Toward understanding closed loop learning control: The importance of basis in searching the phase space

A Thesis Presented

by

Florian Rainer Langhojer

to

The Graduate School

in Partial Fulfillment of the Requirements

for the Degree of

Master of Arts

in

Physics

Stony Brook University

August 2004

Abstract of the Thesis

Toward understanding closed loop learning control: The importance of basis in searching the phase space

by

Florian Rainer Langhojer

Master of Arts

in

Physics

Stony Brook University

2004

This thesis investigates different routes to gaining physical insight from closed loop learning control experiments. We focus on the role of the basis in which pulse shapes are encoded and the algorithmic search is performed. We present simulations and experiments in selective molecular fragmentation using shaped ultrafast laser pulses. We demonstrate reduction of the dimensionality of the search space to one or two degrees of freedom. We show how the search space can then be mapped out in detail along the most important degree of freedom, leading toward a better understanding of the control mechanism.

To my family

Contents

List of Figures

Acknowledgements

First, I would like to thank my advisor Tom Weinacht for giving me the opportunity to work with him. He had great confidence in me, taught me, and guided my efforts. I highly appreciate the substantial help and cooperation from David Cardoza. He initiated me into the details of the lab and made the work very enjoyable. I also want to thank Carlos Trallero for discussions, help, and company.

Furthermore I want to thank everyone with whom I spent great times outside the lab and who made my stay an integrated experience.

I owe a debt of gratitude to everyone who is involved in the exchange program between the physics departments of the University of Würzburg and Stony Brook University for all the efforts to organize my stay here. I am grateful for the financial support from the State of Bavaria ("Stipendium nach dem Bayerischen Begabtenförderungsgesetz"). This work was supported by the National Science Foundation, the American Chemical Society, and the

SUNY Research Foundation.

I owe my warmest thanks to my family for always encouraging me. Without their support I would not have been able to come here and begin this work, nor finish it.

Chapter 1

Introduction

The development of ultrafast lasers opened up a vast field of new applications and new physical research. Ultrashort pulse durations of less than 100 as have been reported[1]. These short timescales render possible the observation and manipulation of ultrafast dynamics in surface physics, biological physics, atomic physics, molecular physics and many more fields of physics and beyond physics. Advances in amplifier technology for ultrashort pulses made very high peak powers of over 1 TW available[2]. Such high power pulses can be used to study the interactions of matter with strong fields. Electric field strengths E much larger than the atomic unit of field are accessible; i.e., $E > e/(4\pi\epsilon_0 r_{Bohr})$.

The progress in laser technology has been accompanied by the development of techniques for pulse shaping, allowing for control over the characteristic variables of the electric field: phase, amplitude, and polarization. By changing

these actively it is possible to access a large range of pulse shapes only limited by the bandwidth of the laser.

One goal of quantum chemistry is to control molecular dynamics. Shaped amplified laser pulses can be used to approach this goal. Modern ultrafast lasers provide timescales and field strengths that are natural to molecules and atoms. However, there is relatively little knowledge about the Hamiltonians that are involved in molecular dynamics and it is in general very difficult to calculate an optimal pulse shape to achieve control over molecular dynamics. In the early 1990s, Rabitz *et al.* [3] introduced the concept of "closed loop learning control". This approach does not require *a priori* knowledge of the molecular system. Instead, a learning algorithm controls the pulse shaper. It uses the effects of the shaped pulses on the molecular system as a feedback in order to create new pulse shapes. After several iterations of this loop the algorithm converges and finds an optimal pulse shape. Learning algorithms such as the Genetic Algorithm (GA) have been remarkably successful in finding optimal tailored laser pulses[4]. Many experimental implementations have shown that this approach is prosperous. Success in controlling vibrational excitation, high harmonic generation, selective molecular fragmentation and other applications has been demonstrated[5, 6, 7, 8, 9]. Our interest is focused on selective molecular fragmentation. There were several recent exciting ex-

perimental demonstrations in this field of laser selective chemistry[10, 11].

While discovering optimal pulse shapes for control has been demonstrated for many different systems, gaining an understanding of the physical control mechanism has generally proven to be more difficult. This is because of the nonlinear dependence of fragment selectivity on pulse shape and the high dimensionality of the parameter space that can be accessed with ultrafast optical pulse shapers. Current pulse shaping technology allows for independent control over the phase, amplitude, and polarization of hundreds of different frequency components in an ultrafast laser pulse[12]. Furthermore, since strong fields are available with amplified pulses, the non-perturbative electric fields in a focused ultrafast laser pulse can open up many new molecular channels through multiphoton absorption. While the high spectral resolution and strong fields have enabled shaped ultrafast lasers to control a wide variety of chemical processes, the high dimensionality of the available phase space has made interpretation very difficult because one does not know *a priori* which dimensions are important and how the different degrees of freedom act in concert.

There have been a few learning control experiments that have gained insight into a particular control mechanism by making detailed comparisons between calculations and experimental data[13, 8]. There have also been experiments that have made use of parameterization of the problem in terms of a few

control variables in cases where the control mechanism is understood[14]. An ultimate goal in this area is to develop a general procedure for automatically gaining Hamiltonian information from systems where the control mechanism is not known *a priori*[15].

The aim of this thesis is to discuss a number of general techniques to reduce the problem complexity and gain insight into the physical mechanisms responsible for control. We discuss simulations and molecular fragmentation experiments.

An initial step in understanding solutions to closed loop optimal control problems is to separate necessary from sufficient elements. Sufficient elements are characteristics of pulse shapes that are not required for control while not having any adverse effects. Necessary elements correspond to pulse shape features that are essential for gaining control. In general, GAs do not distinguish between necessary and sufficient features, and therefore it is difficult to extract control mechanism information from optimal pulses discovered by GAs.

Here, we discuss two mechanisms for separating necessary from sufficient elements of optimal pulses. The first is looking at the variation of pulse shape elements within a population as the algorithm converges, and the second is using a cost functional (penalty) to actively suppress unnecessary deviations in pulse shape from a chosen reference pulse shape (typically an unshaped

pulse).

The use of a cost functional or gene variation may help to illuminate which degrees of freedom are necessary to achieve control. However, the number of degrees of freedom may still be large. One approach to lowering the dimensionality of the problem is to change the basis in which the space of accessible pulse shapes is searched - i.e., transform to a basis in which the problem is separable. Here, we will associate the dimensionality of the problem with the number of necessary degrees of freedom required to achieve a specified control goal. The transformation required may be linear (analogous to forming normal modes for a series of coupled oscillators) or non-linear (analogous to transforming from Cartesian to spherical coordinates). This distinction between linear and nonlinear transformations is subject to future work.

If the optimal pulse shape for achieving a particular control target can be expressed in terms of a few degrees of freedom, then it is feasible to scan along these degrees of freedom to map out a "control surface" in analogy to constructing a potential energy surface (PES) along a subset of coordinates in a polyatomic molecule.

Discovering optimal bases for molecular control is a nontrivial task, and it is in general not obvious what transformation will be most successful in reducing the dimensionality of the problem. However, in this thesis we demonstrate that

two simple, intuitive basis transformations are capable of greatly reducing the dimensionality in two separate molecular fragmentation problems, allowing for scans along important dimensions of the problem. These scans point toward a physical interpretation of the control in each of the systems examined.

Chapter 2

Experimental Setup

The experiments described in this thesis make use of a Ti:Sapphire laser system capable of producing 30 fs pulses with an energy of ≈1 mJ at a repetition rate of 1 kHz[16, 17]. A femtosecond oscillator seeds a chirped pulse multipass amplifier system. We shape the laser pulses and focus them into our interaction chamber. A schematic diagram of the setup of the essential devices for the fragmentation experiments is shown in Fig. 2.1.

Since the development of ultrafast lasers, a number of techniques to shape the electric field of laser pulses have been introduced. We chose to use an Acousto-Optic Modulator (AOM) pulse shaper[18] for several reasons. The phase is written to the spectrum as a smooth function and there are no pixelation artifacts in the shaped pulse. With modern digital to analog computer boards it is also possible to update the pulse shape fast so that we can exploit

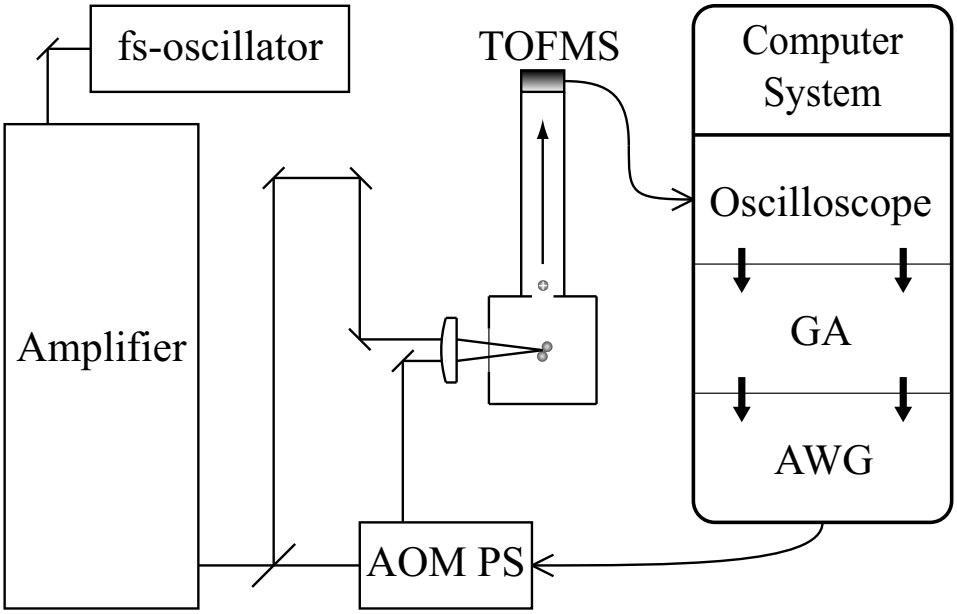

Figure 2.1 Experimental setup. AOM PS: pulse shaper, TOFMS: time of flight mass spectrometer, GA: genetic algorithm, AWG: arbitrary waveform generators

the high repetition rate of the laser system and have a high speed feedback loop for learning control experiments. Furthermore there is a low degree of coupling between the phases of different frequency components and, from a programming point of view, the transfer function is straight forward, i.e., one can write programs that operate on a phase function $\phi(\omega)$ and write this function directly to the pulse shaper. Amplitude shaping is straight forward, though we did not make use of it in the experiments.

The principle of an AOM pulse shaper is like most ultrafast pulse shaping techniques based on manipulating the pulse (phase and amplitude) in the frequency domain. We use an f-$2f$-f configuration as shown in Fig. 2.2. A

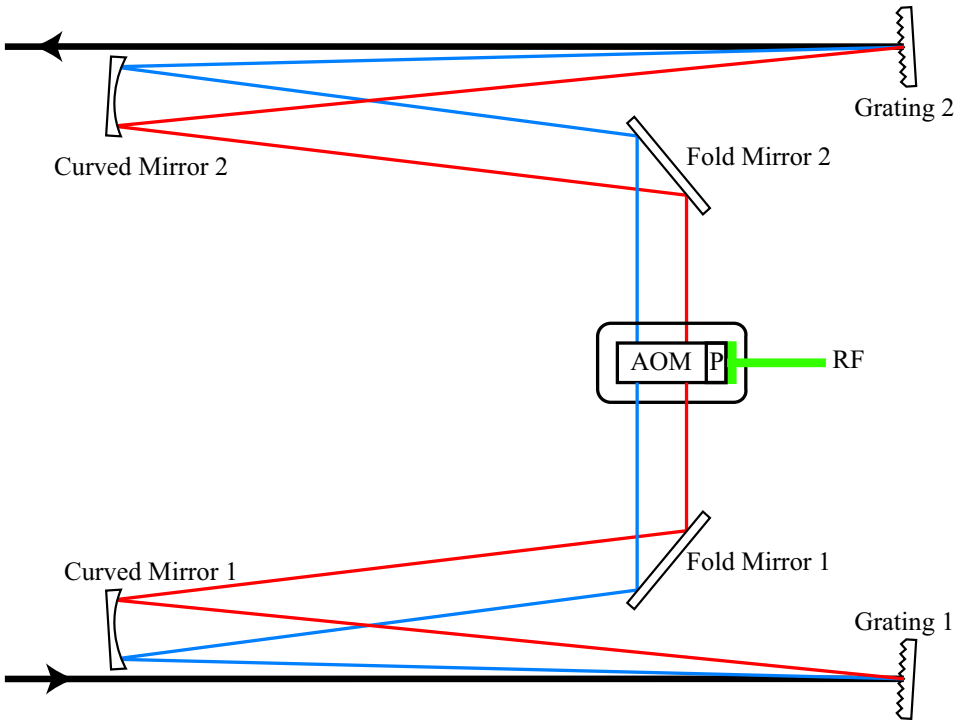

Figure 2.2 AOM pulse shaper. RF: connector for the shaped radio frequency. P: piezo crystal. AOM: acousto-optic modulator crystal.

grating disperses the frequencies of the incoming pulse. The diverging beam is collimated by a curved mirror. The AOM sits in the plane where the laser frequencies are mapped to space. We apply a shaped radio frequency (RF) pulse to the piezo crystal which is attached to the AOM crystal. This launches an acoustic wave across the AOM. The RF pulse is timed with the laser pulse such that the complete RF wave is written to the AOM by the time the laser pulse arrives. The laser "sees" the acoustic wave as transmissive diffraction grating, because the density modulations caused by the acoustic wave along the crystal correspond to modulations in the index of refraction. Our AOM

has a diffraction efficiency of $\approx 50\%$. The diffracted beam is refocused onto the second grating where the frequencies are recombined to form the shaped laser pulse. The overall efficiency of our pulse shaper is $\approx 30\%$.

