mathbf{Q}) + \lambda_2 \nabla \mathbf{h}^{\text{J}}(\mathbf{Q}) + \sum_{\alpha=1}^{N^{\text{con}}} \zeta_{\alpha} \nabla \mathbf{k}^{\alpha}(\mathbf{Q})] \cdot \delta\mathbf{Q} \quad (7a')$$

Equation 7a is the most challenging of the components of eq 7 to solve owing to the presence of the second derivative terms ∇s^{J} , $\nabla \mathbf{g}^{\text{J}}$, and $\nabla \mathbf{h}^{\text{J}}$. Although any approximation that ultimately yields $\nabla L^{\text{J}} = 0$ is satisfactory, the more accurately ∇L^{J} is approximated the more rapidly the iterative procedure will converge. The erratic behavior of \mathbf{g}^{J} , \mathbf{h}^{J} (and \mathbf{h}^{J} , \mathbf{h}^{J} but NOT s^{J}) makes \mathbf{g}^{J} ill-suited for extrapolation or divided difference computation.

When \mathbf{Q} is close to a seam point, there is a unique ray (direction $\bar{\lambda}$) in the $g-h$ plane for which the computed energy difference gradient, $\nabla(\Delta E_{\text{IJ}})(\mathbf{Q})/2$, and coupling vector \mathbf{h}^{J} are to a good approximation $\mathbf{g}^{\text{J},(\bar{\lambda})}$ and $\mathbf{h}^{\text{J},(\bar{\lambda})}$. This intuitive result follows from the degenerate perturbation theory expression for the energies in the vicinity of a conical intersection

$$E_{\pm}(\mathbf{Q}) = s^{\text{J}}(\mathbf{Q})^{x,\text{J}} \pm |(\mathbf{g}^{\text{J}}(\mathbf{Q})^{x,\text{J}} \cdot \delta\mathbf{Q})^2 + (\mathbf{h}^{\text{J}}(\mathbf{Q})^{x,\text{J}} \cdot \delta\mathbf{Q})^2|^{1/2} \quad (11)$$

where $E_+ = E_{\text{J}}$ and $E_- = E_{\text{I}}$. Consequently (see Figure 2) for $\bar{\lambda}$ given by¹¹

$$\bar{\mathbf{g}}^{\text{J}} \cdot \delta\mathbf{Q} = \bar{q} \cos \bar{\lambda} \quad \bar{\mathbf{h}}^{\text{J}} \cdot \delta\mathbf{Q} = \bar{q} \sin \bar{\lambda} \quad (12a)$$

the energy difference gradient at \mathbf{Q} is $\mathbf{g}^{\text{J},(\bar{\lambda})}$, that is

$$\nabla(\Delta E_{\text{IJ}})(\mathbf{Q})/2 = \nabla(\Delta E_{-})(\mathbf{Q})/2 = \cos \bar{\lambda} \bar{\mathbf{g}}^{\text{J}} + \sin \bar{\lambda} \bar{\mathbf{h}}^{\text{J}} \equiv \bar{\mathbf{g}}^{\text{J},(\bar{\lambda})} \quad (12b)$$

Further by construction $\delta\mathbf{Q}$ is orthogonal to $\bar{\mathbf{h}}^{\text{J},(\bar{\lambda})}$

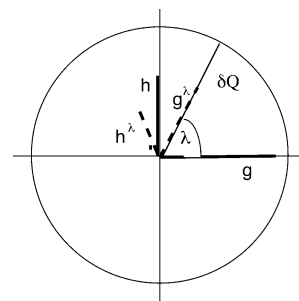


Figure 2. Schematic representation of quantities in eq 11, with $\bar{\mathbf{g}}^{\text{J}}$ denoted \bar{g} , $\bar{\mathbf{h}}^{\text{J}}$ denoted \bar{h} , $\bar{\mathbf{g}}^{\text{J},(\bar{\lambda})}$ denoted $\bar{g}^{\bar{\lambda}}$, and $\bar{\mathbf{h}}^{\text{J},(\bar{\lambda})}$ denoted $\bar{h}^{\bar{\lambda}}$.

$$\bar{\mathbf{h}}^{\text{J},(\bar{\lambda})} \cdot \delta\mathbf{Q} = (-\bar{\mathbf{g}}^{\text{J}} \sin \bar{\lambda} + \bar{\mathbf{h}}^{\text{J}} \cos \bar{\lambda}) \cdot \delta\mathbf{Q} = [\bar{\mathbf{g}}^{\text{J}}(\bar{\mathbf{h}}^{\text{J}} \cdot \delta\mathbf{Q}) + \bar{\mathbf{h}}^{\text{J}}(\bar{\mathbf{g}}^{\text{J}} \cdot \delta\mathbf{Q})] \bar{q} \cdot \delta\mathbf{Q} \quad (12c)$$

as required. Note that because these vectors are not normalized $\bar{\mathbf{g}}^{\text{J},(\bar{\lambda})}$ need not be, and in general is not, perpendicular to $\bar{\mathbf{h}}^{\text{J},(\bar{\lambda})}$.

