Control Goal Selection through Anticorrelation Analysis in the Detection Space

Huyen T. Tran,[†] Dmitri A. Romanov,[‡] and Robert J. Levis*,[†]

Center for Advanced Photonics Research, Department of Chemistry, and Department of Physics, Temple University, Philadelphia, Pennsylvania 19122

Received: April 26, 2006; In Final Form: July 10, 2006

A statistical method is reported to determine the pairs of fragment ions in a mass spectrum that are most susceptible to control by adaptive optimization of the laser pulse shapes in the strong-field regime. The proposed method is based on covariance analysis of the mass spectral fragmentation patterns generated by a set of randomly shaped pulses. The pairs of fragment ions that have higher negative covariances possess a correspondingly higher degree of controllability in an adaptive control experiment, whereas the pairs that have higher positive covariances possess correspondingly lower controllability.

Introduction

Predicting the outcome of the highly nonlinear interaction of an intense laser pulse with a molecule is difficult. The task of predicting the controllability of potential product channels resulting from such interactions is equally daunting. The development of mass spectral sensing technologies in the strongfield regime, however, rely on knowledge of the controllability of product channels.¹ To address this challenge, we present a new method for predicting controllability and test the method using strong-field laser mass spectrometry.

In strong-field mass spectrometry, molecular fragmentation is caused by a laser that has an electric field strength on the order of that which binds the valence electrons to molecules.² The laser—molecule interaction results in energy deposition into all internal degrees of freedom (rotational, vibrational, and electronic excitation), resulting in ionization and photodissociation. Recent reports demonstrate that different laser pulse shapes lead to different molecular fragmentation distributions.^{1,2} The ability to manipulate mass spectral fragmentation patterns using shaped laser pulses establishes the potential for the selective excitation, photodissociation, and ionization of molecules.

The combination of laser pulse shaping, adaptive feedback control, and strong-field mass spectrometry have enabled bond-selective photochemistry. In this method, the spectral phase and amplitude of an intense laser pulse are optimized using feedback from experimental measurements. This adaptive feedback approach has been used in mass spectrometry to increase the ratio of one fragment ion to another. The desired control can be achieved without a priori knowledge of the Hamiltonian of the molecule involved in excitation.¹ Many control experiments on model systems of polyatomic molecules have since been probed.^{3,4}

The interaction of an intense laser pulse with a polyatomic molecule in a strong-field typically results in several product states in the mass spectrum and a correspondingly large number of potentially controllable ion ratios. Determining the highly controllable product states is valuable for elucidating control mechanisms of laser-assisted chemical reactions and for devel-

oping further applications. The knowledge of the controllable ion ratios will provide a statistical means for unveiling important features of the molecular Hamiltonian that are susceptible to effective manipulation by tailored pulse shapes. This provides insight into the alterable patterns of molecular fragmentation. From a practical standpoint, the strong-field laser pulse shapes that control various fragment ratios can substantially facilitate the analytical identification of a molecule in a complex mixture, as may be observed, for example, in toxic chemical detection systems. Selecting controllable product states is relatively easy when the dissociation channels of the sample species are few, as is the case for a reasonably small molecule such as CH₂BrCl, which has only a few prominent peaks in its mass spectrum.³ The task of selecting controllable product states becomes more challenging as the number of fragmentation channels increases. When there are more than two dissociation channels, one approach to controllable product selection involves chemical intuition. For instance, in the case of acetophenone (C₆H₅-CO-CH₃) with four major product states, the ion signals for $OC-CH_3^+$ and $OC-C_6H_5^+$ were selectively manipulated.^{1,6} Despite the positive demonstrations of controllability, chemical intuition may not always be useful in guiding the selection of the most controllable channels. Several other product states with significant ion yield in the mass spectrum of acetophenone have not been tested for control. Thus, there may remain combinations of molecular fragment ions that display equal or higher controllability, in comparison to OC-CH3⁺ and OC-C₆H5⁺ channels.

