On the construction of quasidiabatic state representations of bound adiabatic state potential energy surfaces coupled by accidental conical intersections: incorporation of higher order terms

Dillon, Joseph; Yarkony, David R.; Schuurman, Michael S.
On the construction of quasidiabatic state representations of bound adiabatic state potential energy surfaces coupled by accidental conical intersections: Incorporation of higher order terms

Joseph Dillon, a,b) David R. Yarkony, a,b) and Michael S. Schuurman2,c)

1Department of Chemistry, Johns Hopkins University, Baltimore, Maryland 21218, USA
2Siegert Institute for Molecular Sciences, National Research Council, Ottawa K1N 6C1, Canada

(Received 11 October 2010; accepted 10 November 2010; published online 24 January 2011)

The quadratic vibronic coupling model is an important computational tool for simulating photoelectron spectra involving strongly coupled electronic states in polyatomic molecules. However, recent work has indicated the need for higher order terms, with most of the initial studies focusing on molecules with symmetry-required degeneracies. In this study we report an extension of our approach for constructing fully quadratic representations of bound electronic states coupled by conical intersections, which allows for the inclusion of higher order terms, demonstrated here employing a quartic expansion. Procedures are developed that eliminate unphysical behavior for large displacements, a problem likely to be an endemic to anharmonic expansions. Following work on representing dissociative electronic states, Lagrange multipliers are used to constrain the constructed representation to reproduce exactly the energy, energy gradients, and/or derivative couplings at specific points, or nodes, in nuclear coordinate space. The approach is illustrated and systematically studied using the four lowest electronic states of triazolyl, (CH)2N3.

I. INTRODUCTION

In the multimode vibronic coupling approach, the energies and interactions of N\textsuperscript{state} adiabatic electronic states coupled by conical intersections are represented by a symmetric N\textsuperscript{state} × N\textsuperscript{state} coupled quasidiabatic Hamiltonian, H\textsuperscript{d}, whose matrix elements are polynomials of varying degree in a set of N\textsuperscript{int} = 3N\textsuperscript{at} − 6 internal coordinates. Several forms for H\textsuperscript{d} have been used since the model was introduced over 25 years ago. In the linear vibronic coupling model, the off-diagonal blocks of H\textsuperscript{d} are linear functions of the internal coordinates, while the quadratic terms in the diagonal blocks do not involve terms which couple the vibrational modes. The quadratic vibronic coupling approximation exists in several forms depending on the terms that are added to the diagonal and off-diagonal blocks. In the most general formulation, the fully quadratic vibronic coupling model, all second-order terms are included in both the diagonal and off-diagonal blocks. Regardless of the functional form of H\textsuperscript{d}, point group symmetry is universally employed to significantly reduce the number of nonzero expansion coefficients that require computation.

The accuracy of any vibronic coupling model is related to the topography of the coupled potential energy surfaces being described and the spatial extent of the region of interest. For example, a quadratic model may be adequate to describe a set of potential energy surfaces in which the low-lying seams of conical intersection are close, geometrically and energetically, to the relevant ground state surface extrema, i.e., minima and saddle points. However, the utility of this simple approach deteriorates as the separation between these regions increases. In a recent work, we have systematically studied the limitations of the fully quadratic vibronic coupling model. These shortcomings in the fully quadratic vibronic coupling model can be mitigated by the inclusion of higher order terms in the polynomial expansion. There have been a limited number of studies that have included higher order terms for coupled potential energy surfaces.

In this work, we consider the addition of anharmonic terms to the fully quadratic vibronic coupling model in the case where the conical intersections are accidental and the low-lying seams of the conical intersection are spatially separated from the global minimum on the ground state potential energy surface. This is accomplished using an extension of our previously reported algorithm for determining fully quadratic, quasidiabatic bound-state Hamiltonians (H\textsuperscript{d}). The extended algorithm discussed here permits the use of arbitrary order polynomials. Inclusion of higher order polynomials can lead to unphysical behavior for large displacements, including unbound potential energy surfaces. Techniques to overcome this potentially endemic problem are developed. The algorithm also incorporates a Lagrange multiplier based constraint procedure that requires specific ab initio data to be reproduced exactly, while the remaining data are approximated in a least square sense. This Lagrange multiplier approach was first introduced in our algorithm for describing coupled adiabatic potential energy surfaces for dissociative states which utilized high order polynomials in exponentially decaying functions of internuclear distances. There, the Lagrange multiplier constraints were shown to be
highly effective in reducing the root mean square (rms) error in the fit.

Three types of *ab initio* data are included: energies, energy gradients, and derivative couplings. Since derivative couplings are included in the fit, we will be able to quantify the residual derivative coupling in this representation. The representation is labeled quasidiabatic since it is only diabatic in a least squares sense and, in any case, truly diabatic states, in general, are not possible.\textsuperscript{14,15}

Section II briefly reviews the fitting procedure. Section III presents the numerical study, which considers the low-lying electronic states of the triazolyl radical, (CH)$_2$N$_3$. Section IV summarizes and discusses directions for future research.

Triazolyl provides numerous challenges for a quasidiabatic representation. This molecule falls in the class of nitrogen containing five member ring compounds known as azolyls and can be viewed as a derivative of the well studied Jahn–Teller radical (CH)$_3$, cyclopentadienyl,\textsuperscript{16–18} which has a degenerate $E_u$ ground electronic state derived from $\pi$ orbitals. In triazolyl the minimum energy conical intersection (MEI) and minimum on the ground state potential energy surface, both of which have $C_2v$ symmetry, are relatively well separated. Further, in triazolyl the low-lying electronic states do not arise exclusively from the $\pi$ orbitals. Two additional electronic states, energetically close to the ground state, arising from excitations out of the nitrogen lone pair orbitals, are interspersed within the $\pi\pi^*$ manifold of cyclopentadienyl-like electronically excited states. Consequently, the MEI involves an intersection with an $n\pi^*$ state, in contrast to the $\pi\pi^*$ intersections observed in the cyclopentadienyl and other azolyl antecedents. Additional seams of accidental two and three state conical intersections are also found. This intricate topography of the low-lying electronic states in the triazolyl radical presents an ideal opportunity to assess the benefits that can be obtained from incorporating higher order terms in the vibronic expansion.

