Chapter 9

Dynamic Molecular Alignment

The Quantum Control application to dynamic molecular alignment [153, 154] is of considerable interest because of its many practical consequences. For instance, many chemical and physical processes, ranging from bimolecular reactions [155] to high harmonic generation [156], are directly influenced by the angular distribution of the molecular sample. Furthermore, in many fundamental molecular dissociation or ionization experiments the interpretation of the collected data will become more efficient if the molecules are aligned with respect to a certain axis. Hence, techniques to generate molecular alignment are needed in practice.

Achieving molecular alignment can be classified into two possible modes:

1. **Pendular State** When the envelope of the field changes slowly compared to the timescale of molecular rotation, typically 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 turned on will vanish once it is turned off again. Thus, such experiments cannot be carried out subject to field-free conditions.

2. **Impulsive Alignment** Here, the duration of the applied pulses is much shorter than a rotational period [157]. A wavepacket of rotational states is constructed such that *field-free alignment* can be considerably attained.

Both modes aim at constructing a superposition of as many angular momentum eigenstates as possible. Due to the *uncertainty principle*, a broad distribution in *angular momentum* corresponds to a narrow distribution of the *angular position*. However, it is important to note that both the amplitudes and the relative phases of the composite rotational states have to be under control in order to achieve alignment. This requirement is fulfilled for the pendular state case, since it is an eigenstate of the combined molecule-field
Hamiltonian. However, in the general case, a randomly phased superposition of rotational states will not interfere favorably in attaining molecular alignment.

For the impulsive case, the evolution of the total wavefunction (after the electric field is turned off) repeats with the revival time

\[ T_{\text{rev}} = \frac{1}{2B_{\text{rot}}c} \]  

(9.1)

where \( B_{\text{rot}} \) is the rotational constant of the molecule and \( c \) is the speed of light. Partial revivals can be observed at \( T_{\text{rev}}/2 \) and, possibly, at \( T_{\text{rev}}/4 \), 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: Low field intensities should be applied in order to avoid a scenario in which the molecules are ionized. This aspect also plays a role in keeping the numerical modeling consistent in describing the molecule as a rigid rotator, as discussed in Chapter 8. Therefore, one would like to achieve high alignment while keeping the peak laser intensity as low as possible.

On that note, recent publications have focused on finding pulse shapes other than the FTL pulse that create a high degree of alignment. Leibscher et al. [158, 159] have theoretically shown that in the nonperturbative regime a train of pulses lead to better alignment than a single FTL pulse. For asymmetric molecules, orientation has been found to be optimized by a sequence of kicks as well [160].

Such pulse sequences can be easily constructed and also optimized with respect to the relatively small number of their control parameters. Therefore, they provide an attractive starting point for more complex optimization schemes, where the electric field is defined by a considerably larger number of control parameters. The task of obtaining high-quality solutions in this high-dimensional search space is nontrivial, already when considering only the ground state in the initial distribution. For finite temperatures, the alignment optimization has to be performed simultaneously for a set of initial rotational states, which, together with the large number of electric field control parameters poses a challenging optimization problem.

### 9.1 Numerical Modeling

The numerical modeling of the rotational framework, as presented in Chapter 8, is adopted here fully. The remaining task is the definition of the alignment observable.
9.1. Numerical Modeling

The alignment calculation uses the following components in our basis:

\[
\langle JM | \cos^2 \theta | JM \rangle = \frac{1}{3} + \frac{2}{3} \left( \frac{J(J+1) - 3M^2}{(2J+3)(2J-1)} \right)
\]

\[
\langle JM | \cos^2 \theta | J+2 M \rangle = \frac{1}{2J+3} \sqrt{(J + M + 2)(J + M + 1)(J - M + 2)(J - M + 1)}
\]

\[
\langle JM | \cos^2 \theta | J-2 M \rangle = \frac{1}{2J-1} \sqrt{(J + M)(J + M - 1)(J - M)(J - M - 1)}
\]

We consider a thermal ensemble of diatomic molecules undergoing irradiation at a finite temperature. The latter is set to \( T = 100 \) K, and implemented by means of a Boltzmann averaging which practically corresponds to the density matrix \( \rho \). The molecule under investigation has a rotational constant of \( B_{\text{rot}} = B_g = B_e = 5 \text{cm}^{-1} \). We set the Rabi peak frequency to \( \Omega_{ge} = 180 \times 10^{12} \text{s}^{-1} \).

For the sake of attaining high molecular alignment while keeping the peak field intensity as low as possible, due to the rigid rotator approximation, we introduce a constraint to the optimization procedure, by means of a punishment term to pulses that are too intense. It explicitly reads

\[
I_p = \int E^2(t) \Theta(E^2(t) - I_{\text{thr}}) \, dt \tag{9.3}
\]

with \( \Theta(x) \) as the Heaviside step function.

Thus, the fitness function assigned to a candidate pulse shape is defined by

\[
F = \max_{E(t)} \left( \cos^2 \theta \right) - \beta I_p. \tag{9.4}
\]

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

Figure 9.1 provides an illustrative overview of the numerical process.

9.1.1 Numerical Simulations: Technical Details

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

- In the absence of a laser field, a random phase should yield 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.
Figure 9.1: 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. 6.29) 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 quantum dynamics numerical modeling. The revival structure (right) is the observed simulated behavior of the molecules, upon which the yield value is based.

- The punishment term, as introduced in Eq. 9.3 and in Eq. 9.4, can yield fitness values below the value of 0.4. The probability of a randomly generated pulse, with no specific parameterization, to get penalized is extremely low.

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

- Due to the heavy computational cost of a single simulator evaluation, we are limited in granting function evaluations. We are thus encouraged to employ optimization routines with minimal settings. Moreover, we shall apply experiments with a low number of repetitions.

9.2 Experimental Procedure

In order to preliminarily assess the performance of the algorithms on the given problem, we have conducted 10 independent runs for each of the de-randomized ES comma-variants with the goal of optimizing the alignment of a sample of generic diatomic molecules undergoing irradiation by a shaped femtosecond laser. We limit each run to 10,000 function evaluations, due to the computational cost of the simulator.
9.2. Experimental Procedure

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>DR1</th>
<th>DR2</th>
<th>DR3</th>
<th>CMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVG-Fitness</td>
<td>0.6399</td>
<td>0.6789</td>
<td>0.6534</td>
<td>0.6261</td>
</tr>
</tbody>
</table>

Table 9.1: Dynamic molecular alignment: Attained fitness values, averaged over 10 runs, for the DES comma variants.

9.2.1 First Numerical Results: Comparison of the Algorithms

Table 9.1 summarizes the numerical results of the runs - the averaged fitness value obtained by each optimization routine. Based on our experience with the problem and the algorithms, the yield differences of Table 9.1 are believed to be significant. Moreover, due to the limited number of simulations we do not provide further statistical analysis of the results.

Roughly speaking, the algorithms were observed to perform equally well, with the exception of the DR2 algorithm that managed to obtain a significantly better optimum than the others. While the DR3 algorithm showed the fastest initial fitness increase, it seemed to get stuck in a sub-optimal local trap after $\approx 2,000$ function evaluations. We have found this behavior to be typical for the DR3 algorithm.

The ranking of the algorithms was qualitatively similar for a number of alignment optimization runs employing different parameter settings. Figure 9.2 presents the best pulse-shape solution attained, as obtained by the DR2 routine.

9.2.2 The Complete-Basis-Functions Parameterization

In this section we present a new method for learning a function, based on a representation transformation, which can also be referred to as parameterization. The so-called Complete-Basis-Functions Parameterization was originally derived for the sake of learning the control function of the dynamic alignment problem, i.e. the phase $\phi(\omega)$, but is a general method for learning a generic $n$-variable function. It can reduce the dimensionality of the search space and possibly boost the convergence speed, respectively, as will be explained in detail.

Appendix B provides the reader with the mathematical background on complete-basis functions, and presents the specific functions that are considered in our study. For the sake of consistency and reading clarity, we specify here our notation for a spanned target function $f(x)$:

$$f(x) = \sum_{k=1}^{K_{\text{max}}} c_k \xi_k(x)$$

with $c_k$ as the expansion coefficients, and $\{\xi_k(x)\}_{k=1}^{\infty}$ as the the set of complete-basis functions.
Preliminary: Expanding a Known Function  As we will demonstrate here, finding the expansion 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 evolutionary algorithm, 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 using the Fourier series:

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

This task can be generalized to functions of higher dimensions, and by using other expansions of complete-basis functions. Following the notation of Appendix B, consider a finite number of the expansion coefficients of the cosine and sine functions, \( \{a_k\}_{k=0}^{K_a}, \{b_k\}_{k=1}^{K_b} \), as the decision parameters to be optimized by the evolutionary search. As a preliminary task in this study, we found that the standard-ES (Schwefel approach) converged easily and quickly to the correct coefficients. This elementary fitting problem was simply defined by means of the square-error minimization: The fitness, subject to minimization, was defined respectively as the root-mean-square error function between the original function and its evolving expansion.

Figure 9.3 presents the outcome of learning the triangle function with the standard-ES, using only the first 20 frequencies (\( K_{\text{max}} = K_a + K_b = 40 \))
Figure 9.3: Learning the triangle function by means of the first 20 Fourier frequencies. The plot shows the original triangle function and its Fourier approximation.

of a Fourier series as building blocks for a given function discretization of \( N = 100 \).