Thus far, an alternative method has not been presented to replace chemical intuition for selecting controllable ion pairs. This heuristic approach becomes problematic as ever-morecomplicated molecules are investigated. In a given mass spectrum of *n* features, there are n(n - 1)/2 different ratios of ion pairs. For a typical organic molecule, the values of *n* from 1 to ~100 is generally operational. Increasing *n* rapidly creates a high-dimensional detection space and represents an attractive opportunity for discrimination if controllable channels can be identified. However, because the interaction between a shaped laser pulse and a molecule is highly nonlinear, the prediction of the outcomes of such interaction is almost impossible. In addition, the large number of degrees of freedom of complex molecules further clouds the predictive possibilities. This

^{*} Author to whom correspondence should be addressed. Tel.: 215-204-5241. Fax: 215-204-6179. E-mail: rjlevis@temple.edu.

[†] Center for Advanced Photonics Research, Department of Chemistry. [‡] Department of Physics.

situation was aptly coined "the curse of dimensionality", and it is especially critical in sensing-decision applications where a system is forced to utilize its resources to represent irrelevant portions of the search space, consequently undermining overall performance.^{7,8} For these reasons, narrowing the search space has become an essential issue in detection and adaptive control settings.

There are two search spaces of interest in the experiments presented here-that for the laser pulse shapes and that for the detection states-and each is astronomically large. One straightforward approach to search-space reduction is to reduce the number of parameters encoding the laser pulse. In the case of shaped laser control, this is accomplished by assigning multiple pixels the same retardance value in the spatial light modulator of the pulse-shaping apparatus.^{1,10,11} Optimal evolutionary algorithms that incorporate parametrization with periodic phase functions and multiple pulse construction for molecular dynamics control have also been reported.¹¹ Principal component analysis has been used to determine the most meaningful combinations of encoding parameters in the control space that produce a desired control goal.¹² All of these search space reduction methods focus primarily on laser pulse reconstruction in the control space and not directly on fragmentation distribution from molecular excitation in the detection space.

A complementary approach to reducing the space of available pulse shapes involves selecting an appropriate control goal. An approach for choosing controllable product channels is presented herein. Analysis of the correlation coefficient between pairs of product states is presented as a means for determining the controllable channels. The statistics of fragmentation patterns for a large set of random pulses are used to establish relative correlations between sets of integrated peak areas. These correlations are determined by the covariance statistics for pairs of mass spectral features of interest and are described by correlation coefficients that range from -1 to +1. The correlation coefficients are the normalized covariances reflecting how members of a set of random variables (the values of the measured mass-spectral peak intensity) vary with respect to each other. Most importantly, the correlation coefficients seem to predict the controllability of the ion yield ratio when shaped laser pulses are used for excitation.

We investigate the strong-field ionization/dissociation covariance matrix of dimethylmethylphosphonate (DMMP), $(CH_3O)-PO-(OCH_3)-(CH_3)$. DMMP is a simulant for the chemical warfare nerve agent sarin. The mass spectrum of DMMP reveals 12 major fragment ions, resulting in 66 possible independent peak ratios upon statistical analysis of many laser pulse shapes. Reports reveal that several DMMP product states are controllable.¹³ In the present paper, the correlation matrix for 12 selected fragment ions in the mass spectrum of DMMP is investigated to determine the most controllable product states. To test the predictive value of the correlation coefficient for controllability, corresponding optimization experiments were performed.

Experimental Methods

A Ti:sapphire laser that consisted of an oscillator (KM Labs) and a regenerative amplifier (Coherent, model L-USP-1K-HE) was used to produce ultrashort pulses at a repetition rate of 1 kHz. The wavelength of the laser beam was centered at 800 nm with pulse duration of ~60 fs. When focused to a 100- μ mdiameter spot, the resulting laser intensity was ~10¹³ W/cm². Pulse shaping was accomplished by an all-reflective optical system. The pulse shaper operates with a computer-controlled spatial light modulator (SLM) placed in the Fourier plane of a 4f spectrometer. The SLM contains an absorbing polarizer and two consecutive 128-element arrays of liquid-crystal pixels that alter the phase and amplitude of individual spectral components of the beam by modulating the retardance of each pixel. Superposition of the modulated pixels produces a particular phase and amplitude mask that is applied to a transform-limited laser pulse, resulting in a temporally shaped laser pulse.

