

CHAPTER II

Coherent control of quantum systems with learning algorithms

2.1 Introduction

Since the advent of the laser in the early 60's, researchers have sought to use laser light to solve problems in virtually every scientific discipline. Chemists, in particular, looked upon this intense single-frequency light source as a way to manipulate chemical reactions. The initial idea was simple: simply tune a laser source to the characteristic frequency of a bond you wish to break and turn up the intensity of the laser. The molecule should snap apart and you could use those products to drive and control chemical reactions. However, in the laboratory this idea was a complete failure. Energy that was dumped into specific bonds redistributed itself throughout the molecule on femtosecond to picosecond timescales [25]. Increasing the energy only broke the weakest bond in the molecule and no control seemed possible.

Thus, the dream of manipulating matter with light pulses soured and research interest dwindled. This early work ignored a significant aspect of the molecules they sought to control. Molecules are quantum mechanical systems and their normal

modes are described by waves that can interfere [1, 2]. Shaped light fields can control this interference and thereby control the final state of the system (as illustrated in Figure 2.1); this approach has been dubbed "coherent control". Coherent Control techniques have evolved from focusing solely on chemical systems to a wide range of quantum mechanical systems, and have been applied to semiconductor systems [26, 27, 28, 29, 30, 31], terahertz radiation sources [32, 33, 34], shaping of Rydberg wave functions [6], and single atoms [11].

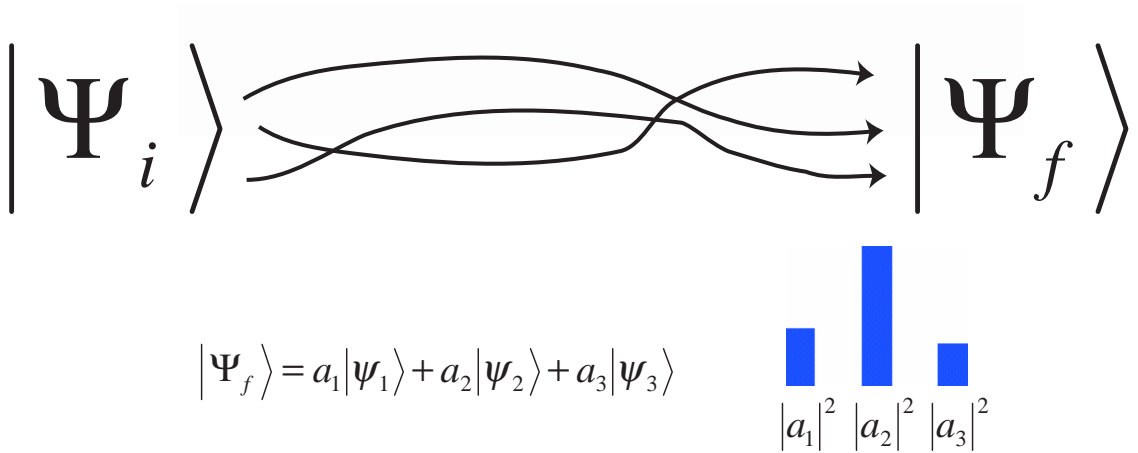


Figure 2.1: Coherent control is accomplished by manipulating interfering pathways in a quantum mechanical system.

2.1.1 A brief history of coherent control

Two major approaches to the control of quantum mechanical systems have been proposed and developed. Both of these ideas exploit wave interferences in quantum mechanical systems and highlight different aspects of this mechanism. The first proposal by Brumer and Shapiro [35] excites two pathways that interfere constructively or destructively depending on the relative phase of two CW lasers. The probability of forming a given product or exciting a particular state depends on the coherent

sum of the two states since they are indistinguishable.

This technique has been utilized by Elliot to control transitions in mercury [36] and thoroughly studied in Robert Gordon's group to control the ionization of HCl, CO, and H_2S [37], photoexcitation of HI [38], and controlling product ratios of the photodissociation of H_2S [39] and CH_3I [40]. Shnitma et al. demonstrated the control of branching ratios in the photodissociation of Na_2 [41], while Dupont et al. demonstrated the control of photoexcited electrons in semiconductors for applications in fast switching [27].