**Proposed Method: Learning an Unknown 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, represented by \( N \) function values, as in our quantum control alignment problem. The inspiration for this method was the initial intuition to the alignment problem, which suggested that the control function should be periodic. Motivated by this intuition, we started to run simulations 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 extended the method to other sets of complete basis functions, and in particular to the sets of functions which are introduced in Appendix B: 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 \( K_{max} \), where preferably \( K_{max} \ll N \). By that we can achieve a dramatic dimensionality reduction of the search space, aiming to boost the convergence speed. The idea is then to apply an evolutionary search to the \( n = K_{max} \)
coefficients of the expansion functions, where a simple transformation is applied for every fitness evaluation. In practice, the required time for additional computation of this transformation is negligible with respect to the objective function evaluation, in most real-world problems.

An ES employing a Fourier auxiliary function has been proposed in the past, known as the FES method [161]. The FES aims at approximating the fitness landscape, and particularly its small attraction basins, by means of the Fourier series. However, the careful reader should notice that our method is based on a different principle. It uses complete-basis functions as a transformation 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 continuous function - and this function is due to be approximated.

Preliminary Calculations

**Quadratic Phase Functions: The α-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 constructed phases:

$$\phi_\alpha(\omega) = \alpha \cdot (\omega - \omega_{\text{central}})^2,$$  \hspace{1cm} (9.5)

where the continuous parameter \( \alpha \) is scanned systematically in the interval \([0, 15]\). Note that these phases are constructed over \( n = 80 \) function values, and given as input to the dynamic alignment simulator as before.

The results of this so-called α-test are presented in Figure 9.4.

As can be clearly seen in the given plot, 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.

**The Initial States Density Test** We set the number of terms in each expansion to \( K_{\text{max}} = 40 \). The following preliminary experiment is meant to compare the natural initial quality of the different parameterizations with respect to the alignment problem. We applied a so-called initial states density test, a statistical fitness measurement of the initialized phase functions in the different parameterizations. For each parameterization in use, i.e., the direct/plain 80-dimensional random phase vector, or the random 40-dimensional coefficient vector for the various polynomials in use, we initialized 1,000 phase functions and calculated their mean fitness and standard deviation. The numerical results are visualized as histograms in Figures 9.5-
9.2. Experimental Procedure

![Graph showing fitness vs. alpha]

Figure 9.4: The $\alpha$-test: The fitness of quadratic phase functions, centered around the central frequency, as defined in Eq. 9.5.

<table>
<thead>
<tr>
<th>Routine</th>
<th>Direct Avg. Fit</th>
<th>Direct 0.6 Eval</th>
<th>Fourier Avg. Fit</th>
<th>Fourier 0.6 Eval</th>
<th>Legendre Avg. Fit</th>
<th>Legendre 0.6 Eval</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,10)-DR2</td>
<td>0.6789</td>
<td>2525</td>
<td>0.4494</td>
<td>N.A.</td>
<td>0.6384</td>
<td>629</td>
</tr>
<tr>
<td>(1,10)-CMA</td>
<td>0.6676</td>
<td>N.A.</td>
<td>0.4542</td>
<td>N.A.</td>
<td>0.6499</td>
<td>515.1</td>
</tr>
<tr>
<td>($\mu, \lambda$)-CMA</td>
<td>0.6261</td>
<td>4962.5</td>
<td>0.6171</td>
<td>4475.8</td>
<td>0.6466</td>
<td>194.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Routine</th>
<th>Bessel Avg. Fit</th>
<th>Bessel 0.6 Eval</th>
<th>Hermite Avg. Fit</th>
<th>Hermite 0.6 Eval</th>
<th>Chebyshev Avg. Fit</th>
<th>Chebyshev 0.6 Eval</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,10)-DR2</td>
<td>0.6299</td>
<td>1390</td>
<td>0.5944</td>
<td>5610</td>
<td>0.4843</td>
<td>N.A.</td>
</tr>
<tr>
<td>(1,10)-CMA</td>
<td>0.6229</td>
<td>2212.9</td>
<td>0.6755</td>
<td>271</td>
<td>0.4979</td>
<td>N.A.</td>
</tr>
<tr>
<td>($\mu, \lambda$)-CMA</td>
<td>0.6232</td>
<td>2719.5</td>
<td>0.6843</td>
<td>118</td>
<td>0.6225</td>
<td>3770.8</td>
</tr>
</tbody>
</table>

9.10, providing the fitness distributions of the various random initializations. See further discussion below.

**Parameterizations: Numerical Results**

In this section we present the numerical results for optimizing the dynamic alignment problem with the different parameterizations - the direct/plain parameterization versus the polynomial-based parameterizations with $K_{max} = 40$ terms. Our runs were based on the following algorithmic kernels:

1. (1,10)-DR2
Figure 9.5: Initial states density test for direct parameterization.

Figure 9.6: Initial states density test for Fourier parameterization.

Figure 9.7: Initial states density test for Legendre parameterization.

Figure 9.8: Initial states density test for Bessel parameterization.

Figure 9.9: Initial states density test for Hermite parameterization.

Figure 9.10: Initial states density test for Chebyshev parameterization.
### 9.2. Experimental Procedure

#### Table 9.3: Parameterizations: Summary of Best Results

<table>
<thead>
<tr>
<th>Parameterization</th>
<th>Best Fitness</th>
<th>0.6 Eval</th>
<th>Routine</th>
<th>Initial States Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct-Param</td>
<td>0.6899</td>
<td>2310</td>
<td>(1,10)-DR2</td>
<td>0.4026 ± 0.018</td>
</tr>
<tr>
<td>Fourier</td>
<td>0.6526</td>
<td>1411</td>
<td>(7,15)-CMA</td>
<td>0.4110 ± 0.019</td>
</tr>
<tr>
<td>Legendre</td>
<td>0.6487</td>
<td>106</td>
<td>(7,15)-CMA</td>
<td>0.3122 ± 0.075</td>
</tr>
<tr>
<td>Bessel</td>
<td>0.6457</td>
<td>6t</td>
<td>(7,15)-CMA</td>
<td>0.2218 ± 0.077</td>
</tr>
<tr>
<td>Hermite</td>
<td>0.6866</td>
<td>31</td>
<td>(7,15)-CMA</td>
<td>0.4558 ± 0.048</td>
</tr>
<tr>
<td>Chebyshev</td>
<td>0.6490</td>
<td>1051</td>
<td>(7,15)-CMA</td>
<td>0.4226 ± 0.023</td>
</tr>
</tbody>
</table>

2. (1,10)-CMA

3. $(\mu_W, \lambda)$-CMA: Following the recommended settings (Eq. 1.47): (7,15) for $n = 40$, versus (8,17) for $n = 80$.

The runs were limited to 10,000 function evaluations. We conducted 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. We consider the yield 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 9.2. Table 9.3 summarizes the best results obtained per parameterization.

#### Analysis and Discussion

An important result that should be pointed out is that all the runs in the various parameterizations 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 parameterizations with respect to the dynamic alignment optimization, as presented in Tables 9.2 and 9.3:

1. **Initial State** The *Hermite* parameterization has clearly the most natural initial representation for the phase function for the given problem, among the various cases, as reflected from the initial states density test.
results (Figures 9.5-9.10 and Table 9.3). Note that the Legendre as well as the Bessel parameterizations have low initial fitness values, even below the direct parameterization, due to the punishment effect. It should be stressed that the standard deviations of the different fitness distributions are reasonably low.

2. **Fitness Values** The Hermite parameterization obtained fitness values as high as the direct parameterization method, though by means of a different algorithm, as will be discussed shortly. As far as we know, the attained yield values in the regime of ≈ 0.69 are the highest cosine-squared alignment values which were ever attained for this particular configuration of the problem. Hence, from the optimization perspective, the proposed parameterization 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 algorithmic kernels. The DR2 obtained the best results for the direct parameterization, but obviously failed to deliver reasonable results for the polynomial-based parameterizations. In most cases, the DR2 does not even converge. The (7,15)-CMA, on the other hand, performed very well with the various polynomial-based parameterizations, and attained fine results also for the direct parameterization. The (1,10)-CMA is clearly inferior with respect to its rank-μ 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 parameterization, and may even introduce a barrier, to some degree, to the global search.

We would like to link this to the conclusions drawn for the QC landscapes of Two-Photon Processes in Chapter 7, where QC landscape analysis stating that first-order information is sufficient for optimizing QC landscapes was experimentally corroborated. The fact that the DR2 algorithm performs so well on the current dynamic alignment landscape, which is a combined OCT/OCE landscape, could be considered as an additional corroboration to this QC landscape analysis.

We shall further explore the performance of the DR2 versus CMA-ES with respect to the direct versus Hermite parameterizations in Section 9.3.

4. **Boosting Convergence Speed** An immediate conclusion from both tables is that the proposed method achieved a significant boost of the convergence speed for all the different polynomial-based parameteri-
zations, in comparison to the direct parameterization. The Hermite parameterization with the \((7, 15)\)-CMA is clearly the fastest routine, and it outperformed the other routines by far. It should be noted that the Legendre as well as the Bessel parameterizations, which have the lowest initial yield values, manage to compensate for that and reach the regime of good solutions (yield > 0.6) rather quickly.