II. THE DIABATIC REPRESENTATION AND ITS DETERMINATION

The diabatic representation, $H^d$, is intended to approximate the *ab initio* energies $E_d(Q)$, the energy gradients $\nabla E_d(Q)$, and the derivative couplings $\Gamma^{(d)}(Q)$ for $N^{\text{state}}$ adiabatic electronic states, for a bound molecule. Here $Q$ is a set of $N^{\text{int}} = 3N^{\text{at}} - 6$ internal coordinates.

A. The quasidiabatic Hamiltonian

$H^d_{\alpha,\beta} = E_d(Q)\delta_{\alpha,\beta} + \sum_{k=1}^{N^{\text{int}}} V_k^{(1),\alpha,\beta} Q_k$

$+ \frac{1}{2} \sum_{k,l=1}^{N^{\text{int}}} V_k^{(2),\alpha,\beta} Q_k Q_l$

$+ \frac{1}{3} \sum_{k,l,m=1}^{N^{\text{int}}} V_k^{(3),\alpha,\beta} Q_k Q_l Q_m + \cdots,$ \hspace{1cm} (1)

where $V^{(k)}$ are to be determined. The adiabatic energies and eigenstates are approximated by the solutions of the electronic Schrödinger equation:

$$[H^a(Q) - I E_0^{(d)}(Q)]d^a(Q) = 0.$$ \hspace{1cm} (2)

B. Equations defining $V^{(k)}$

The equations defining $V^{(k)}$ are obtained by differentiating Eq. (2) and using the derivative of Eq. (1), which yields

$$M^{I,J,(d)}_{k} = M^{I,J,(d)}_{k} Q_{k} - \sum_{\alpha,\beta} d_{\alpha}^{a} d_{\beta}^{a} V^{(1),\alpha,\beta}_{k}$$

$$= \sum_{\alpha,\beta,l} [d_{\alpha}^{a} d_{\beta}^{a} Q_l] V^{(2),\alpha,\beta}_{k}$$

$$+ \sum_{\alpha,\beta,l,m} [d_{\alpha}^{a} d_{\beta}^{a} Q_l Q_m] V^{(3),\alpha,\beta}_{k,l,m} + \cdots,$$ \hspace{1cm} (3a)

where

$$M_{k}^{I,J,(d)}(Q) = \frac{\partial}{\partial Q_k} E_0^{(d)}(Q)$$

and

$$M_{k}^{I,J,(d)}(Q) = \left( E_0^{(d)}(Q) - E_I^{(d)}(Q) \right) \left( d^{I |} \frac{\partial}{\partial Q_k} d^{I} \right),$$ \hspace{1cm} (3b)

where in the second Eq. (3b) $I \neq J$. Inserting Eq. (1) into Eq. (2) gives

$$M_{0}^{I,I,(d)}(Q) = \sum_{\alpha,\beta,k} [d_{\alpha}^{a} d_{\beta}^{a} Q_k] V^{(1),\alpha,\beta}_{k}$$

$$+ \frac{1}{2} \sum_{\alpha,\beta,k,l} [d_{\alpha}^{a} d_{\beta}^{a} Q_k Q_l] V^{(2),\alpha,\beta}_{k,l}$$

$$+ \frac{1}{3} \sum_{\alpha,\beta,k,l,m} [d_{\alpha}^{a} d_{\beta}^{a} Q_k Q_l Q_m] V^{(3),\alpha,\beta}_{k,l,m} + \cdots,$$ \hspace{1cm} (3c)

where

$$M_{0}^{I,I,(d)}(Q) = E_0^{(d)}(Q) - E_I^{(d)}(Q).$$ \hspace{1cm} (3d)

As discussed previously,\textsuperscript{3} $V^{(1),\alpha,\beta}_{k}$ are obtained using analytic gradient techniques.\textsuperscript{19} If $M^{I,J,(d)}$ and $E_0^{(d)}(Q)$ (and hence $M_{0}^{I,I,(d)}$) are replaced by the corresponding *ab initio* determined values, $M^{I,J}$, $E_I$, and $M^{I,I}$, Eqs. (3a) and (3c) assert that an $H^d$ determined energy, energy gradient or interstate coupling gradient (energy difference multiplied by the derivative coupling), is equal to the corresponding *ab initio* quantity. In this manner, Eqs. (3a) and (3c) relate the unknown coefficients $V^{(k),\alpha,\beta}_{k}$ to *ab initio* quantities. Note that to construct Eq. (3a) for $I \neq J$, the signs of the $\alpha$ and $\beta$ indices must be consistent at each geometry (there is no requirement that the sign at different geometries are compatible). This is achieved by adjusting the overall sign of $d^I$, $I = 1 - N^{\text{state}}$ until consistency is achieved. In addition, at points of near degeneracy, the energy gradients may have to be reordered before Eq. (3a) is evaluated for $I = J$. 

C. Constrained pseudonormal equations

In order to determine the above expansion coefficients one requires \textit{ab initio} data at \(N^{\text{point}}\) displaced geometries, which are denoted \(Q(m), 1 \leq m \leq N^{\text{point}}\). Then, the left hand sides of Eqs. (3a) and (3c), \(M_{k,l}^{\alpha}Q(m)\), form a vector \(\mathbf{M}\) of length \(N_{\text{eq}} = [(N^{\text{state}}N^{\text{state}} + 1)/2]N^{\text{int}} + N^{\text{state}}N^{\text{points}}\), where the first and second terms enumerate the number of equations derived from gradient and energy data, respectively. The \(N^{\text{uniq}}\) unique (Ref. 4) elements of \(V^{\alpha}_{(j)\alpha,\beta}\) may also be indexed as a vector \(\mathbf{V}\), and Eqs. (3a) and (3c) can be rewritten as

\[\mathbf{WV} = \mathbf{M},\]

where \(\mathbf{W}\) is a matrix of dimension \(N^{\text{eq}} \times N^{\text{uniq}}\), constructed from the factors in the square brackets on the right hand sides of Eqs. (3a) and (3c).

