# On Parameterizations and Randomized Search Heuristics in Computationally-Expensive Quantum Control Simulations

Ofer M. Shir[1]

[1]Computer Science Department, Tel-Hai College, and Migal Institute, Upper Galilee, Israel

### Abstract

This manuscript introduces the optimization of dynamic molecular alignment by means of shaped femtosecond laser pulses, and analyzes the application of specific Derandomized Evolution Strategies (DES) to this challenging high-dimensional *Quantum Control* problem. With a computationally expensive noise-free simulator, specific DES variants based on low-dimensional parameterization of the electric field are applied to the task and compared among each other. In addition, a new parameterization approach for learning a function, based on complete-basis functions, is introduced and applied to the problem. A simplified variant of the alignment problem is then comprehensively analyzed, on which selected parameterizations and optimization routines are shown to consistently attain optimal and sub-optimal sets of solutions. The combined study of algorithms and parameterizations provides altogether a sharper observation of optimality. Some conclusions concerning first- versus second-order information in DES are drawn. Some of the methods reported here can likely be transferred to more generalized high-dimensional problem domains, where the input solution is specified by a parameterized curve that undergoes spectral transformations.

**Keywords**: Simulation-Based Optimization, Derandomized Evolution Strategies, CMA-ES, Quantum Control, Laser Pulse Shaping, Dynamic Molecular Alignment, Evolutionary Approach to Many-Parameter Physics.

## 1 Introduction

While the usefulness as well as the powerful performance of contemporary Evolutionary Algorithms (EAs) [1] cast no doubt, the determination of efficient evolutionary routines operating in the realm of computationally expensive simulators is much needed. Such simulations are often deployed in physics, and they require robust methods which will drive an optimization campaign into maximal yield with as few simulator evaluations as possible, and without considering approximate surrogate models. This manuscript discusses the deployment of specific EAs, namely Derandomized Evolution Strategies (DES) [2, 3], for the optimization of an application from the field of physics, and investigates the interplay between algorithms to parametrization in the light of convergence speeding-up. While the main focus here is on the application to a specific use-case, this manuscript might also be interesting for researchers who apply EAs to computationally expensive simulations in other domains. The results that we shall describe in this manuscript for the prescribed application domain, can likely be transferred to similar problems, where the input solution is specified by a parametrized curve, and in particular when this curve undergoes spectral transformations.

Controlling the motion of atoms and molecules has been a dream since the early days of Quantum Mechanics. Although this quest initially met with failure, the foundation of the Quantum Control (QC) field in the 1980s, throughout the development of various approaches [4, 5, 6], has finally brought this dream to fruition. Quantum Control, sometimes referred to as Optimal Control or Coherent Control, aims at altering the course of quantum dynamics phenomena for specific target realizations. There are two main threads within Quantum Control, *theoretical* versus *experimental* control, as typically encountered in physics. They have experienced an amazing increase of interest during the past 10 years, in parallel to the technological developments of ultrafast laser pulse shaping capabilities, that made it possible to turn the dream into reality. For a broad field review see [7, 8, 9].

Quantum Control Theory (QCT) [10, 11] aims at manipulating the quantum dynamics of a *simulated system* by means of an external control field, which typically corresponds to a temporal electromagnetic field arising from a laser source. Explicit details concerning the simulated system, expressed by means of its Hamiltonian, are required. Typically, no constraints such as fluence or resolution, are considered. Quantum Control Experiments (QCE) [6], on the other hand, consider the realization of QC in the real-life laboratory, aiming at applying closed learning-loops for altering the course of quantum phenomena. Here, the yield, or success-rate, correspond to a physical measurement, whereas no model Hamiltonian is required. The nature of the optimization is fundamentally different than in QCT, due to the posed laboratory constraints: limited bandwidth, limited fluence, control resolution, proper basis, etc.

The QC application to *dynamic molecular alignment* (for a review, see [12]) is of considerable interest in this context because of its many practical consequences: a multitude of chemical and physical processes ranging from bimolecular reactions [13] to high harmonic generation [14] are influenced by the angular distribution of the molecular sample. Furthermore, in many fundamental molecular dissociation or ionization experiments the interpretation of the collected data becomes much easier when the molecules are known to be aligned with respect to a certain axis. Hence, techniques to generate molecular alignment are much needed.

In this study we consider a numerical modeling of the dynamic molecular alignment, driven by a known Hamiltonian, but designed in a lab-oriented manner. In essence, it is QCT combined with some QCE characteristics, and we set it as our representative use-case of computationally expensive simulations. We choose to restrict this study to noise-free simulations, as we are interested in the physics of the system, rather than conducting an actual simulation of real laboratory experiments.

The goal that we want to achieve in this manuscript is twofold, and it is thus divided into two parts, respectively. The first part of this paper is devoted to a large extent to an investigation of the performance of specific Evolution Strategies variants [2, 3], as well as parametrization, with respect to the given optimization task. Especially, it focuses on a comparison between first-order to second-order Derandomized Evolution Strategies, that correspond to employing a number of strategy parameters scaling *linearly* (individual step-sizes) versus *quadratically* (arbitrary normal mutations, by means of a full covariance matrix) with the search space dimensionality. The second part concentrates on the physical interpretation of the obtained optimal solutions. In particular, we will investigate the underlying fine details that characterize the attained solutions, and deepen into the interplay between optimality of solutions to the optimization aspects of parametrization and scalability.

The remainder of this paper is organized as follows. In Section 2 the dynamic molecular alignment will be introduced, and its numerical modeling will be outlined. Section 3 describes the optimization routines in use, namely Derandomized Evolution Strategies. The experimental procedure of the alignment problem is described in Section 4. Section 5 presents a new parametrization approach for learning a function, based on complete-basis functions, and describes its application to the alignment problem. A simplified variant of the problem is studied in Section 6, in order to allow a better understanding of the optimal parametrization. Finally, a scalability study of the problem is conducted and presented in Section 7. Section 8 draws some conclusions and discusses directions for future research.

## 2  Use-Case: Molecular Alignment

The reader who wishes to abstract from the physics details could simply view our selected use-case for this study as a non-linear problem of finding an optimal functional, with real-world applications. It uses a high-dimensional multi-point representation of that functional, where the dimensions under investigation are in the range of $80 - 1000$, and employs a simulator with a duration of 35s per single evaluation.

### 2.1  Background

Using laser electric fields, molecular alignment can be achieved in two manners: if the envelope of the field changes slowly compared to the timescale of molecular rotation (usually in the picosecond regime), each rotational state of the initial Boltzmann distribution is transformed adiabatically into a *pendular state*. The drawback of this approach is that any alignment produced while the field is on will disappear when it is switched off again, so that experiments with the aligned sample cannot be performed under field free conditions. The opposite strategy is followed in *impulsive alignment*, where the duration of the applied pulses is much shorter than a rotational period [15]. In this case, a wavepacket of rotational states is created that can lead to considerable field-free alignment. Both methods aim at creating a superposition of as many angular momentum eigenstates as possible, as, due to the angular version of the uncertainty principle, a broad distribution in angular momentum corresponds to a narrow angular distribution. However, both the amplitudes and the relative phases of the composite rotational states have to be under control in order to achieve alignment. While for a pendular state, which is an eigenstate of the combined molecule-field Hamiltonian, this requirement is automatically fulfilled, a randomly phased superposition of rotational states will generally not interfere favorably in order to create alignment. For the impulsive case, the evolution of the total wavefunction (after the electric field is turned off) repeats with the *revival time* $T_{rev} = \frac{1}{2Bc}$ (where $B$ is the rotational constant of the molecule and $c$ is the speed of light). Partial revivals can be observed at $T_{rev}/2$ and, to a lesser extent, at $T_{rev}/4$, when one half or one quarter, respectively, of the populated rotational levels have undergone an identical number of rotations. Shaped femtosecond laser pulses that lead to a high degree of alignment manage to maximize the number of rotational states that are in phase at these times. However, they have to fulfill an additional requirement: the description of the molecule as a rigid rotator is strictly valid only for low field intensities, and the higher the applied intensity, the more important other competing channels like dissociation and ionization will become. Therefore, one would like to achieve good alignment while keeping the peak laser intensity as low as possible.

For this reason varuiys publications have focused on finding pulse shapes other than a Fourier transform limited (FTL) pulse that create a high degree of alignment. Leibscher et al. [16, 17] have theoretically shown that in the nonperturbative regime a train of pulses lead to better alignment than a single FTL pulse. Each of the spikes in such a train has to be synchronized with a half- or a full revival time. The suggested scheme has been experimentally realized by measuring enhanced alignment of iodobenzene [18] and nitrogen [19], respectively, when the molecular sample was irradiated by two short pulses with a delay corresponding to the rotational period of the molecule. A further extension has been proposed by Underwood et al. [20], where the two pulses were allowed to have different axes of polarization. For asymmetric molecules, orientation has been found to be optimized by a sequence of kicks as well [21].

Such pulse sequences can be easily constructed and also optimized with respect to the relatively small number of parameters they contain. Therefore, they provide a good starting point for more complex optimization schemes, where the electric field is characterized by many more parameters. To find good solutions in this high-dimensional search space is a nontrivial task already for zero Kelvin temperature ($T = 0$), when considering only the ground state in the initial distribution. For finite temperatures the alignment optimization has to be performed simultaneously for a whole set of initial rotational quantum numbers, which, together with the large number of parameters needed to describe an arbitrary pulse shape, renders the optimization problem even more difficult.

## 2.2   Numerical Modeling

The interaction of a generic linear molecule with a laser field is described within the framework introduced in [22]. Briefly, we calculate the time evolution of a thermal ensemble of molecules quantum mechanically by performing calculations for a range of initial rotational levels, characterized by their rotational quantum number $J_{initial}$ and the projection of the angular momentum on the laser polarization axis $M_{initial}$, respectively, and Boltzmann averaging the results for a given temperature $T$. We take the molecule to be a rigid rotor, which allows a description of its wavefunction solely in terms of the rotational *eigenstates* $|JKM\rangle$ (where $K = 0$ for a diatomic molecule). Two electronic states are taken into account, the ground state denoted by $X$ and an off-resonant excited state denoted by $A$. Hence the wavefunction for a given $M$ is expanded as

$$\Psi_M(t) = \sum_{J=M}^{N_{rot}} \alpha_{XJM}(t)\psi_{XJM} + \alpha_{AJM}(t)\psi_{AJM} \tag{1}$$

with $N_{rot} = 20$ rotational levels, where the $\psi_{XJ}$ correspond to the even Legendre polynomials and the $\psi_{AJ}$ correspond to the odd Legendre polynomials. This expansion was confirmed to give converged results in the present calculations.

The time dependence of the molecular wavefunction is given by

$$i\frac{\partial \Psi_M}{\partial t} = (H_0 + V)\Psi_M(t) \tag{2}$$

The Hamiltonian consists of a molecular part $H_0$ and the interaction with the laser field, given by

$$V = \vec{\mu} \cdot \vec{E(t)} \cos(\omega_0 t) \tag{3}$$

The Eigenenergies of $H_0$ are given by

$$E(J) = hcBJ(J + 1) \tag{4}$$

where $B$ is the *rotational constant* of the molecule. The laser field induces transitions between the rotational states which, in the off-resonant case, occur via subsequent Raman processes. The transitions between $X$ and $A$ were assumed to proceed via the selection rules of the quantum numbers $\Delta J = \pm 1, \Delta M = 0$. By inserting Eq. 1 into Eq. 2 and applying a transformation into a rotating frame, one ends up with a set of differential equations for the expansion coefficients of the wavefunction into dressed states (see [22] for the explicit expressions).