Typical convergence profiles for Hermite versus direct parameterizations are plotted in Figure 9.11.

5. **Physics Interpretation** Aiming at gaining physics insights into the nature of highly-fit phase functions with respect to the alignment problem, we examined the nature of good solutions in the different parameterizations. The idea was to calculate the distributions of the coefficients, and try to identify dominance of certain components (frequencies in the Fourier case). Unfortunately, such dominance could not be identified within the results. The set of attained optimal phases reveals high complexity, which could not be tackled. This provides us with the motivation to explore a simpler variant of the alignment problem in Section 9.3.
Figure 9.12: Optimized pulses and alignment for $I_{thr} = 0.2 \cdot I_{FTL}$, $I_{thr} = 0.25 \cdot I_{FTL}$ and $I_{thr} = 0.3 \cdot I_{FTL}$. Figure courtesy of Christian Siedschlag [162].

<table>
<thead>
<tr>
<th>Intensity</th>
<th>$I_{FTL}$</th>
<th>0.2</th>
<th>0.25</th>
<th>0.3</th>
<th>0.36</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle \cos^2(\theta) \rangle$</td>
<td>0.662</td>
<td>0.673</td>
<td>0.6734</td>
<td>0.689</td>
<td></td>
</tr>
</tbody>
</table>

Table 9.4: Best $\langle \cos^2(\theta) \rangle$ values obtained with the DR2 algorithm over five runs for different values of $I_{thr}$ [162].

9.2.3 Further Investigation

We would like to review here briefly additional calculations for this alignment problem, which were carried out by Siedschlag and Vrakking (see, e.g., [162]).

**Punishment Strength** By decreasing $I_{thr}$, the search algorithm was shown to look for effective pulses with less available peak intensity. The numerical results of additional optimization runs, carried out by the DR2 algorithm, for $I_{thr} = 0.2 \cdot I_{FTL}$, $I_{thr} = 0.25 \cdot I_{FTL}$ and $I_{thr} = 0.3 \cdot I_{FTL}$ are presented in Table 9.4. Note that lowering $I_{thr}$ could slightly improve the attained alignment. Overall, the evolutionary search was able to make up for the smaller peak intensities by redistributing the fluence in a clever way, so to speak. The optimized pulse-shapes for the three lower threshold intensities are presented in Figure 9.12. The three solutions are observed to be remarkably similar.

**Constructed Pulse Trains** Siedschlag and Vrakking [162] also treated the question whether a simple train of pulses that is constructed by an appropriately designed phase function yields results that are comparable to those achieved by the evolutionary approach. In particular, the question ad-
Figure 9.13: A cut through the contourplot of Figure A.7 for $A = 2.26$, for which the largest alignment ($\langle \cos^2(\theta) \rangle = 0.589$) in the two-parameter approach under the condition $I < 0.36 \cdot I_{FTL}$ was achieved [162]. Figure courtesy of Christian Siedschlag.

dressed trains of pulses, which are generated by oscillatory phase functions. Explicitly, the following family of phases was considered:

$$\phi_{osc}(\omega) = A \cdot \sin(\omega \Delta + \alpha) \quad (9.6)$$

The two relevant parameters, $A$ and $\Delta$, were scanned in a search for the pulse that would produce the best alignment; Figure A.7 presents the outcome of that scan. The magnitude of $A$ controls the distribution of the available intensity over the peaks in the pulse train (and hence the peak intensity with respect to the FTL solution), while $\Delta$ corresponds directly to the time delay between two consecutive peaks. Note that the maximally obtained alignment yield in this scan was $A \approx 0.68$ and $\Delta = 1.7$ ps; However, its corresponding peak intensity was too high for the model, i.e., $I > 0.36 \cdot I_{FTL}$.

Figure 9.13 presents a cut of the contourplot scan of Figure A.7, at the maximally obtained yield in the allowed range (0.589). It was concluded in [162] that this approach was not flexible enough to adapt to the finer details of the time-dependent alignment response.

### 9.3 Investigation of Optimality: Zero Kelvin

Here we focus in a simplified variant of the original alignment problem, at zero temperature ($T = 0$ $K$) and with only a single rotational level at the
Chapter 9. Dynamic Molecular Alignment

initial distribution. The numerical modeling of Eq. 8.8 considers now \( M = 0 \) and reads:

\[
|\Psi(t)\rangle = \sum_{J=0}^{N_{rot}} \alpha_{J}^{(g)}(t) |gJ\rangle + \exp(-i\omega_{0}t) \alpha_{J}^{(e)}(t) |eJ\rangle \tag{9.7}
\]

The motivation for this simplification is to allow studying the physical characteristics of the optimal solutions, which would not have been possible for the general case, e.g., tracking the time-dependent population of the rotational levels, given only the ground level at initialization. From the technical perspective, this simplification reduces the simulator evaluation time to approximately 5s on a single P4-HT 2.6GHz processor.

We carried out calculations optimizing field-free molecular alignment starting from \( J = 0 \) for a number of algorithmic approaches and various Rabi peak frequencies. In each case, the same calculation was attempted by means of 20 runs. Each run was limited to 20,000 function evaluations. We restrict the discussion in this section to the best results obtained in each series of 20 trials.

Figure 9.14 presents a comparison between one optimization of dynamic alignment starting from \( J = 0 \), performed using the DR2 algorithm under perturbative conditions \( (\Omega_{ge} = 40 \times 10^{12}\text{s}^{-1}) \) and four optimizations performed under non-perturbative conditions \( (\Omega_{ge} = 160 \times 10^{12}\text{s}^{-1}) \) using both the DR2 and the CMA algorithms, with either a direct/plain parameterization of the phase or with the Hermite parameterization, employing the first \( K_{max} = 40 \) Hermite polynomials. Furthermore, based on our previous observations in this chapter, we employed \((1,10)\)-DR2 or \{(7,15), (8,17)\}-CMA (the latter depends on the parameterization used).

The obtained result at low laser intensity \( (\Omega_{ge} = 40 \times 10^{12}\text{s}^{-1}) \) is simple: A pulse train is observed where the spacing between the peaks is approximately the rotational period of a coherent superposition state consisting of \( J = 0 \) and \( J = 2 \) only \( (T_{rev2} = \frac{1}{6\hbar\omega_{rot}} = 1.1\text{ps}) \). The time-dependent intensity is given by a train of pulses where the largest pulse reaches an intensity of \( 0.36 \cdot I_{FTL} \).

The obtained pulse-shapes at high laser intensity \( (\Omega_{ge} = 160 \times 10^{12}\text{s}^{-1}) \), are considerably more complex and no simple periodicity can be observed. The averaged as well as largest values of \( \langle \cos^{2}(\theta) \rangle \) attained are shown in Table 9.5.

In consistency with the numerical results of the previous section, the highest alignment yield values attained for this particular system were also obtained by the DR2 with plain parameterization as well as by the CMA with Hermite parameterization. Employing the CMA with plain parameterization or the DR2 algorithm with the Hermite parameterization yields a slightly lower values over 20 trials. Based on our experience with the problem and the algorithms, the yield differences of Table 9.5 are believed to be significant.
Figure 9.14: (a) Comparison of an optimization performed employing the DR2 algorithm with $\Omega_{ge} = 40 \times 10^{12} \text{s}^{-1}$ and (b) Four calculations with $\Omega_{ge} = 160 \times 10^{12} \text{s}^{-1}$ employing the DR2 and the CMA algorithms with either a plain or Hermite parameterizations of the control phase function.

<table>
<thead>
<tr>
<th></th>
<th>(1,10)-DR2</th>
<th>${(7, 15), (8, 17)}$-CMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plain Param.</td>
<td>$0.9559 \pm 0.007$ (0.9622)</td>
<td>$0.9413 \pm 0.006$ (0.9508)</td>
</tr>
<tr>
<td>Hermite Param.</td>
<td>$0.9501 \pm 0.004$ (0.9570)</td>
<td>$0.9583 \pm 0.003$ (0.9618)</td>
</tr>
</tbody>
</table>

Table 9.5: Maximizing the cosine-squared field-free molecular alignment starting from $J = 0$ ($T = 0K$) at $\Omega_{ge} = 160 \times 10^{12} \text{s}^{-1}$ over 20 runs with 20,000 function evaluations per run; Mean and standard-deviation values are given, with the maximal value obtained in brackets.
This is supported by inspection of the pulse shapes shown in Figure 9.14. The two most successful optimizations (CMA/Hermite and DR2 Plain) not only share their yield value of \( \langle \cos^2(\theta) \rangle \), but furthermore make use of a pulse shape that is very similar.

### 9.3.1 Conceptual Quantum Structures

The time-dependent population of the 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 obtained 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.

We applied the SWFT routine to the optimal solutions which were found in the various runs under non-perturbative conditions. Figures A.10, A.11, A.12 and A.13 visualize the typical population process of the rotational levels for four typical solutions of the different optimization procedures (2 parameterizations times 2 DES variants). The observed quantum energy levels are indeed as expected from theory.

The results reveal two different conceptual quantum structures, which correspond to optimal and sub-optimal solutions in terms of the alignment yield. The plain-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\(^{th}\) rotational level in the SWFT picture, corresponding to \( J = 6 \). On the other hand, the plain-CMA and Hermite-DR2 procedures obtain solutions with lower yield, which are characterized by a gradually increasing population of the rotational levels.