In general, \(N^{\text{eq}} > N^{\text{uniq}}\) so that the system of equations is overdetermined. Consequently, we partition Eq. (4) into two sets of equations. The first \(N^{\text{bind}}\) equations, which are to be solved in a least squares sense, will be denoted below by the superscript \(\text{lsq}\). The remaining \(N^{\text{eq}} - N^{\text{bind}} = N^{\text{nodes}}\) equations are solved exactly to determine the nodes, denoted below by the superscript \(\text{node}\). Then, the desired solution of the equations that comprise Eq. (4) is obtained by minimizing the Lagrangian:

\[L(V, \lambda) = 1/2 \sum_{j=1}^{N^{\text{lsq}}} \left( M_j - \sum_{i=1}^{N^{\text{uniq}}} W_{ji} V_i \right)^2 + \sum_{j=N^{\text{lsq}}+1}^{N^{\text{eq}}} \lambda_j \left( M_j - \sum_{i=1}^{N^{\text{uniq}}} W_{ji} V_i \right) + 1/2 V^\dagger t V \]

with respect to both the parameters \(V_j\) and the Lagrange multipliers \(\lambda_j\), which gives

\[\begin{pmatrix} W^{\text{lsq}} & -W^{\text{node}} \\ -W^{\text{node}} & 0 \end{pmatrix} \begin{pmatrix} V \\ \lambda \end{pmatrix} = \begin{pmatrix} W^{\text{lsq}} & M^{\text{lsq}} \\ M^{\text{node}} & \lambda \end{pmatrix}. \]

The final term in Eq. (5), \(t = \mathbf{I}\), is a small, positive damping term. It serves to prevent parameters that are only marginally defined using a given data set from producing an ill-behaved fit. Equation (6) and its defining equations, Eqs. (3a)–(3d), determine \(V\). These equations must be solved self-consistently, since the \(d\) are not known until \(V\) are determined.

III. APPLICATION: TRIAZOLYL

A. Electronic structure of the low-lying states of triazolyl

As noted in Sec. I, triazolyl is a substitutional derivative of the cyclopentadienyl radical, \((\text{CH})_2\), and falls in the class of molecules known as azolyls \((N_x(\text{CH})_y)\), with \(x + y = 5\). The cyclopentadienyl radical displays a Jahn–Teller distorted \(2E_i\) ground state that originates from an incompletely filled \(\pi\) orbital. Substitution of \(\text{CH}\) groups with nitrogen atoms reduces the molecular point group symmetry from \(D_{5h}\) to \(C_{2v}\), thereby precluding the possibility of symmetry-required degeneracies. Still a number of these substituted species do exhibit, the cyclopentadienyl paradigm, low-lying \(2A_2\) and \(2B_1\) symmetry electronic states analogous to the components of a degenerate \(2E_i\) state. This analysis was found to be valid in previous investigations of the pyrrolyl, \(10,20,21\) \(N(\text{CH})_4\), and pyrazolyl, \(22,23\) \(N_2(\text{CH})_3\), radicals. In these species, the low-lying electronic states displayed the same general structure: a ground state minimum of \(2A_2\) symmetry and an MEX between the \(2B_1\) and \(2A_2\) electronic states. Although the pyrrolyl radical required only two electronic states to accurately simulate the photoelectron spectrum of the precursor anion, this was not the case in the pyrazolyl radical where interactions with a third electronic state resulting from a nitrogen lone pair excitation were essential.\(22,23\)

The addition of a third adjacent nitrogen atom produces the triazolyl radical, and the growing number of nitrogen lone pair orbitals yields even more substantial deviations from the original scheme. Excitations from these lone pair orbitals give rise to two low-lying \(n\pi\pi^*\) \(2A_1\) and \(2B_2\) electronic states in addition to the ground \(2B_1\) and \(\pi\pi\pi^*\) \(2A_2\) states. In fact, the first excited state is an \(n\pi\pi^*\) transition, with the four lowest electronic states given in order of increasing energy by: \(2B_1 < 2A_1 < 2A_2 < 2B_2\). This intercalation of \(n\pi\pi^*\) and \(\pi\pi\pi^*\) character electronic states indicates that the cyclopentadienyl paradigm has exhausted its utility.

The switch in symmetry and character of the first excited state profoundly changes the structure of the branching space at the MEX, which now includes a component out of the molecular plane. These four low-lying states give rise to two and three state conical intersections, as in the related pyrazolyl radical, \((\text{CH})_3N_2\).\(22–24\)

Thus the electronic structure of triazolyl is considerably more complex than its azolyl antecedent’s pyrrolyl and pyrazolyl. It is precisely for this reason that the triazolyl radical was expected and indeed found to be an ideal test case for the incorporation of higher order terms in the construction of \(H^4\).

B. Characterization of the electronic structure

While it is convenient to employ \(C_{2v}\) symmetry labels to discuss the electronic structure of triazolyl, all \textit{ab initio} calculations were done in the \(C_1\) symmetry. The molecular orbitals were determined employing a double zeta polarization (DZP) basis set\(25\) in conjunction with a four state-averaged MCSCF procedure using an 11 electron, 8 orbital complete active space (CAS). The active space includes the five \(\pi\) orbitals \((3b_1, 2a_2)\) and three lone pair nitrogen orbitals \((2a_1, 1b_2)\). Electron correlation was treated at the multireference single and double configuration interaction (MR-CISD) level of theory. With a frozen core of nitrogen and carbon 1s orbitals, the MR-CISD wave functions are comprised of all single and double excitations relative to the CAS and to the remaining doubly occupied \(\sigma\)-bonds. The resulting MR-CISD expansion is comprised of \(\sim 200 \times 10^6\) configuration state functions. All electronic structure calculations reported in this work employed the COLUMBUS suite of electronic structure codes.\(10,26\) Figure 1 depicts the triazolyl radical and indicates the atom labeling used in this section.
C. Reference structures

Three key C<sub>2v</sub> extrema were identified: a minimum, \( \text{min}^2 \text{B}_1 \) and saddle point, \( \text{ts}^2 \text{A}_1 \) on the ground state potential energy surface, and the minimum energy \( 2 \text{B}_1 - 2 \text{A}_1 \) conical intersection, \( \text{mex}^2 \text{B}_1 - 2 \text{A}_1 \). The two \( \sigma \) planes will be denoted \( \sigma_x \) and \( \sigma_y \) for the molecular and \( \angle \text{NNN} \) bisection planes, respectively. The geometries of the extrema are summarized in Table I. The curvature of the potential energy surface at \( \text{min}^2 \text{B}_1 \) and \( \text{ts}^2 \text{A}_1 \) was established by constructing the \( \text{ab initio} \) Hessians at each critical point. The harmonic frequencies for \( \text{min}^2 \text{B}_1 \) and \( \text{ts}^2 \text{A}_1 \) are reported in Table II.