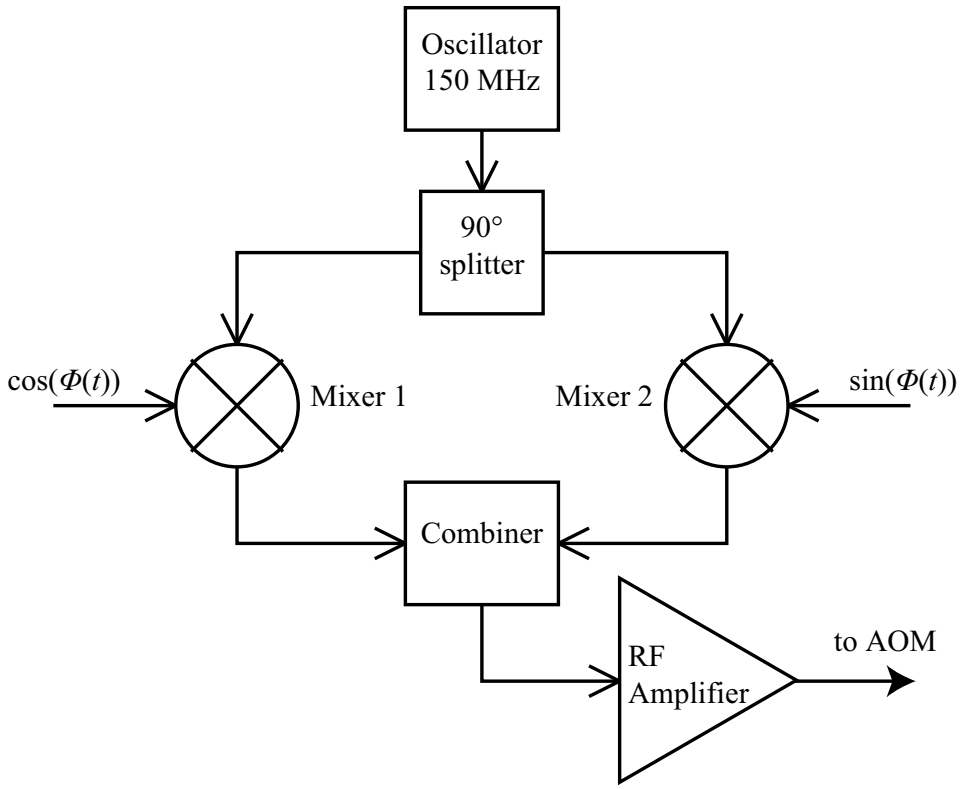

Figure 2.3 RF shaping scheme.

We produce the shaped RF pulse using the scheme shown in Fig. 2.3. A local oscillator puts out a sine-wave of 150 MHz which is evenly split into two components that are phase shifted by 90 degrees. These two components are mixed with the outputs of two computer controlled arbitrary waveform generators. We program the generators to produce voltages that represent

the sine and the cosine of the desired phase function ϕ. The mixed signals are then recombined. We amplify the shaped RF to drive the piezo with a peak power of $\approx$ 2 W. This combination of the splitter, the mixers and the combiner is referred to as I&Q modulator and commonly used in RF technology. Mathematically, the output $U(t)$ of the I&Q modulator can be described as

$$\begin{aligned} U(t) &= \cos(\omega t)\sin(\phi(t)) + \sin(\omega t)\cos(\phi(t)) \\ &= \sin(\omega t + \phi(t)), \end{aligned} \tag{2.1}$$

where ω is the carrier frequency (150 MHz) and $\phi(t)$ is the phase we wish to write to the laser pulse. The RF pulse lasts for the period of time Δt the acoustic wave needs to travel across the AOM crystal. Since the frequencies of the laser are mapped to the length of the AOM, the phase $\phi(t)$ maps to a phase $\phi(\omega_{laser})$ written to the laser pulse. Our pulse shaper allows for a maximum of 60π phase across the spectrum of the laser pulse.

For pump-probe experiments, we make use of a Mach-Zender interferometer (see Fig. 2.1). One arm contains the AOM pulse shaper and the other arm provides an equivalent delay. We program a linear spectral phase, $\phi(\omega) = k(\omega - \omega_0)$, onto the pulse shaper to introduce a pump-probe delay. We are able to scan the time delay by changing k. For learning control experiments we block the unshaped pulses.

The pulses from both arms of the interferometer are focused into a vacuum chamber and interact with an effusive molecular beam[17]. The pressure inside the chamber was kept at $\approx 5\times10^{-6}$ Torr during the experiments. A Time of Flight Mass Spectrometer (TOFMS), which measures the mass to charge ratio of positive ions produced in the laser focus, is connected to the interaction region. A computer-based oscilloscope records the ion spectra.

The GA we use in our simulations and experiments considers an ensemble of pulse shapes (the population) as potential solutions to the problem under study (e.g. selective molecular dissociation). Each of these pulse shapes (individuals) is represented by a series of numbers (genetic code) that encode the time-dependent electric field.

In the laboratory, we directly control the spectral phase. Thus, the genetic code represents the spectral phase $\phi(\omega)$ that is written to the AOM. However, there are many ways of encoding this phase. Each pulse shape in the search space can be expressed as an expansion on a set of basis vectors (genes), where the expansion coefficients are the genetic material (gene values). As discussed below, the choice of basis in which any solution is expressed (and in which the algorithm performs the search) is very important in guiding the interpretation.

Each electric field is allowed to interact with the system (either experimentally or by simulation) and a measure of the success (fitness) of each shaped

laser pulse is collected. In the molecular experiments, the ratios or differences of peak integrals representing different fragmentation channels are typically used to evaluate the fitness of each laser pulse that produces a TOFMS. In the simulations, we send the genetic code to a mathematical function which assigns a fitness to each pulse shape. All pulse shapes are sorted according to their fitness. The top 50% of the pulses are used to generate a new population of pulse shapes through a set of three operators (reproduction). The operators we used for reproduction are elitism (10%), mutation (20%) and two-point crossover (70%)[19]. Elitism allows the very best pulse shapes to remain unchanged and be passed to the next generation, ensuring that there is no loss of important genetic information. The type of mutation we make use of is creep mutation: A small part of the genetic code is assigned random new values with a Gaußian probability distribution centered around the old value. Two-point crossover exchanges a randomly chosen part of the genetic material between two individuals. The new population (generation) of pulse shapes is then evaluated, sorted and used for reproduction. The GA goes through many generations until there is no improvement in the average fitness (convergence).

Chapter 3

Simulations

While learning algorithms are important for effectively searching the phase space, they do not yield insight into molecular physics without additional efforts. Solutions discovered by learning algorithms can often have many features that are not specific to the problem being solved, but are rather a consequence of the way in which the algorithm operates. One must therefore be careful to separate characteristics of the solutions that are dictated by the algorithm from ones that occur in response to the physics in the problem.

To investigate the effectiveness of genetic variation and a cost functional in separating necessary from sufficient elements of optimal solutions, we initially focus on simple simulations that make the interpretation of the two approaches clear. Following this, we look at other simulations to highlight the importance of basis change in reducing dimensionality.

3.1 The Importance of the Cost Functional: Simple Two-Dimensional Simulations

The use of genetic variation as an identifier for unnecessary genetic information is motivated by the idea that a collection of "fit" pulses should have similar necessary features whereas any features that are unnecessary should have significant variation throughout the population. A simple test of this idea is to work on an optimization problem that requires only a subset of the available genes (i.e., the problem has a lower dimensionality than the allowed number of genes). We performed simulations on a very simple problem where the fitness landscape consisted of a two dimensional Gaußian function:

$$F(G_1, G_2) = \mathrm{e}^{-(G_1-1)^2-(G_2-1)^2} \tag{3.1}$$

Here, G_1 and G_2 are the first and the second gene. The goal for the GA was maximizing F, which corresponds to $G_1 = G_2 = 1$. In addition to the first two genes we gave the GA eight unnecessary degrees of freedom. None of these genes had an effect on the fitness. The values of the genes were allowed to vary between -2 and $+2$. The initial population was randomly distributed in a box around $\overrightarrow{0}$ (corresponding to an unshaped pulse) with a spread of 0.2. We added 7% noise, which is roughly the noise level of experimental molecular fragmentation data for low signal averaging.

The outcome of a typical GA run is shown in Fig. 3.1 a). We plot the average values of each of the genes from the last generation. The error bars represent the standard deviation in the last generation. One can see that the algorithm performed well in finding the solution by looking at the first two genes. Their values converged to 1. However, note that, although none of the other genes affect the fitness, the average values for many of them deviate significantly from zero. Not only are the average values for these "junk" genes nonzero, but the distributions of the gene values about the average values are very narrow - much narrower than one would expect for a random distribution of gene values. This makes it very difficult to infer which genes are necessary and which are merely sufficient simply by looking at the distribution of gene values within the population. If one assumed that all genes that deviate significantly from 0 are essential, here, one would yield a dimensionality of more than two, despite the fact that the problem is only two dimensional by design.

We believe that the convergence of unnecessary genes to values different from the mean gene value is inherent to GAs, due to the conjoint action of selection and reproduction. The effect of selection and reproduction is to lower the diversity of junk genes. Selection reduces diversity by eliminating many pulse shapes and reproduction largely does not generate new gene values.

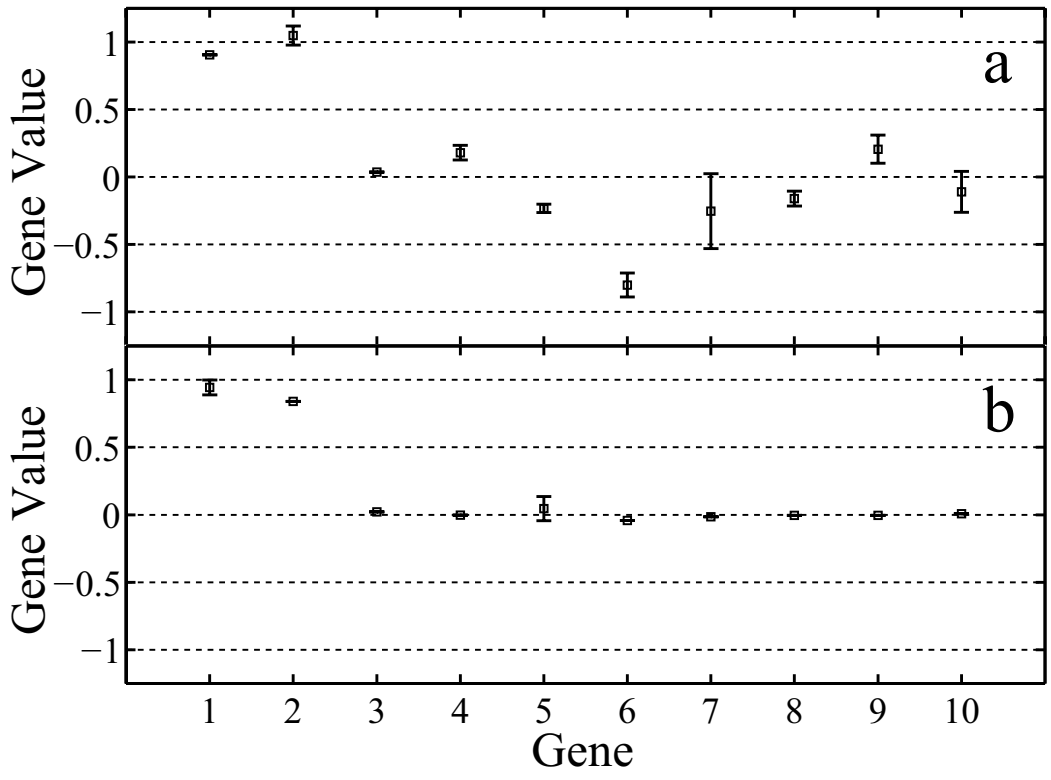

Figure 3.1 Typical outcome of simple algorithmic optimizations (see Eq. 3.1). The data points and error bars are the average values and standard deviations of each of the genes from the last generation. a) without cost functional b) with cost functional.

We found that increasing the population size or mutation rate and range in order to combat this was not effective because such large population sizes and mutation rates are required to maintain normal junk gene distributions that the GA no longer converges on experimentally feasible time scales. If the results of real experiments are to be interpreted through statistical measures of genetic variation, these variations must give a clear picture of the necessary

genetic material for simple problems such as the one discussed here. We argue that because the statistical variation of junk genes only approaches a random distribution in the limit where the algorithm does not rapidly converge, this approach to separating necessary from sufficient information is not generally appropriate.

The second technique we investigated to filter out the important dimensions is the use of a cost functional[20, 6]. Each pulse shape is penalized for deviations from a chosen reference pulse shape. It is implemented by creating a new fitness according to

$$F_n = F - W \sum_{i=1}^{N} |G_i - G_{0i}|^p, \tag{3.2}$$

where F is the original fitness, $\overrightarrow{G_0}$ is a chosen reference, N is the number of genes, W is a weight, and p is chosen to be between 0.25 and 2. The choice of p is discussed in more detail below. By looking at Eq. 3.2, one can see that any deviation from the reference results in a decreased fitness. This penalizes differences from the reference. In our simulations and experiments the reference was $\overrightarrow{0}$. We determined the value for W by trial and error. The cost functional was found to work best if the original fitness F went to $\approx 95\%$ of the fitness found with $W = 0$. In this regime, the cost functional does not prevent the GA from converging, but it is able to suppress unnecessary genetic variations from the reference.