In ref 1 it was shown numerically that $\bar{\mathbf{g}}^{\text{J}}$, $\bar{\mathbf{h}}^{\text{J}}$ are well-behaved along the search path. Thus the use of $\bar{\mathbf{g}}^{\text{J}}$ and $\bar{\mathbf{h}}^{\text{J}}$ transfers the erratic behavior from \mathbf{g}^{J} and \mathbf{h}^{J} to $\bar{\lambda}$. To use this result near the seam, the point \mathbf{Q} should be rotated from its nascent, random, position near $\mathbf{Q}^{x,\text{J}}$ to a preassigned position in the $g-h$ plane in each iteration. This amounts to searching (or walking) along a path parallel to the conical intersection seam, as suggested in the title of this work. To perform this rotation, explicitly second derivative information is required or the wave functions need be determined at the rotated point. These options are costly. However, if $\bar{\mathbf{g}}^{\text{J}}$ and $\bar{\mathbf{h}}^{\text{J}}$ are introduced only after the search produces \mathbf{Q} a point *on* the seam, the rotation in coordinate space can be replaced by a rotation of the degenerate electronic states—compare eqs 8 and 12b and see ref 11. In this way, the rotation is actually avoided, because both E^{T} and the $g-h$ plane (eq 10b) are invariant to this rotation.

The preceding analysis demonstrates the promised result: using E^{T} rather than E_{I} in the Lagrangian in eq 4, and replacing \mathbf{g}^{J} , \mathbf{h}^{J} by $\bar{\mathbf{g}}^{\text{J}}$, $\bar{\mathbf{h}}^{\text{J}}$ once ΔE_{IJ} is sufficiently small, yield a conical intersection search algorithm that is expressed in terms of extrapolatable functions. In this work we illustrate the improved performance that is possible with this approach using a simple updating procedure to approximate the three second derivative matrices in eq 7a.

B. Numerical Implementation: The Updating Scheme.

The three second derivatives in $\nabla \nabla L^{\text{J}}$, $\nabla \nabla E^{\text{T}} = \nabla s^{\text{J}}$, $\nabla \mathbf{g}^{\text{J}}$, and $\nabla \mathbf{h}^{\text{J}}$ are approximated using a simple updating procedure.¹² This approach costs no more than approximating $\nabla \nabla L^{\text{J}}$ by a unit matrix and hence is suitable when N^{int} is not small. It will also serve to illustrate the potential power of the using an extrapolatable function approach.

For a function $\mathbf{g}(\mathbf{Q}) = \nabla \Phi(\mathbf{Q})$, that is, for \mathbf{g} the gradient of Φ

$$\mathbf{g}(\mathbf{Q} + \delta\mathbf{Q}) = \mathbf{g}(\mathbf{Q}) + \nabla \mathbf{g}(\mathbf{Q}) \cdot \delta\mathbf{Q} \quad (13a)$$

so that

$$\mathbf{g}(\mathbf{Q} + \delta\mathbf{Q}) - \mathbf{g}(\mathbf{Q}) \equiv \delta \mathbf{g}(\mathbf{Q}) = \nabla \nabla \Phi(\mathbf{Q}) \cdot \delta\mathbf{Q} \equiv \mathcal{H} \delta\mathbf{Q} \quad (13b)$$

where \mathcal{H} is the Hessian for Φ . Equation 13b permits information obtained from the gradients at successive iterations (i and $i + 1$ below) to be incorporated into \mathcal{H} , as follows

$$\mathcal{H}^{(i+1)} \delta\mathbf{Q} = (\mathcal{H}^{(i)} + \Delta \mathcal{H}) \delta\mathbf{Q} = \delta \mathbf{g}(\mathbf{Q}) \quad (14a)$$

so that

$$\Delta\mathcal{H}\delta\mathbf{Q} = \delta\mathbf{g}(\mathbf{Q}) - \mathcal{H}^{(i)}\delta\mathbf{Q} \equiv \delta\mathbf{g}^p(\mathbf{Q})$$

and

$$\delta\mathbf{Q}^\dagger \Delta\mathcal{H}\delta\mathbf{Q} = \delta\mathbf{Q}^\dagger \delta\mathbf{g}^p(\mathbf{Q}) \quad (14b)$$