The original revival structures for two obtained solutions, representing the two conceptual structures, are given in Figures A.8 and A.9. The optimal family of solutions (Figure A.8) possesses a dramatic revival structure, with a typical strong pulse in the train which lies on the boundary of the punished regime \( \langle I \approx 0.36 \cdot I_{FTL} \rangle \). This strong pulse seems to be essential in giving the molecules the right 'kick', and most likely responsible for the dominant population of the 4\(^{th}\) rotational level in the SWFT picture \( (J = 6) \). The sub-optimal family of solutions (Figure A.9) possesses a revival structure with a smooth exponential envelope, and thus has a gradual building-up of the rotational 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
9.3. The Zero Kelvin Case Study

DES subject to Hermite expansion of the control phase.

9.3.2 Maximally Attained Yield

While this does not constitute a proof, we speculate that within the constraints in the optimization (i.e., the finite pulse bandwidth and energy, as well as the finite resolution of the phase function), both algorithms have found a solution that approaches the best solution that is possible. However, even if the solutions are optimal within the constraints set by the laser bandwidth, the laser pulse energy and the parameterization of the phase, it is clear that the solutions do not approach the maximum alignment that can be supported by the basis of \( N_{\text{rot}} = 20 \) rotational states (see Eq. 9.7) that were used in the calculation. The maximum alignment supported by this basis is the largest eigenvalue of the observable matrix, which was found to be 0.9863. The corresponding eigenvector will be referred to here as the maximal eigenvector or the maximal wavepacket.

We ascribe the difference between this maximum value and the values obtained in the optimizations as being largely due to the finite laser bandwidth in our calculations. The bandwidth and the pulse duration of a laser pulse with a Gaussian shape are related by Eq. 6.30, where the spectral amplitude parameter reads \( c_B = 0.441 \). Thus, for a pulse with a 100fs Fourier-limited duration, the bandwidth is \( \Delta \omega_{\text{laser,FWHM}} = 0.0182 \text{eV} = 147 \text{cm}^{-1} \). When a molecule undergoes a Raman transition from \( J = J_0 \) to \( J = J_0 + 2 \), the energy absorbed from the laser field is \( B_{\text{rot}} \cdot (4J_0 + 6) \). This absorbed energy is the difference between the pump- and dump-photons involved in the Raman excitation. Consequently, the Raman excitation becomes frustrated when \( B_{\text{rot}} \cdot (4J_0 + 6) > \Delta \omega_{\text{laser,FWHM}} \). In our case, with a rotational constant of \( B_{\text{rot}} = 5 \text{cm}^{-1} \), this threshold occurs for \( J_0 \approx 6 \).

As Figure 9.15 shows, the rotational wave packet that displays the largest alignment after the optimization contains only limited contributions from \( J = 8 \) and \( J = 10 \), and none from rotational levels above \( J = 10 \). By contrast, the maximal wavepacket contains contributions all the way up to \( J = 18 \). In this respect, it may appear to be surprising that a high yield of 0.962 can be obtained when the optimized wavepacket differs so much from the maximal wavepacket. In order to assess the crucial influence of the bandwidth constraint on the cut-off of accessible \( J \) values, additional calculations were performed with the original bandwidth doubled, while the fluence was kept fixed (thus corresponding to a 50fs pulse with \( \Omega_{\text{ge}} = 226 \times 10^{12} \text{s}^{-1} \)). These results are also presented in Figure 9.15 as a reference to the calculations with the original bandwidth. The doubling of the bandwidth permitted populating up to \( J = 12 \), and thus produced an enhanced alignment yield of 0.975. Note that the distinction between the two families of solutions, corresponding to the two algorithmic classes, as discussed in Section 9.3.1, can be clearly observed in Figure 9.15.
The difference between the maximal wavepacket and optimized wavepacket is also reflected in the angular probability distribution functions, as presented in Figure 9.16. These probability distribution functions are respectively constructed from the coefficients of the maximal eigenvector as well as the state obtained from the optimized field, based on Eq. 9.7. Even though at the higher bandwidth the discrepancy between the optimally controlled distribution function and the maximally attainable limit appears to be significant, a high alignment value was still obtained.

The explanation for this excellent behavior, despite considerable differences in the composition of the wavefunction, lies in the variational principle (see, e.g., [126]), which states that a first order error in a trial wavefunction (i.e., the wavepacket from the bandwidth limited optimal control field) will produce an extremum eigenvalue (i.e., alignment yield) of second-order error:

$$\frac{\langle \psi | \mathcal{H} | \psi \rangle}{\langle \psi | \psi \rangle} = E_n + \langle \delta | \mathcal{H} | \delta \rangle \langle \delta | \delta \rangle = E_n + \mathcal{O}(\delta^2)$$

(9.8)

9.3.3 Another Perspective to Optimality: Phasing-Up

When a molecule is exposed to a shaped, intense laser pulse the optimization has to accomplish two things. First, the optimization has to create a wavepacket consisting of a large number of rotational states that can serve to align the molecule. Second, the optimization has to prepare the wavepacket with the correct phase relationship between the component wavefunctions, so that during its field-free evolution these components would coherently add-up to generate an optimally aligned wavefunction. While there is no criterion available that allows us to ascertain whether the algorithm has optimized the population distribution, it is possible to investigate the phase relationship of the component wavefunctions in the optimized solutions. Maximum alignment occurs if at some point in time the phases of all component wavefunctions differ from each other by 0 (modulo 2π).

Explicitly, given a wavefunction,

$$\psi = \sum_j a_j^{(t)} \cdot |j\rangle \cdot \exp \left( -i \frac{E_j t}{\hbar} \right),$$

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

$$a_j^{(t)} = r_j^{(t)} \cdot \exp \left( i \varphi_j^{(t)} \right).$$

(9.9)

We thus question 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 routine was implemented in the following manner: It accepts the $a_j^{(t)}$ as input, and aims at optimizing the phases $\varphi_j^{(t)}$. 
Figure 9.15: TOP: The distribution of the maximal and the best optimized wavepackets over the rotational states. Stars represent the maximal wavepacket in the finite rotational basis (i.e., corresponding to the highest-ranked eigenvector of the observable matrix). Diamonds represent the 1st optimized set of solutions (CMA-Hermite / DR2-Plain), and Squares represent the 2nd optimized set of solutions (CMA-Plain / DR2-Hermite); Circles represent calculations with doubled bandwidth and the same fluence (50fs pulse with $\Omega_{ge} = 226 \times 10^{12} \text{ s}^{-1}$), optimized by the DR2 subject to plain parameterization. The figure clearly shows that the limited field bandwidth cuts off the rotational states for the optimized solutions after $J = 10$, when the original bandwidth is used, or after $J = 12$ when the bandwidth is doubled. Furthermore, this plot illustrates the distinction between the two families of solutions for the original bandwidth (i.e., Diamonds versus Squares) arising from the different algorithmic approaches. BOTTOM: The alignment as a function of the overlap of the optimized wavepackets $|\Psi\rangle$ with the maximal eigenvector $|V\rangle$. Note that the overlap for the original bandwidth never exceeds 0.8 in magnitude. Also note the three clusters for the families of algorithmic solutions.
Figure 9.16: Left axis: Normalized angular probability distribution function for the maximal case $|\psi_{\text{max}}(\theta)|^2 \sin(\theta)$, and the optimized control function $|\psi_{\text{opt}}(\theta)|^2 \sin(\theta)$. Right axis: The value of $\cos^2(\theta)$. The constraints prohibit the evolutionary algorithm from attaining the absolute maximal angular probability distribution function; However, the expectation value of the observable $\langle \cos^2(\theta) \rangle_{\text{opt}} = 0.9621$ when using the original bandwidth corresponding to a 100fs Fourier-limited pulse is within 0.025 of the maximum attainable value $\langle \cos^2(\theta) \rangle_{\text{max}} = 0.9863$. When doubling the bandwidth (i.e., basing the shaped laser pulse on a 50fs Fourier-limited pulse) $\langle \cos^2(\theta) \rangle_{\text{opt}}$ increases to 0.975, which is only 0.0113 away from the maximum attainable value.

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 algorithm for the tuning of the 10 relevant phases. Note that a single function evaluation has the duration of $\approx 0.5s$.

We considered 50 different cases of high-quality 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 was calculated:

$$\Delta \phi^{\text{optimal}} = 0.0117$$
2. The 50 test-cases, as originally obtained by the original optimization prior to this optimization procedure, held phases which were 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 was recorded.

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.

To summarize, while we cannot establish whether the optimization has distributed the population in the best possible way, we do observe that the algorithm has properly phased-up all component wavefunctions with respect to each other. This type of coherent alignment of phases was also observed to be optimal in the mechanistic analysis of another state-to-state control application [163].

### 9.4 Evolution of Pulses under Dynamic Intensity

Our observation so far regarding the alignment problem, and in particular concerning its zero-Kelvin variant in the previous section, provides us with the motivation to investigate optimized pulse structures that obtain high alignment yield at different laser intensities, and especially their evolution subject to a slowly-varying laser intensity. This section is a direct experimental continuation to Section 9.3, considering solely the zero-Kelvin alignment variant with two specific algorithmic approaches that were employed for its optimization: the DR2-plain and CMA-Hermite procedures.