Fig. 3.1 b) shows the result of a GA run on the two dimensional problem with cost functional ($p = 0.25$). Even without prior knowledge about the problem, it is obvious that optimization is possible with only the first two genes. The absolute values of the other genes are less than 0.05 after convergence. This is in contrast to Fig. 3.1 a), where many genes have nonzero values. Note that the solutions found by the GA with the cost functional applied do not have exactly the expected values for genes 1 and 2 ($G_1 = G_2 = 1$) but are slightly smaller. This is because the cost functional modifies the fitness landscape by pressuring the solutions toward the reference. It introduces a gradient in all dimensions of the fitness landscape pointing towards $\overrightarrow{G_0}$. One has to be careful to choose a weight for the cost functional, that allows it to restrict unnecessary deviations from the reference, while not overly distorting the fitness landscape. Lower noise in the problem allows for a smaller weight of the cost functional.

The nature of the gradient is determined by the parameter p. Low values (e.g. $p = 0.25$) put more pressure on small gene values, i.e., there is a larger fitness increase for decreasing small gene values than for decreasing large gene values by the same amount. High values (e.g. $p = 2$) put more pressure on large gene values, i.e., there is a larger fitness increase for decreasing large gene values than for decreasing small gene values by the same amount. Of course,

the deviation from zero is not always a good measure for the importance of a gene. Small gene values can be important for control (i.e., changing their value to zero can significantly reduce the fitness). Determining the dimensionality of a solution means setting some threshold value for small gene values to be considered insignificant. After setting such a threshold, one has to justify the choice by checking whether an individual with all gene values, that are smaller than the threshold, set to exactly 0 still yields a fitness of at least 95% of its original fitness. We performed this check for all of our control results.

By adding the cost functional to the algorithm, we are able to reduce the dimensionality of the simple problem to the essential degrees of freedom. This is in contrast to the use of genetic variation as a measure of necessary degrees of freedom, which was shown to yield ambiguous information for the simple toy problem considered here. The performance of the cost functional in more complicated simulations and experiments is discussed below.

3.2 Change of Basis in Simulated Raman Scattering Experiments

Understanding learning control requires more than identifying necessary and sufficient elements of solutions found by the GA. An important goal is to reduce

the dimensionality and diagonalize the search space. In a completely diagonal basis, the Hessian matrix $A_{ij} = \partial_i \partial_j F(x_1, ..., x_N)$ of the fitness function F is diagonal[1][21]. Reducing the dimensionality and diagonalizing the phase space correspond to finding a basis in which the fitness depends on only a few degrees of freedom and the coupling between these degrees of freedom is minimized.

In order to explore the effects of changing the basis in which the GA searches, we simulated the problem of optimizing the nonlinear spectrum of an ultrafast laser pulse, which plays an important role in many coherent control problems[22]. In particular, we concentrated on optimizing a chosen spectral component of $I(t)$. We considered a laser pulse with a spectral bandwidth of 28 nm centered at 778 nm. The GA was able to shape the spectral phase $\phi(\omega)$. For each pulse shape we computed the intensity as a function of time,

$$I(t) = \left| \mathcal{F} \left\{ |E_0(\omega)| \, \mathrm{e}^{\mathrm{i}\phi(\omega)} \right\} \right|^2, \tag{3.3}$$

where $|E_0(\omega)| = \sqrt{I_0(\omega)}$ corresponded to a measured spectrum of our laser pulse. The spectrum $S(\Omega)$ of $I(t)$ was calculated:

$$S(\Omega) = \mathcal{F} \left\{ I(t) \right\}. \tag{3.4}$$

[1]If the fitness can be separated in terms of a sum, i.e. $F(x_1, ..., x_N) = f_1(x_1) + ... + f_N(x_N)$, the Hessian matrix of the fitness function is diagonal. If, however, the fitness function is separable in terms of a product, i.e. $F(x_1, ..., x_N) = f_1(x_1) \cdot ... \cdot f_N(x_N)$, the Hessian matrix of the logarithm of the fitness function is diagonal.

The goal for the GA was increasing the spectral content of $I(t)$ between 4.10 THz and 5.60 THz while minimizing spectral content between 1.85 THz and 3.60 THz and between 6.33 THz and 16.5 THz. The fitness F for the GA was the difference in the integrated spectrum between the first region and the other two regions normalized to the entire spectrum. This optimization serves as an idealized model of perturbative impulsive stimulated Raman scattering (ISRS) excitation of different vibrational modes which depends on $S(\Omega)$[23].

We ran the GA on this problem in two different basis sets. First we used a simple differential basis in which the genes consist of the phase differences between discrete frequencies. The phase $\phi(\omega_i)$ at each sample point ω_i is given by:

$$\phi(\omega_i) = \sum_{j=1}^{i} G_j \qquad \text{if} \qquad \omega = \omega_i. \tag{3.5}$$

These discrete values of ϕ at ω_i are interpolated with cubic spline polynomials to make the phase a smooth function.

The result of such a simulation is shown in Fig. 3.2. We plot the gene values of solutions averaged over 10 separate runs. The error bars represent the standard deviations over these runs. Fig. 3.2 a) illustrates GA runs where we did not apply any cost functional. Fig. 3.2 b) shows the results of GA runs where we applied a cost functional as defined in Eq. 3.2 with $p = 2$. Here, the cost functional was not able to reduce the number of genes that deviated from

zero significantly. This is because the dimensionality of the problem cannot be reduced in this basis. The degrees of freedom are coupled and the problem cannot be treated as multiple single parameter optimizations (i.e., the Hessian matrix is not diagonal). Furthermore, there exists no linear transformation which can transform the problem into a basis in which the dimensionality of the problem is reduced.

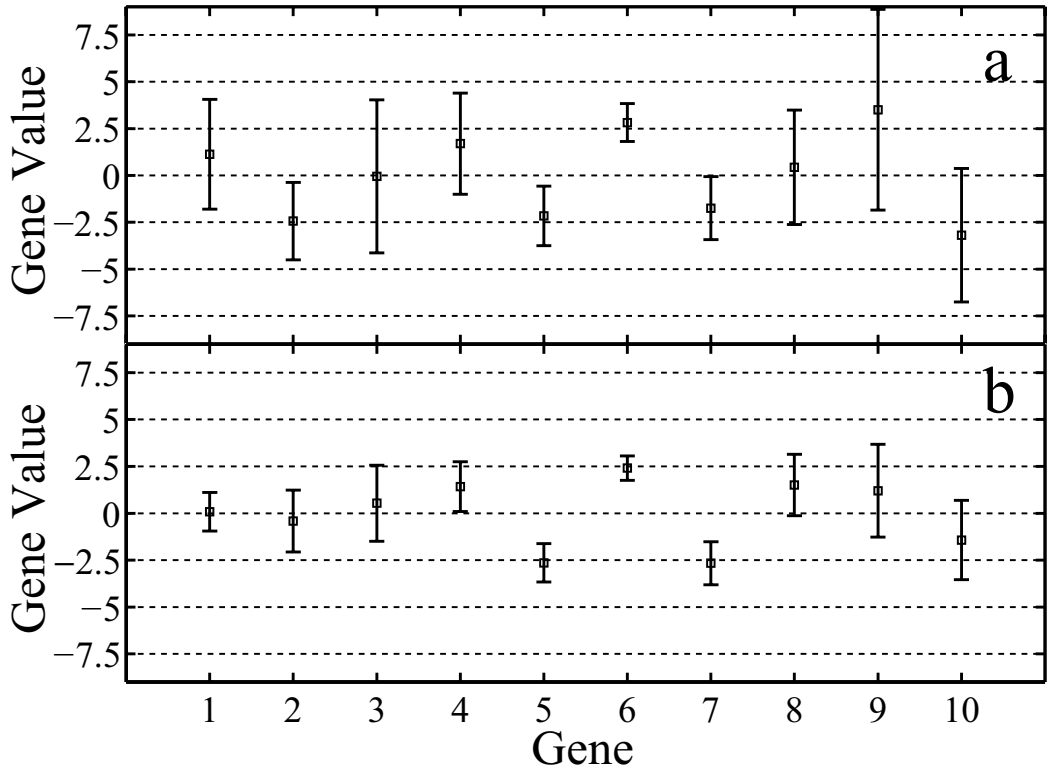

Figure 3.2 Averaged solutions of the ISRS simulations a) without and b) with cost functional. The genes encode differential phase (see Eq. 3.5).

We argue that the quality or suitability of a particular basis can be eval-

uated according to the number of essential degrees of freedom, the degree of coupling between them, and the uniqueness of solutions in that basis. A low dimensionality of the control space is essential for interpretation and further analysis of the control phase space. Low dimensional solutions found in an intuitive basis set can reveal immediate insight into the physical control mechanism. Moreover, it is possible to explore the phase space in interesting regions by scanning along these few essential components. For degrees of freedom that are not coupled, an N dimensional problem is equivalent to N one dimensional problems. For an uncoupled problem the search space is diagonal. Having unique solutions for a given basis aids in interpreting the control mechanism, and ensures reproducibility in finding optima in the phase space. Redundancy in a search space makes the interpretation more complicated. If the gene values of solutions from multiple GA runs largely agree (i.e., have a low standard deviation) one can be fairly sure that the same physical control mechanism is being addressed by these solutions and that there is only one accessible globally optimal pulse shape. Thus, in this case, averaging solutions over multiple GA runs is justified. If, however, different runs show different solutions, there are likely multiple equivalent optima in the search space. As we show for one of the molecular experiments below, averaging over different solutions can be very misleading.

The simple differential basis applied to the power spectrum optimization problem does not meet the criteria stated above. Both the dimensionality of the solutions and the variation of the gene values among different solutions are high.

Here, we demonstrate a connection between an intuitively appropriate basis for control and an understanding of the underlying control mechanism. Knowing that a periodic phase modulation is needed for solutions to this problem, we transformed the problem into a basis which allows the genes to encode for periodic spectral phase. This Fourier Phase Modulation (FPM) basis encodes the phase as

$$\phi(\omega) = \sum_i G_i \cos(k_i \omega). \tag{3.6}$$

The modulation frequency k_i is given by

$$k_i = \frac{2\pi}{N} \frac{M-1}{G-1} i, \tag{3.7}$$

where N is the total number of points for sampling ϕ (600), M (20) is the highest possible number of phase modulation periods, and G (20) is the number of genes. In this basis each gene codes for a periodic phase modulation of a distinct frequency k_i. Typical results from running the GA in the FPM basis are shown in Fig. 3.3. The data in Fig. 3.3 a) is the average solution from 10 runs without cost functional and Fig. 3.3 b) is the average solution from 10 runs with cost functional. The most striking difference between the

solutions found in the two basis sets is the number of nonzero genes when the cost functional is applied. In the FPM basis the cost functional was able to reduce the dimensionality to 1, making use of only gene 4. The same gene was prominent in all runs without cost functional, but other genes were also nonzero and the essential components of the search space were not evident. The cost functional reveals the essential dimensions. We call gene number 4 the essential dimension because it accounts for over 95% of the total fitness increase. The change to the FPM basis made it possible to reduce the dimensionality of the optimization problem using the cost functional. This simple case where the choice of optimal basis is obvious helps guide and interpret the change of basis in more complicated problems - e.g., the FPM basis may be well suited to problems that involve periodicity. We argue below that there will frequently be experimental evidence to guide the choice of a basis which may help to lower the dimensionality of the problem and aid in the interpretation of the control mechanism.

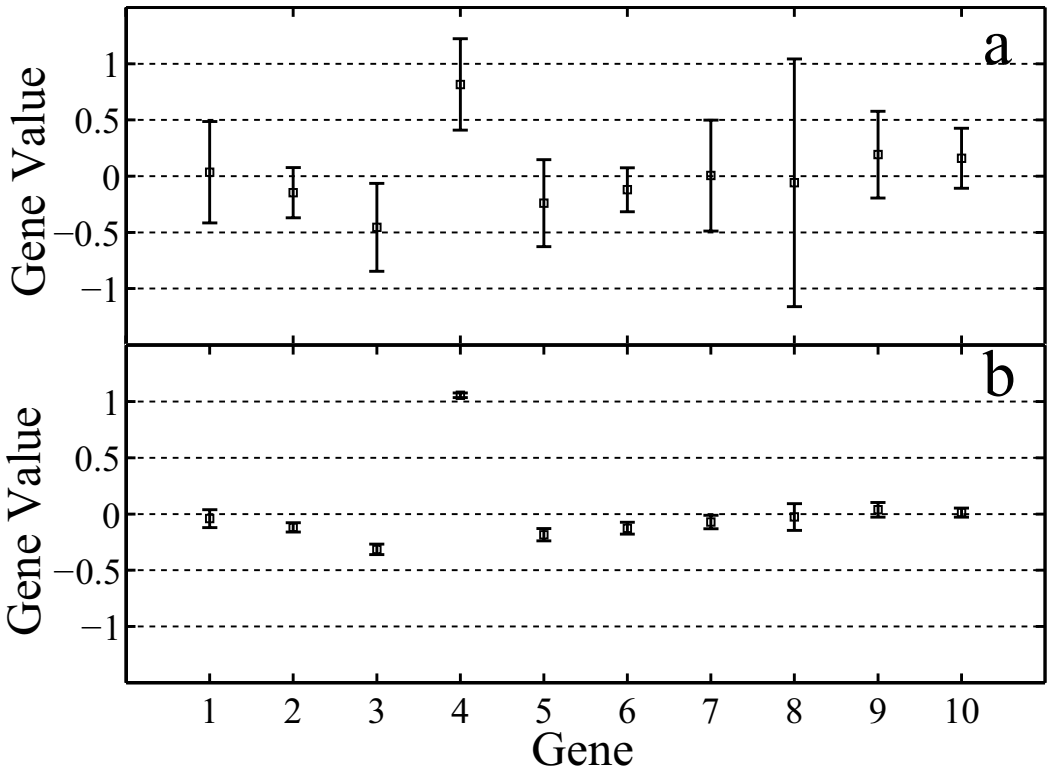

Figure 3.3 Averaged solutions of the ISRS simulations a) without and b) with cost functional. The genes are encoded in the FPM basis (see Eq. 3.6).