Therefore

$$\mathcal{H}^{(i+1)} = (\mathcal{H}^{(i)} + \Delta\mathcal{H}) = \mathcal{H}^{(i)} + e\delta\mathbf{g}^p\delta\mathbf{g}^{p\dagger} \quad (14c)$$

where

$$e = 1/[\delta\mathbf{Q}^\dagger \delta\mathbf{g}^p(\mathbf{Q})] \quad (14d)$$

A detailed discussion of this approach can be found in ref 12.

In this work we will denote by M the update approach used to approximate $\nabla\nabla L^J$, with $M = I$ denoting no updating, that is $\nabla\nabla L^J = I$, and $M = Uab$ indicating ∇s^J is updated at each iteration and $\nabla\mathbf{g}^J$ and $\nabla\mathbf{h}^J$ are updated and included in $\nabla\nabla L^J$ provided $\Delta E_{IJ} < a \times 10^{-b}$. In this scheme $\mathcal{H}^0 = I$. No attempt was made to use conjugate gradient techniques¹² normally included in updating, as that would require repeated determination of the electronic wave functions, the most time-consuming part of a step in the iterative procedure. Uab uses only information normally computed at each step in the iterative procedure. It therefore costs no more (per iteration) than $M = I$.

C. The 1,2²A States of NH₃. The utility of this approach was assessed using the 1¹A–2¹A conical intersections of NH₃. These conical intersections, which are relevant to the recent vibrationally mediated photodissociation in Crim's laboratory,²⁰ have also been determined at the state-averaged multiconfigurational self-consistent field level.^{21,22} In this work two wave functions are used, a first-order configuration interaction (FOCI) wave function [approximately 350 000 configuration state functions (CSFs)] and a second-order configuration interaction (SOC) wave function [approximately 8 million CSFs] each based on an 8 electron–8 orbital active space using standard cc-tzp bases with diffuse functions on nitrogen. All numerical calculations were performed with a prerelease version of the COLUMBUS⁹ suite of electronic structure codes.

Figure 3a reports $E_{1^1A}(M)$, $E_{2^1A}(M)$, and $\Delta E_{1^1A,2^1A}(M)$ and Figure 3b reports NRN(M) along a search path with one R(NH) constrained to $2.0a_0$ and the remaining coordinates optimized using method $M = I$ or Uab . Here the FOCI wave functions were used. The $\Delta E_{1^1A, 2^1A}(I)$ show that $M = I$ quickly locates points on the seam. However, the NRN(I) shows that the $M = I$ method is too slowly convergent to be of practical value. From $E_{1^1A}(U00)$ it is seen that $M = U00$ recovers most of energy lowering missed by $M = I$. Note that $M = I$ reaches the seam earlier in the iterative sequence than Uab ; see $\Delta E_{1^1A, 2^1A}$. This indicates that improving ∇s^J produces a better balance between energy lowering and energy difference reduction. The sequence of results for $(a, b) = (5, 5)$, $(1, 5)$, and $(5, 6)$ is particularly encouraging. Each $(a, b) = (5, 5)$, $(1, 5)$, and $(5, 6)$ deviates from the $(0, 0)$ result at successive higher values of iteration index and each produces more converged results, that is, smaller values of NRN(M) and of $\Delta E_{1^1A,2^1A}(M)$ [for $(1, 5)$ and $(5, 6)$]. In particular $\text{NRN}(U00) > \text{NRN}(U55) > \text{NRN}(U15) \sim \text{NRN}(U56)$ and although $\Delta E_{1^1A,2^1A}(U00)$ is slightly less than $\Delta E_{1^1A,2^1A}(U55)$, $\Delta E_{1^1A,2^1A}(U00) > \Delta E_{1^1A,2^1A}(U15) \sim \Delta E_{1^1A,2^1A}(U56)$.