The envelope of the laser field, which completely determines the dynamics after the transition to the rotating frame has been performed, can be described by:

$$E(t) = \mathbb{R}\left\{ \int A(\omega)\exp(i\phi(\omega))\exp(i\omega t)\, d\omega \right\} \tag{5}$$

In the present calculation the control function is the spectral phase $\phi(\omega)$, which defines the phase at $n$ frequencies $\{\omega_i\}_{i=1}^n$ that are equally distributed across the spectrum of the pulse. These $n$ values $\{\phi(\omega_i)\}_{i=1}^n$ are the decision parameters to be optimized, and upon their calibration they are interpolated onto $2^{14}$ points for the calculation of the electric field in Eq. 5. In order to achieve a good trade-off between high resolution and optimization efficiency, the value of $n = 80$ turned out to be a good compromise. The search space is therefore an 80-dimensional hypercube spanning a length of $2\pi$ in each dimension. Real-laboratory pulse shapers typically use
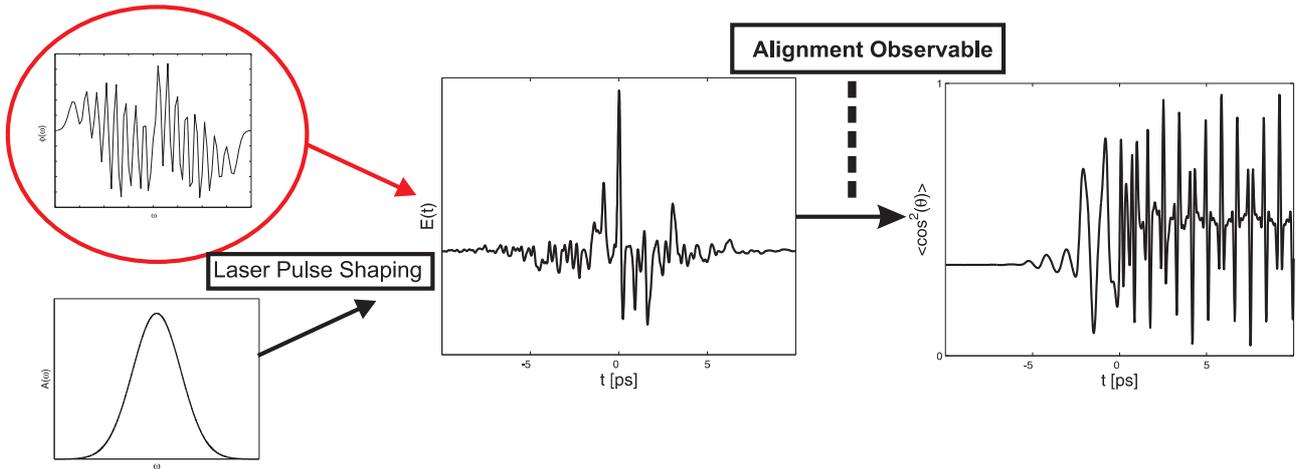
Figure 1: An overview of the numerical process. The control function is the phase (circled, top left), the amplitude function is fixed and approximated by a Gaussian (bottom left). The shaping process (Eq. 5) generates the electric field, $E(t)$ (center). The "Schrödinger Box" of the alignment observable represents the numerical calculation of the interaction between the electric field with the molecules, based on the specified quantum dynamics equations. The revival structure (right) is the observed simulated behavior of the molecules, upon which the yield value is based.

64 or 128 pixels for the spectral phase, so in that respect our simulation reflects a feasible realization. The scalability aspect of the control phase will be addressed in Section 7.

The spectral function $A(\omega)$ is taken to be a Gaussian with a width chosen such that the *full-width-at-half-maximum* (FWHM) length of the *Fourier transform limited* (FTL) pulse (obtained by setting $\phi(\omega) \equiv 0$) equals 100fs. With $B = 5\text{cm}^{-1}$, as it was used in all the optimizations, the pulse is thus much shorter than the rotational period of $T_{rev} = 3.33$ ps. In order to keep the description as general as possible, the peak field strength is not fixed explicitly; instead, we set the peak Rabi frequency $\Omega_{XA}$ for the transition between the $X$ and the $A$ electronic state, which is proportional to the product of peak field strength and the coupling matrix element between $X$ and $A$. Since the target application of the optimal control process is the dynamic alignment of the molecules, the yield of the calculation (i.e., the *objective function*) is set to the *cosine-squared observable value*, subject to maximization. The explicit details will be given in Section 4. Fig. 1 provides an illustrative overview of the numerical process. A typical phase function and a typical laser pulse, obtained by an evolutionary optimization, are given for illustration in Fig. 2.

# 3 Algorithms

On the basis of earlier work on dynamic alignment [23] and due to previous calculations showing that certain DES variants perform better with respect to other EAs on these problems, we restrict our study to these state-of-the-art algorithms.

## 3.1 Evolution Strategies

*Evolution Strategies* (ES) [3] are canonical evolutionary algorithms for continuous function optimization, due to their straightforward real-valued encoding, their specific variation operators, as well as to their high performance on benchmark problems compared to other methods (see, e.g., [1, 24]). A task becomes more suitable for an ES when the dimensionality of the search space increases. In Evolution Strategies, each individual carries, apart from the real-valued array of search parameters, a vector of strategy parameters, which determines the mutation step-size in each direction. Both the search parameters, also referred to as object variables or decision parameters, as well as the strategy parameters, are optimized.

## 3.2 Derandomization

Mutative step-size control tends to work well for the adaptation of a global step-size, but tends to fail when it comes to the individual step-sizes (i.e., *variances* and *covariances* of the joint mutation distribution). This is due to several disruptive effects [25], as well as to the fact that the selection of the *strategy parameters* setting is indirect, i.e. the vector of a successful mutation is not used to adapt the step size parameters, but rather the
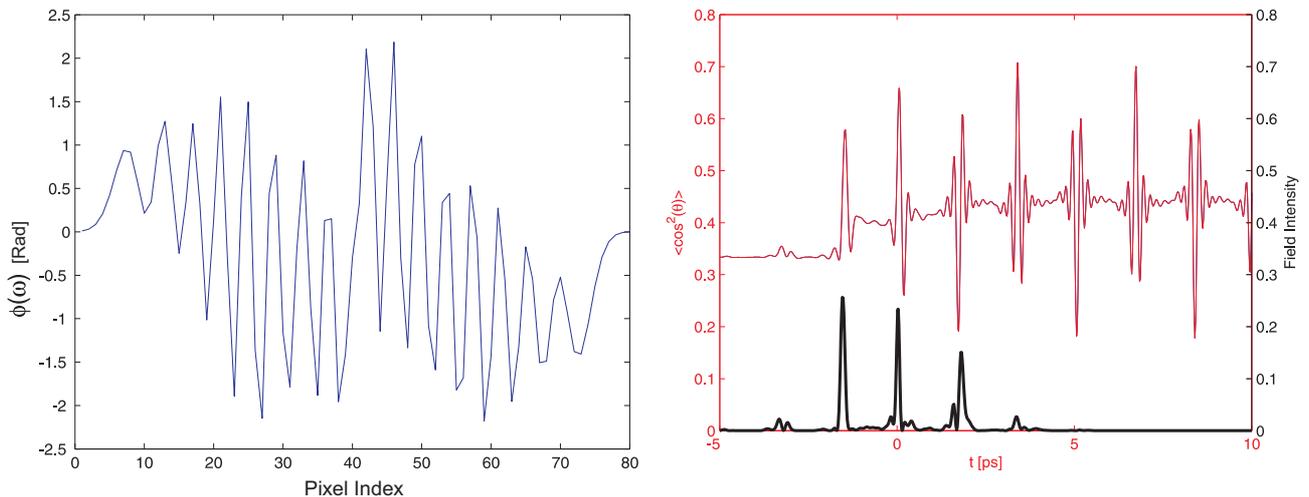
Figure 2: [LEFT] A phase function obtained by an evolutionary search (frequency domain). [RIGHT] Given the spectral phase plotted in the left – thin red line: alignment, quantified in cosine-square (see section 4 for details); thick black line: intensity of the laser pulse (time domain).

parameters of the distribution that led to this mutation vector. The so-called *derandomized mutative step-size control* aims to treat those disruptive effects.

It is important to note that the different DES variants hold different numbers of strategy parameters for adaptation, and this is a factor in the speed of the optimization course. The different algorithms employ a number of strategy parameters scaling either *linearly* or *quadratically* with the search space dimensionality $n$, corresponding to **first-** or **second-order information** of the search landscape. The careful reader should note that we consider here first- or second-order variation information in a stochastic evolutionary algorithm as equivalent to first- or second-order search information of a deterministic algorithm.

We hereby present briefly two derandomized-ES variants that were in use in our calculations: the DR2 algorithm, which historically was the second derandomized variant to be released, and the CMA-ES, the fourth variant in the series.

## 3.3 First-Order DES: The DR2 Algorithm

The second derandomized ES variant [26] aims to accumulate information about the correlation or anti-correlation of past mutation vectors in order to adapt the *global step-size* as well as the *individual step-sizes* by introducing a quasi-memory vector. This accumulated information allows for omitting the stochastic element in the adaptation of the strategy parameters - updating the strategy parameters only by means of successful past mutations, rather than with random steps. This method stores first-order information by means of its $\mathcal{O}(n)$ strategy parameters, where $n$ is the search space dimensionality.

We follow the recommended population size for $(1, \lambda)$ derandomized ES (see, e.g., [27]), and set $\lambda = 10$. By definition, this strategy does not apply recombination. The mutation step for the $k^{th}$ individual, $k = 1 \ldots \lambda$, reads:

$$\vec{x}_k^{(g+1)} = \vec{x}^{(g)} + \delta^{(g)} \vec{\delta}_{scal}^{(g)} \vec{z}_k \qquad \vec{z}_k \sim \vec{\mathcal{N}}(0, 1), \tag{6}$$

where the evolving strategy parameters $\delta^{(g)}$ and $\vec{\delta}_{scal}^{(g)}$ are the *global step-size* and the *variation directional vector*, respectively. We refer the reader to [26] for the detailed description of the defining update steps.

## 3.4 Second-Order DES: The $(\mu_W, \lambda)$ CMA

The $(\mu_W, \lambda)$-**CMA-ES** algorithm (see, e.g., [25, 28]) applies *principal component analysis* (PCA) to the *selected* mutations during the evolution, also referred to as *"the evolution path"*, for the adaptation of the covariance matrix of the distribution. The concept of *weighted recombination* is introduced: applying intermediate multi-recombination to the best $\mu$ out of $\lambda$ offspring with given weights $\{w_i\}_{i=1}^{\mu}$. The result is denoted with $\langle \vec{x} \rangle_W$. $\mathbf{C}^{(g)} \in \mathbb{R}^{n \times n}$ is the covariance matrix, which is eigen-decomposed as $\mathbf{C}^{(g)} = \mathbf{B}^{(g)} \mathbf{D}^{(g)} \left( \mathbf{B}^{(g)} \mathbf{D}^{(g)} \right)^T$. By definition, this method stores second-order information by means of its $\mathcal{O}(n^2)$ strategy parameters. The characteristic generation step for the $k^{th}$ individual, $k = 1, \ldots, \lambda$, reads:

$$\vec{x}_k^{(g+1)} = \langle \vec{x} \rangle_W^{(g)} + \sigma^{(g)} \mathbf{B}^{(g)} \mathbf{D}^{(g)} \vec{z}_k \qquad \vec{z}_k \sim \vec{\mathcal{N}}(0, 1) \tag{7}$$

We refer the reader to [28] for the detailed description of the defining update steps of the strategy parameters.