The shaped beam was focused inside the ion-extraction region of a linear time-of-flight mass spectrometer (TOF-MS) apparatus, using a lens with a focal length of 20 cm and a numerical aperture of 0.05. The DMMP (Sigma-Aldrich, 97%, CAS No. 756-79-6) was introduced into the TOF chamber in the gas phase via a variable leak valve at a pressure of $\sim 2.5 \times$ 10⁻⁵ Torr. The resulting mass spectra were measured using a digital oscilloscope, and 100 mass spectra were averaged per pulse shape. To construct the correlation matrix, the 12 most prominent molecular fragment ions were selected in the DMMP mass spectrum that was produced by a transform-limited pulse. In the optimization experiments, an evolutionary algorithm⁹ that applied an adaptive feedback closed-loop control was used to determine the optimal laser pulse shape. Specifically, the mutation rate was 3%, one crossover per genome was performed, and the elitism factor was 2 for all experiments. Many parametrization schemes have been suggested for efficient control.14 To complete the proof-of-concept optimizations on a manageable time scale, the pixels of the SLM arrays were tied together in blocks of 16. Under these conditions, optimization was typically achieved in \sim 30 generations, with 40 individuals per generation.

Computational Methods

Let the mass spectrum contain *N* prominent peaks. The *N*-peak mass spectral pattern is recorded for a large number *M* of randomly shaped laser pulses. This generates *N* sets of random variables, $\{X_0\}$, ..., $\{X_N\}$, with *M* elements in each set: $\{X_0\} = \{X_{00}, X_{01}, ..., X_{0M}\}$, etc. In these sets, the first subscript indicates a particular mass spectral feature and the second indicates a particular pulse shape. The covariances between mass spectral features then are defined as¹⁵

$$Cov(X_{i}, X_{j}) = E\{[X_{i} - \mu_{i}]|X_{j} - \mu_{j}]\} = \frac{1}{M} \sum_{k=0}^{M} [X_{ik} - \mu_{i}][X_{jk} - \mu_{j}]$$
(1)

where $E\{\dots\}$ indicates the set averaging and μ_i refers to the mean value; $\mu_i = E\{X_i\}$. This definition makes for a symmetric matrix,

$$Cov(X_i, X_j) = Cov(X_j, X_i)$$
(2)

and, in the case of i = j,

$$Cov(X_i, X_i) = E\{X_i^2\} - [E\{X_i\}]^2 = \sigma_i^2 = Var(X_i) \quad (3)$$

and represents the variance of the variable X_i , where σ_i is the standard deviation. The symmetric matrix of correlation coefficients is constructed as

$$\rho_{ij} = \rho(\mathbf{X}_i, \mathbf{X}_j) = \frac{\operatorname{Cov}(\mathbf{X}_i, \mathbf{X}_j)}{\sigma_i \sigma_j} \tag{4}$$

The elements of the matrix ρ (that is, the correlation coefficients) are values lying conveniently in the interval

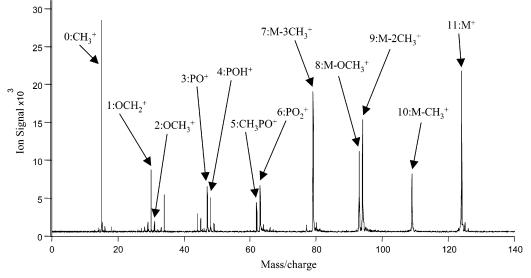


Figure 1. Time-of-flight (TOF) mass spectrum of dimethylmethylphosphonate (DMMP) with a transform-limited laser pulse at a center wavelength of 800 nm, pulse duration of \sim 60 fs, and intensity of \sim 10¹³ W/cm². The indices to various peaks used to generate the correlation matrix are indicated.