The most significant conclusion from Table I, as it pertains to the present study, concerns the proximity of \( \text{mex}^2 \text{B}_1 - 2 \text{A}_1, \text{min}^2 \text{B}_1 \), and \( \text{ts}^2 \text{A}_1 \). Focusing on \( \sigma(\text{N}1-\text{N}2) \) and \( \angle \text{NNN} \), it is seen that although \( \text{mex}^2 \text{B}_1 - 2 \text{A}_1 \) lies between \( \text{min}^2 \text{B}_1 \) and \( \text{ts}^2 \text{A}_1 \), it is appreciably closer to \( \text{ts}^2 \text{A}_1 \) and rather well separated from \( \text{min}^2 \text{B}_1 \). Similar conclusions are obtained from the energies also reported in Table I, where \( \text{mex}^2 \text{B}_1 - 2 \text{A}_1 \) is 3737 (415) cm\(^{-1}\) above \( \text{min}^2 \text{B}_1 \) (\( \text{ts}^2 \text{A}_1 \)).

D. Orthogonal intersection adapted Coordinates

The relative geometric and energetic relationships between the extrema located in the vicinity of \( \text{mex}^2 \text{B}_1 - 2 \text{A}_1 \) are best evinced in terms of the orthogonal\(^\text{27}\) intersection adapted coordinates.\(^\text{28}\) Orthogonal intersection adapted coordinates are an orthogonal transformation of the internal coordinates with a conical intersection as the origin. In this coordinate system, the \( x \) and \( y \) coordinates are collinear with the \( g \) (energy difference gradient) and \( h \) (interstate coupling gradient) directions, respectively, and thus define the branching\(^\text{28}\) or \( g-h \) plane. The remaining coordinates need only be mutually orthogonal and orthogonal to \( x \) and \( y \). At \( \text{mex}^2 \text{B}_1 - 2 \text{A}_1 \), the \( x \) and \( y \) directions transform as \( a_1 \) and \( b_1 \), respectively, with the displacements along the \( y \) direction involving motion out of the molecular plane. These vectors are reported in Fig. 1, while the entire transformation to orthogonal intersection adapted coordinates is provided in the supplemental data.\(^\text{29}\)

Note in \( C_{2v} \) symmetry there are \((6a_1, 2a_2, 2b_1, 5b_2)\) internal coordinates. Expressing the relevant extrema in intersection-adapted coordinates, one finds that \( \text{min}^2 \text{B}_1 \) is displaced from \( \text{mex}^2 \text{B}_1 - 2 \text{A}_1 \) almost exclusively along the \( -x \) direction, while \( \text{ts}^2 \text{A}_1 \) is primarily along the \( +x \) direction. Thus, for each critical point, it is seen that the dominant displacement involves branching plane coordinates. The terms in \( \text{H}^d \) that involve these directions directly, as well as those higher order terms that couple to the branching plane coordinates, are expected to be preeminent in the accurate description of the topography in the vicinity of the \( \text{mex}^2 \text{B}_1 - 2 \text{A}_1 \).

Figure 2 presents a three-dimensional plot of the potential energy surfaces for the four states in question in the \( g-h \) plane of \( \text{mex}^2 \text{B}_1 - 2 \text{A}_1 \). Of particular relevance to this work are the interactions of the first, second, and third excited states in the region of the ground state minimum.

Another significant feature of the \( \text{ab initio} \) data is the two low frequency modes <309 cm\(^{-1}\) at \( \text{min}^2 \text{B}_1 \), reported in Table II. One transforms as \( b_1 \) (an out-of-plane motion similar to \( h \) at the \( \text{mex}^2 \text{B}_1 - 2 \text{A}_1 \)) and the other transforms as \( b_2 \) (an in-plane motion similar to \( h \) found in the \( \text{mex}^2 \text{B}_1 - 2 \text{A}_1 \) in pyrazolyl\(^\text{23}\)). These motions are expected to be highly anharmonic and hard to characterize with the quadratic Hamiltonian.

---

**Table I.** Structures of extrema and corresponding energies (cm\(^{-1}\) relative to \( \text{mex}^2 \text{A}_1 - 2 \text{B}_1 \)) DZP/MRCl calculations.\(^\text{a}\)

<table>
<thead>
<tr>
<th></th>
<th>( \text{min}^2 \text{B}_1 )</th>
<th>( \text{ts}^2 \text{A}_1 )</th>
<th>( \text{mex}^2 \text{A}_1 - 2 \text{B}_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma(\text{N}1-\text{N}2) )</td>
<td>1.3770 1.3770</td>
<td>1.2742 1.2742</td>
<td>1.2955 1.2955</td>
</tr>
<tr>
<td>( \sigma(\text{N}2-\text{C}1) )</td>
<td>1.3102 1.3102</td>
<td>1.3703 1.3703</td>
<td>1.3525 1.3526</td>
</tr>
<tr>
<td>( \sigma(\text{C}1-\text{C}2) )</td>
<td>1.4737 1.4737</td>
<td>1.4032 1.4032</td>
<td>1.4225 1.4225</td>
</tr>
<tr>
<td>( \sigma(\text{C}1-\text{H}1) )</td>
<td>1.0756 1.0756</td>
<td>1.0727 1.0727</td>
<td>1.0733 1.0733</td>
</tr>
<tr>
<td>( \angle(\text{N}2\text{N}1\text{N}3) )</td>
<td>111.3 111.3</td>
<td>122.0 122.0</td>
<td>119.1 119.1</td>
</tr>
<tr>
<td>( \angle(\text{N}1\text{N}2\text{C}1) )</td>
<td>106.6 106.6</td>
<td>101.5 101.5</td>
<td>103.0 103.0</td>
</tr>
<tr>
<td>( \angle(\text{H}1\text{C}1\text{N}2) )</td>
<td>123.1 123.0</td>
<td>121.2 121.2</td>
<td>122.0 122.0</td>
</tr>
</tbody>
</table>