# 4 The Alignment Problem: Practical Observation

In this section we provide the reader with the essential information regarding our setup and our experimental observations for the optimization of dynamic alignment.

## 4.1 The Optimization: Modus Operandi

In order to assess the performance of the algorithms we have conducted 10 independent runs for each routine with the goal of optimizing the alignment of a sample of generic diatomic molecules undergoing irradiation by a shaped femtosecond laser. The calculations were based on the following core mechanisms: **(1)** $(1, 10)$-DR2, **(2)** $(1, 10)$-CMA, and **(3)** $(8, 17)$-CMA (default settings for $n = 80$).

We have used the maximum $\langle \cos^2(\theta) \rangle$ that occurs under field free conditions after the laser pulse, where $\theta$ is the angle between the molecular and the laser polarization axis, as a measure of the alignment. The temperature of the ensemble is set to $T = 100$K. The peak Rabi frequency between the two electronic states $X$ and $A$, that determines the interaction strength, was $\Omega_{XA} = 180 \cdot 10^{12}$ s$^{-1}$; the detuning with respect to the intermediate level was $\Delta = 10^{15}$ s$^{-1}$. The FTL width of the pulse was set to $T_{FWHM} = 100$ fs. With the two-photon Rabi frequency [29] being $\Omega_{2ph} = \frac{4\pi \Omega_{XA}^2}{\Delta}$, an integration of $\Omega_{2ph}$ over the pulse length yields 4.78 two-photon oscillations; thus the intensity is in the non-perturbative regime.

Since we want to achieve a high degree of alignment with a peak intensity as low as possible, an additional constraint was introduced that punishes pulses that are too intense. Towards this end, we have used

$$I_p = \int E^2(t)\Theta(E^2(t) - I_{thr}) \, dt, \tag{8}$$

where $\Theta(x)$ is the Heaviside step function, so that the fitness function assigned to a pulse shape was

$$F = \max_{E(t)} \langle \cos^2(\theta) \rangle - \beta I_p. \tag{9}$$

By choosing $\beta$ large enough, $I_{thr}$ can be used to effectively operate the evolutionary search only on a subset of pulses whose maximum peak intensity approaches the threshold intensity from below. We have typically used $\beta = 1$; unless otherwise specified, we set the threshold $I_{thr} = 0.36 \cdot I_{FTL}$.

## 4.2 Setup

We hereby provide some information about the experimental setup of the dynamic alignment numerical simulation:

- The *cosine-squared alignment* yields a real value in the interval $[0, 1]$, subject to maximization, as was introduced earlier. In the absence of a laser field, an arbitrary measurement yields on average an alignment value of 0.333, due to the isotropic $3D$ space. In the presence of a laser field, a random phase typically obtains alignment values around 0.4. The best alignment result known to us in the given temperature is $\approx 0.7$.

- The punishment term, which was introduced earlier in Eq. 8 and Eq. 9, can yield fitness values below the value of 0.4. The probability of a randomly generated pulse, with no specific parametrization, to get this punishment is extremely low.

- Every fitness evaluation call requires approximately 35 seconds on a single P4-HT 2.6$GHz$ processor.

- **Due to the heavy computational cost of a single fitness evaluation, we are practically limited in granting function evaluations and obliged to apply minimal settings, as will be described**.

## 4.3 Preliminary Numerical Results

Table 1 summarizes the numerical results of the runs - the averaged alignment value obtained by each optimization routine. Due to the limited number of simulations we do not provide further statistical analysis of the results. Based on our experience with the problem and the algorithms, the yield differences of Table 1 are believed to be significant. Fig. 3 presents the best solution found, obtained by the DR2 routine.

| Alignment | (1,10)-DR2 | (1,10)-CMA | (8,17)-CMA |
|---|---|---|---|
| AVG-Fitness | 0.6789 | 0.4676 | 0.6261 |

Table 1: $\langle \cos^2 \theta \rangle$ mean values obtained with the different algorithms over 10 runs.
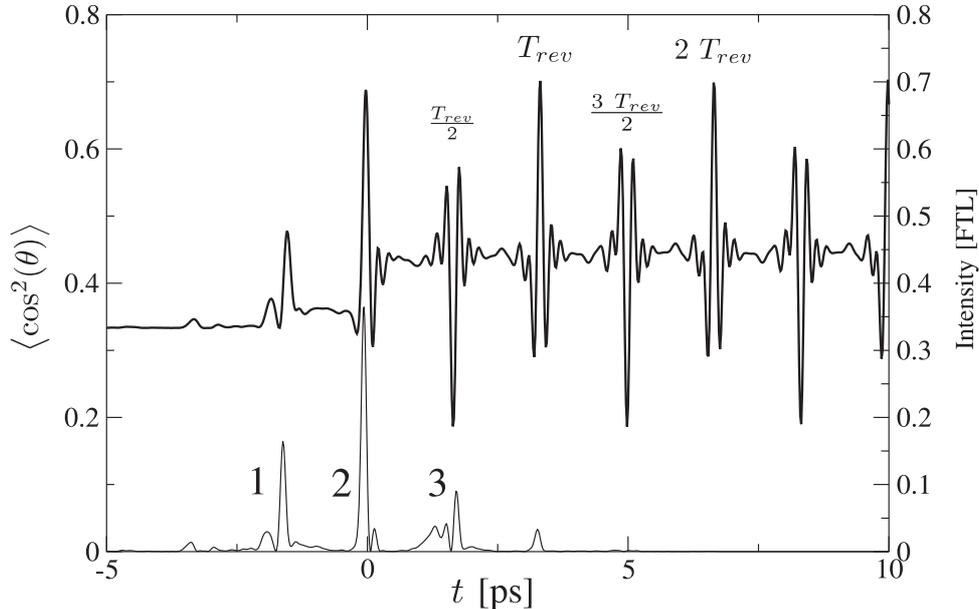


Figure 3: Best solution attained by the DR2. Thick line: alignment; thin line: intensity profile of the optimized laser pulse. The solution consists of three main peaks (see labels).

# 5 The Complete-Basis-Functions Parametrization

We hereby propose a new method for learning a function, based on a representation transformation, which can also be referred to as parametrization. The so-called *Complete-Basis-Functions Parametrization Method* was originally constructed for learning the control function of the dynamic alignment problem, the phase $\phi(\omega)$, but is a general method for learning a generic $n$-variable function. It reduces the dimensionality of the search space and speeds-up the convergence respectively, as will be explained in detail.

The Appendix provides the reader with the mathematical background on complete-basis functions, and presents the specific functions that are considered for this study.

**Preliminary: A Simple Problem**

As we will demonstrate here, finding the expaInversension of a known function by means of a given set of complete-basis-functions, i.e. finding the coefficients of the functions in this basis, is an easy task for a simple EA, and in particular for the standard ES. For simplicity, and without loss of generality, let us assume that the task is to approximate a one-variable function, representing a one-dimensional curve, by means of a trigonometric series. This task can be generalized to functions of higher dimensions, and by using other expansions.
Consider the *expansion coefficients* of the cosine and sine functions, $\{a_k\}_{k=0}^{\infty}$, $\{b_k\}_{k=1}^{\infty}$ (see Appendix for the notation in use), as the decision parameters to be optimized in the evolutionary search. As a preliminary task in this study, we found that the standard ES converged easily and fast to the right coefficients, where this basic fitting problem was simply defined as an *inverse problem* (for a general review on such a problem class, see, e.g., [30]), with the objective function subject to minimization was the root-mean-square error function between the original function and its evolving expansion.

Fig. 4 shows the result of learning the *triangle function* with a standard Evolution Strategy, using only the first 20 frequencies ($n = 40$) of a *Fourier* series as building blocks for a given discretization of $N = 100$ points on the curve.

**Learning a Function**

The idea of spanning a function using a set of complete basis-functions can also be applied for the task of learning an unknown function, which is represented by a parametrized curve, as in our QC alignment problem. The inspiration for this method was the physicists' intuition to the alignment problem, which stated that the control
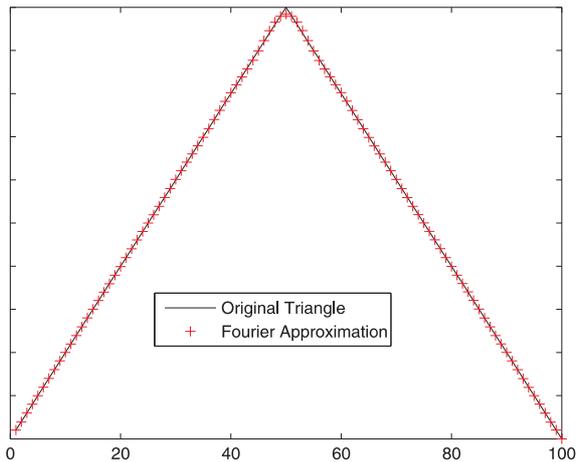
Figure 4: Learning the *triangle function* using the first 20 *Fourier* frequencies. The plot shows the original triangle function and its Fourier approximation, whose curves are parametrized by $N = 100$ points.
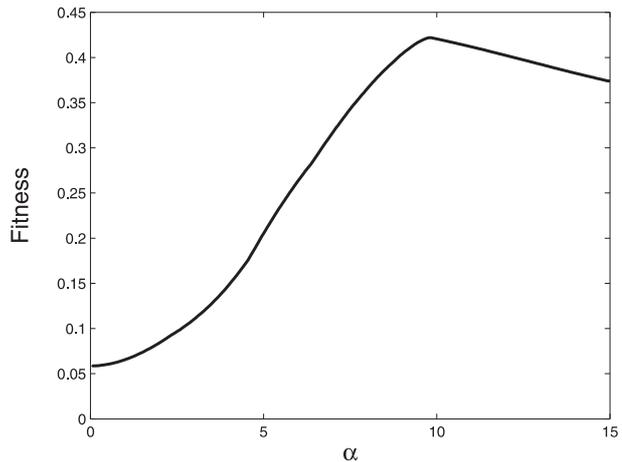
Figure 5: The alpha-test: the fitness of quadratic phase functions, centered around the central frequency.

function should be periodic. Motivated by this intuition, we started to run experiments in which an ES was aiming at learning $\phi(\omega)$ using the *harmonic functions* as building blocks. Rather than learning the interpolated values of the control function, the coefficients of the harmonics (Fourier components) were optimized. Following the success of those experiments, we expanded the method to other sets of complete basis functions, and in particular to the sets of functions which are introduced in the Appendix: The *Legendre Polynomials*, the *Bessel Functions*, the *Hermite Polynomials*, and the *Chebyshev* polynomials.