[-1, 1]. A correlation $\rho_{ij} = +1$ indicates the points (X_{ik}, X_{jk}) lie on a straight line with positive unit slope in the (X_i, X_j) plane; a correlation of -1 indicates that the points lie on a straight line with negative unit slope. All the diagonal elements of the matrix (i = j) are equal to 1. The values of the nondiagonal elements are used to sort the pairs of ion peaks, according to the average correlation of their ion intensities. Because the matrix is symmetric, the values that lie either above or below the diagonal are sufficient for this analysis. These values will reveal and quantify the pattern in variations of ion yield in response to a change in the pulse shape.

A positive correlation occurs when X_i and X_j have a tendency to increase or decrease simultaneously, with respect to their mean values, as the laser pulse shape is varied. Thus, a positive correlation coefficient implies that the peak intensity of the fragment ion X_i increases (decreases) as the intensity of the fragment ion X_j increases (decreases). A correlation coefficient that approaches zero conveys the lack of correlation between X_i and X_j . A negative correlation coefficient implies that an increase (decrease) in X_i coincides with a corresponding decrease (increase) in X_j .

The cases of strong negative correlation $(\rho_{ij} \rightarrow -1)$ are of particular interest for control purposes. This is because many optimization procedures require enhancement of one product state at the expense of another. The ratio between the yield of two product channels that are negatively correlated has an increased potential for such optimal control, because, on average, one channel has a tendency to increase (decrease) when the other decreases (increases), thus providing a measure of discrimination. Quantifying the variation in the yield between two product channels may be useful in predicting the controllability of a given ratio. Moreover, the dynamic range for the control of these product state ratios should also reflect a parallel trend in correlation coefficients (i.e., anticorrelation should predict higher controllability and, thus, discrimination).

Results and Discussion

Figure 1 displays the TOF mass spectrum of DMMP that has been ionized and fragmented using a transform-limited pulse taken with an intensity of $\sim 10^{13}$ W/cm², at a carrier wavelength of 800 nm and with a pulse duration of 60 fs. Out of the many fragments detected, the 12 most prominent ion peaks were

chosen for covariance analysis. These include the following: $CH_3^+ (m/z = 15)$, $OCH_2^+ (m/z = 30)$, $OCH_3^+ (m/z = 31)$, $PO^+ (m/z = 47)$, $POH^+ (m/z = 48)$, $CH_3PO^+ (m/z = 62)$, $PO_2^+ (m/z = 63)$, M- $3CH_3^+ (m/z = 79)$, M- $OCH_3^+ (m/z = 93)$, M- $(2CH_3)^+ (m/z = 94)$, M- $CH_3^+ (m/z = 109)$, and $M^+ (m/z = 124)$, and are labeled in Figure 1 as peaks 0-11, respectively. Statistical analysis of this series of mass spectra provides 66 independent peak ratios that could be used for optimization control. Mass spectra were collected, and the corresponding peak intensities were measured for a large number of random pulse shapes. The values of these peak intensities were used to construct the correlation matrix according to eq 1.

A key objective at this stage in the investigation was to collect sufficient statistics on the fragmentation patterns as a function of the shape of the laser pulse, to construct a matrix that describes the true correlations associated with each pair of mass spectral peaks. This is accomplished by sampling the available pulse shape space with an appropriately large number of randomly shaped pulses. The task of minimizing the sampling space is of primary importance, considering that there are an astronomic number $(>10^{40})$ of achievable pulse shapes. To determine the size of the population necessary to represent the covariance properties of the entire pulse shape space, populations of increasing size were sampled until no major variation in the correlation matrix was detected. The matrix was first recorded for the 12 ions of DMMP using 1000 randomly shaped pulses. The experiment was repeated and a substantially different correlation matrix emerged. The sampling population was increased to 3000 and the correlation matrix was measured twice. Again, the correlation matrix revealed significant variation between runs. Subsequent trials were conducted twice using both 5000 and 7000 randomly shaped pulses. For these larger sampling sets, the values of the corresponding correlation coefficients in all four trials were indistinguishably similar. The results suggest that 5000 randomly shaped pulses are sufficient to make statistically valid predictions, in regard to correlations between the dissociation products of DMMP. The correlation matrix measured for 5000 pulses is expected to be unique to DMMP, because the dissociation dynamics are dependent on the unique molecule-laser interaction Hamiltonian.³