Although this approach has been quite successful, its efficiency is limited because quasi-CW lasers act on only a fraction of the thermal distribution of atoms or molecules in which control is sought. For example, at 1 K, the thermal energy kT is ~ 20 GHz, which is larger than a ns pulse bandwidth. Energy redistribution relies on relatively slow processes such as collisions to restore depleted populations. Furthermore, the decay of coherences limits the amount of energy that can be effectively used for control because these coherences are required for stable wave interference.

An alternative approach proposed by Tannor and Rice suggested using pairs of short pulses to manipulate quantum mechanical wave packets [42, 43]. More generally, this can be viewed as optimal control in which some optimally shaped electromagnetic field is used to sculpt a desired final quantum mechanical state [44]. This concept is illustrated in Figure 2.2. The pulses used in this approach are shorter than the energy redistribution and coherence decay times in molecules, making it highly efficient. A short pulses excites vibrational wave packets that evolve in a

field-free manner, and that achieve a desired target goal at some specified time.

There are generally three types of control that are grouped in this control scheme. The first utilizes pulse pairs that excite two time-delayed quantum mechanical wave packets that interfere to excite specific states in the system. The second, called pump-dump, excites a wave packet on an excited state surface, then a second pulse "dumps" that population to a ground-state level, or some dissociative product channel. The final, more general, approach is to determine a single, shaped optical field that creates an optimal quantum mechanical state.

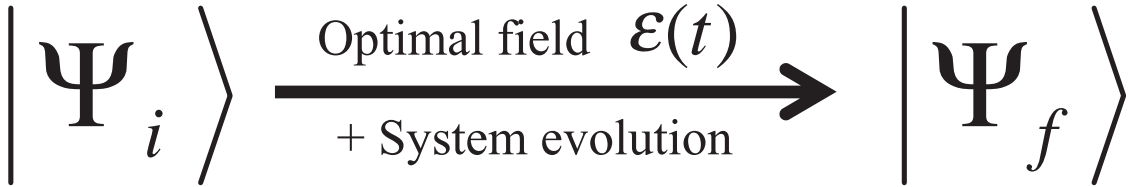


Figure 2.2: Optimal control is a variation of coherent control that uses a shaped optical field that drives a quantum system from some initial state into a desired final state.

Coherent control with pairs of unshaped phase-locked pulse pairs has been demonstrated in a host of molecular [45, 46, 47, 48, 49, 50, 51, 52, 53] and semiconductor [28, 29, 30, 34, 31] systems. Optimally shaped laser fields offer the most versatile and general approach to the control of quantum mechanical systems. The optimal field guides the system from some initial state to the desired final state by working in conjunction with the system evolution. The capabilities of this approach can be extended in the high-field regime as the optimal field itself begins to substantially modify the potential of the system under control and allows for more complex control [10].

The calculation and shaping of optimal fields for control has been demonstrated for selective control of optical phonons in cryogenic solids [4], the control of excited state vibrational wave packets [54, 55], controlling the yield of a chemical reaction [56], the control of 2-photon absorption [5, 57, 58], and for the control of molecular vibrations with overtone excitation [21].

Determining the optimal driving field for controlling quantum systems requires detailed knowledge of the Hamiltonian of the system to be manipulated, ample computer time to calculate the field (where all experimental conditions must be known exactly). The primary difficulty here is that the Hamiltonian for most systems is unknown, particularly for complex chemical systems. Furthermore, it might be difficult and time-consuming to calculate the optimal field given the Hamiltonian. Once found, there is no guarantee that the optimal field will be robust and the control may be very poor on average as the experimental conditions fluctuate. Finally, even if we have the exact optimal field, it can be difficult to reliably deliver the optimal field to the quantum system to be controlled.

2.2 Learning algorithms for coherent control of quantum systems

The difficulty of calculating the optimal field for controlling a quantum system gets progressively more difficult as the system gets more complex. However, a clever approach proposed by Rabitz et al. [3, 59, 2], demonstrated through computational simulations that trial-and-error learning algorithms can in principle be applied to optimally control quantum systems. This approach essentially uses the quantum

system under control as an analog computer, testing each trial solution, and guiding the laser system to discover the optimal field. This approach solves the basic problems of optimal control discussed above. The quantum system knows its own Hamiltonian, and no computation time is required because the system solves for the solution essentially instantly. The solution found is guaranteed to be robust because the control is achieved in a noisy laboratory environment where the experimental parameters may be fluctuating. Finally, since the control field is found *in-situ*, the correct optimal field is automatically delivered to the quantum system under control.

