Ultrafast ring opening in 1,3-cyclohexadiene investigated by simplex-based spectral unmixing

Cite as: J. Chem. Phys. **136**, 054303 (2012); https://doi.org/10.1063/1.3681258
Submitted: 29 March 2011 • Accepted: 12 January 2012 • Published Online: 02 February 2012

J. L. White, J. Kim, V. S. Petrović, et al.

Ultrafast ring opening in 1,3-cyclohexadiene investigated by simplex-based spectral unmixing

J. L. White,[1,2,a] J. Kim,[1,3] V. S. Petrović,[1,3] and P. H. Bucksbaum[1,2,3]
[1]*PULSE Institute, Stanford University, Stanford, California 94305, USA*
[2]*Department of Applied Physics, Stanford University, Stanford, California 94305, USA*
[3]*Department of Physics, Stanford University, Stanford, California 94305, USA*

(Received 29 March 2011; accepted 12 January 2012; published online 2 February 2012)

We use spectral unmixing to determine the number of transient photoproducts and to track their evolution following the photo-excitation of 1,3-cyclohexadiene (CHD) to form 1,3,5-hexatriene (HT) in the gas phase. The ring opening is initiated with a 266 nm ultraviolet laser pulse and probed via fragmentation with a delayed intense infrared 800 nm laser pulse. The ion time-of-flight (TOF) spectra are analyzed with a simplex-based spectral unmixing technique. We find that at least three independent spectra are needed to model the transient TOF spectra. Guided by mathematical and physical constraints, we decompose the transient TOF spectra into three spectra associated with the presence of CHD, CHD^+, and HT, and show how these three species appear at different times during the ring opening. © *2012 American Institute of Physics*. [doi:10.1063/1.3681258]

I. INTRODUCTION

The ring opening of 1,3-cyclohexadiene (CHD) to form hexatriene (HT) belongs to a class of pericyclic reactions[1] found in many biochemical pathways, the most important of which is the biosynthesis of vitamin D_3 from its provitamin dehydrocholesterol. For this reason, CHD has been a model for studies of non-radiative photoisomerization.

The current understanding of CHD-HT isomerization comes from time-resolved multi-photon fragmentation experiments with mass-resolved detection,[2–9] REMPI spectroscopy,[10] resonance Raman excitation spectroscopy,[11] electron diffraction,[12] electron energy loss measurements,[13] and liquid phase transient absorption measurements.[14–16] Experiments that control the CHD:HT branching ratio have been performed in both gas phase[2] and liquid phase.[17,18] The isomerization has also been extensively studied computationally.[19–22] It is now generally accepted that the photoinitiated ring opening is governed by non-adiabatic interactions; at least two conical intersections are found to play a role in the photoinitiated isomerization (Figure 1). When an ultraviolet pulse launches a wavepacket from the S_0 ground state of CHD onto the first excited S_1 state, most of the wavepacket passes through an S_1/S_2 conical intersection onto the lower surface after approximately 55 fs.[9] The wavepacket then accelerates towards another conical intersection between the S_1 and S_0 surfaces where it bifurcates and continues simultaneously towards both the HT and CHD ground states. The isomerization to the cZc conformation of HT is complete in approximately 200 fs.[9] The cZc-HT conformer is a transient which undergoes interconversion to more stable conformers. Theory suggests that the CHD:HT branching ratio is approximately 50:50 for isolated molecules;[19] a CHD:HT ratio of 60:40 was reported for experiments in solution phase.[23]

Differences in the laser-induced photo-fragmentation patterns of CHD and HT have been used to track the time evolution of the conversion to HT following photo-excitation of CHD. Short intense 800 nm laser pulses lead to easily distinguishable fragmentation patterns for these two isomers.[7] The difference in the fragmentation patterns has been attributed to differences in photoabsorption in the cations.[7,24,25] Kotur *et al.* used the fragmentation patterns to guide a control experiment in which they were able to increase the HT yield by shaping the UV launch pulse.[2] Fuss *et al.* reported a time-resolved laser fragmentation experiment in which they fit the ion signals to exponentials in order to relate the temporal evolution of certain fragments to population passing through specific locations along the isomerization pathway.[7] Fuss *et al.* note that each time constant appears in several mass peaks, but ultimately conclude that monitoring masses 79 and 80 (corresponding to C_6H_7 and C_6H_8) is adequate in order to follow the isomerization. Kosma *et al.* monitor masses 79 and 80 when performing a similar experiment with 13 fs resolution.[9]

In the work of Fuss *et al.*,[7] Kotur *et al.*,[2] and Kosma *et al.*,[9] typically a pair of time-of-flight (TOF) mass peaks is used to follow the isomerization between the two molecules. Decrease in the parent peak ion count (i.e., mass 80) is associated with the disappearance of the CHD isomer, while the increase of the H ion count is associated with the formation of HT, as is the increase in the count of ion fragments with two C atoms. In all these examples, the choice of fragments is based on empirical evidence. When only two fragments are used to follow the isomerization, one can discriminate between at most two species (unless other parameters are also observed, such as time evolution). Here we will use the term *species* to refer not only to chemical species, but to any molecular or ionic entity resulting in a distinctive fragmentation pattern. If more than two species are present (perhaps in the form of transient spectral features or unanticipated reaction products) a two-mass peak analysis

[a]Author to whom correspondence should be addressed. Electronic mail: wlj@stanford.edu.

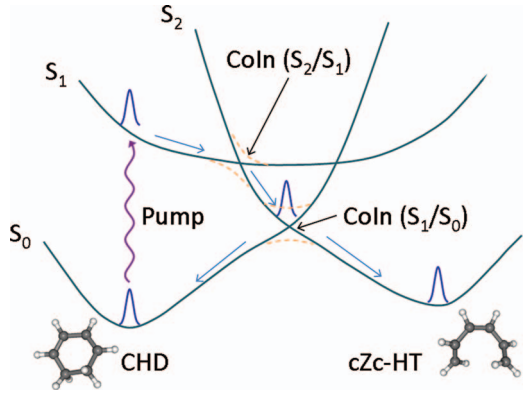

FIG. 1. Cartoon of CHD and HT potential energy surfaces along the reaction coordinate.