Chapter 4

Controlling Molecular Fragmentation

We performed the simulations presented in the previous chapter in order to understand the effects of a cost functional and the change of basis. In this chapter we describe how we used this understanding to learn about the physical mechanisms which form the basis of our control experiments on three different molecules.

In selective fragmentation experiments performed on 1,1,1-trifluoroacetone (CH_3COCF_3), 1,1,1-trichloroacetone (CH_3COCCl_3) and bromoiodomethane (CH_2BrI) we show that it is possible to separate necessary from sufficient features in the optimal control pulse by changing basis and using the cost functional. Changing basis allows for a low dimensional control space in which a control surface may be mapped out and point toward the physical mechanisms responsible for control.

4.1 Experiments on Trifluoroacetone

In our fragmentation experiments on trifluoroacetone we controlled the CF_3^+ / CH_3^+ ratio. We concentrate here on our results when the goal was to maximize CF_3^+ at the expense of CH_3^+. Typical TOF ion spectra are shown in Fig. 4.1. Fig. 4.1 b) is the fragmentation spectrum resulting from an unshaped laser pulse. Fig. 4.1 a) is the spectrum obtained with a shaped laser pulse after a closed loop optimization.

The most natural fitness function for achieving control over channel branching is the ratio of the products. However, the simple ratio is not an optimal fitness function because it is especially susceptible to noise at low signal levels. Instead, we used the difference between the ion yields of CF_3^+ and CH_3^+ as fitness. This proved to be much more stable. In the case of our experiments on CH_3COCF_3, we found that the optimal pulse shapes for optimizing differences and ratios were equivalent. Of course, this is not always true and one has to choose the fitness function carefully.

For an unshaped laser pulse the CF_3^+/CH_3^+ ratio was 1.2. When the goal for the algorithm was minimizing this ratio, it was possible to drive this ratio to 0.6 with a low intensity nearly transform-limited pulse. Aiming to maximize the ratio, we were able to yield $CF_3^+/CH_3^+ = 3.0$. Correlated with the increased CF_3^+ yield was an increase in the CH_3CO^+ yield. The other peaks in the

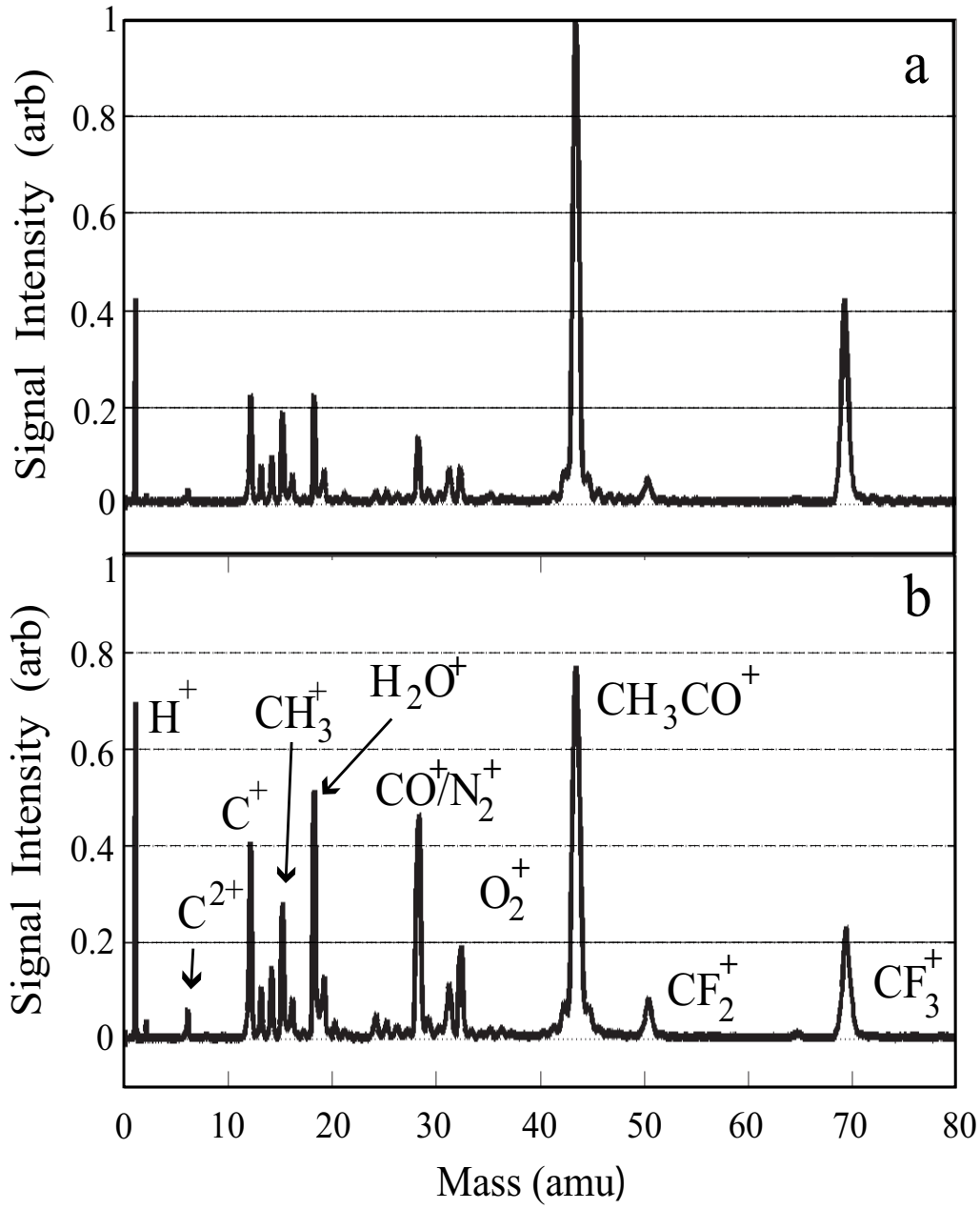

Figure 4.1 Trifluoroacetone fragmentation spectra a) after learning control optimization and b) for an unshaped pulse. Both panels have the same absolute scale. N_2^+, O_2^+, H_2O^+ and the other unlabeled peaks are background signals and seen before the injection of the molecular sample.

spectrum decreased. We were able to produce similar control results many times.

In Fig. 4.2 a) we plot an average of solutions found by the GA over multiple runs in the simple difference basis (see Eq. 3.5) with the cost functional suppressing unnecessary deviations from a transform-limited pulse. The genetic code for each data set represents an average of the fittest five pulses from multiple GA optimizations. The error bars in these figures represent the standard deviation of gene values over the separate optimizations. One can see that many genes were necessary for control. As the large error bars show, even with the cost functional on, there is significant spread in the genetic code of the solutions. Thus, according to our requirements for a good basis, the differential basis is not well suited to this fragmentation problem.

In determining what might constitute an appropriate basis for the trifluoroacetone experiments, we looked at the optimal solutions, which we characterized using second harmonic generation FROG[24]. In Fig. 4.3 we plot the reconstructed electric field of a typical optimal pulse in a) time and b) frequency domain representation. In time, the shaped pulse shows three prominent pulses separated by ≈ 90 fs. This corresponds to a frequency of ≈ 11 THz.

We observed such a well defined periodicity for all of the solutions. This, as well as the results of the ISRS simulations, led us to use the FPM basis

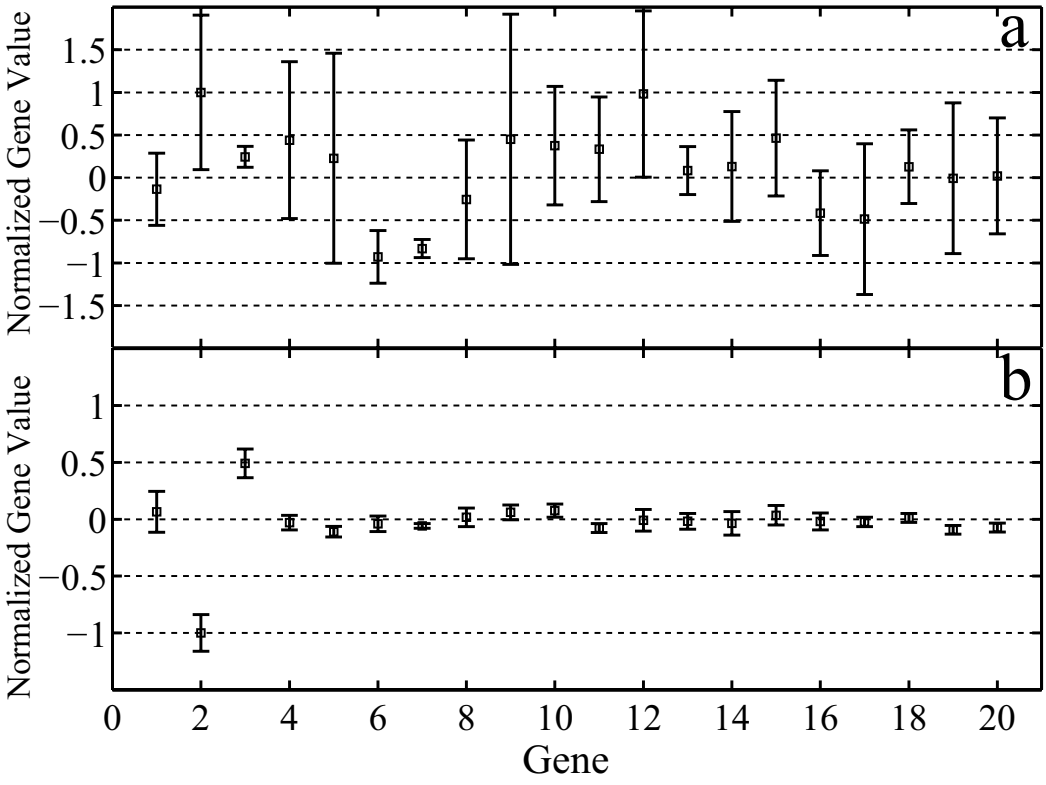

Figure 4.2 Averaged solutions when optimizing CF_3^+ at the expense of CH_3^+ in trifluoroacetone a) in the differential phase basis b) in the FPM basis.

(see Eq. 3.6) for this molecular experiment. The solution found by the GA in this basis is presented in Fig. 4.2 b). We average over several runs. The error bars are the standard deviations over the separate optimizations. Only two gene values are significantly different from zero. These are the only degrees of freedom needed for control. To prove this, we programmed pulses with only genes 2 and 3 kept at their optimal values and all other genes set to 0 onto the pulse shaper. The fitnesses calculated from the resulting TOF ion spectra

were at least 95% of the original fitnesses for all of the optimal pulses used to produce Fig. 4.2 b). So, this FPM basis reduces the dimensionality of the search space for this problem to two dimensions. The reproducability was excellent in this basis set. Note that the error bars in Fig. 4.2 b) are much smaller than those in Fig. 4.2 a).

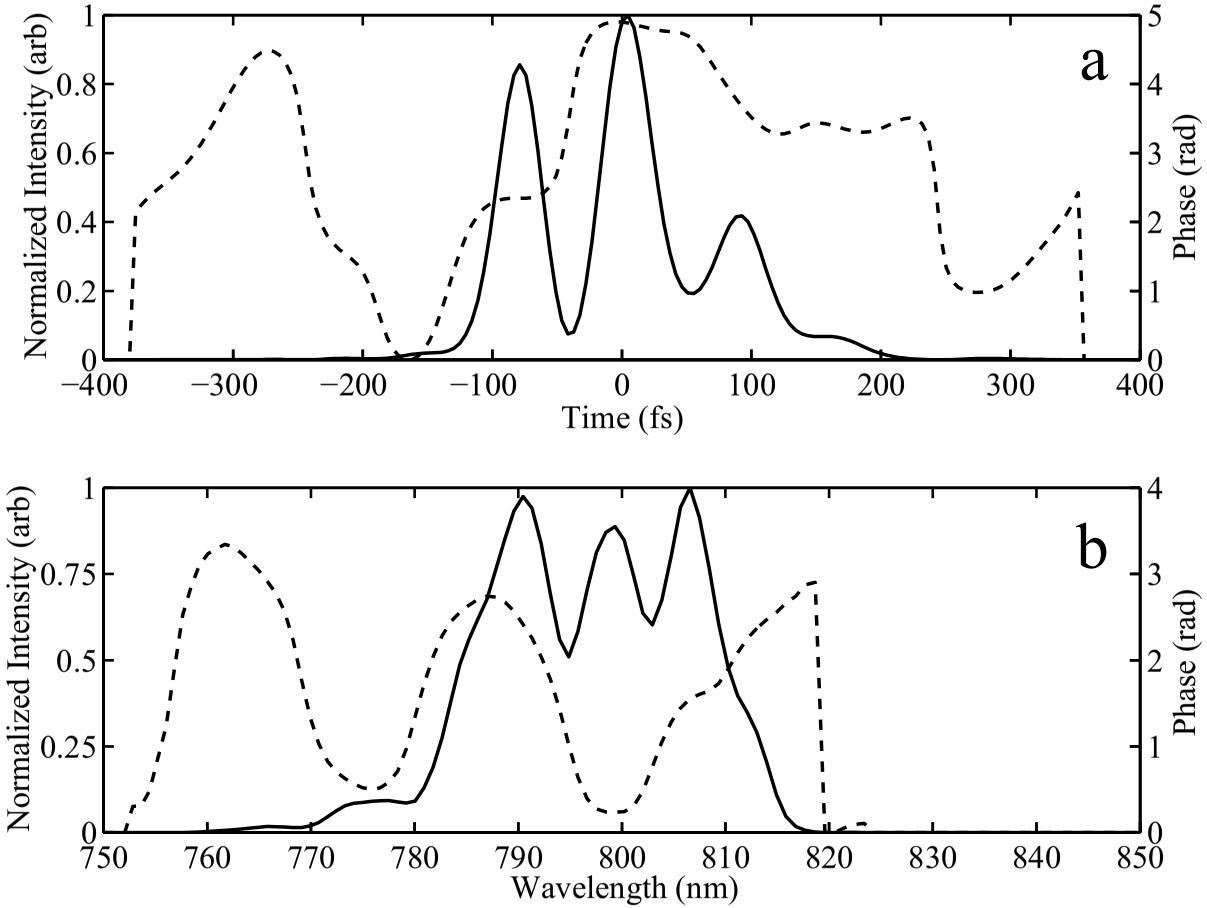

Figure 4.3 Reconstructed a) time and b) frequency representation of a solution when optimizing the CF_3^+ fragmentation channel of trifluoroacetone. The intensities are plotted with solid lines and the dashed lines show the phases. Note that there is an uncertainty of ≈ 20 nm in the absolute scale for the wavelengths shown in b).