Figure 2.3 illustrates the basic concept of a learning loop. The learning algorithm tests large groups of trial pulse shapes on the quantum system and evaluates how well the pulse shape allows the system to evolve to the desired outcome. The learning algorithm then uses the results of the experiments to determine new pulse shapes to try, and repeats the loop until the system is in the target quantum state. As the learning algorithm iterates through each loop, it learns to control the quantum system better until eventually the algorithm "teaches" the laser how to control the quantum system.

2.3 An overview of learning algorithms

The type of learning algorithm employed to control quantum systems is typically an evolutionary algorithm. These algorithms are so named because the inspiration for their design is based loosely on the principles of biological evolution. The basic steps of reproduction, natural selection, and diversity by variation are all used in a typical evolutionary algorithm. The reproduction stage mixes genetic information

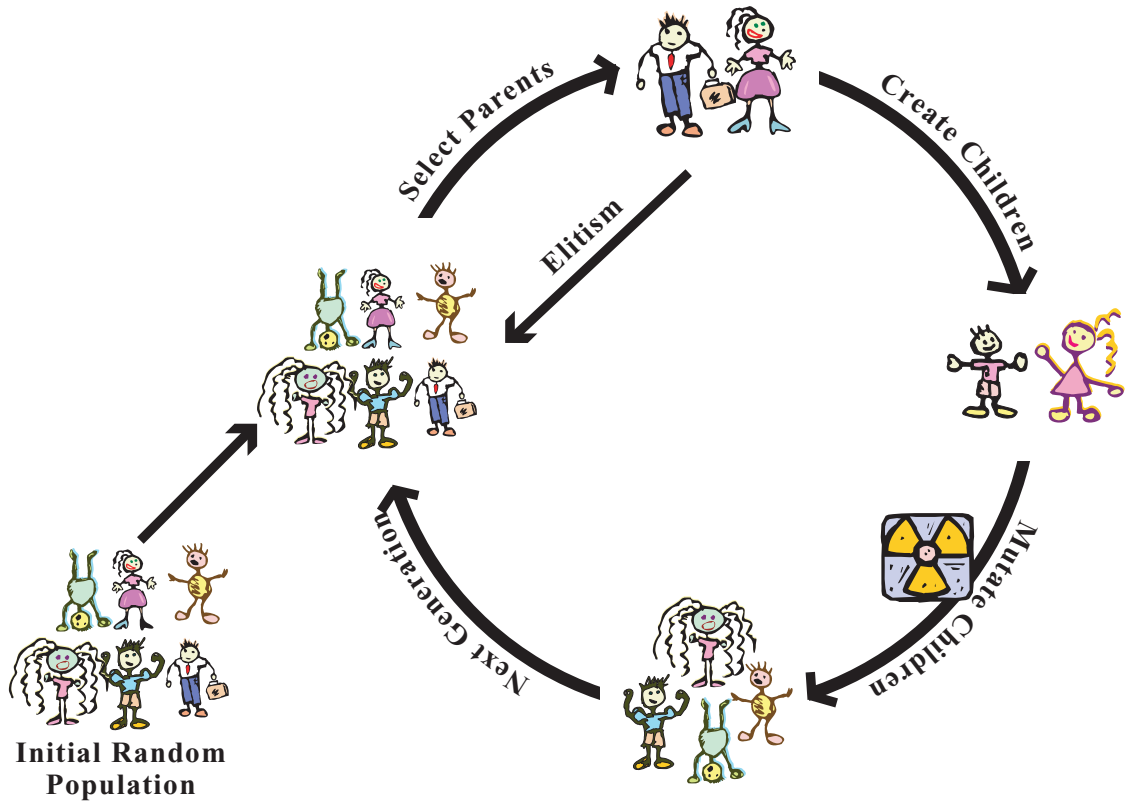


Figure 2.3: An evolutionary algorithm can be used to teach a laser system to control a quantum system. The algorithm begins with a random population. Each member of the population is a trial pulse shape; for our system, a trial pulse shape is represented by the set of voltage levels applied that distort a deformable mirror. Each trial pulse shape is tested in the system under control, and then a fitness value is calculated that quantifies the quality of the control. A fraction of the fittest members of the population are duplicated, forming the set of children. Those children are mutated so that the solutions are perturbed. The parents and mutated children are combined and the process repeats. After a number of iterations, the algorithm converges to an optimal control field (the best trial pulse shape).