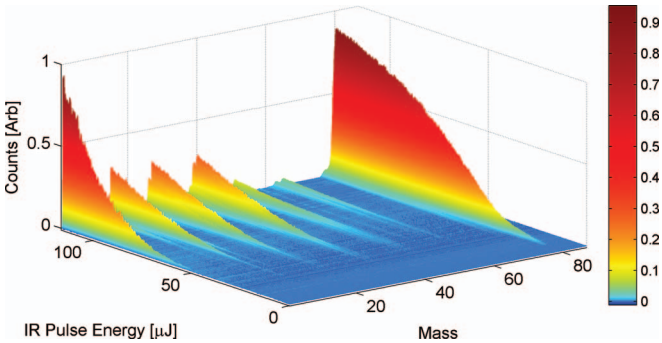

FIG. 2. CHD TOF mass spectrum as a function of IR pulse energy in an IR-only experiment. For pump-probe experiments, the IR pulse energy was chosen such that multiple daughter ions were present.

must underfit the data. More information contained in other peaks in the mass spectra could be used.

In this paper, we demonstrate an analysis based on geometric spectral unmixing to determine the number of fragmentation patterns and estimate their TOF spectra without prior knowledge of the number of species present.[26] The method benefits from the richness and redundancy of information contained in the *complete* time-of-flight fragmentation mass spectrum and not only a subset of peaks. This procedure is especially useful in cases where one (or more) of the species involved is transient or otherwise not available for direct study. We apply this spectral unmixing technique to a time-resolved study of CHD-HT isomerization and find that the fragmentation patterns of *three species* compose the TOF spectra collected during the UV-initiated isomerization of CHD. We assign the species to CHD, cZc-HT, and CHD^+ and estimate their laser fragmentation time-of-flight mass spectra. We then estimate their contribution to the fragmentation mass spectrum as a function of probe delay and discuss the time dependence of the underlying processes.

II. EXPERIMENT

Experiments were performed in an effusive beam of 1,3-cyclohexadiene (Aldrich, 97%, no further purification), skimmed and expanded at room temperature (its vapor pressure is approximately 10 mbar.[27]). A commercial Ti:sapphire laser produced 2.5 mJ pulses with 80 fs FWHM duration centered at 800 nm wavelength at a repetition rate of 1 kHz. A 100 μJ portion of this beam (referred to below as the IR beam) was delayed by -400 fs to +600 fs (20 fs steps) and used as the fragmentation pulse. Another portion of the Ti:sapphire beam (600 μJ) was used to create the UV excitation pulse by Type I frequency doubling in β-barium borate (BBO) (400 μm) and subsequent Type I frequency summing of 400 nm and 800 nm pulses in BBO (300 μm). The resulting UV pulse had < 30 μJ pulse energy and $\sim$120 fs duration FWHM centered at 266 nm. The UV pulse energy was set between the multiphoton ionization and fragmentation thresholds for CHD.

The IR pulse energy was chosen such that when CHD was fragmented, many different mass fragments were present but not every CHD molecule was being ionized or fragmented. Having many different fragments is advantageous when unmixing spectra. Figure 2 shows how the TOF spectrum of CHD varies as a function of the IR pulse energy (0$\sim$120 μJ) in an IR-only experiment. For pump-probe experiments, the IR pulse energy was $\sim$100 μJ.

In pump-probe experiments, the two laser beams propagated collinearly through the chamber and intersected the gas jet at 90°. Overlap of the foci in the direction of propagation was achieved by detuning a telescope in the IR beam path. At the focus, the UV beam was larger than the IR beam. In the interaction region, the UV and IR beam intensities can be estimated by the ionization threshold as well as the geometry of the beams and other similar factors. We estimate that the UV beam intensity was $\sim 10^{11}$ W/cm^2 and the IR beam intensity $\sim 10^{13}$ W/cm^2.

Ions were extracted with a 750 V/cm external field (always present) in the direction perpendicular to both the laser beams and molecular beam and collected at a multichannel plate detector after propagating 11 cm. A Poisson statistical analysis of the signal leads to an estimate that several hundred molecules were dissociated or ionized on each laser pulse, the majority of them contributing to the parent ion group. The strong IR field and the large number of molecules fragmented on each shot mean that each fragmentation spectrum is the result of summing over all possible dissociation/ionization channels. The UV pump pulse alone was responsible for approximately half of the parent ions.

III. UNMIXING SPECTRA

The TOF spectral data in our experiment are the result of various processes, such as: (1) UV-only production of CHD^+; (2) fragmentation of CHD, CHD^+ or HT by the IR probe laser; or (3) fragmentation of transient intermediates. Our analysis addresses several questions: (1) How many separate physical species or processes contribute to the ion spectra? (2) Does each have a characteristic TOF spectral signature? (3) What is the underlying physics?

Experimental spectral data are often mixtures of a small set of pure spectra from several distinct species, or *endmembers*. Many statistical analysis techniques have been developed to *unmix* a data set containing a mixture of distinct

species to yield pure spectra, called *endmember spectra,* under a variety of assumptions, including cases where the number and identities of the endmember spectra are initially unknown. A review of spectral unmixing from a remote sensing perspective is given in Ref. 26 and from a chemometrics perspective in Ref. 28 (where it is referred to as multivariate curve resolution). Here we follow an intuitive geometric unmixing method known in both fields.[29,30] We determine the number of endmembers needed to model the experimental TOF spectra, estimate the endmember spectra (using both physical and geometric knowledge), and finally unmix the experimental spectra to determine their composition as a function of pump-probe delay.