4.2 Experiments on Trichloroacetone

The success in finding a good basis for trifluoroacetone led us to investigate a similar molecule. 1,1,1-Trichloroacetone (CH_3COCCl_3) is closely related to 1,1,1-trifluoroacetone; chlorine, the next halogen in the periodic table, substitutes fluorine.

In analogy to the experiments on trifluoroacetone we aimed for increasing the yield in CCl_3^+ at the expense of CH_3^+. Since the CCl_3^+ signal was much smaller than the CH_3^+ signal, we used a modified difference as fitness function: $F = C \cdot A - B$ where A was the CCl_3^+ and B the CH_3^+ signal. C was chosen to be between 5 and 10. This increases the emphasis on the objective of enhancing A relative to the objective of decreasing B.

Fig. 4.4 shows typical ion spectra for a) an optimized and b) an unshaped pulse. While unshaped pulses produced a CCl_3^+/CH_3^+ ratio of 0.086, optimized pulses increased this ratio to 0.55. The shaped pulses enhanced the total counts of the CCl_3^+ fragment and decreased the total counts of all other fragments. In contrast to our findings for trifluoroacetone, the CH_3CO^+ signal decreased for the optimized pulse shape. This is a strong argument for a control mechanism that is different from the one in trifluoroacetone (see discussion in sections 5.1 and 5.2).

When we ran the algorithm in the differential basis (see Eq. 3.5) we always

found solutions with a high number of nonzero genes. The cost functional was not able to decrease the dimensionality in this basis. However, given the fact that this molecule is similar to trifluoroacetone, we were hoping that the FPM basis (see Eq. 3.6) constitutes a good basis. Assuming that vibrational dynamics are the underlying physical mechanism for the control in both molecules, one would expect the characteristic timescales for trichloroacetone to be longer than the ones for trifluoroacetone. Due to the higher mass of the chlorine atoms, the vibrational frequencies of trichloroacetone are expected to be smaller than for trifluoroacetone. In the FPM basis, the genes with higher indices encode for lower frequencies of intensity modulations. When we ran the optimizations in the FPM basis and also applied the cost functional the GA repeatedly found low dimensional solutions where only 1 to 2 genes were needed. However, the positions of the important genes were not reproduced over different runs. Here, the change to the FPM basis provided reduction in dimensionality, although it did not provide reproducibility. It is important to note that there are cases where a particular basis is not perfectly suited but can still reveal some physical insight. For this problem, we found that the solutions were trains of pulses with a range of frequencies in $I(t)$ ($\approx$ 1.7 THz to $\approx$ 12 THz).

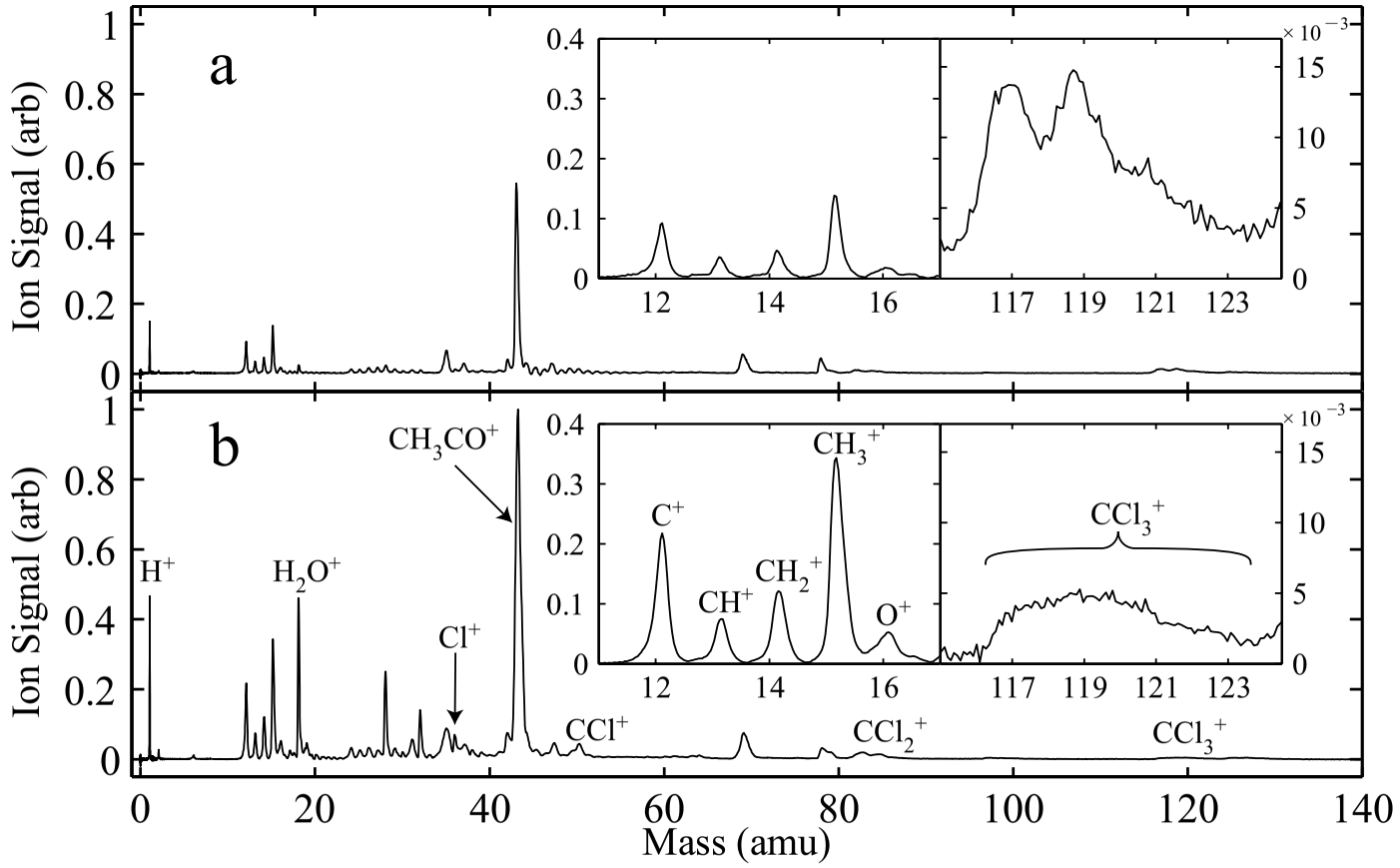

Figure 4.4 Trichloroacetone fragmentation spectra a) after learning control optimization and b) for an unshaped pulse. Both panels have the same absolute scale. Note that chlorine naturally occurs as two isotopes of masses 35 and 37 amu. Therefore all fragments which include Cl show isotope splitting. H_2O^+ and all the unlabelled peaks are background signals.

4.3 Experiments on Bromoiodomethane

The third molecule on which we performed fragmentation control experiments is bromoiodomethane (CH_2BrI). This molecule is the subject of a recent theoretical coherent control study and experimental nanosecond dissociative ionization experiments[25, 26, 27]. Ultrafast coherent control experiments have also been performed on a closely related molecule[28]. Here, our control goal was enhancing the breakage of the stronger C−Br bond while minimizing the breakage of the weaker C−I bond. As a fitness function we used a modified ratio $F = A/(A + B + C)$, where A was equal to the integrated signal from CH_2I^+ and B was equal to the integrated signal from CH_2Br^+. The small constant C is chosen to prevent noise from producing artificially high fitness increases at low signal levels. We also tried optimizing the simple ion signal difference ($F = A - B$). However, here we saw that solutions when optimizing differences did not always agree with solutions when optimizing ratios.

An ion spectrum produced by an unshaped laser pulse can be seen in Fig. 4.5 b). The ratio of the number of CH_2I^+ fragments to the number of CH_2Br^+ fragments is 0.11. After running the optimization algorithm we produced a ratio of 0.27. The spectrum corresponding to the optimal pulse is shown in Fig. 4.5 a).

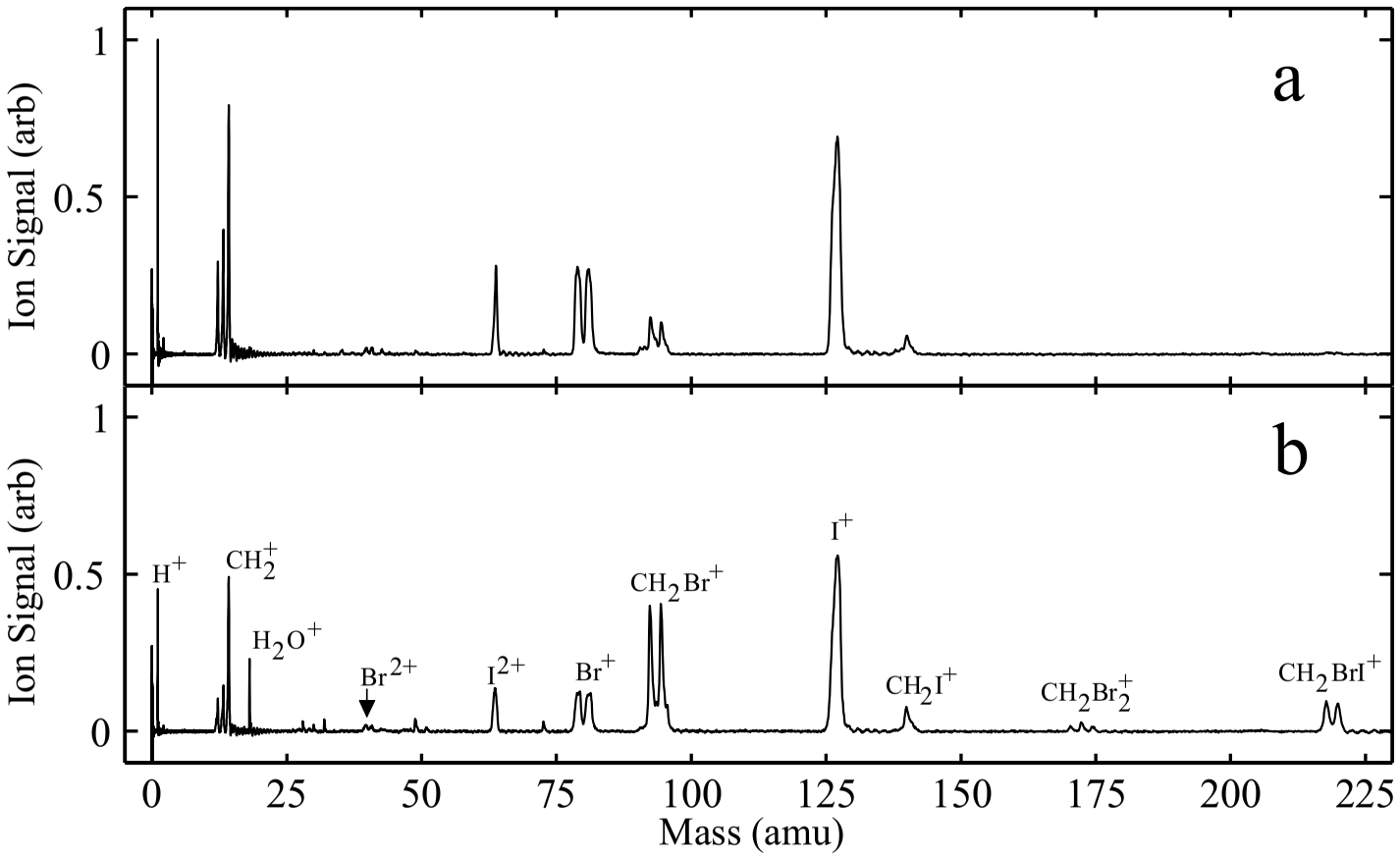

Figure 4.5 Bromoiodomethane fragmentation spectra a) after learning control optimization and b) for an unshaped pulse. Both panels have the same absolute scale. H_2O^+ and all the peaks of masses between 25 and 75 except for Br^{2+} and I^{2+} are background signals and seen before injection of the molecular sample.

Two representative solutions found in the simple difference basis (see Eq. 3.5) are plotted with different symbols in Fig. 4.6 a). The GA repeatedly found low dimensional solutions. This simple basis satisfies the criteria of low dimensionality. At least 95% of the fitness increase is due to few genes. When we performed the checks for small gene values being insignificant, we typically needed to keep 2 to 4 genes. However, solutions found during different GA optimizations do not agree and furthermore they correspond to separate optima in the search space. Fig. 4.7 illustrates this. We plot the CH_2I^+/CH_2Br^+ ratio during a scan between the two solutions shown in Fig. 4.6 a). These two solutions are represented as the vectors $\overrightarrow{G_1}$ and $\overrightarrow{G_2}$, respectively. The pulse shapes $\overrightarrow{G}(s)$ for this scan were calculated according to:

$$\overrightarrow{G}(s) = \overrightarrow{G_1} + s \cdot \left(\overrightarrow{G_2} - \overrightarrow{G_1}\right) \tag{4.1}$$

We varied the parameter s from 0 to 1 and thus scanned from the solution $\overrightarrow{G_1}$ to the solution $\overrightarrow{G_2}$. In between the two solutions the ion yield ratio dropped to less than 0.2. This corresponds to only 56% of the maximum increase for an optimized pulse. This drop allows us to infer that the two solutions are distinct optima in the search space rather than two points on a large shallow optimum. Averaging the two solutions would yield a very poor solution and would clearly lead to a misinterpretation of the control space.