from two members of a population to produce a "child" (new trial solution) with a combination of traits from the two parents. Natural selection is the process that chooses individuals that propagate to the population that forms the next generation. In nature, this is a combination of the fitness (or degree to which an individual has successfully adapted to his environment) and chance that determines what genetic information is passed to the next generation. Finally, there is diversity in genetic information due to the variation in members of a population. Mutation of genetic information is the primary way that new genetic information is introduced into a population.

2.3.1 Terminology of Evolutionary Algorithms

Before delving into a discussion of evolutionary algorithms (EA), it is important to define the components of the algorithms. The DNA gives the genetic information of a member of the population, which can carry either phenotypic (characteristics of the individual) or genetic (a coded form of the phenotype) information. The type of information contained in the DNA varies according to the type of EA used. A population is a set of members, which are described by their DNA. The fitness of an individual is a measure of how well the individual has adapted to the environment. For the purposes of controlling a quantum system, the fitness is determined by evaluation some observable of the quantum system. This fitness function can be viewed as a mountainous landscape (i.e., fitness landscape) and a "good" solution corresponds to a tall mountain peak. There is may be a global optimum, which is simply the tallest peak in the landscape. The process of selection determines which

members of the population are carried over to the next generation. It follows that selection also determines what information from the previous experiments is retained by future generations. Recombination and mutation operations serve to perturb the DNA content of members of the population in order to explore the parameter space of the fitness landscape and avoid a local minima.

2.3.2 An overview of evolutionary computation

The foundation for current evolutionary algorithms come from work done just as computers began to emerge as a tool in large research facilities. However, the low computational speed and memory of these machines hampered this early work. The idea of automatic programming was first presented in the work of Friedberg [60] at IBM. The idea was to use a limited set of instructions from which a computer algorithm could select to construct a program that would convert an input to a desired output. This approach used no selection mechanism and turned out to do worse than a pure random search. The next attempt was by Bremermann [61] to perform evolutionary optimization. He sought to optimize functions by recombination and mutation. However, his algorithm did not converge well and was largely ignored. Both automatic programming and evolutionary optimization, although flawed, provided the basis for future developments in evolutionary algorithms.

Using the lessons of Friedberg and Bremermann, Fogel tried another approach he called "evolutionary programming" (EP) [62]. This new approach used selective pressure to push a population towards a solution. The original implementation compared a parent and a child (a mutated version of the parent) and kept the better

one for the next generation. He later expanded the number of members to form real populations and introduced recombination operators. An EP represents the DNA (or set of information that describes the trial solution as implemented in a learning system) as simply the control knobs (i.e., parameters of the system that we control) that manipulate the device that produces the trial solution we wish to analyze in our apparatus. The DNA elements are mutated with a normally distributed random variable weighted by a scaled fitness value. Recombination is not used in EP because it is seen as secondary to mutation [62]. The members of the population that form the next generation are taken from the union of the set of parents and children from a tournament (tournament selection operator). This work was ignored until the early 70's when genetic algorithms and evolutionary strategies were developed independently in the US and Germany.

Evolutionary Strategies (ES) were developed by Bienert, Rechenberg, and Schwefel at the Technical University of Berlin for the optimization of fluid dynamics problems (e.g., optimization of nozzle shapes) [63]. The initial implementation compared a parent (μ) and a child (λ) in a manner similar to EP. Later versions made use of large populations including both parents and children ($\mu + \lambda$), referred to as *elitism*, or only the children made by the parents (μ, λ). Schwefel also introduced the self-adaptation of algorithm operators so that the operators can be optimally adjusted throughout the optimization [64].