The key intuition underlying this approach is that the experimental spectra will be contained in a simplex whose vertices are the pure spectra. An (n-1)-simplex is an (n-1)-dimensional analog of a triangle or tetrahedron and has n vertices. It is the smallest convex set containing the vertices. A key property of simplices is that any point in the simplex can be written as a linear combination of the vertices in which the coefficients are nonnegative and sum to one.

An experimental spectrum with m spectral channels can be considered a vector in an m-dimensional vector space. We expect that the experimental spectra will be linear combinations of just a few ($n \ll m$) endmember spectra plus noise. Using this linear mixture model, each experimental spectrum S_{exp} can be written

$$\mathbf{S}_{exp} = \sum c_i \mathbf{E}_i + \varepsilon, \tag{1}$$

where the E_i are the (possibly unknown) endmember spectra, c_i is the fraction of the ith endmember spectrum found in S_{exp}, and ε is noise. We require $c_i \geq 0$ on physical grounds. Since the solution to Eq. (1) is not unique when the E_i are unknown, one can further require that $\Sigma\, c_i = 1$ to facilitate the interpretation of the c_i as proportions of the endmember spectra E_i. It is always possible to satisfy these constraints by choosing vertices E_i that enclose the data. We then have

$$S_{exp} \sum c_i E_i + \varepsilon, \quad c_i \geq 0, \quad \sum c_i = 1. \tag{2}$$

Equation (2) is the convex geometry model (CGM), and it is equivalent to requiring that the data (neglecting noise) lie inside a simplex whose vertices are the endmember spectra E_i. Sometimes one of the vertices corresponds to a constant (possibly all zero) background; this background endmember is called the *shade endmember* (from its use to account for lighting differences in remote sensing applications) and its coefficient expresses the relative abundance of the constant background in a spectrum. In chemometrics, it is common to divide each spectrum by its size (the sum of the counts in the spectrum, also called the 1-norm) in order to map spectra with the same shape to the same point; this scaling has the special property that it preserves the CGM and Eq. (2) still holds although the c_i will be different.[29] We choose not to scale our spectra for a number of reasons. The most important reason is that differences in spectrum size contain physically relevant information. Spectrum size reflects the number of ions collected and so is a measure of the fragmentation cross section. Moreover, the presence of transient resonances would mean that the relative abundance of each species determined from scaled spectra would be wrong. Additionally, our collected spectra at each delay do not vary greatly in spectrum size. The lack of variation is not surprising since, after averaging, the number of molecules in the interaction region and the laser pulse energies are both nearly constant. Finally, scaling the spectra is sensitive to the choice of constant background.

One way to estimate unknown endmember spectra is to determine the vertices of a simplex enclosing the experimental spectra. The purer the purest spectra inside the simplex, the better this geometry-based endmember estimation will be. In this paper, we use a combination of physical and geometric arguments to estimate the endmembers. Algorithms for automated endmember extraction exist,[30–33] but we do not use them in the present work.

To determine the number of endmembers, we first use a dimension reduction technique to determine the effective dimensionality of the data. The aim is to identify and retain meaningful dimensions while discarding the rest. In short, we seek to approximate our experimental spectra with a small set of basis spectra, but we do not yet require these basis spectra to be pure. Here we use principal component analysis (PCA) to reduce the dimensionality of the data, but other methods such as the minimum noise fraction[34] transform are also widely used. PCA is a standard statistical technique for determining, within a high-dimensional space, the subspace that "contains" the data.[35] (More correctly, PCA determines an *affine* subspace, that is, a translation of a subspace.) PCA identifies the subspace by retaining the p dimensions (linear combinations of spectral channels) along which the data had the highest variance and eliminating dimensions along which the variance is small. Intuitively, the high variance directions correspond to significant changes in the TOF spectrum, while low variance directions are uninteresting. The p-dimensional subspace thus determined contains the largest fraction of the variance of the data of any p-dimensional subspace, and projecting the original spectra into this subspace gives the best p-dimensional approximation to the original spectra in a least-squares sense. The subspace can be found as follows. We begin with the collection of experimental spectra, each of which can be regarded as an m-dimensional vector where m is the number of spectral channels. Next, the mean spectrum (the vector resulting from taking the arithmetic mean of all the experimental spectra) is subtracted from each experimental spectrum. Then we diagonalize the covariance matrix (of the spectra). The eigenvalues give the variance of the experimental data along the direction of the corresponding eigenvector; the eigenvectors are ordered from greatest eigenvalue to least and are called principal components (PCs). The choice of how many dimensions (PCs) to retain is made by examining the eigenvalues. Often (but not always), there are p eigenvalues that are much larger than the rest, indicating that p is the effective dimensionality of the data and is a natural choice for the number of dimensions to retain. In our analysis, each PC is a linear combination of spectral channels and so can be regarded as a spectrum. The first PC is the direction of largest variance in the data, the second PC is the direction of second largest variance (orthogonal to the first PC), and so on for subsequent PCs. The PCs form an orthonormal basis

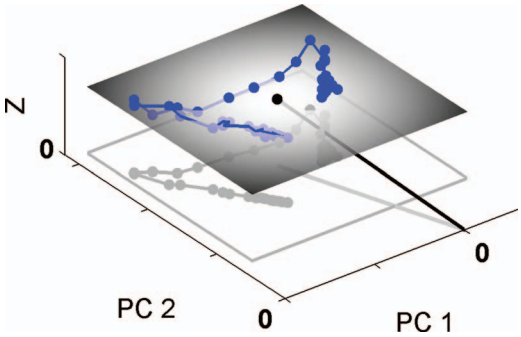

FIG. 3. Geometry of the UV-800 pump-probe PC affine subspace containing the TOF spectra collected as CHD isomerizes. The spectra have been projected into the three dimensions determined by the first two PCs and the mean spectrum (black dot; a black line connects the mean spectrum and the origin). The Z direction is the component (normalized to have unit length) of the mean spectrum that is orthogonal to the first 2 PCs. A region in the PC space is indicated by the gray plane. Each experimental spectrum is shown as a blue dot, and blue lines connect adjacent delays (separated by 20 fs). Three spectra are required to model the data (for instance, the mean vector and two PCs, or three pure spectra). The projection onto the Z = 0 plane is shown in light gray. All axes are drawn at the same scale.