The second example, which uses the SOC treatment and $M = U16$, focuses on the switch to the orthogonal representation when ΔE_{IJ} is small. Figure 4a reports $\Delta E_{1^1A,2^1A}$ and NRN. This figure evinces the good convergence obtained in the previous example with $\text{NRN} < 2 \times 10^{-5}$ after 13 iterations. The vertical double headed arrow indicates that at iteration 7, $\Delta E_{1^1A,2^1A} < 10^{-6}$, at which time use of the orthogonal representation begins.

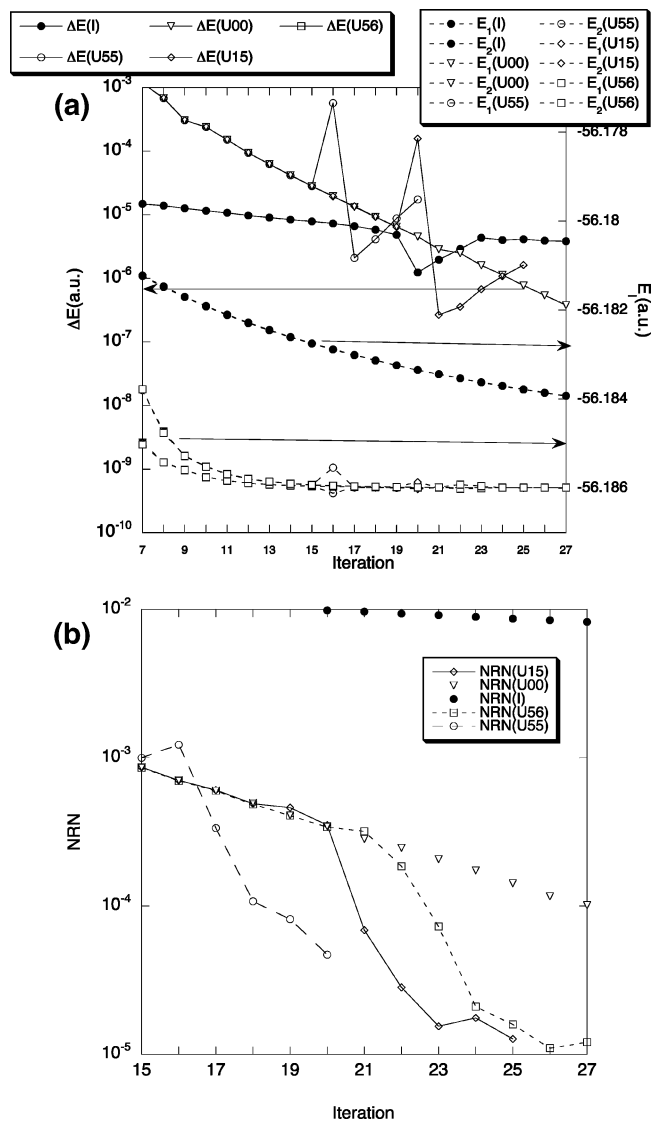


Figure 3. Plot of (a) $E_1(M)$, $E_2(M)$, and $\Delta E_{12}(M)$ and (b) NRN(M) for $M = I, Uab$. Here $M = I$ implies $\nabla\nabla L$ set to I and Uab implies $\nabla\nabla L$ approximated by ∇s^J , which in turn is approximated by updating, with $\nabla\mathbf{g}^J$ and $\nabla\mathbf{h}^J$ included and updated provided $\Delta E_{IJ} < a \times 10^{-b}$.

The affect of this change is seen in Figure 4b, which reports g^J_2 , h^J_2 (the second component of the indicated vectors, which is representative), the associated Lagrange multipliers λ_1 , λ_2 , and for comparison, s^J_2 . Here and below $I = 1^1A$, $J = 2^1A$. Prior to iteration 7, g^J_2 , h^J_2 show considerable variation, reflecting, at least for the iterations with $\Delta E_{1^1A,2^1A}$ small, the erratic behavior of \mathbf{g}^J and \mathbf{h}^J . This behavior abruptly changes at iteration 7, where $\mathbf{g}^J \rightarrow \bar{\mathbf{g}}^J$ and $\mathbf{h}^J \rightarrow \bar{\mathbf{h}}^J$ and so become slowly varying functions of the iteration index. Note that s^J is slowly varying throughout. These results support the viability of the proposed numerical approach.

Though far from a conclusive demonstration of the efficacy of this approach, the results are quite encouraging. Performance can be expected to improve even further when more sophisticated extrapolation techniques are used.