The DNA (or set of variables that produces a trial solution) in an ES simply uses the control knobs directly just as in the EP. The mutation operator adds a

normally distributed random variable, with some variance, to each component of the DNA. The variance is also modified on each iteration so that the mutation rate self-adapts. A recombination operator then splits the DNA into pieces and exchanges information. This recombination operator has been viewed by some as a macro-mutation operator. The members of the population of the next generation are then simply the set of children, or the combined set of children and parents in the case where elitism is included. The optimal convergence rate has been found to correspond to a ratio of children to parents equal to seven ($\frac{\mu}{\lambda} = 7$) [63]. There is no theoretical optimum for the number of parents; however, setting μ too small results in a path oriented search, whereas a larger μ takes advantage of a population, but consumes more computation and experimental time.

John Holland, working independently at the University of Michigan, developed the genetic algorithm (GA) in parallel with the ES development in Germany [63]. The DNA in a GA is not simply the control knobs, but a transformed version of the control knobs (or variables we apply to our "solution generator apparatus"). Instead, each variable in the DNA is transformed into a new representation. The most commonly used representation for DNA in a GA is to convert the control knobs to their binary numeric representation. For example, if the "control knobs" that would be used in an evolutionary program or an evolutionary strategy were the set $\{2, 4, 9, 3, 4\}$, then in a 4-bit binary encoding, the DNA would be $\{00100100100100110100\}$. Due to the binary nature of the representation, a Gaussian mutation operator is no longer appropriate, and as a result, the mutation operator used in a GA is random bit flips.

The mutation operator is seen by Holland as secondary, so the mutation probability is set to be small in a typical GA implementation.

Recombination on the other hand, is quite evolved in a GA. The initial operator was simply a one-point crossover, similar to ES, where the genetic information is exchanged between two parents at a single point. The obvious extension of this is to exchange several blocks of DNA (multipoint crossover) by splicing the DNA at several locations. More advanced versions of these recombination operators have been made to self-adapt. The simplest, segmented crossover, is multipoint crossover with an adaptable number of crossover points [65]; they also used shuffle crossover which mixes up the ordering of the DNA blocks. The most advanced recombination operator, punctuated crossover, dynamically varies the number and locations of crossover points [66]. Instead of choosing the member for the next generation to be only the top performers, the members are selected by a roulette wheel selection. The idea is that each member of a population (trial solution) is assigned a probability that is proportional to its fitness value. A fraction of the current population is chosen to proceed to the next generation by a random number generator that favors more "fit" solutions that have larger fitness values.

2.4 Experimental demonstrations of Learning Control

A number of experiments have recently demonstrated the use of shaped pulses for control of quantum systems with a learning algorithm. Bardeen et al. [67] demonstrated that a learning algorithm can determine that a pulse with positive chirp is optimally effective in avoiding saturation of a molecular transition and the control

in I_2 in both the gas and solid phases. Gerber et al. [9] demonstrated that molecular dissociation could be controlled through the use of pulses with a complex shape determined through a learning algorithm. Bucksbaum et al. [6] demonstrated the use of iterative algorithms to "sculpt" Rydberg atom wavefunctions into the desired configuration, and also to control Stokes scattering in molecular systems. Weinacht et al. demonstrated the control of Raman scattering in liquids [68]. Leone et al. [8] demonstrated the time-shifting of rotational wave packet dynamics in Li_2 . These experiments represent systems at the two extremes of complexity. In the case of one-and two-photon absorption or molecular excitation, the physical reasons behind the optimum solutions are straightforward to understand. In the case of vibrational excitation or dissociation of polyatomic molecules, the pulse shapes obtained through optimization are complex and extremely difficult to interpret.

In contrast, as will be discussed in Chapter 3, the case of high-harmonic generation represents a quantum process that is highly nonlinear, but that nevertheless has proven to be both accessible to experiment and theoretically tractable. The optimal laser pulse for coherent x-ray generation can be explained as a new type of "intra-atomic" phase matching [13], that enhances the constructive interference of the x-ray emission from different electron trajectories driven by adjacent optical cycles for a particular wavelength (i.e. harmonic order). This intra-atomic phase matching allows us to selectively increase the brightness of a single harmonic order by over an order of magnitude, essentially channeling the nonlinear response of the atom to a particular order of nonlinearity. Furthermore, the arbitrary control over

the shape of the driving pulse allows us to spectrally narrow a given harmonic order very effectively, resulting in a bandwidth of the harmonic peak that is likely to be at or near the time-bandwidth limit for such a short x-ray pulse. Finally, optimization of a single harmonic without suppressing adjacent harmonics can increase the brightness of some harmonic orders by factors of ≈ 30 .