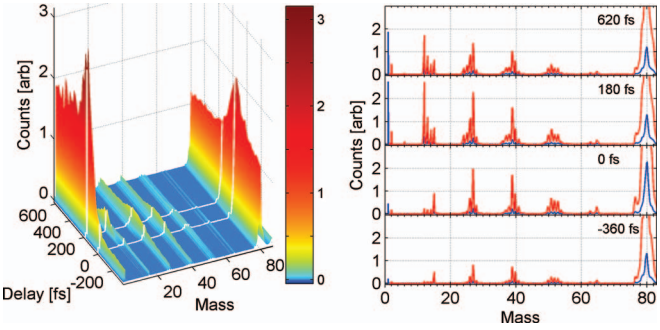

FIG. 4. (Left) TOF mass spectra as a function of pump-probe delay. Lines have been added to highlight the delays corresponding to the maximum yields of masses 1 and 80 (at 180 fs and 0 fs, respectively). (Right) TOF mass spectra for several delays, including the earliest delay and those highlighted in white on the left-hand plot. The blue lines are the raw spectra and the red lines show the spectra magnified eight times (the hydrogen peak is not shown magnified).

for the subspace and the simplex containing the data lies in this space. To reduce the dimensionality of the data, we approximate each experimental spectrum by the mean spectrum plus some linear combination of the first p PCs (a translation plus a vector in the subspace spanned by the first p PCs; see Figure 3). For each experimental spectrum, the coefficients of the linear combination are given by the projection of the mean-subtracted experimental spectrum onto the first p PCs.

If p PCs are retained, then there are at most n = p + 1 endmembers (simplex vertices). In addition to helping to estimate the number of endmembers, dimension reduction also makes it easier to visualize the simplex (and how the spectra are mixed). Figure 3 shows the geometry of the affine subspace for our TOF spectra; in this case the data falls into a two-dimensional subspace (p = 2) and the simplex can have at most 3 vertices (a triangle, n = 3).

IV. RESULTS

Figure 4 shows the TOF mass spectra as a function of pump-probe delay. Each mass spectrum is an average of 600 laser shots. The parent peak (mass 80) is always present but reaches a maximum at zero delay. The H^+ peak is almost absent at negative delays and reaches its maximum around 180 fs delay, after which it decreases to a constant value.

When PCA is applied to the collection of TOF mass spectra recorded at different delays in the pump-probe experiment, only two principal components account for more than 98% of the variance in the spectra (Figure 5 inset). This implies that the collected TOF mass spectra very nearly lie in a two dimensional affine subspace of the detection space. This affine subspace does not contain the origin. Consequently three linearly independent vectors are needed to span that affine subspace (for example, the mean vector and the two most significant PCs, as shown in Figure 3), suggesting that three endmember spectra are needed to explain the data.

Figure 5 shows the experimental mass spectra projected onto different pairs of principal components ordered according to decreasing variance. When projected onto the PC1-PC2 plane, the TOF spectra trace out a characteristic "cat face" (enlarged in Figure 6) as a function of pump-probe delay. Multiple pump-probe experiments taken weeks apart and under slightly different experimental conditions all exhibit this characteristic cat shape and fall into the span of the same three endmembers.[36] The positions of the "ears" allow immediate identification of two characteristic times: the peak in the parent ion (left ear) and the H^+ signal maximum (right ear). The two maxima are separated by 180 fs, comparable to the findings of previous investigations.[2,7] At long delay times, the product signal is constant, resulting in a compact cloud of

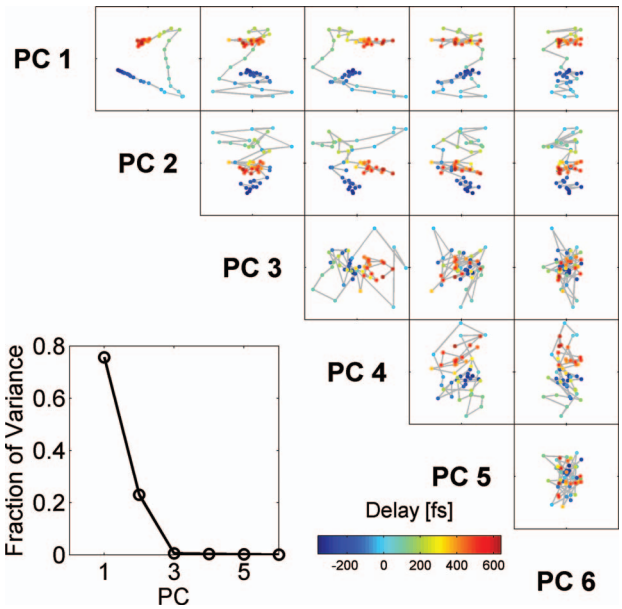

FIG. 5. The TOF spectra projected onto different pairs of PCs (directions of greatest variance). Each point corresponds to a spectrum at a single delay. Lines connect spectra at adjacent delays (20 fs steps). Different pair plots are not shown at the same scale. (Inset) The fraction of variance accounted for by first six PCs. The fraction of variance is proportional to the corresponding eigenvalue of the covariance matrix.