**Table II.** Harmonic frequencies (cm\(^{-1}\)) at \( \text{min}^2 \text{B}_1 \) and \( \text{ts}^2 \text{A}_1 \) determined from MR-SDCI and \( \text{ab initio} \) \( \text{H}^d \) calculations.

<table>
<thead>
<tr>
<th></th>
<th>( \text{min}^2 \text{B}_1 )</th>
<th>( \text{ts}^2 \text{A}_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_1 )</td>
<td>3408.9 3409.2</td>
<td>3440.2 3439.0</td>
</tr>
<tr>
<td>( b_1 )</td>
<td>1558.0 1594.0</td>
<td>1493.3 1486.0</td>
</tr>
<tr>
<td>( b_2 )</td>
<td>1344.6 1334.6</td>
<td>1330.7 1331.1</td>
</tr>
<tr>
<td>( b_3 )</td>
<td>1099.7 1119.4</td>
<td>1227.8 1229.2</td>
</tr>
<tr>
<td>( b_4 )</td>
<td>1052.0 1050.0</td>
<td>1117.8 1126.2</td>
</tr>
<tr>
<td>( b_5 )</td>
<td>933.6 921.0</td>
<td>937.3 947.0</td>
</tr>
</tbody>
</table>

\(^\text{a}\)Each structure has \( C_{2v} \) symmetry, distances in angstrom, and angles in degrees. \( \text{H}^d \) refers to \( \text{ab initio} \) \( \text{H}^d \), see text for details.
E. Construction of $H^d$

1. *Ab initio* data

The *ab initio* data, including energies, energy gradients, and derivative couplings, used to build $H^d$ were determined in the vicinity of the three extrema noted above. The points, $Q(m)$, consist of 41 points near $\text{me}^x\text{A}_1-^2\text{B}_1$, 52 points near $\text{mi}^2\text{B}_1$ and 52 points near $ts^2\text{A}_1$.

The 41 points near $\text{me}^x\text{A}_1-^2\text{B}_1$ were obtained via displacements in the above described orthogonal intersection adapted coordinates. They include the origin and three displacements ($\pm d_x$, $(+d_y)$, with $ldi = 0.015$ a.u. Here and below, $(d_j, e_k)$ denotes a displacement of magnitude $d$ and $e$ in coordinates $j$ and $k$, respectively. To these points are added coupled branching/seam space displacements, denoted ($\pm d_x$, $\pm e_k$, $\pm d_y$), $ldi = 0.007$, $let = 0.01$, and $k = 1–13$. The final 11 points consist of the displacements, $(+d_x, +e_k, -d_y)$, $(-d_x, -e_k, e_y)$ for $k \in a_1$ coordinates and $(+d_x, -e_k, +d_y)$ for $k \in b_1$, where $ldi = 0.007$ and $let = 0.01$.

For the two sets centered about $\text{mi}^2\text{B}_1$ and $ts^2\text{A}_1$, *ab initio* data were obtained at each stationary point as well as along displacements in symmetry-adapted valence internal coordinates. Twenty one displacements in which $ldi = 0.010$ consist of $\pm d$ displacements along the $a_i$ directions and $+d$ displacements along the remaining directions. As above, positive and negative displacements in the non-$a_i$ directions are symmetry related and thus only the positive displacements were computed. For these 22 points, energies, energy gradients, and derivative couplings for all four electronic states were determined.

In addition to the 22 points in near $\text{mi}^2\text{B}_1$ and $ts^2\text{A}_1$, *ab initio* data were obtained at 30 (15 forward and 15 backward) displacements of length $ldi = 0.001$ along natural internal coordinates, defined in the supplemental data.

2. $H^d$ behavior of $H^d$ for large displacements

In the construction of quasidiabatic $H^d$ using polynomials including cubic and quartic terms, it was found that while the description of the fitted *ab initio* data was quite good, for larger values of some of the coordinates, the $H^d$ determined potential energy surfaces displayed the wrong curvature, leading to unbound potentials. The origin of this complication arises from the fact that the number of unique parameters rapidly increases with the order of the polynomials employed, and although the defining equations are globally overdetermined, data from poorly defined regions may lead to parameterizations, which extrapolate poorly. This behavior is not totally unexpected since *ab initio* data in these geometric regions are not computed and may even be unattainable owing to limits of the multireference configuration wave functions used to obtain the data. Quadratic $H^d$ are less susceptible to this problem as there are many fewer parameters to determine.

A method was developed to eliminate this problematic behavior. First, a damping or smoothing factor, $t$, was incorporated into the fitting procedure, as indicated in Eq. (6). Here $t = t_y$, which was chosen to be $5 \times 10^{-5}$ but similar results were obtained for $t \sim 10^{-7}$, suppresses poorly defined coefficients, thereby eliminating a significant portion of the interpolation points, was not included in the fit. This data set includes energies, energy gradients, and derivative couplings for all four electronic states at ten additional points that lie along a linear synchronous transit path starting near $\text{mi}^2\text{B}_1$, connecting to $\text{me}^x\text{A}_1-^2\text{B}_1$, continuing to $ts^2\text{A}_1$, and finishing beyond the saddle point. These points are included to assess the ability of $H^d$ to reproduce data not included in the fit. The internal coordinates for these points are included in the supplemental data.
TABLE III. Dimensions of equations used to define \( V^{(m)} \) and breakdown of origin of higher order terms.\(^a\)