Another basis with intuitive physical meaning is a polynomial expansion.

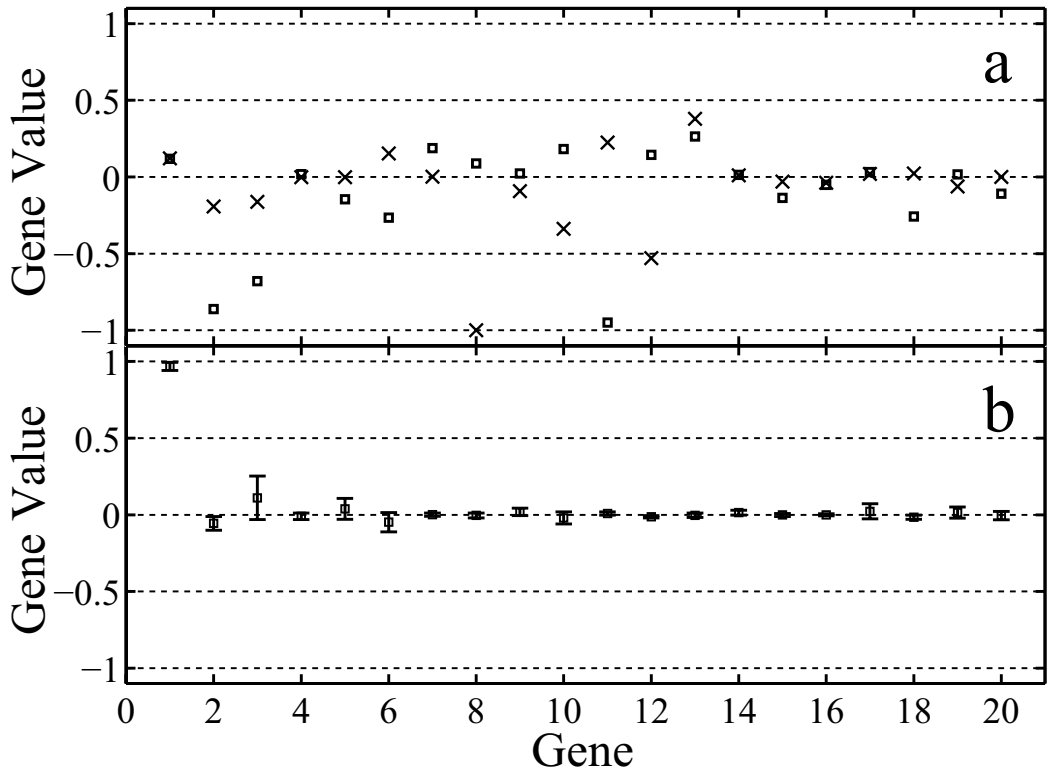

Figure 4.6 Solutions found when optimizing CH_2I^+ at the expense of CH_2Br^+ in bromoiodomethane. a) Two separate solutions found in the differential phase basis b) Averaged solutions found in the polynomial basis.

It is an ideal expansion for linear dispersion and provides for simple control over pulse duration. We expand the phase written to the pulse shaper about the center frequency ω_0:

$$\phi(\omega) = \sum_{i=2}^{N+2} G_i a_i (\omega - \omega_0)^i \tag{4.2}$$

N is the number of genes. The GA operates on the genes $\{G_i\}$ with values between -1 and 1. The coefficients a_i are determined by the pulse shaper

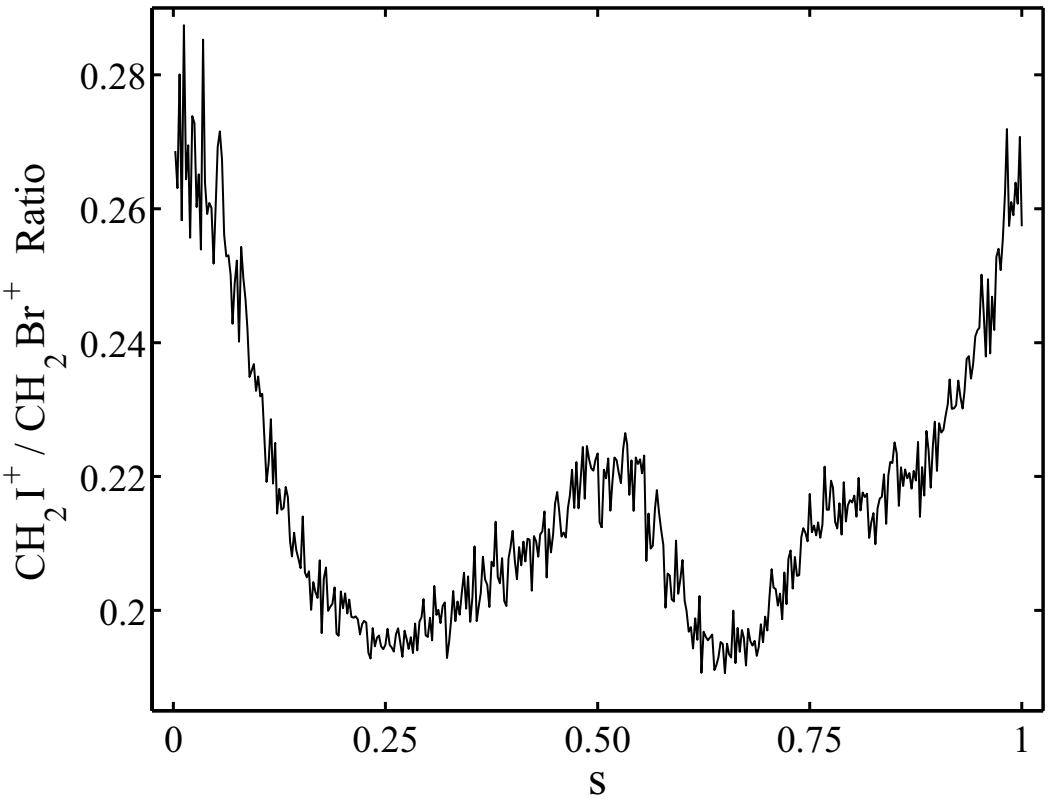

Figure 4.7 CH_2I^+/CH_2Br^+ ratio during a scan according to Eq. 4.1 from the solution marked with squares (at $s = 0$) to the solution marked with crosses (at $s = 1$) in Fig. 4.6 a).

resolution limits for each order i separately.

Fig. 4.6 b) illustrates the solutions found in the polynomial basis. We average over several solutions. The error bars are the standard deviations over the different optimizations. Mainly second order and some fourth order phase were required to optimize the ion ratio. We wrote pulse shapes to the AOM where we only kept the second order term and we still measured 95%

of the optimal increase in CH_2I^+/CH_2Br^+. Therefore this polynomial basis did extremely well at decreasing the dimensionality of the control problem. It also proved to be highly repeatable, justifying an average over several different solutions.

We used FROG to characterize the solutions found in the polynomial basis. Fig. 4.8 shows reconstructed a) time and b) frequency representations of an optimal pulse intensity and phase. The time duration (FWHM) is ≈ 1 ps. As expected, the quadratic phase term is by far the dominant pulse feature. The $I(t)$ is close to being Gaußian. The deviations are mainly due to a non-Gaußian spectrum $I(\omega)$.

We also used the FPM basis on the CH_2BrI control problem. We found 2 to 4 necessary genes. But solutions of different optimizations were not consistent. This illustrates that different molecules with different control mechanisms require different bases. The FPM basis allows for simple encoding of periodic "kicks". We are currently investigating the molecular dynamics of trifluoroacetone and on how control over the fragmentation of this molecule can be related to periodic intensity modulations. The first gene of the polynomial basis codes for linear frequency chirp and for pulse duration. As discussed below, low dimensionality in this basis leads to a different physical picture than in the case of the trifluoroacetone.

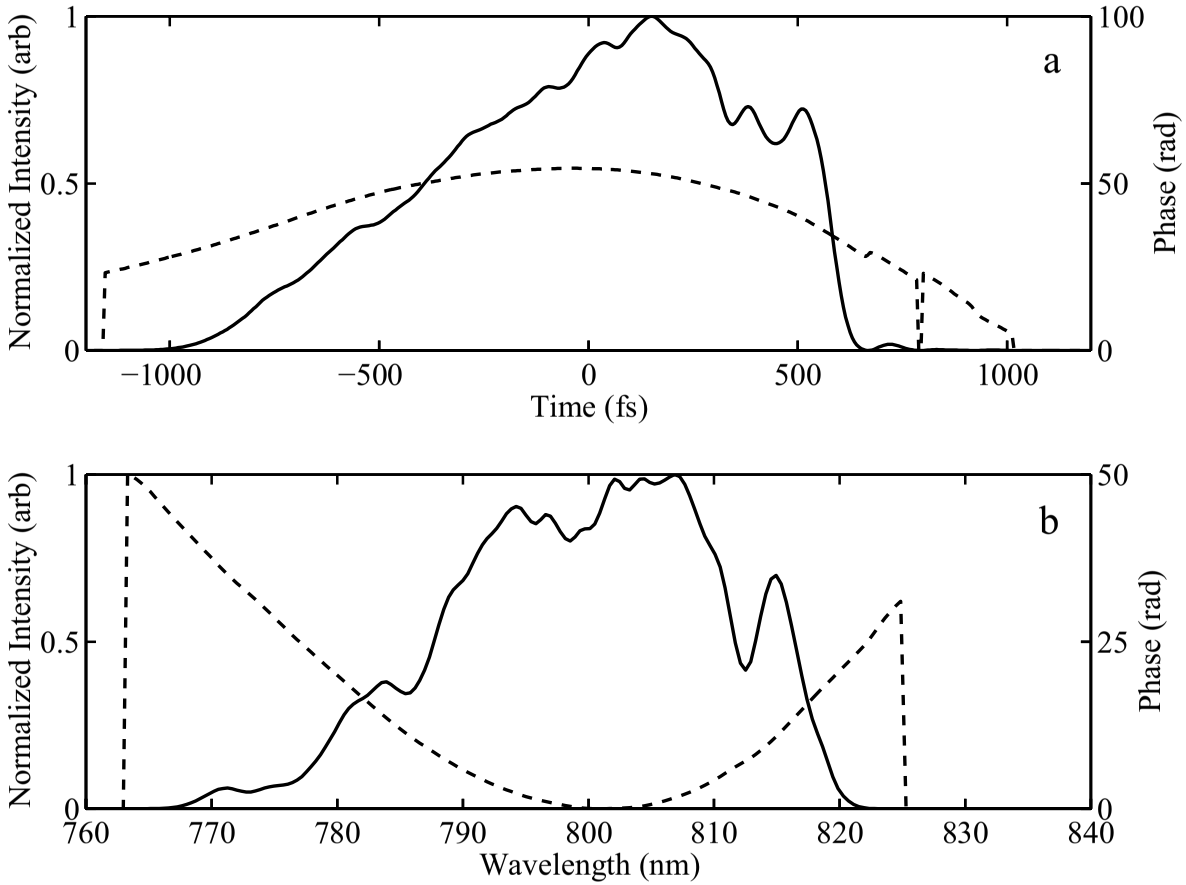

Figure 4.8 Reconstructed a) time and b) frequency representation of a solution when optimizing the CH_2I^+ fragmentation channel of bromoiodomethane. The intensities are plotted with solid lines and the dashed lines show the phases. Note that there is an uncertainty of $\approx$ 20 nm in the absolute scale for the wavelengths shown in b).

Chapter 5

Exploring Interesting Regions of the Search Space

One approach to understanding the control dynamics is to exhaustively map out the phase space of all possible pulse shapes. This approach is in general not feasible due to the exponential scaling of the phase space with dimension. If the full resolution of our pulse shaper is exploited (over 100 resolution points), and one uses a coarse 8 bit encoding for each phase and amplitude, then there are over $2 \times 2^{800} = 10^{241}$ pulse shapes to be tried. By lowering the dimensionality, the interesting regions of the phase space (those surrounding extrema and connecting the unshaped pulse with the solution in the fitness landscape) can be explored systematically. Here we present a few scans along essential vectors in representations where the dimensionality of the problem was shown to be greatly reduced.

5.1 Pump-Probe Experiments on CH_3COCF_3

For trifluoroacetone we found that the FPM basis, where the genes correspond to distinct periodicities in $I(t)$ of the laser pulse, reduced the dimensionality. This implies that the most important feature for control is the temporal spacing of a series of pulses. Thus, we examined the fragmentation as a function of time delay between a pair of transform-limited laser pulses. Fig. 5.1 shows the CF_3^+ fragment yield (heavy line) together with the $I(t)$ of an optimized pulse (thin line). The magnitude of the modulations in the ion signal proved to be very sensitive to the intensities of the pump and the probe pulses. We saw the most distinct features at focused intensities of 1.7×10^{14} W cm^{-2} for the pump and 1.4×10^{14} W cm^{-2} for the probe pulse. This is very close to the intensities where the control data was taken. The largest peak in the optimized pulse was 1.6×10^{14} W cm^{-2}. The only other fragment that showed correlated modulations was CH_3CO^+. The features in the optimal pulse match up with increases in the CF_3^+ signal.