Dissociation patterns can vary greatly from pulse shape to pulse shape. For this reason, the components of a mass spectrum

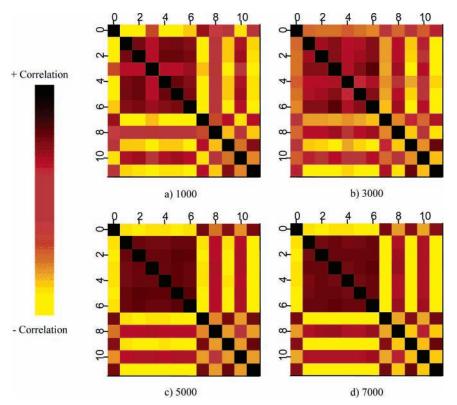


Figure 2. Normalized correlation matrixes of the ratios of 12 fragment ion intensities collected from (a) 1000, (b) 3000, (c) 5000, and (d) 7000 randomly shaped pulses.

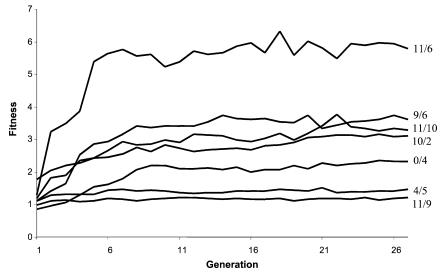


Figure 3. Fitness as a function of generation for the optimization of ion ratios for peaks corresponding to 11/6, 9/6, 11/10, 10/2, 0/4, 4/5, and 11/9.

that resulted from an individual pulse shape were normalized. For any given mass spectrum, the product states chosen for statistical analysis were divided by the area occupied under the sum of those states. Following normalization, the 12 ions were inserted into a matrix with dimensions of 12×12 , to compute their correlations. A false color image map is used to represent the correlation values for the selected ions of DMMP in the matrix. Figure 2 shows the correlations for populations of 1000, 3000, 5000, and 7000 iterations. The color scheme gradient represents the full dynamic range of the correlation coefficients; yellow (light) corresponds to -1, black corresponds to +1.

To test the hypothesis that correlation predicts controllability, adaptive control of the peak ratios was conducted on selected pairs of fragment ions. These sets contain the M^+ , $M^-(CH_3)^+$, $M^-(2CH_3)^+$, PO_2^+ , CH_3PO^+ , POH^+ , OCH_3^+ , and CH_3^+ ions.

Figure 3 is a plot of the fitness (ratio) as a function of generation for these ion ratios. To quantify the "controllability" of a given pair of ions, we define the controllability as the ratio of the fitness values for the initial and final generations. Thus, high controllability corresponds to a large difference between the random and optimized laser pulse shapes. (This value is distinct from the dynamic range which is defined by the minimum ratio of A/B to the maximum ratio of A/B.) The highest controllability was observed for the M⁺/PO₂⁺ratio (after ~30 generations of optimization). The correlation coefficient for this ion pair was -0.777. The M $-(2CH_3)^+/PO_2^+$ ion pair had a correlation coefficient of -0.657, and this pair showed a reduced controllability compared to M⁺/PO₂⁺. The ion ratios M⁺/M $-CH_3^+$, CH₃⁺/POH⁺, M $-CH_3^+$ /OCH₃⁺, and POH⁺/CH₃PO⁺ had coefficients of -0.505, -0.295, 0.304, and 0.356, respectively, and