Furthermore, it was thought that the learning algorithm was not sufficiently robust to operate on high-order or highly complex quantum systems. Our work demonstrating control of high harmonic generation showed that a high order quantum system could be controlled, and that control could also be well understood.

2.5 Experimental Apparatus

For our work, we used a broad-bandwidth, short-pulse-optimized Ti:sapphire amplifier system into which a closed-loop pulse shaping apparatus was incorporated [69]. By careful design of these amplifier systems, pulses as short as 15 fs (approximately 6 optical cycles) FWHM can be generated at high repetition-rates and high pulse energies (up to 7 kHz with pulse energies > 1 mJ). In such laser systems, low energy pulses of duration ~ 10 fs are generated by a broad-bandwidth Ti:sapphire oscillator [70], stretched in time to lower their peak intensity, and then amplified in two amplifier crystals prior to re-compression. This type of laser system is ideal for inclusion of a simple, phase-only, pulse shaper into the beam before amplification, since phase modulations introduced by the shaper will propagate without distortion in the high-energy, amplified, laser pulse.

We used a new type of phase-only pulse shaper for this work, incorporating a

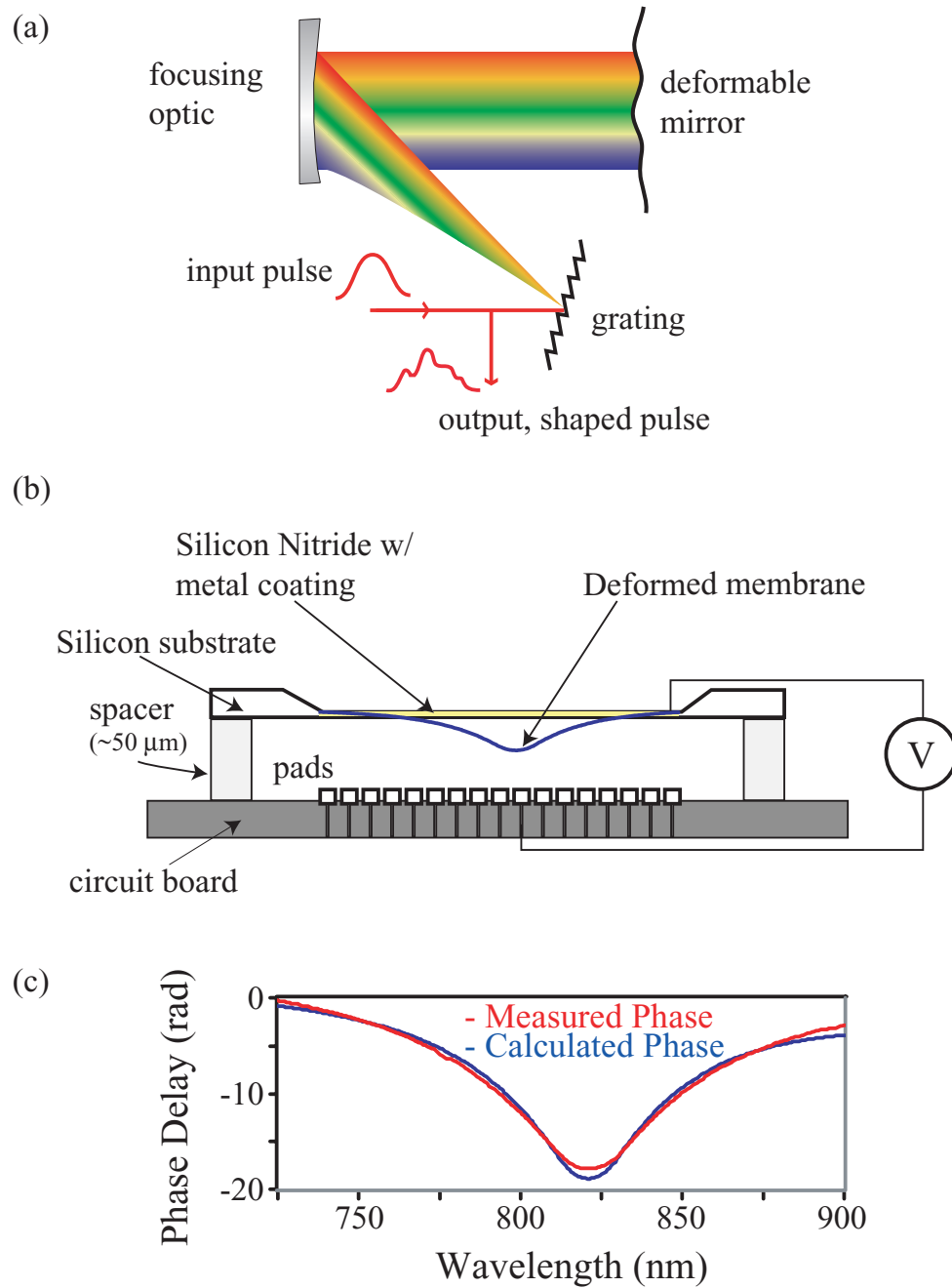


Figure 2.4: A deformable mirror that folds a zero-dispersion stretcher (a) can be used to adjust the relative delay of colors (and thereby spectral phase) of an optical pulse. The mirror (b) is constructed by growing a layer of Si_3N_4 (600 nm) on a Si substrate and back-etching a window which is over-coated with a metal. This mirror is suspended over an array of pads patterned on a PC board. When a voltage is applied to one of the pads, the mirror deforms due to electrostatic attraction, changing the spectral phase (c) [71].

micro-machined deformable mirror [69, 71]. This simple shaper (illustrated in Figure 2.4) works by separating the color components of the ultrashort light pulse (which span ~ 80 nm bandwidth centered at 800 nm) using a grating, then reflecting them from the deformable mirror. Subsequently, the color components are reassembled to form a collimated, temporally shaped, beam. Altering the exact shape of the mirror can then control the relative arrival time of each color component in the pulse. Thus, the pulse shaper manipulates the phase of the pulse in the spectral domain, reshaping the pulse shape and phase in the time domain, while conserving the pulse energy. The mirror itself is a smooth silicon-nitride surface incorporating 19 actuators that deform the mirror — thus it is possible to precisely control the pulse shape, without introducing artifacts due to discrete pixellation. Furthermore, by