#### 9.4.1 Evolutionary Algorithms in Dynamic Environments

From the algorithmic perspective, the optimization framework becomes now an evolutionary search subject to a dynamic environment [71].

Evolutionary Algorithms are natural candidates for optimization in dynamic environments, due to the straightforward analogy with organic evolution, which occurs in a continuously varying environment. Typical approaches for dynamic environments include the promotion of diversity, the use of multi-populations, the introduction of memory-based components, or the assignment of so-called scouts that maintain information about the search space.

Evolution Strategies are a particularly good choice, for their built-in mutative self-adaptation mechanism. The standard-ES has been demonstrated to perform well under a dynamic environment of a time-varying sphere model ("a landscape with catastrophes"), using a comma strategy and with
no recombination (see, e.g., [164]). The mutative self-adaptation mechanism played a crucial role, in allowing a rapid adjustment of the evolving individuals to the time-dependent location of the global maximum: The optimal mutation strategy parameters were learned successfully, without exogenous control. Other empirical studies extended this model to continuously moving peaks, and reported on satisfying adaptation of the standard-ES [165]. Arnold and Beyer considered specific derandomized Evolution Strategies, and showed theoretically that the step-size adaptation mechanism works perfectly well on a moving-sphere problem [166]. In light of these findings, we find our candidate derandomized ES variants perfectly suited for the current optimization task.

9.4.2 Dynamic Intensity Environment: Procedure

In order to observe, and possibly understand how the optimal laser pulse shape evolves from the simple pulse train obtained for $\Omega_{ge} = 40 \times 10^{12}s^{-1}$ (Figure 9.14 (a)), into a much more complicated pulse-shape for $\Omega_{ge} = 160 \times 10^{12}s^{-1}$ (Figure 9.14 (b)), a series of calculations were conducted where $\Omega_{ge}$ was increased linearly as a function of the generation number. In these calculations, the molecule was initially exposed to a shaped laser field with $\Omega_{ge} = 40 \times 10^{12}s^{-1}$, and over 10,000 generations this value linearly increased to $\Omega_{ge} = 180 \times 10^{12}s^{-1}$. This was immediately followed by a linear decrease of the intensity over additional 10,000 generations, back to the initial value of $\Omega_{ge} = 40 \times 10^{12}s^{-1}$. Note that a generation involves 10 or 15 function evaluations, for the DR2-plain or CMA-Hermite procedures, respectively. Furthermore, we consider two control resolutions for the plain-parameterization, $n_1 = 80$ versus $n_2 = 160$, in order to test the algorithmic performance in these two search space dimensions.

The analysis of the dynamic intensity environment is discussed next at several levels.

Intensity Milestones: Dynamic vs. Static Optimization

Figure 9.17 presents the best evolution runs of the DR2-plain optimization procedure for $n_1 = 80$ and $n_2 = 160$ pixels, respectively. It contains four curves, which correspond to the evolution progress in the ramped-up and ramped-down laser intensity environments of the two different runs. Note that the ramped-down curves of the two runs merge. The ramped-up curves differ significantly in the initial learning periods, due to the different search space dimensionality, as expected.

Following the initial learning period of the optimization procedure, a smooth increase is observed in the alignment yield $\langle \cos^2(\theta) \rangle$, as a function of the laser intensity. The best $\langle \cos^2(\theta) \rangle$ value, as reported in the static high intensity case (Table 9.5), is successfully recovered: A $\langle \cos^2(\theta) \rangle$ value
of 0.962 was obtained at $\Omega_{ge} = 160 \times 10^{12}\text{s}^{-1}$. Thus, the dynamic environment does not hamper the optimization performance given a desired target intensity, as long as the initial learning period is passed.

Figure 9.18 presents a comparison between the pulse-shape attained by the DR2 during a dynamic-intensity run at the milestone of $\Omega_{ge} = 160 \times 10^{12}\text{s}^{-1}$, to the equivalent optimized pulse-shape attained in the static optimization procedure at the same Rabi frequency milestone, previously shown in Figure 9.14. Several conclusions may be drawn from this comparison. While the $\langle \cos^2(\theta) \rangle$ yield value is similar for both calculations (as well as in further calculations using this approach), the pulse shapes are dramatically different. Evidently, the pulse shape that the algorithm finds is heavily influenced by the way that the adaptation of the pulse intensity steered the calculations through the search landscape. This behavior is consistent with theoretical analysis of Quantum Control landscapes and their level sets [131, 134].

Evolution of Pulses

We devote this section to the exploration of the pulse shapes obtained in the dynamic intensity environments. Our experimental procedure has essentially an asymmetric nature due to its two stages: The first stage of ramping the intensity from low-to-high requires a learning phase (see Figure 9.17), whereas when reversing the process and bringing the intensity back down the optimization starts from a converged result. Thus, highly optimized solutions can be maintained throughout the latter excursion, and the transition from high-to-low intensity can be continuously observed. This process is illustrated both in Figure 9.19 and in Figure 9.20. In the latter, a sequence of pulses are shown, starting from pulses at low intensity (top-left corner), where the learning process takes place, moving along the snapshot gallery in a matrix-indexing-order fashion, to the center of the plot where the intensity is in its maximal regime, before reducing to a lower intensity again for the pulses shown in the lower-right part of the plot. These latter pulse-shapes are very simple pulse trains, with a pulse separation of $1/(3B_{rot}c) = 2.2$ps. Such a pulse train is very different from the pulse train obtained for the static problem (Figure 9.14), where a pulse separation of 1.1ps was observed in the static calculation at $\Omega_{ge} = 40 \times 10^{12}\text{s}^{-1}$. Nevertheless, the alignment observed at the end of the optimization of Figure 9.20 reaches a value of $\langle \cos^2(\theta) \rangle = 0.548$, which compares rather well with the value of 0.550 obtained in Figure 9.14. At these low intensities, as previously observed at high intensity, vastly different pulse shapes are able to produce similar optimized values of $\langle \cos^2(\theta) \rangle$. These solutions are on a level set, but the present calculations do not reveal if these solutions are on connected (i.e., continuously morphable from one level set to another), or disconnected components of the level set. At low intensity, the $1/(6B_{rot}c) = 1.1$ps pe-
Figure 9.17: Evolution course of the best DR2-plain runs for phase resolutions of \( n_1 = 80 \) and \( n_2 = 160 \) pixels, on the **ramped-up** intensity (dashed or dotted, respectively) versus the ramped-down intensity (reversed direction, solid curves that merge for both runs). Each direction corresponds to \( 10^5 \) generations (\( 10^6 \) function evaluations).

Figure 9.18: Comparison of pulse shapes that were obtained in optimizations employing the DR2-plain procedure, when using a fixed \( \Omega_{ge} = 160 \times 10^{12} \text{s}^{-1} \) (bottom, and see Figure 9.14), or – at this same value of \( \Omega_{ge} = 160 \times 10^{12} \text{s}^{-1} \) – in the course of an optimization where \( \Omega_{ge} \) was linearly varied from \( 40 \times 10^{12} \text{s}^{-1} \) to \( 180 \times 10^{12} \text{s}^{-1} \).
Figure 9.19: Intensity dependence of the alignment $\langle \cos^2 \theta \rangle$ and the laser pulse shape from the \textbf{ramped-up} dynamic intensity environment, subject to a linear increase: $\Omega_{ge} := 40 \times 10^{12}\text{s}^{-1} \rightarrow 180 \times 10^{12}\text{s}^{-1}$. Snapshots are taken at - (a) $54 \times 10^{12}\text{s}^{-1}$, (b) $110 \times 10^{12}\text{s}^{-1}$, (c) $166 \times 10^{12}\text{s}^{-1}$ - and analyzed respectively.
period observed in Figure 9.14 and the \(1/(3 B_{\text{rot}} c) = 2.2 \text{ps}\) period observed in Figure 9.20 correspond to a laser interaction that occurs once per period \(T_{\text{rev}} = 1/(6 B_{\text{rot}} c) = 1.1 \text{ps}\) of the \(J = (0, 2)\) coherent superposition state (Figure 9.14), or every second period (Figure 9.20). This can easily be observed in Figure 9.19, where the temporal behavior is shown for the laser pulse shape and the induced dynamic alignment for \(\Omega_{ge} = 54 \times 10^{12} \text{s}^{-1}\), \(\Omega_{ge} = 110 \times 10^{12} \text{s}^{-1}\), and \(\Omega_{ge} = 166 \times 10^{12} \text{s}^{-1}\). As the intensity is increased, higher rotational states begin to contribute to the rotational wavepacket and the \(T_{\text{rev}} = 1/(2 B_{\text{rot}} c) = 3.3 \text{ps}\) rotational period begins to assert itself. This is a consequence of the energy differences between rotational levels \(J_0\) and \(J_0 + 2\), being multiples of \(2B_{\text{rot}}\) for all values of \(J_0\). In the latter half of the pulse \((t > 0)\), additional narrowly spaced pulses come into play, being spaced by \(T_{\text{rev}}/4 = 1/(8 B_{\text{rot}} c) = 0.8 \text{ps}\). The occurrence of these new peaks comes at the expense of the peak at 2.2 ps, which is considerably weakened in the calculation at \(\Omega_{ge} = 110 \times 10^{12} \text{s}^{-1}\) (Figure 9.19(b)), and is completely absent in the calculation at \(\Omega_{ge} = 166 \times 10^{12} \text{s}^{-1}\) (Figure 9.19(c)). In the latter calculation a new peak has appeared at a delay of 3.3 ps, corresponding to the full revival of the rotational wavepacket formed.