Assuming that the desired discretization is up to a resolution of $N$ points in the interval, we limit the number of elements in the expansion series to $n$, where preferably $n \ll N$. By that we can achieve a dramatic dimensionality reduction of the search space, aiming to achieve a convergence speeding-up. The idea then is to apply an evolutionary search for the $n$ coefficients of the building functions, where a simple transformation is applied for every fitness evaluation. In practice, the time for additional computation of this transformation is negligible with respect to the time for objective function evaluation in every real-world problem.

An ES using a Fourier auxiliary function has been proposed in the past, known as the FES method [31]. The FES aims to approximate the fitness landscape, and especially its small attraction basins, using the Fourier series. However, the careful reader should notice that our method is based on a different principle. It uses complete-basis functions for the approximation of the decision parameters themselves, rather than the fitness landscape, which is left untouched. It strongly relies on the fact that these decision parameters represent a parametrized curve to be approximated/learned.

## 5.1 Preliminary Calculations

### 5.1.1 Quadratic Phase Functions – The $\alpha$-Test

Since we are about to investigate representations of low-order polynomials, we would first like to address the question whether there exists a trivial extremum which would become a local trap for such phase functions. Hence, we calculated the fitness of constructed quadratic phase functions, centered around the central frequency. Explicitly, we considered the following family of phases:

$$\phi(\omega)_\alpha = \alpha \cdot (\omega - \omega_{central})^2, \tag{10}$$

where the continuous parameter $\alpha$ is sampled systematically in the interval $[0, 15]$. Note that these phases are constructed over $n = 80$ function values, and given to the cosine-squared evaluator as before.

The results of this so-called $\alpha$-*test* are presented in Fig. 5. Evidently, most of the quadratic phase functions attain extremely low fitness values, due to large punishment terms, and they never exceed the fitness value of 0.45. This eliminates the existence of a trivial quadratic solution for the problem.

### 5.1.2 The Initial States Density Test

We set the number of terms in each expansion to $K_{max} = 40$. The following preliminary experiment is meant to compare the natural initial quality of the different parametrization. We applied a so-called *initial states density*

8

Table 2: Parametrization: Averaged Performance

| Routine | Direct | | Fourier | | Legendre | |
|---|---|---|---|---|---|---|
| | Avg. Fit. | 0.6 Eval | Avg. Fit. | 0.6 Eval | Avg. Fit. | 0.6 Eval |
| (1,10)-DR2 | 0.6789 | 2325 | 0.4494 | N.A. | 0.6384 | 629 |
| (1,10)-CMA | 0.4676 | N.A. | 0.4542 | N.A. | 0.6409 | 515.1 |
| $(\mu,\lambda)$-CMA | 0.6261 | 4962.5 | 0.6171 | 4475.8 | 0.6466 | 194.5 |
| Routine | Bessel | | Hermite | | Chebyshev | |
| | Avg. Fit. | 0.6 Eval | Avg. Fit. | 0.6 Eval | Avg. Fit. | 0.6 Eval |
| (1,10)-DR2 | 0.6299 | 1390 | 0.5944 | 5610 | 0.4843 | N.A. |
| (1,10)-CMA | 0.6229 | 2212.9 | 0.6755 | 271 | 0.4979 | N.A. |
| $(\mu,\lambda)$-CMA | 0.6232 | 2719.5 | 0.6843 | 118 | 0.6225 | 3770.8 |

Table 3: Parametrization: Summary of Best Results

| Parametrization | Best Fitness | 0.6 Eval | Routine | Initial States Density |
|---|---|---|---|---|
| Direct-Param | 0.6899 | 2310 | DR2 | $0.4025 \pm 0.017$ |
| *Fourier* | 0.6526 | 1411 | CMA-(7,15) | $0.4110 \pm 0.019$ |
| *Legendre* | 0.6487 | 106 | CMA-(7,15) | $0.3122 \pm 0.075$ |
| *Bessel* | 0.6457 | 61 | CMA-(7,15) | $0.2217 \pm 0.077$ |
| *Hermite* | 0.6866 | 31 | CMA-(7,15) | $0.4557 \pm 0.047$ |
| *Chebyshev* | 0.6490 | 1051 | CMA-(7,15) | $0.4226 \pm 0.023$ |

*test*, a statistical fitness measurement of the initialized phase functions in the different parametrization. For each parametrization in use, i.e. the direct 80-dimensional random phase vector, or the random 40-dimensional coefficient vector for the various polynomials in use, we initialized 1000 phase functions and calculated their mean fitness and standard deviation. The numerical results are visualized as histograms in Fig. 6-11, providing the fitness distributions of the random initializations.

## 5.2   Parametrization: Numerical Results

In this subsection we present the results of the runs of the different parametrization methods - the direct parametrization versus the polynomial-based methods with $K_{max} = 40$ terms. Our runs were based on the following core mechanisms: **(1)** $(1,10)$-DR2, **(2)** $(1,10)$-CMA, and **(3)** $(7,15)/(8,17)$-CMA (default settings for $n = \{40,80\}$, respectively). The runs were limited to $10,000$ function evaluations. We performed 10 runs per method.

We consider the performance criteria of the various methods as the following:

- The mean fitness values per method over the 10 runs.

- The averaged number of evaluations per method until the fitness value of 0.6 was reached during the runs ("*fixed target*" perspective). We consider the value of 0.6 as the lower bound of the regime of good solutions.

- The results of the *initial states density test*, as was introduced earlier: the averaged initial fitness values per method, with the standard deviation.

We provide a table of results, which consists of the numerical values of the specified performance criteria per method. It is given as Table 2. Table 3 summarizes the best result obtained per parametrization.

## 5.3   Analysis and Discussion

An important result that should be noted is that **all** the runs in the various parametrization have converged into a highly fit phase function with at least one optimization routine, i.e., all the given complete-basis functions are capable of spanning a good phase function with $K_{max} = 40$ terms.

Furthermore, we would like to analyze shortly the experimental results of the various parametrization with respect to the dynamic alignment optimization, as presented in Tables 2 and 3:

1. **Initial state**. The *Hermite* parametrization has clearly the most natural **initial** representation for the phase function in our given problem, among the various cases, as reflected from the *initial states density test* results (Fig. 6-11 and Table 3). Note that the *Legendre* and the *Bessel* parametrization have low
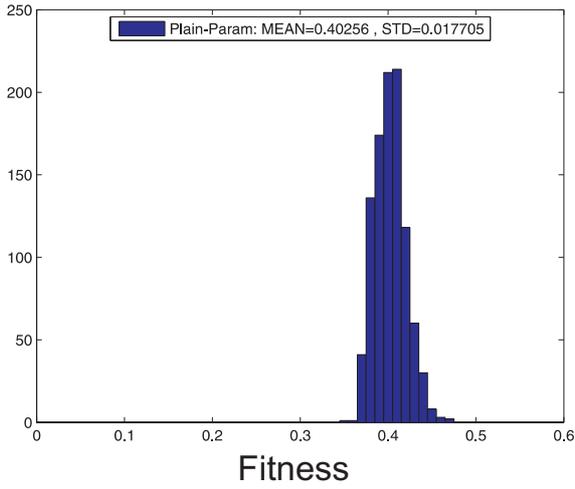
Figure 6: Initial states density test for the **direct** parametrization.
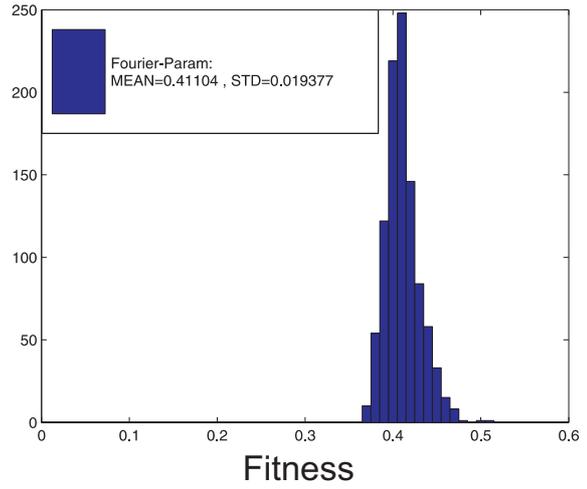


Figure 7: Initial states density test for the **Fourier** parametrization.
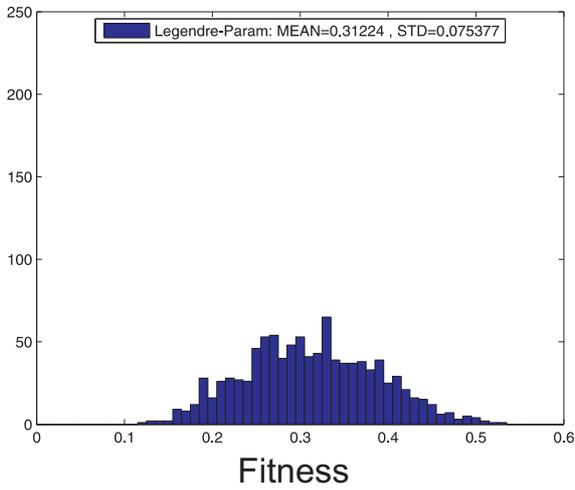


Figure 8: Initial states density test for the **Legendre** parametrization.
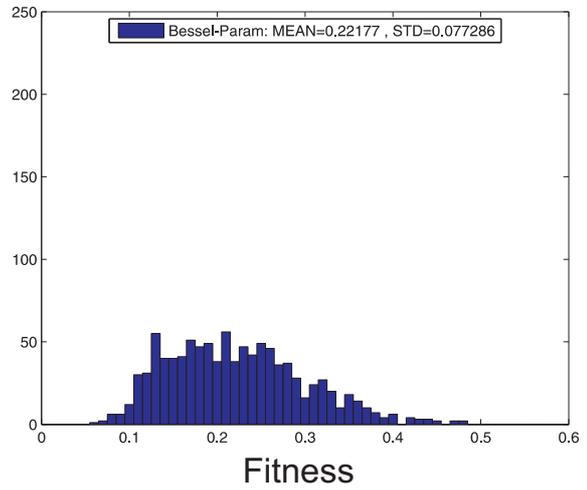


Figure 9: Initial states density test for the **Bessel** parametrization.
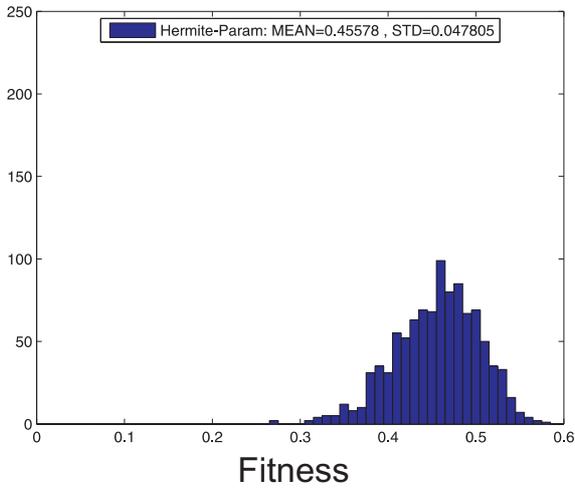


Figure 10: Initial states density test for the **Hermite** parametrization.
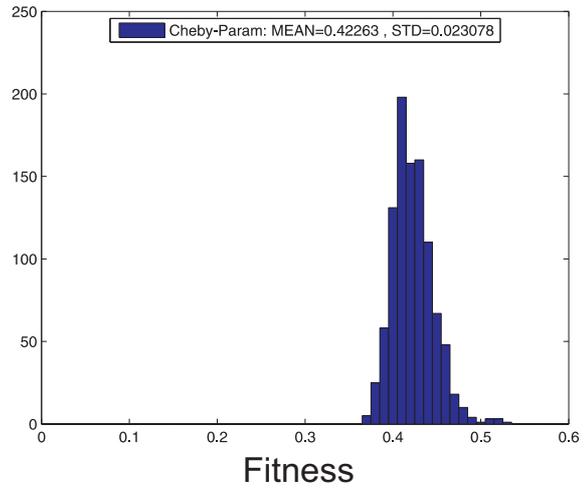


Figure 11: Initial states density test for the **Chebyshev** parametrization.

initial fitness values, even below the direct parametrization, due to the punishment effect. It should be stressed that the standard deviations of the different fitness distributions are reasonably low.