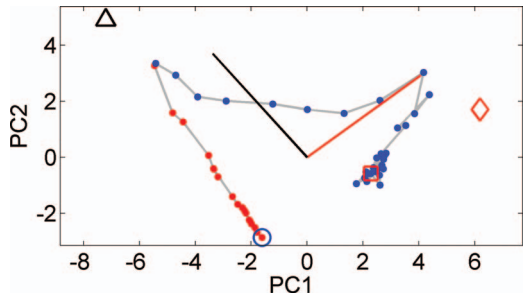

FIG. 6. Experimental spectra and endmember spectra projected into the PC1-PC2 plane. The spectra are mean-subtracted, and the origin indicates the location of the mean experimental spectrum in the PC1-PC2 plane. The CHD (blue circle), CHD^+ (black triangle), and HT (red diamond) endmembers and the assumed 50:50 CHD:HT spectrum (red square) are also shown. The directions corresponding to an increase in H^+ (red line) and mass 80 (black line) are also shown. The red dots indicate spectra collected at negative delay (IR precedes UV) and the blue dots indicate spectra collected at positive delay. Adjacent spectra are separated by 20 fs.

red points. Looking at other PC pair plots in Figure 5, we see that the experimental spectra exhibit much less structure when projected onto higher PCs. For example, the projection onto PC3 may be noise since the projection is not associated with a particular delay; the noise may due to laser pulse energy fluctuations. A possible exception is PC4, which appears to be related to UV-IR overlap around zero delay. However, the PC4 contribution to the spectra is very small (<0.3% of the variance) and we choose to neglect it. By eliminating higher dimensions, which typically have little structure, PCA acts to filter out noise.

When the spectra are viewed as vectors in a vector space, the trajectory is closely related to the underlying physical processes. For example, the straight edges of the data projection in PC space are significant features characteristic of conversion between two species when the total number of molecules is held constant. In our experiment, the number of molecules in the interaction region and the laser pulse energies are both approximately constant. If the cross-section for fragmenting a given endmember is constant as well (i.e., the probability of fragmenting a given species is constant), then the spectra resulting from the mixing of n endmembers fall into an n-1 dimensional space. For example, the spectra resulting from mixing two endmembers fall on a line connecting the two endmembers and for three endmembers the mixed spectra lie on the plane defined by the three endmembers. Thus straight lines are characteristic of the conversion of exactly one species into another and planarity is characteristic of conversions among three species.

A. Endmember spectrum estimation

The unmixing analysis can be taken one step further to estimate the pure spectra (position the endmembers in the plane of the first two PCs). First, a constant background due to the fragments created by the UV pulse alone is subtracted from the TOF spectra. The choice of this background affects the endmember spectra, but not the locations where they pierce the PC plane or their proportions. In order to estimate the CHD endmember, we assume that only neutral CHD is present at the earliest delay (when the probe comes ~360 fs before the pump) and we assign the earliest mass spectrum (see Figure 4) to pure CHD.

In order to estimate the HT endmember spectrum, we assume that among molecules that absorbed one UV photon, the CHD:HT ratio is 50:50 at the longest probe delays (~600 fs). This is the branching ratio estimated by theory.[19] We further assume that the total number of molecules is independent of delay and that the ion yield per CHD molecule is the same at the earliest and longest delay. We then use the CHD endmember and the 50:50 CHD:HT spectrum to place the HT endmember as shown near the right ear; this endmember corresponds to having all excited molecules eventually isomerizing to HT. (A CHD:HT branching ratio of 60:40 as measured in liquid phase[23] would place the HT endmember 25% farther away from the CHD endmember. On the other hand, 100% branching to HT, as reported in,[37] would place the HT endmember at the longest probe delay.) This HT endmember is based on the ion yield per HT molecule at long times. (We defer discussion of the transient H^+ peak, which may be the result of an enhanced ion yield.)

Rather than relying on the spectrum at greatest delay to be 50:50 CHD:HT, we could have used purely geometric arguments to estimate the HT endmember. For example, we could find the smallest simplex enclosing the data and assign the vertex closest to the "right ear" to the HT endmember. The CHD:HT branching ratio at long delays would be roughly 50:50. However, it is not clear *a priori* that the experimental spectra contain a high enough proportion of HT to permit reasonable estimation of the enclosing simplex, and the branching ratio arrived at this way would largely reflect the transient increase in the H^+ signal.

We now place the third endmember, which we call the CHD^+ endmember for reasons explained below. Following our conjecture that straight lines in the PC1-PC2 plot correspond to conversion between two endmembers, we posit that the third endmember should roughly lie on the line connecting the earliest spectrum (CHD endmember) with the cross-correlation peak (left cat ear). We additionally require all time-of-flight mass peaks in the endmember to be positive. These considerations result in the placement of the CHD^+ endmember spectrum as shown in Figure 6. The CHD^+ endmember is named for its close association with an increase in mass 80 ions (the direction indicated by the black line in Figure 6), although it may be more correct to think of it as a "cross-correlation endmember" arising from processes that occur when the laser pulses are overlapped in time.

The positions of the three endmembers in the PC plane are shown in Figure 6. They form a simplex roughly enclosing the data. The composition of a mixed spectrum can be easily determined from its position. The fraction of each endmember spectrum varies linearly from 100% at the endmember vertex to 0% at the far side of the triangle.

In Figure 6 we have also plotted the projections of the vectors corresponding to the H^+ direction (red line) and the parent ion direction (black line) onto the PC1-PC2 plane. This allows one to easily follow the behavior of different peaks during the delay scan. The parent ion count reaches a maximum near T_0 (near the CHD^+ endmember spectrum)

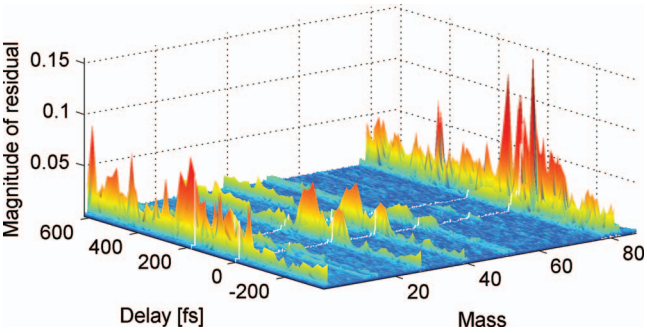

FIG. 7. Magnitudes of residuals left after fitting the experimental spectra to three endmembers. The vertical scale is the same as in Figure 4. The delays corresponding to the maximum yields of masses 1 and 80 are highlighted.

and the H^+ count reaches a maximum when the trajectory is closest to the HT endmember spectrum. The spectra at longer time delays contain more H^+ than the early time delays, reflecting a plateau in the count of hydrogen ions after the maximum. The H^+ and parent ion counts are shown in Figure 10 for comparison.