We thus conclude that the optimal pulses observed in the simulations arise as a result of an interplay between the temporal structure that is required to optimize the transfer from \(J = 0\) to \(J = 2\), leading to peak separations that are a multiple of \(1/(6 B_{\text{rot}} c)\), and the temporal structure that is required to optimize the transfer from there to higher rotational levels, which leads to peak separations that are multiples of \(1/(8 B_{\text{rot}} c)\).

**Step-Size and Phase Trajectories**

Figure 9.21 presents the calculation of the Euclidean distance between evolving control phase functions that are determined sequentially as optimal every 100 generations (i.e., between following best-individuals), as well as the global step-size of the mutation operator in those time stamps. Dramatic changes between control phases are observed in the initial learning period, as expected. This is followed by a trend of mild changes, with several bursts of \(\approx 2\pi\) variations. We propose the so-called *wrapping effect* as an explanation for these \(\approx 2\pi\)-jumps: The control phase function is subject to \([0, 2\pi]\)-periodic boundary conditions, that are enforced by wrapping a phase value. Upon examination of the phase space, it is indeed confirmed that these bursts are caused by a boundary wrapping of a phase function value (its index varies). We thus conclude that the variations in the phase space are consistently mild subject to the dynamic laser intensity. This is consistent with the step-size behavior (presented in log\(_{10}\) scale), which stays in the order of \(10^{-2}\) after the learning period, with expected fluctuations.

Interestingly, following the initial learning period, the algorithm "stays in the neighborhood", which seems to be sufficient for determining optimal
9.4. Evolution of Pulses under Dynamic Intensity

Figure 9.20: Evolution of laser pulses subject to linearly increased followed by linearly decreased laser intensity, $\Omega_{\text{pe}} := 40 \times 10^{12}\text{s}^{-1} \rightarrow 180 \times 10^{12}\text{s}^{-1} \rightarrow 40 \times 10^{12}\text{s}^{-1}$, presented as snapshots of optimized pulse shapes at specific intensity milestones. The order follows a matrix-indexing fashion. The pulse-shapes obtained in the end of the process, i.e., after the ramping-down to the regime of low-intensity (bottom right) are a simple pulse train with pulse separation of $1/(3B_{\text{rot}}c) = 2.2\text{ps}$. 

$\Omega_{\text{pe}} = 1.2 \times 10^{12}\text{s}^{-1}$
controls for the continuously changing laser intensity. This means that high alignment yield at different laser intensities corresponds to a neighborhood of the control space.

9.5 Scalability: Control Discretization

In this section we aim at exploring the scalability of the alignment problem with respect to the control resolution. So far, the latter has been fixed in our calculations to $n = 80$. In particular, we would like to study the trade-off between the control resolution, 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. Due to computational considerations, we choose to conduct the scalability calculations on the zero-Kelvin variant of the alignment problem. Also, we select the DR2 subject to the plain parameterization as our optimization kernel for this investigation.
9.5. Scalability: Control Discretization

![Graph showing fitness vs. n]

Figure 9.22: Best, mean and worst cosine-squared alignment values obtained by the DR2 for each parameterization, over 10 runs of 20,000 function evaluations each (see legend).

9.5.1 Numerical Observation

We apply the DR2 algorithm to the optimization task in the following manner: 10 runs per control discretization, with \( n = \{80, 100, 120, \ldots, 680, 700\} \), and additionally with \( n = \{800, 900, 1000\} \). Each run is limited to 20,000 function evaluations.

Figure 9.22 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 presented. 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 obtain high quality solutions in the regime of 0.94 even for \( n = 400 \).

A typical evolution run for \( n = 100 \) is given in Figure 9.23. 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 succeed in tackling the problem within the limited number of function evaluations. A typical run for \( n = 700 \) is presented in Figure 9.24.
Figure 9.23: A typical DR2 evolution run for $n = 100$, with 20,000 function evaluations. Successful learning is observed after $\approx 5,000$ evaluations.

Figure 9.24: A typical DR2 evolution run for $n = 700$, with 20,000 function evaluations. No successful learning is observed.
9.5. Scalability: Control Discretization

![Graph showing fitness over function evaluations](image)

Figure 9.25: DR2 evolution run, for \( n = 1000 \), with 100,000 function evaluations. The best cosine-squared alignment value found was \( f^* = 0.9583 \).

**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 posed a limitation on the search and did not allow a successful learning of the decision parameters and convergence into a good solution.

We have conducted another series of runs, limited now to 100,000 function evaluations, for the extreme case of \( n = 1000 \). We were surprised to find out that some of the runs did succeed in converging successfully into fine solutions of high yield values. In particular, we would like to point out a run which attained 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 in Figure 9.25. A rough observation reveals that the DR2 'takes-off' into a convergence pathway only after \( \approx 50,000 \) function evaluations, and then it needs additional 30,000 function evaluations to reach saturation. This numerical observation indicates that the learning task of the decision parameters in this problem is still feasible in higher dimensions of the control function, as long as the granted number of function evaluations is sufficiently large. From the algorithmic perspective, the employed DES variant, the DR2 algorithm, tackled successfully this 1000-dimensional problem. However, from
the physics perspective, such a high-resolution parameterization does not seem to pay-off, as far as the cosine-squared observable is concerned, and there seems to be no justification to employ discretization of the control phase function with more than \( n=80 \) pixels.

9.6 Intermediate Discussion

Our calculations so far, especially in Sections 9.3 and 9.4, show that it is possible to encounter high diversity of optimal solutions in constrained numerical simulations of Quantum Control, and moreover, that the examination of such rich sets of solutions can become an important aspect of the control experiments. The diversity of successful controls likely contains useful dynamical information, and may also provide the decision maker with a list of choices to consider for weighing in other ancillary control criteria, e.g., multi-criterion decision making. The present calculations optimizing dynamic molecular alignment in a diatomic molecule exposed to an intense, shaped laser field, provide compelling evidence that the absolute value of the quantity that is being optimized (i.e., the fitness) is the true measure of success, and that the same value of the fitness may be achievable by widely differing laser pulse shapes that share only a limited number of common features. Each of these solutions has the potential of carrying valuable information about the underlying physics, where some of the solutions provided key information on the dynamics of the alignment process. Viewed in this sense, the uniqueness of the fitness value, and the diversity of the solutions that can lead to accomplishment is a blessing in disguise.

We also showed that the optimized alignment yield attained a value which was very close to the maximal possible yield in the current framework, even when the constraints on the optimization translated into a significant distortion of the resultant wavepacket. By relaxing specific constraints, we showed that it was possible to enhance the observable alignment further toward the maximal attainable alignment possible for the rotational basis set used. This outcome leads to the optimistic conclusion that high yields may be obtained, even when \textit{a priori} it seems that the system is subject to severe constraints for constructing the wavepacket. As discussed, the origin of this behavior can be understood in terms of the \textit{variational principle}, as well as the physical observable involving an integration over the wavefunction which hides some of its discrepancies.

As a direct implementation of these conclusions, we would like to complete our work on the optimization of dynamic molecular alignment by means of two additional aspects - multi-objective optimization, as well as niching.
9.7 Multi-Objective Optimization

As further investigation of the alignment problem, we would like to extend our single-criterion optimization approach to a Pareto Optimization approach. As previously introduced in Chapter 5, Pareto Optimization aims at attaining the efficient set for a given multi-objective optimization problem and its corresponding Pareto front. In particular, we are interested in removing the punishment approach to high-intensity pulses, and rather consider the fluence of the pulse as an independent objective, subject to minimization. Thus, the observable’s yield remains as an objective, while we choose to define the total-SHG signal of the electric field as the secondary objective subject to minimization.

Formally, we aim at finding the Pareto front for the following bi-criteria problem:

\[
\begin{align*}
    f_1 &= \max_{E(t)} \langle \cos^2(\theta) \rangle \rightarrow \max \\
    f_2 &= \int_{-\infty}^{\infty} |E(t)|^4 dt \rightarrow \min
\end{align*}
\]

(9.10)

In order to select an appropriate optimization method, the following characteristics of the objective functions in the application problem are of importance: Based on our accumulated experience with the problem in its single-criterion form, we assume that the functions \( f_1 \) and \( f_2 \) are continuous in most points, highly nonlinear and multimodal. Nothing is known yet about the shape of the Pareto front for the application problem. Analytical techniques and methods based on differential calculus are likely to fail in this problem, because of the complexity of the integral equations.

9.7.1 Choice of Methods

We choose to apply the NSGA-II, as presented earlier (Section 5.1.2), to the current task. Due to the duration of the simulator evaluation, we would like to consider a specific metamodel that may allow for the acceleration of the calculations.

Metamodel-Assisted NSGA-II In order to accelerate stochastic optimization algorithms in the presence of time consuming function, metamodels have been frequently proposed (see, e.g., [167, 168, 169]). A metamodel is an approximation of an objective function that is learned from a set of evaluations.