not altering the spectrum of the pulse, we avoid possible pulse distortions due to nonlinear self-phase modulation in the amplifier. The deformable mirror used in this work was capable of deflecting up to $4 \mu\text{m}$ (or 20π at 800 nm), which compensates for the dispersion accumulated by propagation through 1 cm of fused silica [71]. The exact shape of the pulse, including the amplitude and phase of the electromagnetic field, can be measured using the second-harmonic generation frequency resolved optical gating (SHG FROG) technique [72].

The laser system was adjusted to produce a transform-limited laser pulse that demonstrates the utility of our evolutionary algorithm, and provides a benchmark pulse shape for comparison with our coherent control experiments described in chapters 3 and 4. We used an evolutionary algorithm that starts with a population of

~ 100 members, each of which corresponds to a particular set of voltages applied to the 19 mirror actuators (the DNA). The fitness of the corresponding pulse shape is then measured experimentally. The fitness is simply a quantitative measure of the desirability of a population member, and for pulse compression, we use the intensity of the frequency doubled laser pulse. The best solutions (largest fitness values) are selected as parents, which determine future populations (generation) of the algorithm. Several copies of each parent form the set of children. The children are mutated with a Gaussian noise function to perturb the solutions. The parents and mutated children are combined to form the population of the next generation. The process is then repeated until the fitness changes by an insignificant amount between generations; at this point, the process is said to have converged. This typically occurs in 50 to 100 iterations, with about 100 population members tested for each iteration. Details of this algorithm can be found in Appendix A, and settings of the algorithm used to control HHG are given in Chapter 3, while those used to control molecules are given in Chapter 4.

2.6 Summary

In this thesis, I describe the construction and application of a learning machine capable of controlling complex, nonlinear quantum mechanical systems in realistic (i.e., noisy) laboratory environments. This approach is extremely powerful in that the algorithm determines how to best control the system. Although this approach has been criticized in the past, Chapters 3 & 4 show that it is very robust and useful. By allowing the learning machine to search the parameter space for an optimal solution,

we remove experimenter bias and we show that the learning algorithm discovers new, unknown, and unexpected solutions. In other words, the learning algorithm truly can learn.