IV. Summary and Conclusions

The search path to an energy-minimized point of conical intersection involves quantities that behave erratically as the conical intersection is approached. This erratic behavior, which is a consequence of the proximity to the conical intersection, precludes the routine use of extrapolation techniques that can

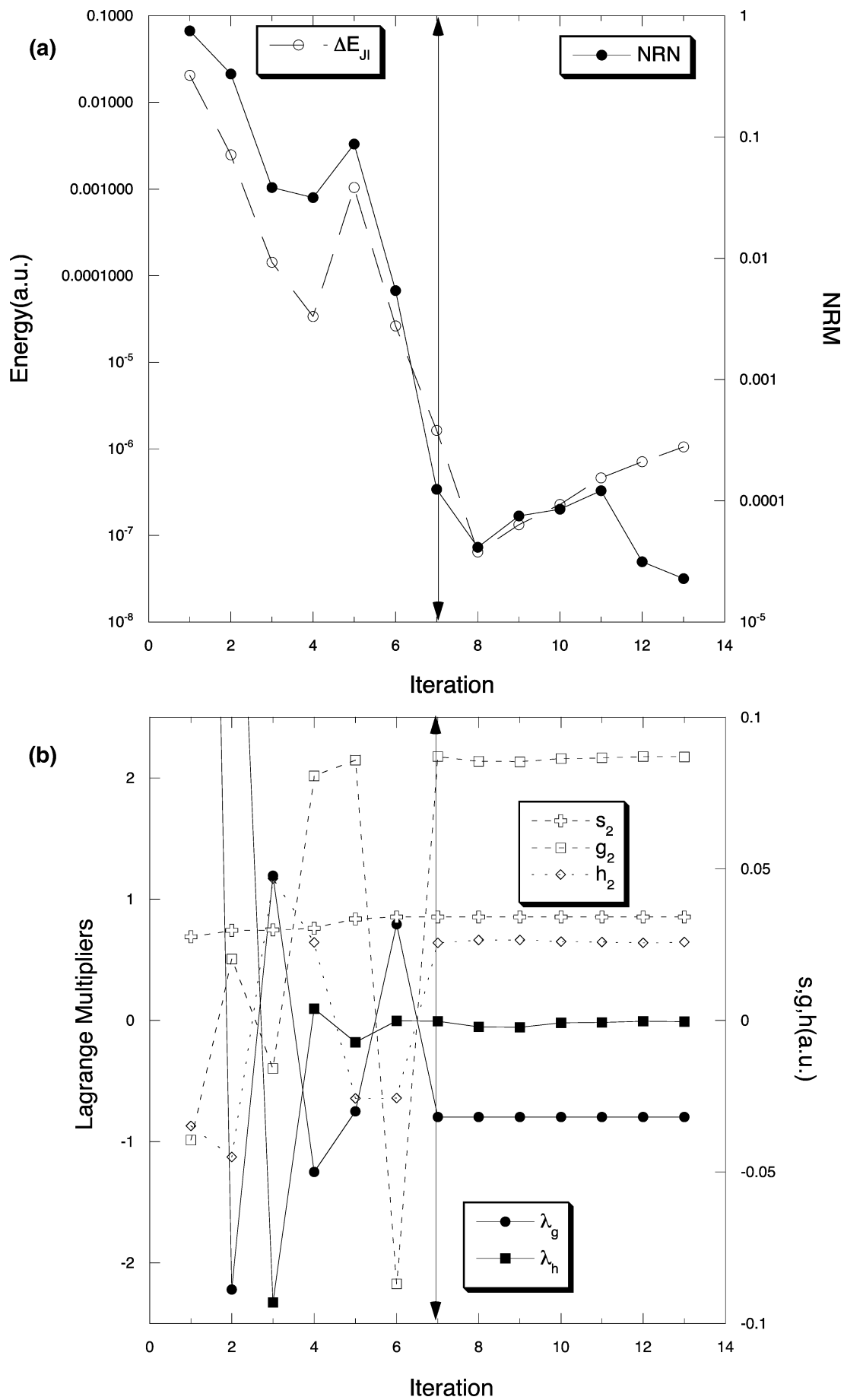


Figure 4. Plot of (a) $\Delta E_{Jl}(M)$ and $NRN(M)$ and (b) s^U_2 , g^U_2 , h^U_2 , with superscript suppressed λ_1 , and λ_2 for $M = U16$.

greatly reduce the time to solution of the iterative procedure. Because the erratic behavior is associated with the conical intersection itself, this behavior might appear to be intrinsic to this type of search. However, this is not the case. Here we have used extrapolatable functions¹¹ to reformulate our two-state energy-minimized conical intersection search procedure to employ functions that are smoothly varying along the search path even as the conical intersection seam is approached. As a result, a simple updating procedure was shown to be effective in improving the performance of this algorithm, opening the way for the use of more sophisticated extrapolation techniques.