More explicitly, given a set of points \( \tilde{x}^{(1)}, \ldots, \tilde{x}^{(k)} \in \mathbb{R}^n \), and the corresponding evaluations of the objective functions at these points, \( \tilde{f}^{(1)} = f \left( \tilde{x}^{(1)} \right), \ldots, \tilde{f}^{(k)} = f \left( \tilde{x}^{(k)} \right) \), the metamodel can be used to compute an approximation, denoted by \( \hat{f}(\tilde{x}) \approx f(\tilde{x}) \), for any point \( \tilde{x} \in \mathbb{R}^n \), in a duration which is considerably shorter than the precise evaluation. As expected,
metamodules tend to be more precise near the training points.

Kriging\(^1\), also referred to as \textit{Gaussian random field models}, is a particular type of interpolation model that has been frequently applied for metamodeling [167, 168, 169]. The statistical motivation for this method is that the deterministic objective functions are considered to be realizations of a Gaussian random field \( \mathcal{G} \). This assumption makes it possible to compute a measure for the uncertainty of predictions, i.e., each prediction value is associated with a standard deviation that can be used for computing two-sided \textit{confidence intervals}.

It is typically assumed that these random variables \( \mathcal{G}_{\tilde{x}} \) are correlated by means of a spatial correlation function,

\[ c : \mathbb{R}^n \times \mathbb{R}^n \to [-1, 1], \]

i.e., a correlation function that depends only on the positions of the random variables in space. In our study we shall use a correlation function of the form:

\[ c(\tilde{x}, \tilde{x}') = \exp\left(-\theta |\tilde{x} - \tilde{x}'|^2\right) \]

The correlation function of the Gaussian random field is estimated from the given data, or given \textit{a-priori}. In this study we apply leave-one-out cross-validation to determine an appropriate value of \( \theta \), as suggested in [170]. After the correlation function is estimated, the prediction is made. For this purpose, the conditional Gaussian distribution at the given input vector \( \tilde{x} \in \mathbb{R}^n \) is computed.

A practical implementation of Kriging has been described by Emmerich [101], and it was successfully employed in engineering design optimization [100, 169, 171]. Multi-objective problems were typically approached by learning metamodules for each objective function separately, in an implementation known as \textit{local Kriging}. We omit here its derivations, and refer the reader to [101].

In the metamodel-assisted NSGA-II [171], Kriging metamodules are used to pre-evaluate the set of offspring solutions and select favorable variants among it for precise evaluation. The uncertainty information can be used to facilitate search in less explored regimes of the landscape.

Algorithm 9 outlines the general Metamodel-Assisted Evolutionary Algorithm (MA-EA), as described by Emmerich [101]. The difference to the generic Evolutionary Algorithm can be summarized as follows:

- All \textit{precisely evaluated points} are stored in a database, denoted by \( D_t \) (cf. lines 4 and 9).

\(^1\)Kriging originates from \textit{geostatistics}, and is named after the mining-engineer Krige.
9.7. Multi-Objective Optimization

Figure 9.26: Outputs of Gaussian Random Field Metamodels using a \( \mathbb{R} \to \mathbb{R} \) mapping example. Three points, \( \bar{x}^{(1)}, \bar{x}^{(2)}, \) and \( \bar{x}^{(3)} \) have been evaluated here. The result of each approximated evaluation at a point \( \bar{x}' \) is represented by the mean value, \( \hat{y} \), and by the standard deviation, \( \hat{s} \), of a 1D Gaussian distribution. Figure courtesy of Michael Emmerich [101].

- The algorithm filters out less promising solutions (cf. line 8) and thereby reduces the offspring population size. The remaining solutions are then precisely evaluated and considered in the subsequent selection.

There are many possibilities to design filters for that purpose. In this study we restrict ourselves to constant output size filters. The size of the resulting filtered set will be denoted by \( \nu \) and the corresponding MA-EA will be termed a (\( \mu + \nu < \lambda \))-EA. All filters will be rank-based, i.e. they sort the offspring population with respect to some criterion, a so-called filter criterion.

We offer a 3D visualization in Figure 9.27 in order to gain some intuition into the different concepts of filters in the bi-criteria case. In the latter, the Pareto-front approximation of the current population is depicted, as well as three offspring individuals, namely \( \bar{x}_1, \bar{x}_2 \) and \( \bar{x}_3 \). The offspring individuals have been evaluated with the Kriging metamodel, and thus their precise values are not yet known, but rather the defining parameters of 2D Gaussian random variables, \( G_{\bar{x}_1}, G_{\bar{x}_2}, \) and \( G_{\bar{x}_3} \) are also visualized in the diagram by means of their probability density functions.

Four different criteria have been discussed by Emmerich [101] for assigning a yield value to a search point \( \bar{x} \), which is based on the prediction provided
by the defining parameters of the Gaussian predictor $G_x$:

- **Mean Value** Non-dominated / crowding distance sorting, based on the expected value for $G_x$ given by $\hat{f}(\tilde{x})$.
- **Lower Confidence Bound (LCB)** Non-dominated / crowding distance sorting on the lower bound edge of the confidence interval of $G_x$.
- **Probability of improvement (PoI)**: The probability that the realization of $G_x$ is non-dominated. It can be computed via integration over the non-dominated set.
- **Expected Improvement (EI)** The expected increase in the dominated hypervolume for $G_x$ is measured.

**Modus Operandi**

We applied the following algorithmic kernels to the Dynamic Molecular Alignment:

- **NSGA-II**: The classical variant by Deb [98, 172].
- **Metamodel-Assisted EA with Probability of Improvement (PoI-EMOA)**.
- **Metamodel-Assisted EA with Expected Improvement (EI-EMOA)**.

The parameterization of these methods is $\mu = 50, \nu = 0.2 \cdot \lambda$, with two different settings for $\lambda$: $\lambda = 250$ and $\lambda = 50$. The parameters of the mutation operator and recombination operator have been chosen as described by Deb.
Algorithm 9 $(\mu + \lambda)$-MA-EA

1: $t \leftarrow 0$
2: $P_t \leftarrow \text{init}() \{P_t \in S^\mu: \text{Set of solutions}\}$
3: Evaluate($P_t$)
4: $D_t \leftarrow P_t$
5: while $t < t_{\text{max}}$ do
6: $G_t \leftarrow \text{Generate}(P_t) \{\text{Generate } \lambda \text{ variations}\}$
7: $\text{Metamodel\_evaluate}(G_t) \{\text{Metamodel is derived from } D_t\}$
8: $Q_t = \text{Filter}(G_t)$
9: $D_{t+1} \leftarrow D_t \cup Q_t$
10: $P_{t+1} \leftarrow \text{Select}(Q_t \cup P_t) \{\text{Rank and select } \mu \text{ best}\}$
11: $t \leftarrow t + 1$
12: end while

[98]. Due to implementation considerations, in practice both objectives were minimized, and therefore we assign:

$$f_1 \rightarrow \text{max} \implies -f_1 \rightarrow \text{min}$$

9.7.2 Numerical Observation

Figures A.14, A.15 and A.16 present the results of our calculations, where the 20%, 50% (median), and 80% attainment surfaces are plotted. Each one of them refers to 5 runs with 20,000 evaluations per run. In order to make the curves easier to be distinguished, we zoomed in a box around the knee point of the Pareto front approximations.

Discussion

The results clearly indicate that there is a conflict between the two objectives, as suspected. Thus, Pareto optimization is an appropriate tool for solving this problem. The fact that a convex Pareto front has been observed suggests that good compromise solutions are likely to be found. We observe a sharp increasing flank at both ends of the approximated Pareto front. Regions of fair trade-offs range from about $-0.6$ to $-0.4$ in the $(-f_1)$ coordinate.

There are significant differences in the behavior of the multi-objective EA variants. The best coverage of the Pareto front has been achieved by the ExI-EMOA. This variant is the only variant that found solutions for $f_1$ above 0.58. The highest value found was 0.6184. The PoI-EMOA resulted in approximations with lower spread. However, the precision of this EMOA variant was better in the regions covered. This result is consistent with some theoretical findings reported in [171], as well as with their numerical assessment on artificial problems reported there. The expected improvement measure puts emphasis on exploring unknown regions, while the probability
of improvement have the tendency to carry out better exploitation of visited regions. Overall, the metamodel-assistance seems to be a valuable ingredient for this problem, as can be seen by comparing the results of the NSGA-II with those of the metamodel-assisted EMOA.

A more detailed analysis of the metamodel-based approximations was performed, in order to assess whether the metamodeling worked as expected from theory. The results are displayed in Figures 9.28 and 9.29, for one of the runs with the ExI-EMOA ($\lambda = 250$). The $f - \hat{f}$ plots indicate that in the whole range of function values the results obtained with the metamodel were strongly correlated with the true function values. The error bandwidth for $f_1$ is about 10\% of its range versus 15\% for $f_2$ with respect to its range. These results correspond to results in similar studies in metamodel-assisted optimization [171]. Moreover, the lower confidence bounds, denoted by $f_{lb}$, have been compared to the outcome of the precise evaluations, $f$. Here, the 95.45%-lower confidence bounds, as computed by the Kriging method, have been assessed for their validity (see Figures 9.28 and 9.29). The results are in conformity with theory for $f_1$. However, some outliers for $f_2$ in the region
of $f_{lb}$ from 0.15 – 0.2 should be reported. However, these outliers did not seem to hamper the algorithmic performance.