2. **Fitness values**. Some of the polynomial-based parametrization obtained fitness values as high as the direct parametrization method, and in particular the *Hermite* parametrization, which obtained the best results among all the methods. As far as we know, the obtained values are the highest *cosine-squared alignment* values which were ever attained. Hence, from the optimization perspective, the proposed parametrization method does not hamper the feasibility to obtain the maximally-attained yield within the limit of function evaluations.

3. **DR2 vs. CMA**. There is a clear trend regarding the two optimization routines. The DR2 obtains the best results for the direct parametrization, but obviously fails to deliver reasonable results for the special parametrization. In most cases, the DR2 does not even converge. The $(7, 15)$ CMA-ES, on the other hand, performs very well with the various polynomial parametrization, and gets fine results also for the plain parametrization. The $(1, 10)$ CMA-ES is clearly inferior with respect to its rank-$\mu$ weighted-recombined sibling. Our proposed explanation for this trend is the strong correlations between the polynomials' coefficients, which make the covariance matrix an essential component for successful optimization. On the other hand, it seems that the covariance matrix is not an essential component for the direct parametrization, and may even introduce a barrier, to some degree, to the global search. It is suggested by Quantum Control Theory that first-order information is sufficient for climbing-up the control landscape, as will be discussed in Section 8. Moreover, we further explore the performance of the DR2 versus CMA-ES with respect to the direct versus Hermite parametrization in Section 6.

4. **Convergence speeding-up**. An immediate conclusion from both tables is that the proposed method achieved a significant speeding-up of the convergence for all the various special parametrization, in comparison to the direct parametrization. The *Hermite* parametrization with (7,15)-CMA is clearly the fastest routine, and it outperforms the other routines by far. It should be noted that the *Legendre* as well as the *Bessel* parametrization, which have the lowest initial values, manage to compensate for that and reach the regime of good solutions rather quickly (fitness of 0.6). Typical convergence profiles for Hermite versus direct parametrization are plotted in Fig. 12.

5. **Physics interpretation**. Aiming to get some physics insights into the nature of highly-fit phase functions with respect to the laser shaping problem, we examined the nature of good solutions in the different parametrization. The idea was to calculate the distributions of the coefficients, and try to identify the dominance of certain components (frequencies in the Fourier case). Unfortunately, such dominance could not be determined within the results. The set of optimized phases reveals high complexity, which could not be tackled. This provides us with motivation to explore a more simple variant of the alignment problem in Section 6.

# 6 Optimal Routines versus Parametrization

Here we focus on a simplified variant of the original alignment problem, at zero temperature and with only a single rotational level at the initial distribution. The motivation for this simplification is to allow studying the physics nature of the optimal solutions, which would not have been possible for the general case, e.g., tracking the time-dependent population[1] of the rotational levels, given only the ground level at initialization.

## 6.1 Procedure

We choose to reduce the optimization procedure to the two optimal parametrization, direct and Hermite, subject to two core mechanisms: $(1, 10)$ strategy for the DR2; $(7, 15)$-CMA for $n = 40$ and $(8, 17)$-CMA for $n = 80$. Due to the simplification of this variant, a duration of a single function evaluation was reduced to 6s on a single processor. The peak Rabi frequency was set here to $\Omega_{XA} = 160 \cdot 10^{12} \text{ s}^{-1}$.

## 6.2 Numerical Results

Table 4 presents the mean values and the standard deviations of the cosine-squared alignment, obtained after 20 runs of $20,000$ function evaluations. The yield differences are believed to be significant also in this case. As can be observed, the DR2 routine clearly outperforms the CMA in the *direct* parametrization. The DR2 fails to obtain high-quality solutions for the *Hermite* parametrization, whereas the CMA does succeed in this task, with

---

[1]The careful reader should note that *'population'* is used here exclusively in the context of *quantum mechanics*, e.g., populating quantum levels.
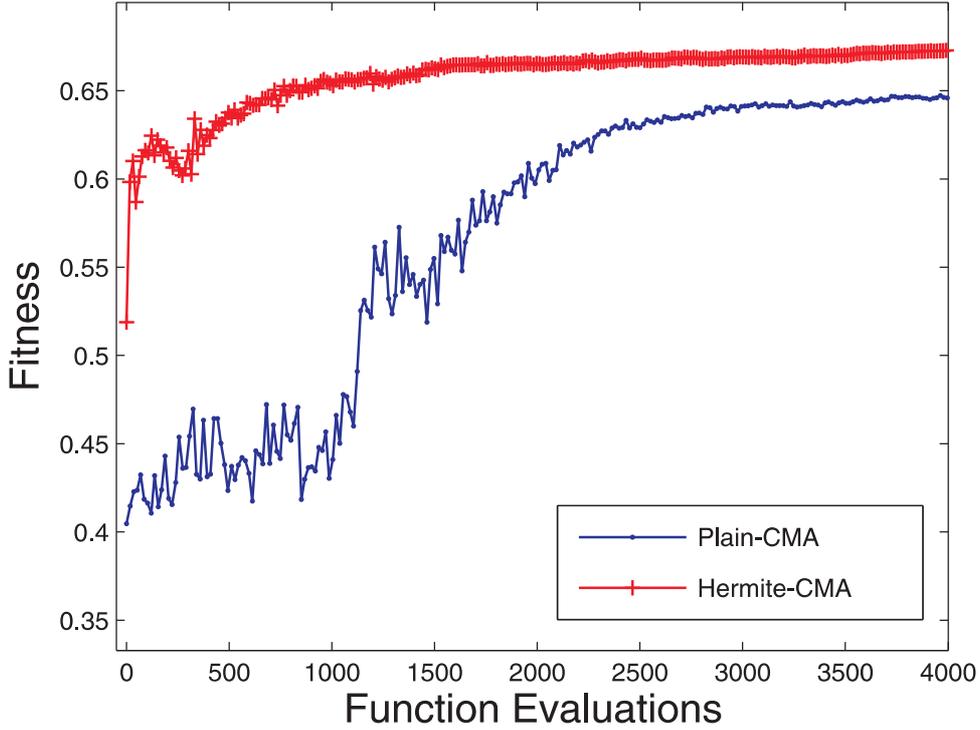
Figure 12: The speeding-up effect: typical convergence profiles of the $(\mu, \lambda)$ CMA-ES for the Hermite versus the direct parametrization.

| | $(1, 10)$ DR2 | $\{(7, 15),\ (8, 17)\}$ CMA |
|---|---|---|
| *Plain* Parametrization | $0.9559 \pm 0.0071\ (0.9622)$ | $0.9413 \pm 0.0058\ (0.9508)$ |
| *Hermite* Parametrization | $0.9501 \pm 0.0043\ (0.9570)$ | $0.9583 \pm 0.0026\ (0.9618)$ |

Table 4: Maximizing the cosine-squared alignment over 20 runs with $20,000$ function evaluations per run; *mean* and *std*; the maximal value obtained is in brackets.

highly-satisfying results. Essentially, this is consistent with the previous results on the general problem of finite temperature. It suggests that the *Hermite* parametrization introduces strongly correlated decision parameters with respect to this landscape as well, whereas the *direct* parametrization can be tackled successfully by a first-order strategy. We may conclude that the *Hermite* parametrization is slightly better, but not dramatically superior on this simplified variant, especially due to the fact that it requires a second-order DES.

## 6.3   Investigation of Optimal Solutions

An optimal solution is represented by its control function $\phi(\omega)$, and the electric field respectively, but one can also examine the resulting revival structure of the molecules. **Only due to our simplified variant, i.e. $J_{initial}^{max} = 0$ at initialization, it is possible to study the *population* of the rotational levels as a function of time**. Otherwise, in the general case, all levels are initially populated, and a thermal averaging is applied. Explicitly, the wavefunction can be expressed as a superposition of those levels,

$$\psi = \sum_j a_j^{(t)} \cdot |j\rangle \cdot e^{-i\frac{E_j t}{\hbar}}. \tag{11}$$

The expectation of the cosine-squared alignment (the objective function, or the *physical observable*) is calculated directly from these complex amplitudes $a_j^{(t)}$, whereas the *population* of the rotational levels is $\left| a_j^{(t)} \right|^2$. This population of rotational levels can be analyzed in a fairly simple technique, known as the *Sliding Window Fourier Transform* (SWFT), which provides us with a **powerful visual tool**. Given the revival structure of an optimal solution, a sliding time window is Fourier transformed, to produce the frequency picture through the alignment process. This windowing creates a transformation which is localized in time. Due to the *quantization* of the rotational levels, only certain frequencies (or *energy* levels, respectively) are expected to appear.

12

**Analysis Results**

We applied the SWFT routine to the optimal solutions which were obtained in the various runs. Fig. 15, 16, 17 and 18 visualize the typical population process of the rotational levels for four typical solutions of the different optimization procedures (2 parametrization *times* 2 DES variants). The observed quantum energy levels are indeed as expected from theory.

The results reveal two different conceptual physical structures, which correspond to optimal and sub-optimal solutions in terms of the physical observable's yield, i.e., the cosine-squared alignment. The Direct-DR2 as well as the Hermite-CMA procedures obtain the best solutions, which share the same structure - they are characterized by the dominant population of the $4^{th}$ quantum energy level in the SWFT picture. On the other hand, the Direct-CMA and Hermite-DR2 procedures obtain inferior solutions, which are characterized by a gradually increasing population of the energy levels.

The original *revival structures* for two optimal solutions, representing the two conceptual structures, are given in Fig. 13 and 14. The *optimal* family of solutions (Fig. 13) presents a dramatic revival structure, with a typical strong pulse in the train which lies on the boundary of the punished regime ($I \approx 0.36$). This pulse seems to be essential in giving the molecules the right 'kick', and most likely responsible for the dominant population of the $4^{th}$ quantum energy level in the SWFT picture. The *sub-optimal* family of solutions (Fig. 14) attains a revival structure with a smooth exponential envelope, and thus has a gradual building-up of the quantum energy levels in the SWFT picture, respectively. It typically contains a train of medium pulses and lacks a dominant one.