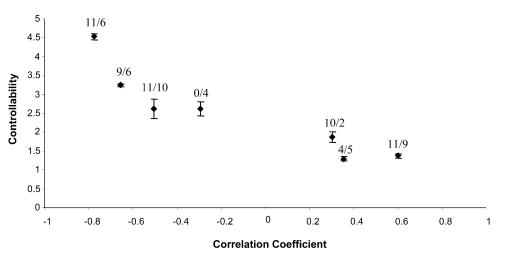


Figure 4. Plot of the controllability as a function of correlation value for a series of ion ratios (11/6, 9/6, 11/10, 10/2, 0/4, 4/5, and 11/9), shown as a function of increasing correlation coefficient. The error bars represent the standard deviation from two optimization experiments for each ion ratio.

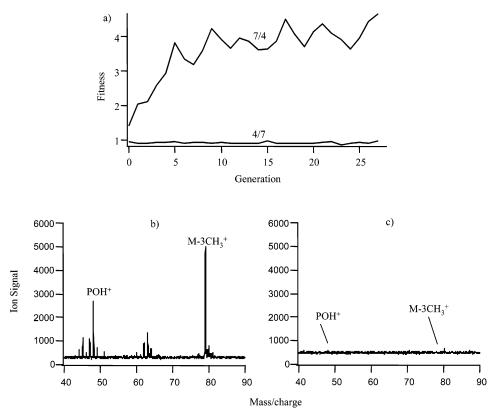


Figure 5. (a) Fitness of $M-(3CH_3^+)/POH^+$, corresponding to ion pairs 7/4 and 4/7. The other panels depict the TOF mass spectra space for (b) 7/4 and (c) 4/7 after optimization.

showed a trend of decreasing controllability. The most limited control was observed for the $M^+/M^-2CH_3^+$ ratio whose correlation coefficient was 0.602. Figure 4 is a plot of the controllability of the optimized peak ratios, as a function of the correlation coefficient. Figure 4 demonstrates that controllability is highest for the most negatively correlated ions and decreases monotonically as the correlation coefficient increases.

On the basis of covariance analysis, limited control should be exhibited on ion pairs with correlation coefficients approaching +1. However, we observed, in this experiment, that some of the ion pairs with positive correlation coefficients were moderately controllable. For example, the ratio $M-CH_3^+/$ OCH_3^+ has a correlation coefficient of 0.304 and exhibits a controllability of a factor of ~2. This represents an ~100% increase in controllability for this ion ratio. A possible explanation for good control with positive correlation coefficient is that the correlation matrix is obtained by sampling many pulse shapes, the majority of which do not lead to controllability for more-positive correlation coefficients. However, the adaptive control experiment is biased toward favorable search space by exploring mutation, crossover, and cloning of the best pulse shapes that are advantageous for a desired outcome. The genetic algorithm is biased toward exploration of dimensions in the control space that are favorable for a particular control goal. For the ion pairs with positive correlation coefficients, which demonstrate little to no control, despite adaptive feedback optimization, the control progress may be extremely gradual, because of (i) undersampling of the desirable pulse shapes during optimization procedures or (ii) the fact that the ratio is simply not controllable. These control outcomes highlight the fact that correlation analysis is a statistical search tool for determining the average differences in the response of ion intensities to a random set of excitation pulses.