Acknowledgment. I thank M. Dallos and Hans Lischka for access to a prerelease copy of COLUMBUS, which was used for all electronic structure calculations used in this study. The support of the National Science Foundation through grant CHE 0206834 is gratefully acknowledged.

References and Notes

- (1) Bernardi, F.; Olivucci, M.; Robb, M. A. *Chem. Soc. Rev.* **1996**, 25, 321.
- (2) Köppel, H.; Domcke, W.; Cederbaum, L. S. *Adv. Chem. Phys.* **1984**, 57, 59.
- (3) Yarkony, D. R. *J. Phys. Chem. A* **2001**, 105, 6277.
- (4) Lengsfeld, B. H.; Yarkony, D. R. *Nonadiabatic Interactions Between Potential Energy Surfaces: Theory and Applications*. In *State-Selected and State to State Ion-Molecule Reaction Dynamics: Part 2 Theory*; Baer, M., Ng, C.-Y., Eds.; John Wiley and Sons: New York, 1992; Vol. 82, p 1.
- (5) Lischka, H.; Dallos, M.; Szalay, P.; Yarkony, D. R.; Shepard, R. *J. Chem. Phys.* **2004**.
- (6) Bearpark, M. J.; Robb, M. A.; Schlegel, H. B. *Chem. Phys. Lett.* **1994**, 223, 269.
- (7) Manaa, M. R.; Yarkony, D. R. *J. Am. Chem. Soc.* **1994**, 116, 11444.
- (8) Lischka, H.; Dallos, M.; Sheppard, R.; Scazlay, P.; Yarkony, D. R. Work in progress.
- (9) Lischka, H.; Shepard, R.; Shavitt, I.; Pitzer, R.; Dallos, M.; Müller, T.; P. G.Szalay; Brown, F. B.; Alhrichs, R.; Böhm, H. J.; Chang, A.; Comeau, D. C.; Gdanitz, R.; Dachsel, H.; Erhard, C.; Ernzerhof, M.; Höchtel, P.; Irle, S.; Kedziora, G.; Kovar, T.; Parasuk, V.; Pepper, M.; Scharf, P.; Schiffer, H.; Schindler, M.; Schüller, M.; Zhao, J.-G. COLUMBUS, An ab initio Electronic Structure Program, 5.9 ed., 2003.
- (10) von Neumann, J.; Wigner, E. *Phys. Z.* **1929**, 30, 467.
- (11) Yarkony, D. R. *Faraday Discuss.* **2004**, 127.
- (12) Murtagh, B. A.; Sargent, R. W. H. *Comput. J.* **1970**, 13, 185.
- (13) Yarkony, D. R. Molecular Structure. In *Atomic, Molecular and Optical Physics Handbook*; Drake, G. L., Ed.; AIP: New York, 1996; p 357.
- (14) Yarkony, D. R. *J. Phys. Chem. A* **1997**, 101, 4263.
- (15) Atchity, G. J.; Xantheas, S. S.; Ruedenberg, K. *J. Chem. Phys.* **1991**, 95, 1862.
- (16) Yarkony, D. R. *Acc. Chem. Res.* **1998**, 31, 511.
- (17) Goldstein, H. *Classical Mechanics*; Addison-Wesley: Reading, MA, 1950.
- (18) Yarkony, D. R. *J. Chem. Phys.* **2000**, 112, 2111.
- (19) Yarkony, D. R. *Rev. Mod. Phys.* **1996**, 68, 985.
- (20) Bach, A.; Hutchison, J. M.; Holiday, R. J.; Crim, F. F. *J. Chem. Phys.* **2003**, 118, 7144.
- (21) McCarthy, M. I.; Rosmus, P.; Werner, H.-J.; Botschwina, P.; Vaida, V. *J. Chem. Phys.* **1987**, 86, 6693.
- (22) Rosmus, P.; Botschwina, P.; Werner, H.-J.; Vaida, V.; Engelking, P. C.; McCarthy, M. I. *J. Chem. Phys.* **1987**, 86, 6677.