We would like to emphasize the fact that we obtained the same family of optimal solutions, representing a single quantum structure, from two different optimization approaches: The first employs a first-order DES subject to direct pixelation of the control phase, while the other employs a second-order DES subject to Hermite expansion of the control phase. We claim that this might suggest that the regime of the global optimum has been reached.

**Another Perspective to Optimality – Phasing-Up**

The coefficients $a_j^{(t)}$ are complex numbers, and as such can be expressed in their *polar* representation:

$$a_j^{(t)} = r_j^{(t)} \cdot e^{i\varphi_j^{(t)}} \tag{12}$$

To our best understanding, their absolute value is governed by the interaction strength (the peak Rabi frequency between the two electronic states $X$ and $A$, $\Omega_{XA}$). Those absolute values are the plotted bars in the *SWFT figures*.

We question, however, whether given a certain population - does the optimization routine produce the optimal set of phases $\varphi_j^{(t)}$? In order to answer this question, a simple optimization procedure was implemented in the following manner: It accepts the $a_j^{(t)}$ as input, and aims to optimize the phases $\varphi_j^{(t)}$ such that the cosine-squared alignment is maximized. Practically, it uses a subroutine from the general alignment code for the evaluation, and applies the CMA-ES (Section 3.4) for the tuning of the 10 phases. Note that a single function evaluation has the duration of $\approx 0.5$s.

We considered 50 different cases of highly-fit solutions to the alignment problem (all solutions have cosine-squared-alignment values in the regime of 0.95) - for each test case 100 independent optimizations were run, aiming to tune the phases. The experimental results are clear and sharp. They are presented at two levels:

1. In all 100 runs for all 50 test-cases - the best solution has always **synchronized phases**. There are different phase values per run, but it does not make a difference for the cosine-squared alignment, as long as the populated levels hold that same phase value. Explicitly, the Sigma-RMS of the phases is calculated:

$$\Delta\varphi^{optimal} = 0.0117$$

2. The 50 test-cases, as originally obtained by the original optimization prior to this optimization procedure, held phases which are not far from being synchronized,

$$\Delta\varphi^{DR2} = 0.0566$$

and indeed, the optimizations did not improve the cosine-squared alignment dramatically: always less than 1% improvement.

We consider this a very strong result - the evolutionary optimization routine managed to tackle the fine-tuning of the quantum control problem, behind the complex transformations and the so-called Schrödinger black-box.

Alignment and Revival Structure of two obtained solutions (Fig. 13 and 14). Thin red line: alignment; thick black line: intensity of the laser pulse.
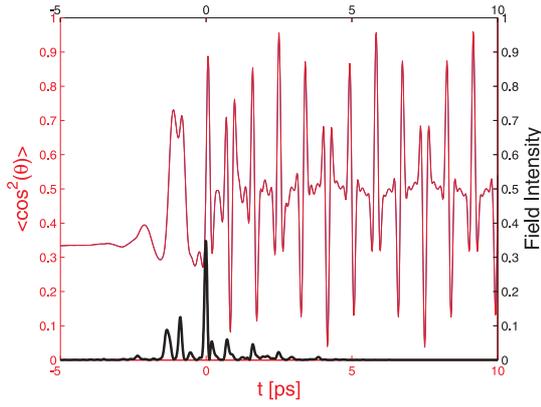


Figure 13: A typical optimal solution, obtained by the DR2-plain; Alignment yield: $< \cos^2(\theta) >= 0.9622$.
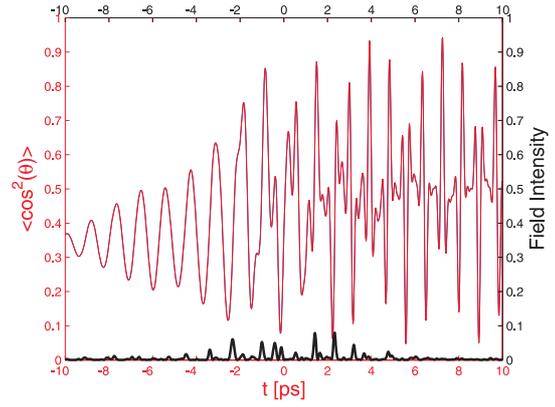
Figure 14: A typical sub-optimal solution, obtained by the CMA-plain: A smooth exponential envelope is observed; Alignment yield: $< \cos^2(\theta) >= 0.9505$.

Each of the SWFT figures (Fig. 15 - 18) represents a *Fourier transform* applied to the revival structures of the optimal solutions (the thin-red alignment curves of Fig. 13-14). The values are log-scaled, and represent how high the rotational levels of the molecules are populated as a function of time. Thus, an exponential envelope (Fig. 14) is represented by a gradual building-up of frequencies (Fig. 16). Note that the quality of the laser pulse cannot be measured in those plots.
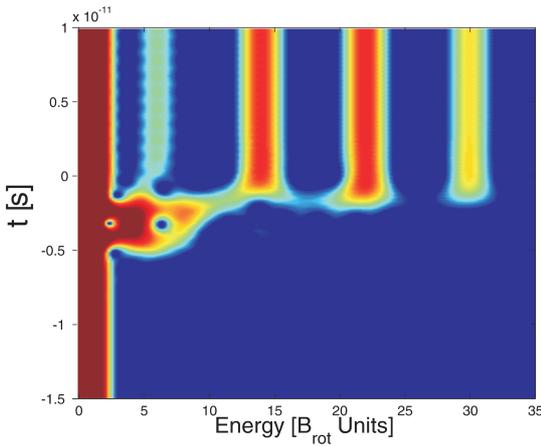


Figure 15: DR2 with plain-parametrization: The $4^{th}$ rotational level is mostly populated after the interaction.
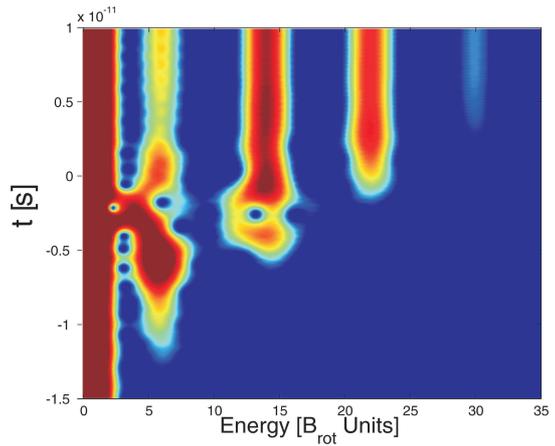
Figure 16: CMA with plain-parametrization: All five first rotational levels are populated gradually after the interaction.
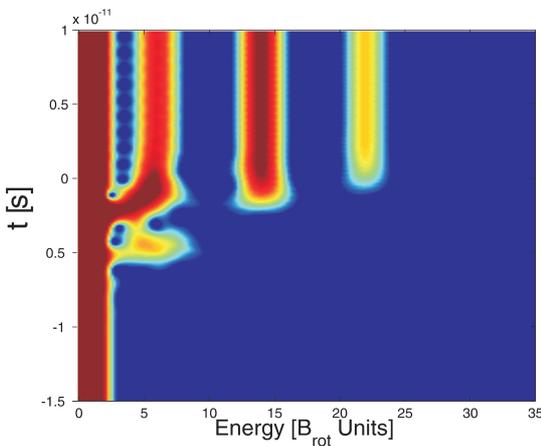


Figure 17: DR2 with Hermite-parametrization: The four first rotational levels are populated gradually after the interaction.
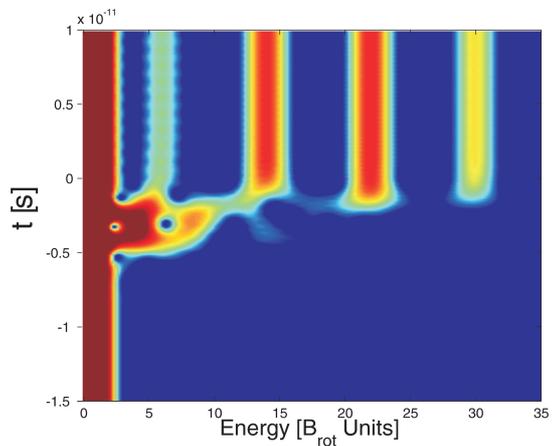
Figure 18: CMA with plain-parametrization: $4^{th}$ rotational level is mostly populated after the interaction.

14
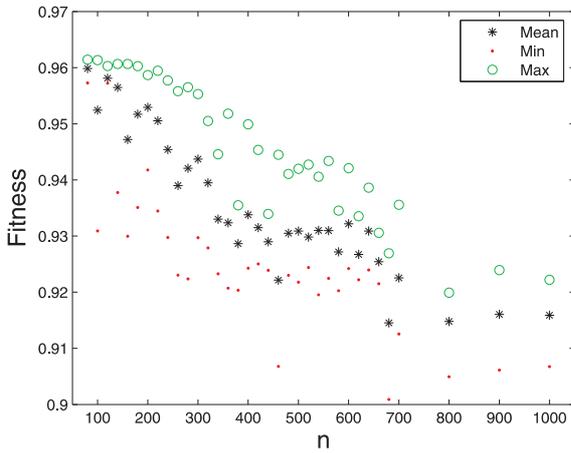
Figure 19: Best, mean and worst cosine-squared alignment values obtained by the DR2 for each parametrization, over 10 runs of $20,000$ function evaluations each.
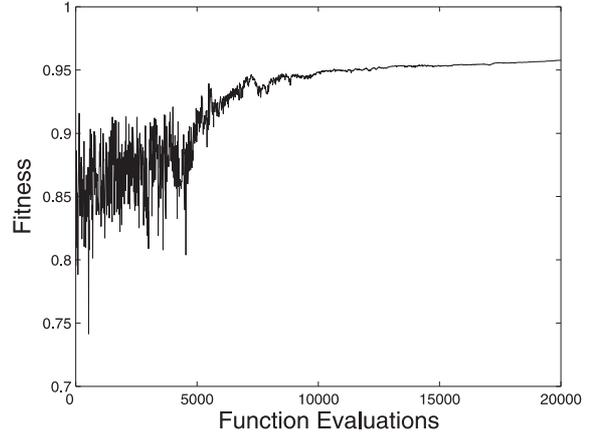


Figure 20: A typical DR2 evolution run for $n = 100$, with $20,000$ function evaluations. Successful learning is observed after $\approx 5,000$ evaluations.
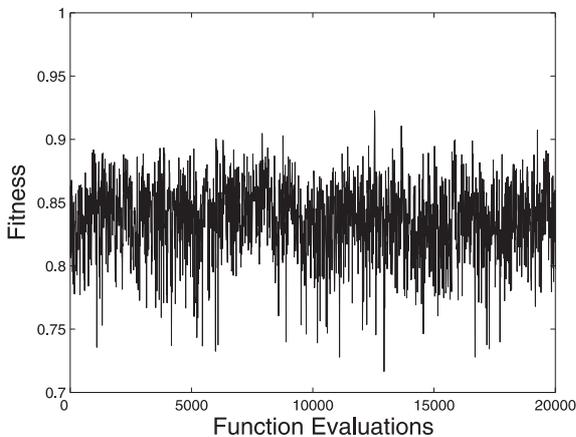


Figure 21: A typical DR2 evolution run for $n = 700$, with $20,000$ function evaluations. No successful learning is observed.
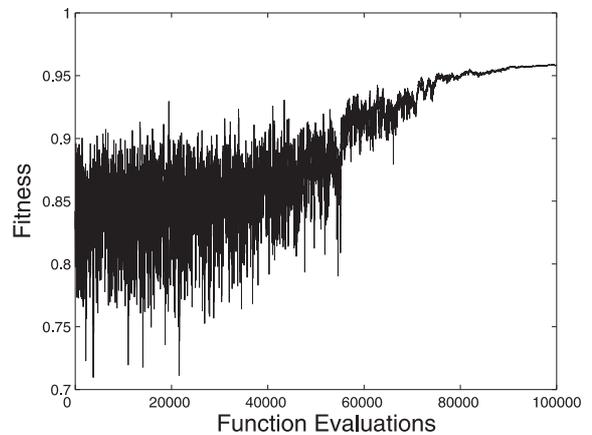


Figure 22: DR2 evolution run, for $n = 1000$, with $100,000$ function evaluations. The best cosine-squared alignment value found was $f^* = 0.9583$.