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( N^{\text{ord}} )</th>
<th>( N^{\text{eq}} )</th>
<th>( N^{\text{link}} )</th>
<th>( \text{Diag} )</th>
<th>( \text{Off1} )</th>
<th>( \text{Off2} )</th>
<th>( \text{Off3} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>2</td>
<td>13896</td>
<td>324</td>
<td>ord = 2</td>
<td>42</td>
<td>22</td>
<td>34</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>13896 (^b)</td>
<td>822</td>
<td>ord = 2</td>
<td>42</td>
<td>22</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ord = 3</td>
<td>30</td>
<td>18</td>
<td>15</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ord = 4</td>
<td>52</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
</tbody>
</table>

\(^a\)Linear terms, \( V^{(1)} \), are determined exactly and are not included in \( N^{\text{link}} \). Diag denotes diagonal blocks (1,3,6,10), off1 denotes off-diagonal blocks (2,9), off2 denotes off-diagonal blocks (4,8) and off3 denotes off-diagonal blocks (5,7).

\(^b\)For (4) 28 (four energies at seven points) tie-up energy equations were added to Eq. (6) as described in the text.

of the “turn-over” behavior. A similar term was employed in the construction of \( H^d \) for dissociative potential energy surfaces.\(^{13}\) Significantly, the inclusion of this term had a negligible impact on the ability of \( H^d \) to reproduce energy, energy gradients, and derivative coupling information, where \( ab \) initio data were available, while drastically improving regions where turn-over occurred. Additionally, when damping was not completely effective, in those regions which remained poorly behaved, approximate energy data were determined from a well-behaved quadratic \( H^d \). This approach takes advantage of the fact that only qualitatively correct energetics are required for regions so far removed from the points of interest. In this study, this corresponded to points approximately 1.5–3 a.u. and \( 10^7–10^8 \) cm\(^{-1} \) from the MEX, where the underlying MR-CISD calculations became unreliable. Denoted “tie-up” equations, these expressions were weighted by a factor between 0.1 and 0.01 to minimize the effect this data had on \( H^d \) in the regions of actual \( ab \) initio data while still compelling the higher order \( H^d \) to be bound.

3. The \( H^d \)

In this study four \( H^d \) were constructed. For each \( H^d \), to provide a correct description of the region of conical intersection, the origin was chosen as \( m \text{ex}^2 B_1–2 A_1 \). The characteristics of these \( H^d \) are summarized in Table III and discussed below. Each \( H^d \) is denoted by a symbol \( (N^{\text{ord}}) \) where \( N^{\text{ord}} \) is the order of the polynomial used and \( a = n, c, \) or \( f \), where \( n \) indicates that a nascent fitting procedure was performed without modification, \( c \) indicates that Lagrange multiplier constraints were included to require the \( H^d \) derived ground state energy at \( m \text{ex}^2 B_1 \) and at \( t \delta A_1 \) agree with the \( ab \) initio value, while \( f \) indicates that damping and tie-up equations were included in the construction of \( H^d \) to ensure correct large displacement behavior. The four expansions constructed are denoted (2n), (2c), (4n), and (4f). The expansion coefficients comprising \( H^d \) (4f) are given in the supplemental data. The second-order description includes all nonzero by \( (C_{2v}, \text{ vibronic symmetry}) \) first- and second-order terms. There are 324 second-order terms. A fully quartic \( H^d \), including all coefficients, would consist of almost \( 10^4 \) parameters. Because this large number of parameters would preclude \( H^d \) from being subsequently used in our spectral simulation programs, as well as dramatically increase the \( ab \) initio data needed to determine these parameters, a carefully selected subset of the third- and fourth-order terms are chosen in which there at most two unique symmetry indices and at least one coordinate must be a branching space coordinate (\( x \) or \( y \)). This yielded a total of 198 cubic and 300 quartic coefficients and a total of 822 unknown parameters to be determined. These parameters counts would be 1764, 7852, and 9940, respectively, if all symmetry nonzero anharmonic terms were included. On the basis of the above selection procedure, these coefficients arise as follows. Each diagonal block of symmetry \( A_1 \) (numbered \( a = 1, 3, 6, 10 \)) consists of 42(42), 30(102), and 54(866) second-, third- and fourth-order coefficients, respectively. The total number of unrestricted coefficients for the higher order terms is included in parentheses. There are three types of off-diagonal blocks \( b = 2.9, c = 4.8, \) and \( d = 5.7 \) whose electronic state direct products carry the spatial irreducible representations, \( B_1 \) for \( b \), \( A_2 \) for \( c \) and \( B_2 \) for \( d \). This reflects the previously noted ordering of the states at \( m \text{ex}^2 \) \( B_1–2 A_1 \), being \( 2B_1, 2A_1, 2A_2, 2B_2 \). The \( a(b) c \) blocks consist of 22(34)(22), 18(15)(6), and 14(14)[14] second-, third-, and fourth-order coefficients, respectively.

As stated earlier, a key issue in higher order polynomial fits is whether there are too many free parameters for the available data as this can lead to oscillations in the \( H^d \) derived potential energy surfaces, resulting in poor interpolation. Table III summarizes the number of equations \( [N^{\text{ord}}] \) required to determine the unique number of \( V^{(m)} [N^{\text{link}}] \) for \( N^{\text{state}} = 4 \) and \( N^{\text{ord}} = 2, 4 \). As can be seen from this table, the system of linear equations Eq. (4) is significantly overdetermined. The issue of interpolation is further discussed below.