The optimal pulse we show as thin line in Fig. 5.1 has a periodicity of $\approx$ 170 fs. This is in contrast to the solutions presented in Fig. 4.2 and Fig. 4.3 where the periodicity is $\approx$ 90 fs. As stated before, we were able to produce the solutions involving periodicities of $\approx$ 90 fs under constant conditions many times. However, under different conditions (e.g. different value for p in the

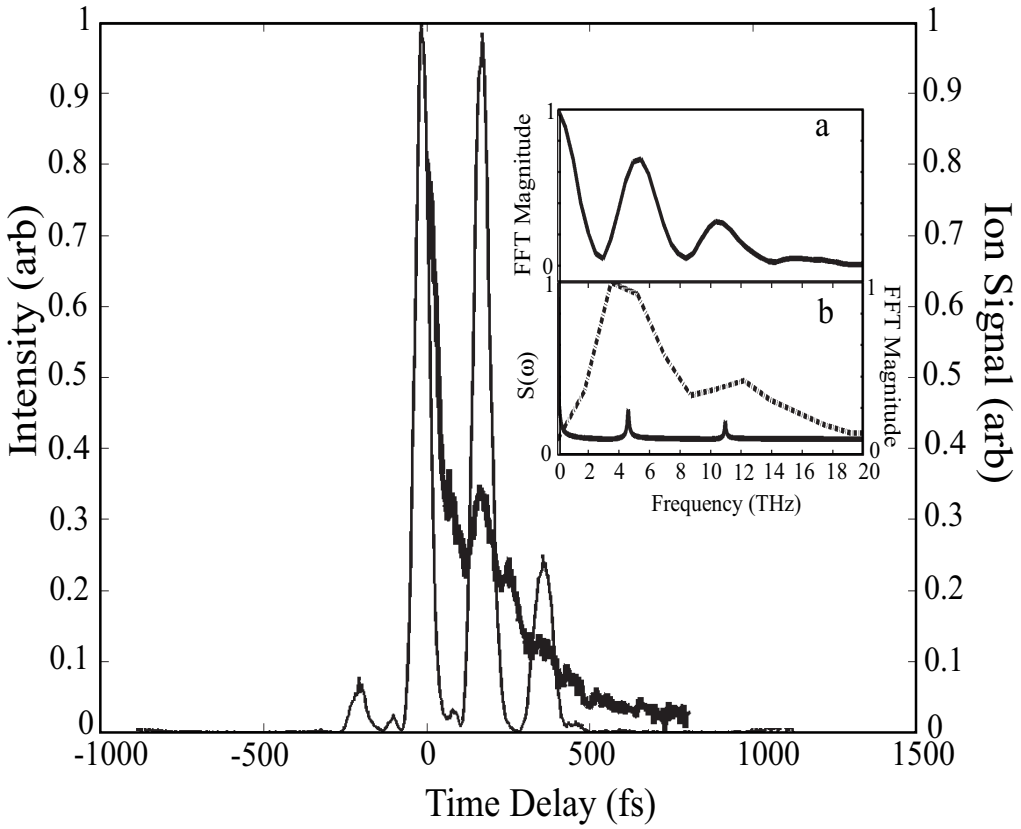

Figure 5.1 Ion signal of CF_3^+ during pump-probe scan (thick line) and optimized pulse $I(t)$ (thin line). In the inset: a) The Fourier transform of the $I(t)$. b) Solid line: Amplitude S of frequencies ω found in the pump-probe data by filter diagonalization analysis. Dotted line: The Fourier transform of the pump-probe data.

cost functional, different fitness function, and/or different laser spectrum), we also found some solutions similar to the one shown in Fig. 5.1.

In addition to the time domain representation of the pump-probe data and the optimal pulse shape, we show as insets the Fourier transform of the pump-probe data and the optimal pulse as well as a filter diagonalization spectrum[29] of the pump-probe data. The inset of Fig. 5.1 shows that the Fourier components of the $I(t)$ (inset a) of a solution agree with the frequencies found by filter diagonalization and Fourier transform of the pump-probe data (inset b). It is important to note that we found two frequency components in the pump-probe data. The higher frequency component at ≈ 11 THz is close to being the second harmonic of the lower frequency component at ≈ 5 THz. In order to optimize the CF_3^+ yield, the GA can make use of both, high and low frequency components, by matching the frequencies of the modulations in the shaped $I(t)$ to the frequencies of the molecular dynamics. Inset a) of Fig. 5.1 shows that this particular optimal pulse contains both Fourier components. The pulses discussed in section 4.1 (see also Fig. 4.3) show only frequency components at 11 THz.

We believe that the observed dynamics in the pump-probe data are vibrational dynamics on an ionic state PES. There are several reasons for this.

1. Preliminary electronic structure calculations show that the vibrational

frequencies for the neutral molecule are much faster than the ones we observe.

2. We account the prominent modulations in the pump-probe data to the excitation of a large wavepacket. Typically the amplitudes of such a wavepacket are much higher when launched by ionization than when launched by Raman scattering off a neutral molecule.

3. The intensities at which we took the pump-probe and the control data are larger than the saturation intensities for single ionization of molecules similar to trifluoroacetone[30].

Electronic structure calculations indicate that the C−C bond between CH_3CO and CF_3 is sufficiently weakened by ionization such that breaking of this bond is virtually assured. Therefore, control of the CF_3^+ signal is actually control over whether or not the CF_3 fragment carries charge. The natural fragmentation of the trifluoroacetone cation will lead to the production of CH_3CO^+ and CF_3. Our experimental data show no evidence for $CH_3COCF_3^+$ which is consistent with previous work[31]. One would expect, then, that the production of CF_3^+ arises from either 1) double ionization of the parent molecule, 2) control of fragmentation of the parent cation to yield CH_3CO and CF_3^+, or 3) subsequent ionization of the CF_3 fragment.

The pump-probe data show distinct modulation in the CF_3^+ signal which is indicative of control over the parent cation and rules out post-fragmentation ionization as an important mechanism. Preliminary structure calculations indicate that modulations in the pump-probe data are consistent with motion of the C−C bond between the CF_3 and the CH_3CO parts of the cation. We believe that the CF_3^+ production is influenced predominantly by stretching and bending motion of this C−C bond. Cooperation between vibrational wave packet motion along an ionic PES and the pulse structure can lead to enhanced CF_3^+ production through at least one of the two first mechanisms listed above.

Additional experimental evidence (see section 5.2) suggests double ionization of the parent molecule as the most likely path to production of CF_3^+. In this scenario we interpret the modulation of the CF_3^+ signal as indicating that wave packet motion along the C−C bond modulates the probability for double ionization of the parent molecule. There is little experimental evidence for the production of CF_3^+ via a single ionization mechanism. However, at this point, we cannot yet rule out the possibility. This mechanism requires the laser pulse to dress the ionic PES so as to allow the wave packet to tunnel to dissociation into CF_3^+ and other fragments[32]. We are continuing to investigate the extent each of these mechanisms accounts for the control.

5.2 Intensity Dependence of the CH_3COCF_3 Fragmentation

Our control experiments aimed for optimizing selective molecular fragmentation by means of shaping the phase of laser pulses. Pulse shaping always involves changing the intensity profile. A shaped pulse has a lower peak intensity than the corresponding unshaped (transform-limited) pulse. Thus it is essential to separately determine what role intensity plays in the control.

Fig. 5.2 shows integrated ion signals during a scan where we varied the intensity of transform-limited laser pulses. The thick line represents the CF_3^+ / CH_3^+ ratio which we tried to control in our learning loop optimization experiments. The graph illustrates that transform-limited pulses of low intensity ($\approx 1 \times 10^{14}$ W/cm^2) lower the CF_3^+/CH_3^+ ratio to ≈ 0.7. As mentioned above, also the solutions to this problem found by the GA are low intensity nearly transform-limited pulses. The CF_3^+/CH_3^+ ratio produced by these optimized pulses was 0.6. Therefore the main portion of this control lies in the intensity dependence of the fragmentation. On the other hand, this measurement of the intensity dependence indicates that maximizing the CF_3^+/CH_3^+ ratio is a more complex control problem. The maximum of CF_3^+/CH_3^+ is 1.2 in these measurements, whereas the GA was able to increase the ratio to 3.0. This proves

that the optimized pulse shapes achieve the control goal not by decreasing the intensity. Instead, the control is a genuine effect of pulse shape, i.e., the structure of the shaped pulse is responsible for control.

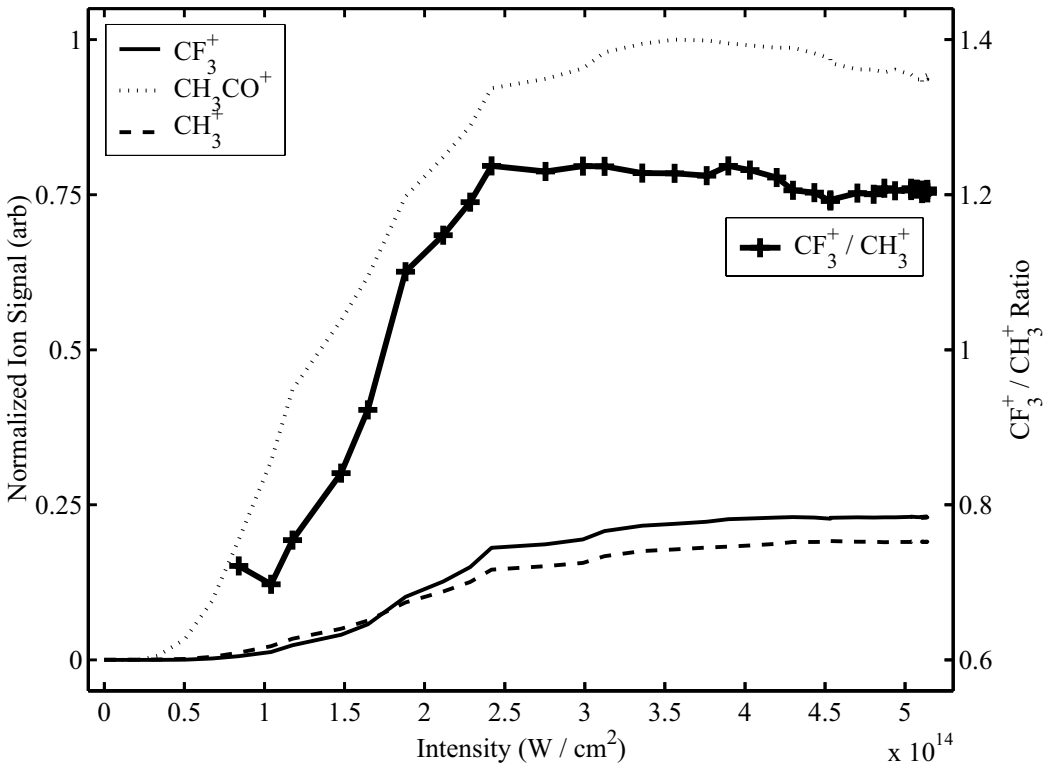

Figure 5.2 Integrated fragment ion signals of trifluoroacetone (thin lines) and CF_3^+/CH_3^+ ratio (thick line) as a function of intensity.

Fig. 5.2 also corroborates our interpretation of the control mechanism. The CH_3CO^+ ion appears at much lower intensities than the CF_3^+ and any other ion. This and the absence of a $CH_3COCF_3^+$ ion in all of our data indicates that a singly ionized molecule virtually always dissociates into a CH_3CO^+

ion and a neutral CF_3 fragment. CF_3^+ ion production only starts at higher intensities. An increase in CH_3CO^+ is correlated with the increase in CF_3^+. This is consistent with our interpretation of the CF_3^+ production as a result of double ionization. In this intensity scan, CF_3^+ is only produced at high intensities where double ionization of the parent molecule by a single pulse is likely. Additional data shows that the threshold intensity for the formation of C^{2+} is very close to the threshold intensity for the CF_3^+ formation. The occurrence of C^{2+} is a strong indication for double ionization. According to our interpretation, the double ion would then dissociate into CF_3^+ and CH_3CO^+ leading to a correlated increase in both signals, which we actually observe. Thus, these measurements provide strong evidence for the double ionization scheme mentioned in section 5.1.

5.3 Pump-Probe Experiments on CH_3COCCl_3

In our control experiments on trichloroacetone the FPM basis reduced the dimensionality. Thus, we also used pump-probe spectroscopy on this molecule and studied the fragmentation spectrum as a function of the time delay between two transform-limited pulses. We found some enhancement in the CCl_3^+ signal for time delays that were significantly longer than the pulse durations, i.e., where the pulses had no temporal overlap. We did not find distinct mod-

ulations in the data. Yet, the pump-probe data showed that there are some molecular dynamics caused by multiple pulses which can enhance the CCl_3^+ production. We believe that the GA makes use of these dynamics in the control experiments by producing trains of pulses. However, since there are no distinct frequencies in the CCl_3^+ signal, a range of frequencies solved the control problem and the GA was able to find many different solutions with only 1 or 2 nonzero genes.

5.4 Quadratic Phase Scan on CH_2BrI

For bromoiodomethane we found that a polynomial expansion is a suitable basis. Most of the control was captured by the first gene which encoded the quadratic term of the phase. An important cut through the search space is therefore a scan that varies the quadratic phase (or linear chirp[33]) of the laser pulse. The ratio of the ion signals of CH_2I^+ to CH_2Br^+ during such a scan is shown in Fig. 5.3. One can see that going from a transform-limited (unshaped) pulse to increasing chirp magnitude (positive and negative) the ratio increases monotonically. Note, however, that there is some asymmetry about zero chirp. The ion ratio goes higher for positive chirp than for negative chirp. This is consistent with the fact that the GA always found solutions with the first gene being at a large positive value.