From the physics point of view the obtained result is interesting, since it shows the nature of the trade-off between the alignment’s observable and the intensity of the electric field, expressed here by means of the second harmonic generation signal. The importance of the intensity criterion is likely to govern the decision of the expert on the trade-off surface, which is to look for solutions with relatively good $f_1$ values in the region of fair trade-offs.

### 9.8 Application of Niching

We shall apply here our DES niching algorithms to the zero-Kelvin variant of the dynamic molecular alignment. Following the application of niching to the population transfer problem in the rotational framework, as described in Section 8.3, we take into consideration the diversity measure issue, and fully adopt the conclusions drawn in Section 8.3.1.

#### Modus Operandi

We consider here three niching strategies:

1. The $(1, \lambda)$-DR2 – for being the best method to perform on this problem, and also as a representative of first-order strategies.

2. The $(1, \lambda)$-CMA – as a representative of second-order information strategies.

3. The $(1 + \lambda)$-CMA – as a representative of elitist strategies.

We conduct 10 runs per method, searching for $q = 3$ niches, subject to plain parameterization of the control phase at $n = 80$ pixels. Each run was limited to 15,000 function evaluations per niche.

#### 9.8.1 Numerical Observation

The calculations are discussed at several levels.

#### Niche-Radius

Following the derivation done for the niche radius in the population transfer problem in Section 8.3.2, we conducted preliminary runs with a niche-radius of $\rho = 110$. However, it performed poorly, in an equivalent way to its initial performance on the population transfer problem: The DR2 as well as the CMA-comma failed, and the CMA-plus obtained good solutions only for the first niche.
Table 9.6: Three niches obtained in 10 runs – averaged yield values (in parentheses - best value attained) – for the three employed niching strategies.

<table>
<thead>
<tr>
<th>Ranked-Niches</th>
<th>DR2</th>
<th>CMA</th>
<th>CMA+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best niche</td>
<td>0.9417 (0.9605)</td>
<td>0.8553 (0.9029)</td>
<td>0.9517 (0.9585)</td>
</tr>
<tr>
<td>2\textsuperscript{nd}-best niche</td>
<td>0.8477 (0.9552)</td>
<td>0.8229 (0.8561)</td>
<td>0.9493 (0.9525)</td>
</tr>
<tr>
<td>3\textsuperscript{rd}-best niche</td>
<td>0.8054 (0.8558)</td>
<td>0.7966 (0.8161)</td>
<td>0.9365 (0.9484)</td>
</tr>
</tbody>
</table>

Table 9.7: Niches correlation for the niches obtained in 10 runs – averaged cross-correlation values, as defined in Eq. 8.16.

<table>
<thead>
<tr>
<th>Niches Correlation</th>
<th>DR2</th>
<th>CMA</th>
<th>CMA+</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c_{1,2} )</td>
<td>0.6784</td>
<td>0.6952</td>
<td>0.6312</td>
</tr>
<tr>
<td>( c_{1,3} )</td>
<td>0.6288</td>
<td>0.6905</td>
<td>0.6062</td>
</tr>
<tr>
<td>( c_{2,3} )</td>
<td>0.7593</td>
<td>0.6951</td>
<td>0.6414</td>
</tr>
</tbody>
</table>

We managed to get satisfying results for \( \tilde{\rho} = 55 \), as will be reported here. Thus, consider all the reported results here as obtained with \( \tilde{\rho} = 55 \).

\textbf{Success-Rate}

The \textit{cosine-squared alignment} of the three methods, for the three obtained niches, is presented in Table 9.6. We can observe a clear trend - the CMA+ mechanism outperformed the other mechanisms, with consistent location of three good niches on average. However, the DR2 mechanism managed to obtain the top-quality solutions for the best as well as for the 2\textsuperscript{nd}-best niches, in consistency with our previously reported results. The latter typically failed to locate a 3\textsuperscript{rd} good niche. The CMA comma-strategy, on the other hand, simply failed in obtaining satisfying niching results on this landscape.

\textbf{Niches Cross-Correlation}

We calculated the cross-correlation coefficients for the obtained pulse-shapes of the different niches, as defined in Eq. 8.16. The results of these calculations are presented in Table 9.7. We may state that the pulse-shapes of the different niches are weakly correlated to one another. In particular, it is interesting to note the low correlation values of the the CMA+ kernel.
Laser Pulse Designs
Our definition of a distance measure to this problem has been proved to be successful. The obtained pulses in the time-domain had indeed different characteristics, representing different conceptual laser-pulse designs. Three niches, obtained in a typical CMA+ run, are plotted by means of their pulse intensities and revival structures in Figures A.17, A.19 and A.21.

Conceptual Quantum Structures Revisited
We would like to offer an additional analysis for our niching solutions. Figures A.18, A.20 and A.22 provide the SWFT picture for the obtained solutions. It can be observed that these three solutions represent the same conceptual quantum structure of states population. This SWFT observation reinforces our conclusions concerning the correlation between the employed optimization routine in combination with the applied parameterization to specific conceptual quantum structures, as drawn in Section 9.3.1. Therefore, we do not find it surprising that all three obtained pulses share the same 'behind-the-scenes physics', due to the fact that they were all obtained with the same algorithmic kernel (e.g., CMA+), subject to the plain parameterization. This observation does not contradict our primary conclusion that the niching process has been successful in locating three different pulse shapes in the temporal domain, as initially required. It simply reveals an additional, well-hidden, degeneracy among the solutions. In the next section we shall offer a way to remove this second degeneracy completely.

Removing the Second Degeneracy
Given the additional degeneracy which was encountered in the SWFT space, one can further develop a problem-specific diversity measurement. In this case, our idea is to consider the wavepacket space, and more explicitly, to evaluate the differences between the population of rotational levels, $|a_j(t)|^2$, as the measurement of diversity between niches. The implementation itself is straightforward, due to the fact that the vector of population coefficients is given by the alignment-routine. Since the coefficients are normalized, subject to the normalization postulate of Quantum Mechanics, it is fairly simple to estimate the niche radius in this case.

Niche Radius: Wavepacket Space  According to the Quantum Mechanics normalization postulate, the wavepacket coefficients in the $N$-dimensional Hilbert space are normalized:

$$
\sum_{j=1}^{N} |a_j(t)|^2 = 1
$$
In the wavepacket treatment for removing the second degeneracy, these coefficients play the role of the decision parameters, as far as the diversity measurement is concerned.

The calculation of $r$ of Eq. 3.3 simply reads:

$$r = \frac{1}{2} \sqrt{\sum_{j=1}^{N_{\text{rot}}} |a_j(t)|^2} = \frac{1}{2}$$

With $q = 3$ and $N_{\text{rot}} = 20$, Eq. 3.5 yields:

$$\rho = \frac{2}{3^{\frac{1}{2}}} \approx 0.47 \quad (9.11)$$

Thus, we set it to $\rho = 0.5$. We choose to employ only the CMA+ kernel in this case, subject to plain as well as Hermite parameterizations, aiming to show feasibility of the defined diversity measure.

This newly-defined diversity-measurement for the alignment problem has been observed to be successful. The obtained pulses in the temporal domain had indeed different characteristics, and in particular their shapes differed in a satisfying manner. We consider here the results obtained when the Hermite parameterization was employed. The best niche obtained in every run was typically of the optimal class known to us: Both the cosine-squared alignment yield, as well as the pulse shape and the population profile, were associated with the best solutions reported previously. The second-best niche was a representative of a sub-optimal set of solutions: It had a lower value of cosine-squared alignment yield and a different profile of population. However, note that the third-best located niche was not typically an interesting solution, as it had dramatically lower alignment values in comparison to the first two niches. The temporal pulse-shapes themselves were very weakly correlated.

Typical solutions of best and second-best niches are plotted in Figures A.23 and A.25, with their corresponding SWFT pictures in Figures A.24 and A.26.

Discussion

We would like to summarize our numerical observation of the applied niching algorithms to the dynamic molecular alignment problem. Niching with the CMA+ kernel performed best, while always obtaining three niches of high-quality laser pulses. The DR2 found the best solution, in consistency with our previously reported observations, but did not perform well on the secondary niches. The CMA-comma failed to obtain satisfying niching results.

The original calculation of the niche radius was not successful at the practical level, as reported for the population transfer problem. After introducing a factor of 0.5 to the original value, the niching process was observed
to be successful. The obtained pulse-shapes were typically weakly correlated, as required.

As far as the algorithmic performance is concerned, we adopt the conclusions drawn for the application of niching to the population-transfer problem. We thus ascribe the failure in practice of the originally calculated niche radius, as well as the compromised performance of the comma-strategy kernels on the secondary niches, to the highly constrained nature of the landscape when underposed to a radius-based niching framework.

Furthermore, we have applied a physics numerical assessment, at the quantum rotational picture, with the so-called SWFT technique. The latter has supported previous observations concerning the correlation between optimization routines in combination with parameterizations to conceptual quantum structures. This observation revealed that all three niches of a given run, which differ sufficiently at the laser-pulse design level (temporal domain), typically share the same conceptual quantum structure at the SWFT picture (wavepacket space). We offered another diversity measure, which relies on the physics information, in order to remove this second degeneracy. This approach indeed succeeded in that, and obtained multiple solutions corresponding to different conceptual designs.