F. Accuracy of the \( H^d \)

1. General comparisons

Table IV provides a coarse comparison of the \( H^d \) derived potential energy surfaces reporting the rms energy dif-
ference of the four \( H^d \) broken down by region, that is, points in the vicinity of \( \min^2 B_1, ts^2 A_1, \) and \( \mex^2 B_1 - 2 A_1, \) denoted \( \Delta(\min), \Delta(ts), \) and \( \Delta(mex) \) respectively, and the curvature of the surfaces, as indicated by \( N_{\min} \) and \( N_{ts}, \) the number of imaginary frequencies obtained from a Hessian computed at \( \min^2 B_1 \) and \( ts^2 A_1, \) respectively. From this table, it is seen that for each \( H^d \) the region of \( \mex^2 B_1 - 2 A_1 \) is best described, followed by \( ts^2 A_1 \) and finally \( \min^2 B_1. \) This is not surprising given that the origin of the expansion is \( \mex^2 B_1 - 2 A_1 \) and given the relative proximity of \( ts^2 A_1 \) and \( \min^2 B_1 \) to \( \mex^2 B_1 - 2 A_1, \) noted above. These results can be modified by including more \( ab \ initio \) data in problematic regions, constraining certain data to be reproduced exactly, or excluding and/or weighting data in a particular regions to improve the accuracy of \( H^d \) in a desired region. This approach allows for the accurate description of a particular region at the expense of, perhaps, limiting accuracy in other regions. The efficacy of that approach, in which the accuracy of a particular region is emphasized, is highly dependent on the case being considered and is beyond the scope of the present work. That issue, which we have discussed previously, will be considered in a subsequent treatment, which reports the photoelectron spectrum of triazolide. Here we are concerned with obtaining a uniformly accurate description of the \( ab \ initio \) data from which \( H^d \) was derived through the incorporation of higher order terms.

2. Reproduction of energy gradients and derivative couplings

The results presented in Table IV appear to show that the \( (2c) H^d \) determined results represent a marked improvement over the \( (2n) H^d \) derived surfaces in their ability to reproduce the ground state minimum. The more precise analysis that follows shows that not only is that improvement illusory, but it also evinces fundamental limitations in the second-order description, which were not found in studies of the related azolyls, pyrazolyl, and pyrrolyl.

Figures 3 and 4 show the comparison of the energy gradients and derivative couplings determined from the \( (2n) \) and \( (4f) H^d \) to the \( ab \ initio \) results from which they were derived. In each figure each nonzero \( ab \ initio \) determined datum is plotted (on the ordinate) against its \( H^d \) determined value (on the abscissa). In the case of perfect agreement of \( H^d \) and \( ab \ initio \) determined datum, the point falls on the diagonal. A 20% error bounding line around the diagonal is also plotted. Nearly equivalent results for the \( (2c) \) and \( (4n) H^d \) (compared respectively to \( (2n) \) and \( (4f) \), respectively) are presented in the supplementary material.

Regarding the quality of the potential energy surfaces determined from a second-order \( H^d \): The quadratic \( (2n) H^d \) can be constrained [yielding \( (2c) H^d \)] to reproduce the energies near the distant \( \min^2 B_1 \) extremum. This reduces the rms error of the \( (2c) H^d \) derived energies such that it is comparable to that of the quartic fits. However, significant errors in the reproduction of the energy gradient and derivative coupling data persist. Additional constraints including the energy gradient and derivative coupling data at the extrema might be expected to improve the accuracy of the quadratic fits. However, these attempts were frustrated by the fact that improvements in one region were universally accompanied by significant degradation in the reproduction of the energy gradients and derivative couplings in another. The lack of flexibility in these low-order expansions precludes the possibility of accurately reproducing the three regions of the potential energy surface required here.

It is clear from a cursory examination of Figs. 3 and 4 that the inclusion of cubic and quartic anharmonic terms leads to a quantitative improvement in the reproduction of the \( ab \ initio \) potential energy surfaces in the regions of import. Interestingly, it was found that the quartic \( H^d \) displays a more pronounced improvement in the description of the energy gradient data in comparison to the derivative coupling data. This observation can be rationalized by the fact that there are several different seams of conical intersections occurring between the four electronic states in question, only one of which (the lowest energy MEX) is explicitly included in constructing \( H^d \).

3. Characterization of the extrema

The accurate reproduction of the frequencies at the two extrema on the ground state potential energy surface is of particular interest if \( H^d \) is to be employed for subsequent spectral simulations. From Table IV it is seen that only the \( (4f) \) (847 cm\(^{-1}\)) and the \( (4n) \) (851 cm\(^{-1}\)) determined \( H^d \) give the correct number of imaginary frequencies at \( ts^2 A_1 \) and are in good accord with the \( ab \ initio \) value (851 cm\(^{-1}\)) shown in Table IV. Likewise, only the \( (4n) \) and the \( (4f) \) determined \( H^d \) give the appropriate curvature at \( \min^2 B_1 \), reproducing the \( ab \ initio \) computed low-frequency modes (204, 304 cm\(^{-1}\)) to high accuracy (\( \Delta \omega = 7, 71 \) cm\(^{-1}\) and \( \Delta \omega = 18, 1 \) cm\(^{-1}\) for \( (4n) \) and \( (4f) \), respectively). The superiority of the quartic fits in the reproduction of the ground state energy gradients is seen clearly in the comparison of Figs. 3(a) and 3(e). The differences between the \( (4n) \) (supplemental data) and \( (4f) \) \( H^d \) determined results in this regard are negligible.