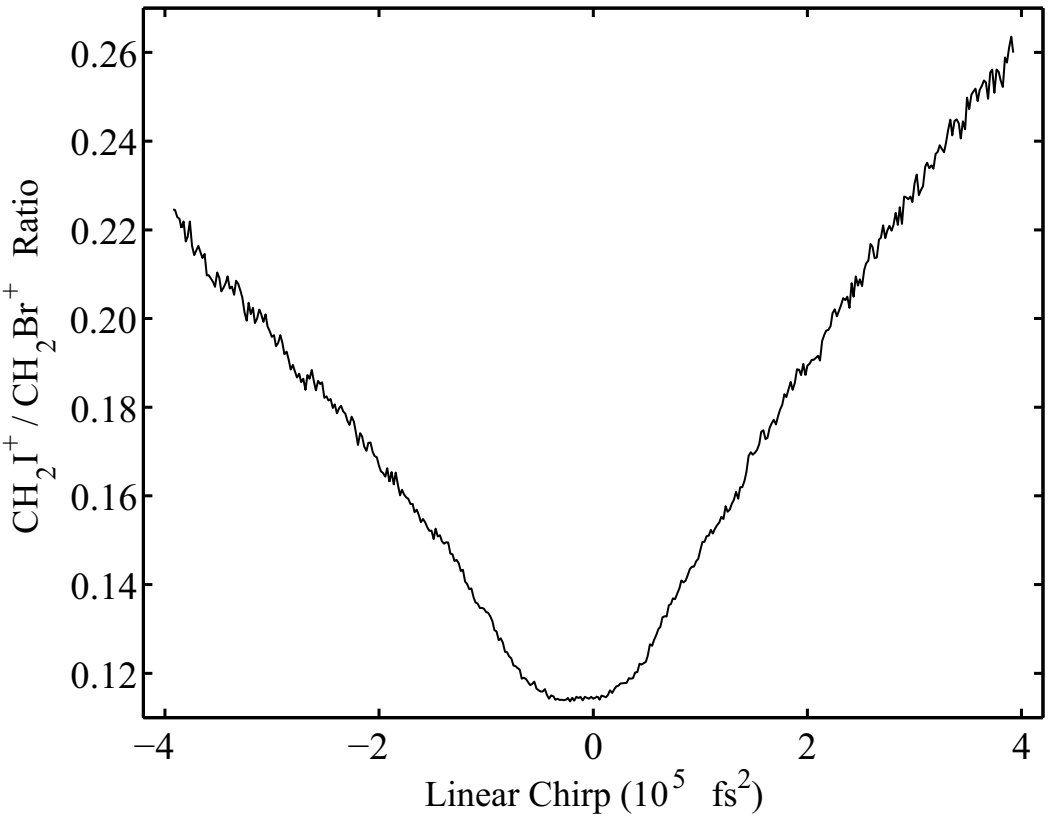

Figure 5.3 CH_2I^+/CH_2Br^+ ratio during a scan of linear chirp. We made use of the entire chirp range of our pulse shaper. The maximum chirp values correspond to pulse durations of 1 ps.

The observation of pulse duration as controlled by quadratic phase driving the branching ratio is similar to that of Itakura *et al.*[34]. They find that pulse duration can control the $CH_2OH^+/C_2H_4OH^+$ branching ratio in dissociative ionization experiments they performed on ethanol. Their measurements found that quadratic phase could drive the ratio from 0.12 to 0.45 when varying the pulse duration from 32 fs to 760 fs. They interpret their findings in terms of

the holding time of a vibrational wave packet on the ionic PES dressed by the light field. The dressing of the molecular PES by the light field steers the wave packet toward the CH_2OH^+ channel and therefore the longer the PES is dressed by the light fields, the higher the branching ratio they measured. Our observations are quite similar, although there is an asymmetry in the branching ratio with respect to the sign of the chirp. This is an indication that the control is not completely explained by the pulse duration. It is also not certain that the wave packet motion in the case of the CH_2BrI is on an ionic surface. The details of the control mechanism are still being investigated, but what is important for this thesis is that the reduction of the dimensionality in the CH_2I^+/CH_2Br^+ control allows for a low dimensional parameter scan and points in the direction of an established physical picture.

5.5 Intensity Dependence of the CH_2BrI Fragmentation

To ensure that the control over the CH_2I^+/CH_2Br^+ ratio is not dominated by the decreasing intensity that goes along with increasing chirp, we also scanned the fragmentation spectrum as a function of intensity for an unshaped pulse. Fig. 5.4 shows integrated ion signals of such a scan. Where the signals of

CH_2I^+ and CH_2Br^+ are sufficiently above the noise level we also plot the CH_2I^+/CH_2Br^+ ratio. The maximum of this ratio is ≈ 0.12 in these measurements, whereas we were able to increase this number to 0.27 by shaping the laser pulses. Therefore, the control we achieved in the experiments on bromoiodomethane is not a simple intensity effect.

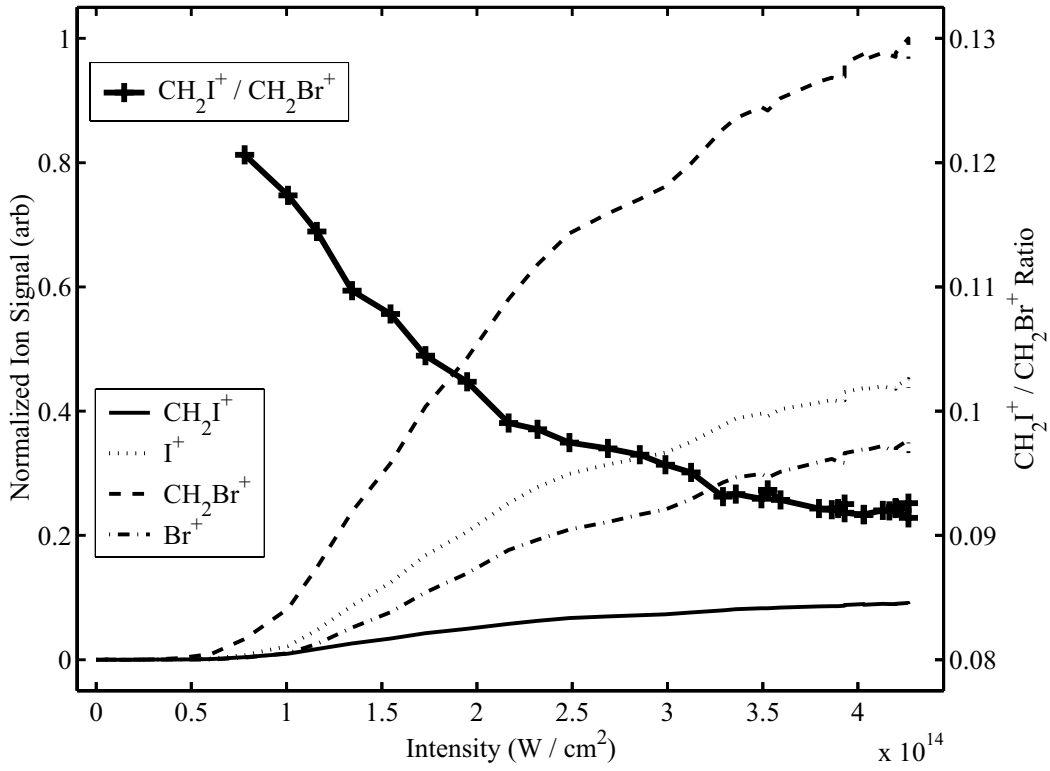

Figure 5.4 Integrated fragment ion signals of bromoiodomethane (thin lines) and CH_2I^+/CH_2Br^+ ratio (thick line) as a function of intensity.

Chapter 6

Conclusions

The focus of this thesis is on developing tools for systematically gaining insight into the physical mechanisms underlying shaped pulse control of molecular fragmentation.

We argue that separating necessary from sufficient information in optimal pulse shapes is an important first step in interpreting solutions found by the learning algorithm. We have shown that when using a GA, one cannot in general use genetic variation as an identifier for necessary information. A cost functional however, can be very effective at isolating necessary information if the number of essential dimensions is less than the number of available ones.

Furthermore, we have shown that a nonlinear change of basis can reduce the dimensionality of the search space to a few degrees of freedom. Ultimately one would like to automate the selection of new bases and search for optimal

bases in a similar way that the GA searches for optimal pulses. What we have shown here is that there are simple bases which can capture the essential features of effective control pulses in one or two dimensions. This is in some sense surprising because of the complexity of the molecular fragmentation control problems discussed here. Our ability to reduce dimensionality with simple basis changes highlights the fact that complexity is not inherent to the problem and a change of representation can reformulate the problem in a much simpler form.

The reduction in dimensionality of the control problems allowed us to scan the control goal and construct a control surface in analogy to a PES for a polyatomic molecule. This helped us to gain insight into the physical control mechanism. Ongoing efforts focus on molecular structure calculations for trifluoroacetone with the goal of comparing them to the experimental results presented here.

Bibliography

[1] N.A. Papadogiannis, B. Witzel, C. Kalpouzos, and D. Charalambidis, Phys. Rev. Lett. **83**, 4289 (1999).

[2] U. Keller, Nature **424**, 831 (2003).

[3] R.S. Judson, and H. Rabitz, Phys. Rev. Lett. **68**, 1500 (1992).

[4] H.A. Rabitz, M.M. Hsieh, and C.M. Rosenthal, Science **303**, 1998 (2004).

[5] M. Dantus, and V.V. Lozovoy, Chem. Rev. **104**, 1813 (2004).

[6] T.C. Weinacht, R. Bartels, S. Backus, P.H. Bucksbaum, B. Pearson, J.M. Geremia, H. Rabitz, H.C. Kapteyn, and M.M. Murnane, Chem. Phys. Lett. **344**, 333 (2001).

[7] A. Flettner, T. Pfeifer, D. Walter, C. Winterfeldt, C. Spielmann, and G. Gerber, Appl. Phys. B **77**, 747 (2003).

[8] R. Bartels, S. Backus, E. Zeek, L. Misoguti, G. Vdovin, I.P. Christov, M.M. Murnane, and H.C. Kapteyn, Nature **406**, 164 (2000).

[9] J.L. Herek, W. Wohlleben, R.J. Cogdell, D. Zeidler, and M. Motzkus, Nature **417**, 533 (2002).

[10] R.N. Zare, Science **279**, 1875 (1998).

[11] N.E. Henriksen, Chem. Soc. Rev. **31**, 37 (2002) and references therein.

[12] T. Brixner, N. H. Damrauer, G. Krampert, P. Niklaus, and G. Gerber, J. Opt. Soc. Am. B **20**, 878 (2003).

[13] C. Daniel, J. Full, L. González, C. Lupulescu, J. Manz, A. Merli, Š. Vajda, and L. Wöste, Science **299**, 536 (2003).

[14] T. Hornung, R. Meier, and M. Motzkus, Chem. Phys. Lett. **326**, 445 (2000).

[15] J.M. Geremia, and H. Rabitz, Phys. Rev. Lett. **89**, 263902 (2002).

[16] S. Backus, C.G. Durfee, M.M. Murnane, and H.C. Kapteyn, Rev. Sci. Instrum. **69**, 1207 (1998).

[17] P. Nürnberger, Master's Thesis, Stony Brook University (2003).

[18] M.A. Dugan, J.X. Tull, and W.S. Warren, J. Opt. Soc. Am. B **14**, 2348 (1997).

[19] L. Davis, *Handbook of Genetic Algorithms*, (Van Nostrand Reinhold, New York, 1991).

[20] Y. Ohtsuki, G. Turinici, and H. Rabitz, J. Chem. Phys. **120**, 5509 (2004).

[21] J.M. Geremia, E. Weiss, and H. Rabitz, Chem. Phys. **267**, 209 (2001).

[22] P.H. Bucksbaum, Nature **396**, 217 (1998).

[23] S.A. Malinovskaya, P.H. Bucksbaum, and P.R. Berman, Phys. Rev. A **69**, 013801 (2004).

[24] R. Trebino, K.W. DeLong, D.N. Fittinghoff, J.N. Sweetser, M.A. Krumbugel, B.A. Richman, and D.J. Kane, Rev. Sci. Instrum. **68**, 3277 (1997).

[25] T. Takayanagi, and A. Yokoyama, Bull. Chem. Soc. Jpn **68**, 2225 (1995).

[26] D.G. Abrashkevich, M. Shapiro, and P. Brumer, J. Chem. Phys. **116**, 5584 (2002).

[27] L.J. Butler, E.J. Hintsa, S.F. Shane, and Y.T. Lee, J. Chem. Phys. **86**, 2051 (1987).

[28] N.H. Damrauer, C. Dietl, G. Krampert, S.H. Lee, K.H. Jung, and G. Gerber, Eur. Phys. J. D **20**, 71 (2002).

[29] D. Neuhauser, J. Chem. Phys. **93**, 2611 (1990).

[30] S.M. Hankin, D.M. Villeneuve, P.B. Corkum, and D.M. Rayner, Phys. Rev. A **64**, 013405 (2001).

[31] S. Anand, M.M. Zamari, G. Menkir, R.J. Levis, and H.B. Schlegel, J. Chem. Phys. A **108**, 3162 (2004).

[32] X.P. Tang, S.F. Wang, M.E. Elshakre, L.R. Gao, Y.L Wang, H.F. Wang, and F.A. Kong, J. Phys. Chem. A **107**, 13 (2003).

[33] R. Trebino, *Ultrafast Optics Files*, http://www.physics.gatech.edu/gcuo /UltrafastOptics/PhysicalOptics/OpticsI16.ppt

[34] R. Itakura, K. Yamanouchi, T. Tanabe, T. Okamoto, and F. Kannari, J. Chem. Phys. **119**, 4179 (2003).