## 7  Scalability

In this section we aim at exploring the scalability of the alignment problem with respect to a specific ES variant (DR2). We are interested in studying the trade-off between the resolution of the control function, which allows fine-tuning of the electric field, to the success-rate of the evolutionary learning process, subject to a fixed number of function evaluations. We choose to conduct a scalability test on the simplified variant of zero-temperature.

We apply the DR2 algorithm to the optimization task in the following manner - 10 runs per discretization, with $n = 80, 100, 120, ..., 700$, as well as to the cases $n = 800, 900, 1000$. Each run is limited to $20,000$ function evaluations. Fig. 19 presents the numerical results of these calculations. The best, mean and worst fitness values obtained by the DR2, after $20,000$ function evaluations, for each discretization, are plotted.

As can be observed, the best fitness value is attained for $n = 80, 100$; as the dimension $n$ increases, there seems to be a weak trend of fitness values decrease, but the DR2 still manages to find optimal solution in the regime of 0.94 even for $n = 400$. A typical evolution run for $n = 100$ is given as Fig. 20. As can be observed from this plot, a successful learning is obtained after $\approx 5,000$ function evaluations. In higher dimensions, i.e. $n \geq 500$, the DR2 does not manage to tackle the problem within the limited number of function evaluations. A typical run for $n = 700$ is given as Fig. 21.

**Granting Additional Function Evaluations**  Given the numerical results of the previous section, we were interested in the question whether the fixed number of function evaluations limits the search and does not allow a successful learning of the decision parameters and a convergence into a good solution. We have conducted a series of runs, now limited to $100,000$ function evaluations, for the extreme case of $n = 1000$. Surprisingly

to some extent, we found that some of the runs did manage to converge successfully into fine solutions of high fitness values. In particular, we would like to mention a run which found a solution with cosine-squared alignment value of $f^* = 0.9583$, a value which is close to the highest value known to us for this variant of the problem. The plot of that specific evolution run is given as Fig. 22. This plot reveals that the DR2 manages to 'take-off' into a convergence pathway only after $\approx 50,000$ function evaluations, and then it needs further $30,000$ function evaluations to reach saturation.

This numerical result indicates that the learning task of the decision parameters in this problem is still feasible in higher dimensions, as long as the granted number of function evaluations is sufficiently increased. Clearly, the specific Evolution Strategy in use has tackled successfully the 1000-dimensional problem from the optimization perspective.

**However, from the *physics* perspective, such a high-resolution parametrization does not seem to pay-off, as far as the cosine-squared observable is concerned, and there seems to be no justification for allowing higher discretization of the control function.**

# 8    Conclusions and Outlook

We would like to briefly summarize this manuscript. We have examined the application of specific ES variants to a specific use-case representing computationally expensive QC simulations, namely the molecular alignment problem. To this end, we have numerically optimized the alignment of a thermal ensemble of rigid rotors by applying pulse shaping techniques resembling the common experimental procedure of femtosecond pulse shaping (QCE). In a comparison of the different variants, we observed that a first-order technique, the so-called DR2 algorithm, outperformed the CMA-ES, which is a second-order technique. We offer the following explanation. QCT proves [32] that "*given a controllable theoretical quantum system, there is always a trap-free pathway up to the top of the control landscape from any location, allowing locating an optimal control with first-order (gradient) information*". Even though this theorem is valid for unconstrained theoretical QC landscapes, it may as well suggest that first-order information is sufficient to optimize other QC landscapes.

We developed a general method for the learning task of any parametrized curve by means of complete-basis functions. We applied a couple of preliminary tests with respect to the application of the method to the alignment problem. The so-called $\alpha$-test has succeeded in justifying this method and did eliminate the possibility of convergence into a trivial attractor. The *initial states density test* contributed a perspective of the nature of the various parametrization with respect to the given objective function. By performing a series of numerical simulations, with a fixed number of coefficients, we managed to show that our method qualified as a robust optimization method, and in particular achieved a significant convergence speeding-up for the given problem. The obtained solutions were at least as good as any other solutions known to us, but were achieved faster. From the algorithmic perspective, they required a second-order optimization routine, such as the CMA-ES.

Finally, we considered a simplified variant of the original problem, in order to further investigate the nature of optimal solutions. By applying two optimization routines, subject to two different parametrization, two typical physics structures were revealed, corresponding to an optimal and sub-optimal sets of solutions. The combination of algorithms and parametrization resulted in a sharper observation of optimality. This confirmed the multi-modality of the search space, as predicted by theory [33] and provided us with strong physics intuition with respect to the shape of optimal pulses. Our new observation suggests that the regime of optimal solutions has been found.

The combination of noise-free numerical modeling with specific EAs and parametrization accomplished a successful interdisciplinary study with fruitful results in both sides. The absence of noise was a blessing in this context, and we believe that it resulted an important physical view of QCT with some QCE characteristics. Upon transferring these algorithms to the laboratory for the sake of experimental optimization of QC systems [34], it is likely that additional study of robustness to noise will be required.

# Appendix: Complete-Basis Functions

Here is a brief summary of the fundamental mathematical concepts behind the Complete-Basis-Functions Parametrization, as presented in section 5. This part is mainly based on Abramowitz [35] and Kaplan [36].
Let $f(x)$ be given in the interval $a \leq x \leq b$, and let

$$\xi_1(x), \xi_2(x), \ldots, \xi_k(x), \ldots \tag{13}$$

be functions which are all piecewise continuous in this interval.

The set $\{\xi_k(x)\}_{k=1}^{\infty}$ is called *complete* if it can span any piecewise continuous function $f(x)$, e.g.,

$$f(x) = \sum_{k=1}^{\infty} c_k \xi_k(x), \tag{14}$$

where the coefficients $c_k$ are given by:

$$c_k = \frac{1}{B_k} \int_a^b f(x)\, \xi_k(x)\, dx, \quad B_k = \int_a^b [\xi_k(x)]^2\, dx \tag{15}$$

The convergence is guaranteed by the so-called *completeness theorem*. Explicitly, the series

$$R_m = \int_a^b \left( f(x) - \sum_{k=1}^m c_k \xi_k(x) \right)^2 dx = \int_a^b (f(x) - S_m(x))^2\, dx \tag{16}$$

converges to *zero* for sufficiently large $m$:

$$\lim_{m \to \infty} R_m = 0, \tag{17}$$

where we denoted the sequence of *partial sums* as $S_m(x)$:

$$S_m(x) = \sum_{k=1}^m c_k \xi_k(x) \tag{18}$$

By definition, the convergence of the series of functions is equivalent to the convergence of $S_m(x)$ to $f(x)$:

$$\lim_{m \to \infty} S_m(x) = f(x) \tag{19}$$

## The Fourier (Trigonometric) Series

A *trigonometric series* is an expansion of a periodic function in terms of a sum of *sines* and *cosines*, making use of the orthogonality property of the harmonic functions. Without loss of generality, let us consider from now on the interval $[0, L]$. Let $f(x)$ be a single-valued function defined on that interval, then its *trigonometric series* or *trigonometric expansion* is given by:

$$\tilde{f}(x) = \frac{1}{2} a_0 + \sum_{k=1}^{\infty} a_k \cos\left( \frac{2\pi k}{L} \cdot x \right) + \sum_{k=1}^{\infty} b_k \sin\left( \frac{2\pi k}{L} \cdot x \right) \tag{20}$$

If the coefficients $a_k$ and $b_k$ satisfy certain conditions, then the series is called a *Fourier series*.
If $f(x)$ is periodic with period $L$, and has continuous first and second derivatives for all $x$ in the interval, it is guaranteed that the trigonometric series of $f(x)$ will converge uniformly to $f(x)$ for all $x$; This is referred to as satisfying the *Dirichlet* conditions. We shall refer in this study to the *trigonometric* series as the *Fourier* series.

## Other Sets of Functions

If one is indeed interested in periodic functions, there is no natural alternative but using the trigonometric series. However, if one is concerned with other representations of a general function over a given interval, a great variety of other sets of functions is available, e.g.:

- **Legendre polynomials**, $P_k(x)$:

$$
\begin{aligned}
P_k(x) = \frac{(2k-1)(2k-3)\cdots 1}{k!} &\left\{ x^k - \frac{k(k-1)}{2(2k-1)} x^{k-2} + \right. \\
&\left. + \frac{k(k-1)(k-2)(k-3)}{2 \cdot 4(2k-1)(2k-3)} x^{k-4} - \cdots \right\}
\end{aligned}
\tag{21}
$$

which can also be defined via Rodrigues' formula:

$$P_0(x) = 1 \quad P_k(x) = \frac{1}{2^k k!} \frac{d^k}{dx^k} \left( x^2 - 1 \right)^k, \ \ k = 1, 2, \ldots \tag{22}$$

If $f(x)$ satisfies the *Dirichlet conditions* mentioned earlier, then there will exist a Legendre series expansion for it in the interval $-1 < x < 1$.
For illustration, the first 10 *Legendre polynomials* are plotted in Fig. 23.

- **Bessel Function of the First Kind and of Order** $l$, $J_l(x)$:

$$J_l(x) = \sum_{k=0}^{\infty} \frac{(-1)^k\, x^{l+2k}}{2^{l+2k} \cdot k! \cdot \Gamma(l+k+1)} \tag{23}$$

with $\Gamma(\alpha)$ as the Gamma Function. Given a fixed $l \geq 0$, the functions $\{ \sqrt{x} J_l(\lambda_{lk} x) \}_{k=1}^{\infty}$ form an orthogonal complete system over the interval $0 \leq x \leq 1$.

17
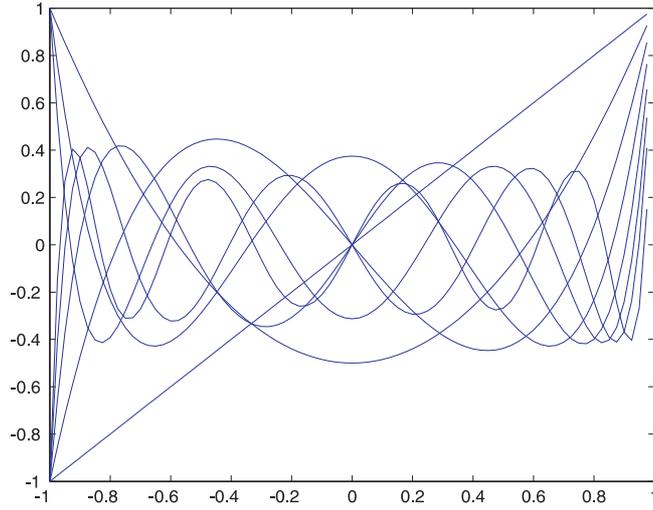
Figure 23: The First 10 *Legendre* Polynomials.