4. \( H^d \) behavior of \( H^d \) for large displacements

Potential energy surfaces determined employing second-order \( H^d \) are generally bound, such that no region of the surface tends to \( -\infty \) with increasing displacement size. More care in this regard is required when constructing higher order expansions. While the nascent quartic \( H^d \) determined potential energy surfaces appeared initially to be remarkably accurate, complications became apparent upon closer inspection. In particular, large displacements along a number of coordinates were found to yield unbound potentials. The large displacement behavior of the \( (4n) H^d \) determined potential energy surfaces and the resolution of the evident problems provided by the \( (4f) \) description are demonstrated in Fig. 5. Depicted are superimposed slices of the potential energy surfaces determined using \( (4n) \) and \( (4f) H^d \) along problematic coordinate directions. The surface cuts along these directions illustrate that in the \( (4n) \) based potential energy surfaces the undesirable turn-over behavior is present in regions where
FIG. 3. (2n) H\textsuperscript{d} determined gradients (plates (a) [D\textsubscript{0}], (b) [D\textsubscript{1}], (c) [D\textsubscript{2}], (d) [D\textsubscript{3}]) and (4f) H\textsuperscript{d} determined gradients (plates (e) [D\textsubscript{0}], (f) [D\textsubscript{1}], (g) [D\textsubscript{2}], (h) [D\textsubscript{3}]), where the state label is given in square brackets. Here, D\textsubscript{n} denote the n\textsuperscript{th} doublet state with 0 denoting the ground state. The \textit{ab initio} value is plotted on the abscissa and H\textsuperscript{d} determined value on ordinate. The symbol X denotes points near \textit{mex}^2\textsuperscript{B}_1\textsuperscript{−2}\textsuperscript{A}_1, O are near \textit{min}^2\textsuperscript{B}_1, and \Delta near \textit{ts}^2\textsuperscript{A}_1.
FIG. 4. Comparison of \((2n)\) and \((4f)\) \(H_d^d\) determined derivative couplings. Rows 1 and 2 (plates (a) and (d) \([D_0-D_1]\), (b) and (e) \([D_0-D_2]\), (c) and (f) \([D_0-D_3]\)) are the \((2n)\) and \((4f)\) \(H_d^d\) values, respectively, as are rows 3 and 4 (plates (g) and (j) \([D_1-D_2]\), (h) and (k) \([D_1-D_3]\), (i) and (l) \([D_2-D_3]\)), where the state labels for the couplings are given in square brackets. The ab initio value is plotted on the abscissa and \(H_d^d\) determined value on ordinate. The symbol X denotes points near \(m^2B_1^2\) or \(m^2A_1\), O are near \(m^2B_1\), and \(\Delta\) near \(ts^2A_1\).
there is no \textit{ab initio} data. However, the inclusion of a damping factor, in conjunction with the tie-up equations, yields (4f) $H_d$. The potential energy surfaces determined from this representation have the accuracy of the (4n) description where it is successful as well as has appropriate large displacement behavior similar to that of the (2n) $H_d$ based potential energy surfaces. Thus, the damping and tie-up procedures permit the beneficial effects of the higher order polynomials to be straightforwardly incorporated without the debilitating large displacement problems.

5. Interpolation and extrapolation

To this point, we have considered only the ability of the $H_d$ to reproduce \textit{ab initio} information included in the data sets employed in their construction. A more stringent test would involve evaluating the ability of the $H_d$ to reproduce the potential energy surfaces at points external to these sets, but still in the expected domain of accuracy. This issue is addressed in Figs. 6(a) and 6(b) which report, respectively, the (2n) and (4f) $H_d$ determined energies for the four lowest energy states along a linear synchronous transit path from the region of $\min^2 B_1$ to that of $ts^2 A_1$. First, note that neither figure indicates any large oscillations or fluctuations of the interpolated data. This is particularly encouraging for the (4f) results. Further note that the errors for the interpolated points mirror those for the points included in the fit. Compare these figures with Table IV. For example, near $\max^2 B_1 \rightarrow A_1$, the origin for the plots, (2n) $H_d$ derived energies reproduce the \textit{ab initio} data almost exactly to within 1 cm$^{-1}$ but as one steps away from $\min^2 B_1 \rightarrow A_1$ toward $\min^2 B_1$ ($ts^2 A_1$), the error, which can be seen in the inset of Fig. 6(a), increases to $\sim 200$ cm$^{-1}$ (60 cm$^{-1}$) for all \textit{ab initio} points included around the minimum (transition state). From Fig. 6(b), the results of the (4f) fit are vastly superior to the (2n) fit, reducing the energy differences to less than 25 cm$^{-1}$, which again is comparable to the $\Delta s$ in Table IV. This is a remarkable result considering the large range of geometric configurations incorporated in the fit and the fact that the energy range for the four states in the domain is on the order of $10^5$ cm$^{-1}$. Finally note that the accuracy of the interpolated results for the (2c) and (4n) $H_d$ derived energies mirrors, as above, the accuracy of the fit results in Table IV and consequently is not discussed here.

IV. SUMMARY AND CONCLUSIONS

In this work we have extended a previously introduced general algorithm for determining fully second-order quasidiabatic representations, $H_d$, of bound-state potential energy surfaces coupled by conical intersections to quasidiabatic...
FIG. 6. Linear synchronous transit path described in text, with the (2n) $\mathbf{H}^4$ and (4n) $\mathbf{H}^6$, shown in panels (a) and (b), respectively. The $\mathbf{H}^4$ calculated values are denoted by O, while the $ab\ initio$ data are marked by X. (\(\bigcirc\), \(\bigtriangleup\), and \(\boxdot\)) denote $\mathbf{H}^4$ determined results for states $D_0$, $D_1$, $D_2$, and $D_3$, respectively. The insets show the energy difference between the $\mathbf{H}^4$ and $ab\ initio$ values. Note that the scale of the inset in panel (b) is 1/8 of that shown in panel (a).

representations based on higher order polynomials. A method to ensure that the determined potential energy surfaces have the appropriate curvature along large displacements from the origin was developed, thereby avoiding unphysical behavior in these regions. Using triazolyl (CH)$_2$N$_3$ as an example, $\mathbf{H}^4$ were constructed for four coupled states using polynomials through fourth order and compared with lower order approximations. Using a small subset of higher order coefficients, <10$^3$ opposed to ~10$^4$ in the full quartic representation, an $\mathbf{H}^4$ was constructed that reproduced the $ab\ initio$ data from which it was derived to a very high degree of accuracy. In particular, this $\mathbf{H}^4$ provided a uniformly accurate representation of the $ab\ initio$ data in three isolated regions: the minimum energy crossing, a ground state transition state, and a ground state minimum. This uniformly accurate representation was not achievable using the standard fully quadratic representation. The inclusion of damping in conjunction with tie-up equations eliminated the spurious large displacement behavior which is likely to be endemic to $\mathbf{H}^4$ based on higher order polynomials.

In a future study, the $\mathbf{H}^4$ reported here will be used as the basis for a determination of the photoelectron spectrum of triazolide, which has been measured but not successfully analyzed. 

ACKNOWLEDGMENTS

This work was supported by Air Force Office of Scientific Research Grant No. FA9550-09-10038 to D.R.Y.

29. See supplementary material at http://dx.doi.org/10.1063/1.3523344 for this data.