One might assume that the predicted high controllability of the ratio A/B suggests that the inverse ratio B/A should be also easily controllable, and thus enhance the dynamic range. However, equal controllability of the forward and inverse ratios is not observed in all cases. Some of the fragment ions that have negative correlation values were only controllable in one direction (e.g., A/B and not B/A). For example, the fragment ions $M-(3CH_3)^+$ and POH⁺ that were anticorrelated (with a correlation coefficient of -0.847) were only controllable for the case of enhancing the $M-(3CH_3)^+/POH^+$ ratio but not for enhancing the POH⁺/M $-(3CH_3)^+$ ratio. The fitness as a function of generation (for these two experiments) is plotted in Figure 5a. For these two ions, the results indicate that the dynamic range for control is restricted, to some degree. Two inter-related reasons may account for the asymmetry in control. First, the controllability of A/B and B/A is dependent on the position in the search space at the start of the optimization procedure. Randomly generated phase and amplitude masks of the initial pulse almost invariably increase laser-pulse duration, in comparison with transform-limited input laser pulse, consequently reducing the laser intensity. This typically biases the first few generations of the optimization to regions of detection space where fragmentation is small and fragment masses are more evenly distributed. Second, the ion intensities sought for the "reverse" optimization may be even smaller than those of the first generation in the available laser pulse shape space. In this case, noise often becomes the major factor, interfering in the control process. These two aspects of the reverse control suppression are well-illustrated by the case for the $M-(3CH_3)^+$ and POH⁺ ions. Figures 5b and 5c show mass spectral fragmentation patterns of DMMP generated by shaped pulses optimized for the forward and reverse experiments. Optimization of the POH⁺/M $-(3CH_3)^+$ ratio was not possible, because the intensity of the fragment ion POH⁺ was so small that, under various pulse shapes, the feature was obscured by noise. These causes of the control asymmetry are further corroborated by the observed positive correlation between the controllability numbers and the values of the final peak intensity ratio. In this regard, an experiment can be suggested for conducting forward and reverse optimization using shaped laser pulses that are restricted to have fixed intensity. We relegate these analyses to future investigations.

Conclusions

A method was presented to determine controllability in various channels of a quantum system. This method is applied to the problem of controlling the fragmentation distribution of molecules that have been ionized using intense laser pulses. A covariance matrix is calculated for the ion product channels of interest, and correlation coefficients were computed for each pair of ions. The highest degree of control was observed for the ion pairs with the highest anticorrelation coefficient. Controllability of the ion ratios decreased as the correlation coefficients of product state pairs became more positive. This represents a method for rapidly determining controllable channels in a complex system.

Acknowledgment. We acknowledge the generous support of the MURI program as managed by the Army Research Office, and the support of DARPA as managed through the Air Force Office of Scientific Research. We thank Alexei Markevitch, John Brady, and George Heck for discussions.

References and Notes

(1) Levis, R. J.; Menkir, G. M.; Rabitz, H. Science 2001, 292, 709.

(2) Levis, R. J.; Rabitz, H. A. J. Phys. Chem. A 2002, 106, 6427.

(3) Damrauer, N. H.; Dietl, C.; Krampert, G.; Lee, S. H.; Jung, K. H.; Gerber, G. *Eur. Phys. J. D* **2002**, *20* (1), 71.

(4) Anand, S.; Zamari, M. M.; Menkir, G.; Levis, R. J.; Schlegel, H. B. J. Phys. Chem. A **2004**, 108, 3162.

(5) Graham, P.; Menkir, G.; Levis, R. J. Spectrochim. Acta, Part B 2003, 58, 1097.

(6) Scott, D. W. Multivariate Density Estimation; Wiley: New York, 1992.

(7) Priebe, C. E.; Marchette, D. J.; Healy, D. M. *IEEE Trans. Pattern Anal. Mach. Intell.* **2004**, *26*, 699.

(8) Goldberg, D. Genetic Algorithms in Search, Optimization and Machine Learning; Addison-Wesley: Reading, MA, 1989.

(9) Pearson, B. J.; White, J. L.; Weinacht, T. C.; Bucksbaum, P. H. Phys. Rev. A 2001, 6306, 3412.

(10) Zeidler, D.; Frey, S.; Kompa, K. L.; Motzkus, M. Phys. Rev. A 2001, 6402, 3420.

(11) White, J. L.; Pearson, B. J.; Bucksbaum, P. H. J. Phys. B: At. Mol. Opt. Phys. 2004, 37, L399.

(12) Palliyaguru, L. 2006, Submitted.

(13) Weinacht, T. C.; Bucksbaum, P. H. J. Opt. B: Quantum Semiclassical Opt. 2002, 4, R35.

(14) Papoulis, A. Probability, Random Variables, and Stochastic Processes, 2nd Edition; McGraw-Hill: New York, 1991.