The residuals from fitting the experimental spectra to the simplex are shown in Figure 7. Any choice of endmember spectra that spans the same space will fit the data equally well, and PCA gives a bound on the goodness-of-fit achievable. It is not possible to reduce the residuals by simply choosing the endmembers differently.

The endmember spectra are shown in Figure 8. These spectra depend on the choice of background. The spectra shown are the vector difference of the endmember location in the PC plane and the background due to the UV pulse. The spectra shown include contributions from a background of CHD molecules that do not absorb a UV photon to initiate the ring opening. Nevertheless, differences among the spectra are clear. The HT endmember spectrum is the only one with a significant contribution from H^+, and other light fragments including both C^+ and C^{++} are also more common in the HT spectrum. The CHD endmember spectrum contains relatively few light fragments (methyl, mass 15, is a notable exception). The CHD^+ endmember spectrum is primarily composed of CHD^+ (mass 80) ions, as might be expected. We note, however, that even when the mass 1 and mass 80 ions are removed from the data set, the spectra still fall into a plane and trace out a triangular "cat face." In other words, it is possible to separate spectra from these three species solely on the basis of the intermediate mass fragments, even if the mass 1 and mass 80 components are neglected.

It would be interesting to compare these estimated endmember spectra with the spectra of pure species subjected to the same laser conditions. However, vibrationally hot cZc-HT is a transient. Other conformers of HT might be used for comparison, but we did not have a sample available.

The maximum in the mass 80 signal and the corresponding maximum in the CHD^+ endmember spectrum contribution at zero time delay are likely due to the temporal overlap of the pump and probe pulses at this point. This is consistent with a multiphoton ionization pathway for CHD involving one UV photon together with multiple IR photons. In like manner, the maximum in H^+ near a delay of 200 fs suggests a peak in HT production, but it could be due instead to a transient increase in the photofragmentation cross section for HT as suggested by Kotur *et al.*[2] and Fuss *et al.*[7] They point out that a transient resonance is suggested by the lateness of the H^+ maximum (200 fs) and the opportunity for resonances as the wavepacket accelerates downhill. The present analysis allows us to explore whether this apparent transient resonance is a new fragmentation channel, accompanied by its own distinct fragmentation pattern, or whether it is a transient increase in the contribution from the same HT endmember spectrum already identified.

At its maximum, the H^+ signal is approximately 50% higher than at long delays. If a distinct transient process (with a distinct fragmentation pattern) is responsible for the increase, the experimental spectra near the HT maximum (right cat ear) should leave the span of the endmembers. However, our analysis shows no evidence of a distinct fragmentation pattern associated with the transient increase in H^+. Three endmembers adequately model the transient increase in the H^+ signal. An examination of other dimensions shows that they contribute little to the transient H^+ maximum. Figure 9 shows the deviation from the PC plane in all dimensions. The deviation from planarity, taking into account all dimensions, is small compared to variation within the plane and the data

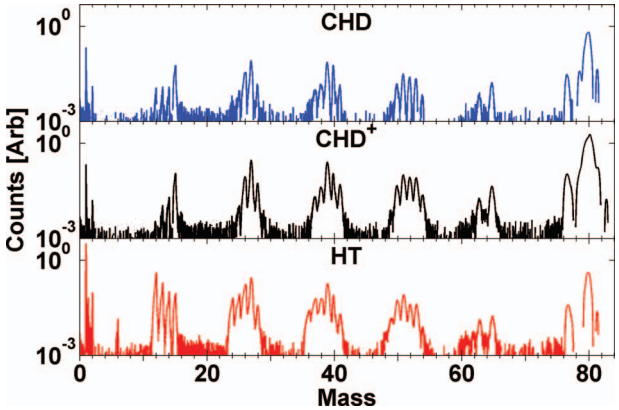

FIG. 8. Endmember spectra determined by using the UV-only background as the origin and plotted on a logarithmic scale. A CHD background due to molecules that do not absorb a single UV photon has not been subtracted.

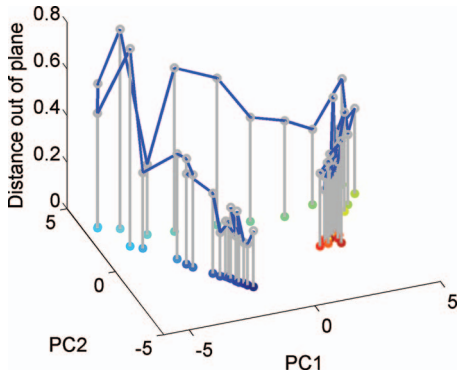

FIG. 9. Residuals from fitting experimental spectra to the 2D affine space formed from the mean vector and the first two principal components. The vertical axis is exaggerated by a factor of 10.

do not leave the simplex in the region of the H^+ maximum. Furthermore, the next most significant PCs, PC3 and PC4, appear to be related to noise and spectral features around zero delay, respectively.

In contrast, a transient enhancement of the same fragmentation route for HT will leave the right cat ear in the span of the endmembers, but generally bent out of the plane to indicate an enhanced HT spectrum. This is consistent with the data. The large constant background and noise in our measurement make it difficult to determine the precise geometry of any out-of-plane excursion (Figures 7 and 9).

We identify H^+ as a signal of HT production. In particular, we find that the maximum in H^+ corresponds to a maximum in the entire HT endmember signal. One explanation for the constancy of the fragmentation patterns is that the resonance is enhancing the transfer of population to a state that is a precursor to fragmentation.