- **Hermite polynomials**, $H_k(x)$:

$$H_k(x) = (-1)^k \exp\left\{x^2\right\} \frac{d^k}{dx^k} \left(\exp\left\{-x^2\right\}\right), \ k = 0, 1, \ldots \tag{24}$$

The *Hermite polynomials* form a complete set of functions over the infinite interval $-\infty < x < \infty$, with respect to the weight function $\exp\left(-\frac{1}{2}x^2\right)$.

- **Chebyshev polynomials of the First Kind**, $T_k(x)$:

$$T_k(x) = \frac{k}{2} \sum_{r=0}^{\lfloor k/2 \rfloor} \frac{(-1)^r}{k-r} \binom{k-r}{r} (2x)^{k-2r}, \ k = 0, 1, \ldots \tag{25}$$

The *Chebyshev polynomials of the First Kind* form a complete set of functions over the interval $[-1, 1]$ with respect to the weight function $\frac{1}{\sqrt{1-x^2}}$.

## Higher Dimensions

An expansion by means of a complete set of functions can be generalized for higher dimensions. For illustration, let us consider the two-dimensional case of the *trigonometric series*. The functions $\cos(\frac{2\pi k}{L} \cdot x) \cdot \cos(\frac{2\pi l}{L} \cdot y)$, $\sin(\frac{2\pi k}{L} \cdot x) \cdot \cos(\frac{2\pi l}{L} \cdot y)$, $\cos(\frac{2\pi k}{L} \cdot x) \cdot \sin(\frac{2\pi l}{L} \cdot y)$, and $\sin(\frac{2\pi k}{L} \cdot x) \cdot \sin(\frac{2\pi l}{L} \cdot y)$ form an orthonormal complete system of functions in the box $[(0,0), (0,L), (L,0), (L,L)]$. Given a function in that domain, $f(x,y)$, its expansion can then be written in the form:

$$
\begin{aligned}
f(x,y) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \lambda_{kl} \cdot \Bigg\{ & a_{kl} \cos(\frac{2\pi k}{L}x) \cos(\frac{2\pi l}{L}y) + \\
+ b_{kl} \sin(\frac{2\pi k}{L}x) \cos(\frac{2\pi l}{L}y) & + c_{kl} \cos(\frac{2\pi k}{L}x) \sin(\frac{2\pi l}{L}y) + \\
+ & d_{kl} \sin(\frac{2\pi k}{L}x) \sin(\frac{2\pi l}{L}y) \Bigg\}
\end{aligned}
\tag{26}
$$

## Corollary

An infinite series of complete basis functions converges to any "reasonably well behaving" function. Hence, it is straightforward to approximate a given function with a finite series of those functions, i.e., by cutting its tail from a certain point. In principle, the sum $S_m(x)$ (Eq. 18) can always be found to a desired degree of accuracy by adding up enough terms of the series. For practical applications, the corollary is that every function can be approximated using a series of complete basis functions, to whatever desired or practical accuracy. Moreover, this corollary can be easily generalized to any desired dimension.

# References

[1] T. Bäck, *Evolutionary Algorithms in Theory and Practice.* New York, NY, USA: Oxford University Press, 1996.

[2] T. Bäck, C. Foussette, and P. Krause, *Contemporary Evolution Strategies.* Natural Computing Series, Springer-Verlag Berlin Heidelberg, 2013.

[3] M. Emmerich, O. M. Shir, and H. Wang, *Handbook of Heuristics*, ch. Evolution Strategies, pp. 1–31. Cham: Springer International Publishing, 2018.

[4] D. Tannor and S. Rice, "Control of Selectivity of Chemical Reaction via Control of Wave Packet Evolution," *Chem. Phys.*, vol. 83, 1985.

[5] P. Brumer and M. Shapiro, "Control of Unimolecular Reactions using Coherent Light," *Chem. Phys. Lett.*, vol. 126, no. 6, 1986.

[6] R. S. Judson and H. Rabitz, "Teaching Lasers to Control Molecules," *Phys. Rev. Lett.*, vol. 68, pp. 1500–1503, Mar 1992.

[7] W. S. Warren, H. Rabitz, and M. Dahleh, "Coherent Control of Quantum Dynamics: The Dream Is Alive," *Science*, vol. 259, pp. 1581–1589, Mar 1993.

[8] H. Rabitz, R. de Vivie-Riedle, M. Motzkus, and K. Kompa, "Whither the Future of Controlling Quantum Phenomena?," *Science*, vol. 288, pp. 824–828, May 2000.

[9] P. Nuernberger, G. Vogt, T. Brixner, and G. Gerber, "Femtosecond Quantum Control of Molecular Dynamics in the Condensed Phase," *Phys Chem Chem Phys.*, vol. 9, no. 20, pp. 2470–2497, 2007.

[10] A. P. Peirce, M. A. Dahleh, and H. Rabitz, "Optimal Control of Quantum-Mechanical Systems: Existence, Numerical Approximation, and Applications," *Phys. Rev. A*, vol. 37, Jun 1988.

[11] S. Shi and H. Rabitz, "Quantum Mechanical Optimal Control of Physical Observables in Microsystems," *Chem. Phys.*, vol. 92, Jan 1990.

[12] H. Stapelfeldt and T. Seideman, "Colloquium: Aligning Molecules with Strong Laser Pulses," *Rev. Mod. Phys.*, vol. 75, pp. 543–557, Apr 2003.

[13] B. Friedrich and D. Herschbach, "Steric Proficiency of Polar 2-Sigma Molecules in Congruent Electric and Magnetic Fields," *Phys. Chem. Chem. Phys.*, vol. 2, pp. 419–428, 2000.

[14] N. Hay, R. Velotta, M. Lein, R. de Nalda, E. Heesel, M. Castillejo, and J. P. Marangos, "High-Order Harmonic Generation in Laser-Aligned Molecules," *Phys. Rev. A*, vol. 65, p. 053805, Apr 2002.

[15] T. Seideman, "Revival Structure of Aligned Rotational Wave Packets," *Phys. Rev. Lett.*, vol. 83, pp. 4971–4974, Dec 1999.

[16] M. Leibscher, I. S. Averbukh, and H. Rabitz, "Molecular Alignment by Trains of Short Laser Pulses," *Phys. Rev. Lett.*, vol. 90, p. 213001, May 2003.

[17] M. Leibscher, I. S. Averbukh, and H. Rabitz, "Enhanced Molecular Alignment by Short Laser Pulses," *Phys. Rev. A*, vol. 69, no. 1, p. 013402, 2004.

[18] C. Z. Bisgaard, M. D. Poulsen, E. Peronne, S. S. Viftrup, and H. Stapelfeldt, "Observation of Enhanced Field-Free Molecular Alignment by Two Laser Pulses," *Phys. Rev. Lett.*, vol. 92, no. 17, 2004.

[19] K. F. Lee, I. Litvinyuk, P. Dooley, M. Spanner, D. Villeneuve, and P. Corkum, "Two-Pulse Alignment of Molecules," *J. Phys. B.*, vol. 37, no. 3, pp. L43–L48, 2004.

[20] J. G. Underwood, B. J. Sussman, and A. Stolow, "Field-Free Three Dimensional Molecular Axis Alignment," *Phys. Rev. Lett*, vol. 94, no. 14, p. 143002, 2005.

[21] C. M. Dion, A. Keller, and O. Atabek, "Optimally Controlled Field-Free Orientation of the Kicked Molecule," *Phys. Rev. A*, vol. 72, no. 2, p. 023402, 2005.

[22] F. Rosca-Pruna and M. J. Vrakking, "Revival Structures in Picosecond Laser-Induced Alignment of I2 Molecules," *Journal of Chemical Physics*, vol. 116, no. 15, pp. 6579–6588, 2002.

[23] O. M. Shir, C. Siedschlag, T. Bäck, and M. J. Vrakking, "Evolutionary Algorithms in the Optimization of Dynamic Molecular Alignment," in *Proceedings of the 2006 IEEE World Congress on Computational Intelligence*, pp. 9817–9824, IEEE Computational Intelligence Society, 2006.

[24] A. Auger and N. Hansen, "Performance Evaluation of an Advanced Local Search Evolutionary Algorithm," in *Proceedings of the 2005 Congress on Evolutionary Computation CEC-2005*, (Piscataway, NJ, USA), pp. 1777–1784, IEEE Press, 2005.

[25] N. Hansen and A. Ostermeier, "Completely Derandomized Self-Adaptation in Evolution Strategies," *Evolutionary Computation*, vol. 9, no. 2, pp. 159–195, 2001.

[26] A. Ostermeier, A. Gawelczyk, and N. Hansen, "Step-Size Adaptation Based on Non-Local Use of Selection Information," in *Parallel Problem Solving from Nature - PPSN III*, vol. 866 of *Lecture Notes in Computer Science*, pp. 189–198, Springer, 1994.

[27] A. Ostermeier, A. Gawelczyk, and N. Hansen, "A Derandomized Approach to Self Adaptation of Evolution Strategies," *Evolutionary Computation*, vol. 2, no. 4, pp. 369–380, 1994.

[28] N. Hansen and S. Kern, "Evaluating the CMA Evolution Strategy on Multimodal Test Functions," in *Parallel Problem Solving from Nature - PPSN VIII*, vol. 3242 of *Lecture Notes in Computer Science*, (Berlin Heidelberg), pp. 282–291, Springer-Verlag, 2004.

[29] A. F. Linskens, I. Holleman, N. Dam, and J. Reuss, "Two-Photon Rabi Oscillations," *Phys. Rev. A*, vol. 54, pp. 4854–4862, Dec 1996.

[30] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*. Philadelphia, PA, USA: SIAM, 2005.

[31] K.-S. Leung and Y. Liang, "Evolution Strategies with a Fourier Series Auxiliary Function for Difficult Function Optimization," in *IDEAL*, vol. 2690 of *Lecture Notes in Computer Science*, pp. 303–312, Springer, 2003.

[32] T.-S. Ho and H. Rabitz, "Why do Effective Quantum Controls Appear Easy to Find?," *Journal of Photochemistry and Photobiology A: Chemistry*, vol. 180, Jun 2006.

[33] M. Demiralp and H. Rabitz, "Optimally Controlled Quantum Molecular Dynamics: A Perturbation Formulation and the Existence of Multiple Solutions," *Phys. Rev. A*, vol. 47, pp. 809–816, Feb 1993.

[34] O. M. Shir and T. Bäck, "Sequential experimentation by evolutionary algorithms," in *Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion*, GECCO '20, (New York, NY, USA), p. 957–974, Association for Computing Machinery, 2020.

[35] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. National Bureau of Standards, Applied Math. Series 55, Dover Publications, 1965.

[36] W. Kaplan, *Advanced Calculus*. Reading, MA, USA: Addison-Wesley, third ed., 1983.