We also see no evidence for a distinct fragmentation pattern associated with vibrationally hot CHD. At long delays, theory predicts a mixture of vibrationally hot CHD and HT. However, the mixed CHD/HT spectra are well-modeled by the HT endmember and the CHD endmember (the spectrum at the earliest delay). We conclude that the difference between the spectra of vibrationally hot CHD and ground state CHD is much smaller than the difference between the spectra of CHD and HT.

B. Unmixing Spectra and Proportion Determination

We now unmix the experimental spectra to determine the coefficients c_i in Eqs. (1) and (2). The endmember coefficients only depend on the position of the experimental spectrum within the simplex (following projection into the simplex). The coefficients are shown as a function of pump-probe delay in Figure 10 (left). For comparison, the H^+ and parent ion signal are shown in Figure 10 (right).

The experimental mass spectra are well-modeled by three fragmentation patterns. The CHD^+ coefficient reaches a maximum near zero delay. The HT coefficient closely follows the H^+ signal. In particular, the maximum in the H^+ peak is part of an overall increase in the contribution from the HT fragmentation pattern. The transient increase in the HT coefficient around ~200 fs delay is indicative of an increase in the number of HT molecules being detected and is consistent with a resonance, as discussed above. Interestingly, the characteristic HT fragmentation pattern is unchanged during the transient increase.

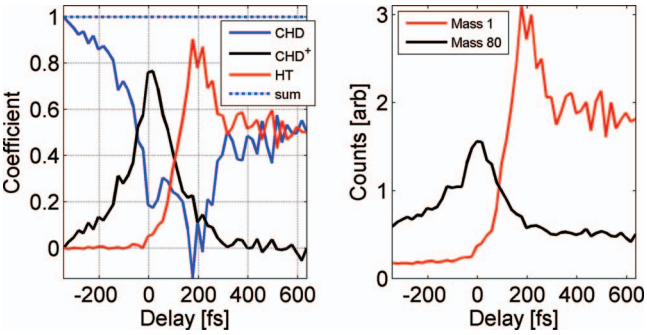

FIG. 10. (Left) Endmember coefficients vs. delay. The spectra were projected into the PC plane before calculating the coefficients, which forces the coefficients to sum to one. The CHD:HT ratio at long times is fixed at 50:50. (Right) The parent peak (black) and H^+ peak (red) vs. delay. A constant background due to the UV pulse has been subtracted.

The noise on the H^+ lineout is clearly reflected in the HT endmember coefficient and also appears on other mass peaks associated with HT. This noise is largely due to fluctuations in the number of HT molecules being detected. We believe that the fluctuations are mostly due to UV pulse energy fluctuations.

The relationship between the coefficients and the proportion of each species present is determined by the ion yield per endmember molecule and the total number of molecules present. The majority of the molecules are not excited and contribute an unexcited CHD background. Thus, when calculating proportions, we only consider molecules that undergo single photon UV excitation. In other words, the proportions are relative to the number of single photon UV-excited molecules. The CHD and HT endmembers were chosen in a way that explicitly determines their proportions. The HT endmember corresponds to the case where all UV-excited molecules have isomerized to HT. If we assume that the total number of molecules is constant and the fragmentation cross-section for each endmember is independent of delay, then for each mixed spectrum the endmember spectrum coefficients are equal to the proportions. These assumptions are the same as the assumptions used to determine the CHD and HT endmembers, but now applied to all delays. The proportions thus determined are for excited molecules that interact with the probe.

The above assumptions are not met in two cases. The first is the transient HT signal increase, which is believed to be due to a resonance, as discussed earlier. Second, the assumptions used to place the CHD^+ endmember do not explicitly determine the proportion of molecules that contribute to the CHD^+ endmember spectrum among molecules that undergo isomerization. Specifically, when both colors are present, the presence of multiple processes makes it hard to separate out the contribution of the unexcited CHD background.

V. CONCLUSION

We have used a simplex-based spectral unmixing technique to analyze TOF spectra collected during a pump-probe experiment that monitored the photoinitiated isomerization of CHD into HT. The technique allows the determination of the number of endmember spectra (pure spectra) necessary to model the data, estimation of endmember spectra, and estimation of their proportions in the experimental spectra (unmixing). We determined that transient TOF spectra from the CHD-HT ring-opening reaction could be fit with three endmembers. The experimental spectra fall into a plane, which is characteristic of conversion among three species. Using a combination of physical knowledge and the simplex-based spectral unmixing technique, we estimated the endmember spectra and assigned them to CHD, CHD^+, and HT. We find that H^+ is due almost entirely to HT; in particular the increase in H^+ around 180 fs delay is part of an overall

increase in the HT fragmentation pattern. We determined that the three fragmentation patterns represented by the three endmembers were sufficient to model the transient increase in the H^+ yield at this delay.

The reduction of the entire data set to mixtures of only three parent spectra is noteworthy, since the dynamics of the CHD-HT ring opening reaction is believed to involve nonradiative transitions between two excited states and a conical intersection to the ground state of CHD or HT. In particular, we find large spectral differences between CHD and HT but no discernible differences between the spectra of ground state CHD and vibrationally excited CHD. It is possible that the fragmentation pattern is largely determined by more general geometrical features, such as whether the ring is open or closed.

We have presented a simple interpretation of the data based on a consideration of the first two principal components. Given sufficient signal to background ratios, additional principal components may be employed to identify further, distinct endmember spectra associated with either specific regions on the potential energy surface (PES) or distinct fragmentation channels. Examination of PC4 displayed in Figure 5 suggests that there might be additional channels.

The unmixing technique presented is widely applicable. Determination of the number of fragmentation patterns involved is vital to avoid underfitting the data and is easily accomplished using a dimension-reduction technique such as principal component analysis. The ability to estimate endmember spectra is especially useful when pure spectra are not available (for example, unstable species or spectra corresponding to certain regions on a PES). A particular strength of the method is that, when viewed as points in a vector space, the geometry of the spectra is closely related to underlying physical processes, such as conversion between species or changes in fragmentation cross section. Low-dimensional models of experimental spectra are easily visualized and important features, such as the mixing of two endmembers, are readily identified. Finally, it is easy to create maps revealing how the spectra change as a function of experimental parameters, such as pump-probe delay.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation.

[1]R. B. Woodward and R. Hoffmann, J. Am. Chem. Soc. **87**, 395 (1965).
[2]M. Kotur, T. Weinacht, B. J. Pearson, and S. Matsika, J. Chem. Phys. **130**, 134311 (2009).
[3]W. Fuss, P. Hering, K. L. Kompa, S. Lochbrunner, T. Schikarski, W. E. Schmid, and S. A. Trushin, Ber. Bunsenges. Phys. Chem. **101**, 500 (1997).
[4]W. Fuss, T. Schikarski, W. E. Schmid, S. Trushin, and K. L. Kompa, Chem. Phys. Lett. **262**, 675 (1996).
[5]S. A. Trushin, W. Fuss, T. Schikarski, W. E. Schmid, and K. L. Kompa, J. Chem. Phys. **106**, 9386 (1997).
[6]W. Fuss, S. Lochbrunner, A. M. Muller, T. Schikarski, W. E. Schmid, and S. A. Trushin, Chem. Phys. **232**, 161 (1998).
[7]W. Fuss, W. E. Schmid, and S. A. Trushin, J. Chem. Phys. **112**, 8347 (2000).
[8]M. Garavelli, C. S. Page, P. Celani, M. Olivucci, W. E. Schmid, S. A. Trushin, and W. Fuss, J. Phys. Chem. A **105**, 4458 (2001).
[9]K. Kosma, S. A. Trushin, W. Fuss, and W. E. Schmid, Phys. Chem. Chem. Phys. **11**, 172 (2009).
[10]M. Merchan, L. Serrano-Andres, L. S. Slater, B. O. Roos, R. McDiarmid, and Y. Xing, J. Phys. Chem. A **103**, 5468 (1999).
[11]M. K. Lawless, S. D. Wickham, and R. A. Mathies, Acc. Chem. Res. **28**, 493 (1995).
[12]C. Y. Ruan, V. A. Lobastov, R. Srinivasan, B. M. Goodson, H. Ihee, and A. H. Zewail, Proc. Natl. Acad. Sci. U.S.A. **98**, 7117 (2001).
[13]U. Killat, Zeitschrift Fur Physik **270**, 169 (1974).
[14]S. Pullen, L. A. Walker, B. Donovan, and R. J. Sension, Chem. Phys. Lett. **242**, 415 (1995).
[15]S. H. Pullen, N. A. Anderson, L. A. Walker, and R. J. Sension, J. Chem. Phys. **108**, 556 (1998).
[16]S. Lochbrunner, W. Fuss, W. E. Schmid, and K. L. Kompa, J. Phy. Chem. A **102**, 9334 (1998).
[17]E. C. Carroll, B. J. Pearson, A. C. Florean, P. H. Bucksbaum, and R. J. Sension, J. Chem. Phys. **124**, (2006).
[18]E. C. Carroll, J. L. White, A. C. Florean, P. H. Bucksbaum, and R. J. Sension, J. Phys. Chem. A **112**, 6811 (2008).
[19]H. Tamura, S. Nanbu, T. Ishida, and H. Nakamura, J. Chem. Phys. **124**, 084313 (2006).
[20]A. Hofmann and R. de Vivie-Riedle, J. Chem. Phys. **112**, 5054 (2000).
[21]P. Celani, F. Bernardi, M. A. Robb, and M. Olivucci, J. Phys. Chem. **100**, 19364 (1996).
[22]P. Celani, S. Ottani, M. Olivucci, F. Bernardi, and M. A. Robb, J. Am. Chem. Soc. **116**, 10141 (1994).
[23]H. J. C. Jacobs and E. Havinga, in *Advances in Photochemistry* (Wiley, New York, 1979), Vol. 11, p. 305.
[24]H. Harada, S. Shimizu, T. Yatsuhashi, S. Sakabe, Y. Izawa, and N. Nakashima, Chem. Phys. Lett. **342**, 563 (2001).
[25]J. D. Hays and R. C. Dunbar, J. Phys.Chem. **83**, 3183 (1979).
[26]N. Keshava and J. F. Mustard, IEEE Signal Process. Mag. **19**, 44 (2002).
[27]H. Y. Afeefy, J. F. Liebman, and S. E. Stein, in *NIST Chemistry WebBook, NIST Standard Reference Database Number* 69, edited by P. J. Linstrom and W. G. Mallard (NIST, Gaithersburg MD, 2010).
[28]J. H. Jiang, Y. Z. Liang, and Y. Ozaki, Chemom. Intell. Lab. Syst. **71**, 1 (2004).
[29]J. H. Jiang, Y. Z. Liang, and Y. Ozaki, Chemom. Intell. Lab. Syst. **65**, 51 (2003).
[30]M. Berman, A. Phatak, R. Lagerstrom, and B. R. Wood, J. Chemom. **23**, 101 (2009).
[31]M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, and J. F. Huntington, IEEE Trans. Geosci. Remote Sens. **42**, 2085 (2004).
[32]M. E. Winter, Proc. SPIE **3753**, 266 (1999).
[33]M. D. Craig, IEEE Trans. Geosci. Remote Sens. **32**, 542 (1994).
[34]A. A. Green, M. Berman, P. Switzer, and M. D. Craig, IEEE Trans. Geosci. Remote Sens. **26**, 65 (1988).
[35]I. T. Jolliffe, *Principal Component Analysis*, (Springer, Berlin, 2002).
[36]See supplementary material at http://dx.doi.org/10.1063/1.3681258 for a figure showing how different pump-probe experiments fall in the span of the same three endmembers.
[37]S. Deb and P. M. Weber, in *Annual Review of Physical Chemistry*, edited by S. R. Leone, P. S. Cremer, J. T. Groves, and M. A. Johnson, (Annual Reviews, Palo Alto, 2011), Vol. 62, p. 19.