Chapter 7

Two Photon Processes

7.1 Introduction

The field of non-linear optics describes optical phenomena which are observed when high intensity light passes through media. The non-linearity is due to the interaction between the light, typically a laser field, and a dielectric media, whose field-induced polarization responds non-linearly to the incident electric field.

Given the temporal intensity of the electric field, \( I(t) \), its non-linear signal of the \( k^{th} \) order is modeled for \( k > 1 \) as:

\[
Signal_{NL}^{(k)} \propto \int_{-\infty}^{\infty} I^k(t) dt, \tag{7.1}
\]

corresponding to the interaction of \( k \) photos.

The field of non-linear optics offers a variety of popular Quantum Control applications. Second-order variants, which correspond to two-photon processes, are particularly attractive because of their easy implementation in the laboratory, as well as their known mathematical formulation. Two-photon processes can be utilized to explore experimental Quantum Control landscapes, and also can form a realistic testbed for global optimization algorithms.

This chapter is devoted to the formal definition of two-photon processes, their mathematical description, and to the application of optimization routines to their signal-maximization problems in the laboratory.

7.2 Second Harmonic Generation

Second harmonic generation (SHG) or frequency doubling is a two-photon process in which an electric field interacts non-linearly with a material and generates an output photon with double the energy of two input photons. The total energy of the output light is proportional to the integrated squared
intensity of the primary pulse, as expected from a second-order non-linear
process.

The time-dependent profile of the laser field is exactly as given in Eq.
6.29. The SHG signal is then defined by:

\[ SHG_t \equiv S_t = \int_{-\infty}^{\infty} I(t)^2 \, dt = \int_{-\infty}^{\infty} |E(t)|^4 \, dt, \quad (7.2) \]

i.e., integration over time of the intensity. SHG is a process that turns out to
be a good test case in the laboratory, and its investigation contributes to the
understanding of other processes. This is because the SHG is a measure of
the pulse duration, and this property is useful as an auxiliary characteristic.
From the theoretical point of view, the SHG is a simple test function, with
some interesting mathematical properties that will be fully derived here, but
yet not an easy optimization task for global optimizers.

### 7.2.1 Total SHG

In order to gain a better insight into the problem, we provide here the reader
with some of its mathematical properties. Especially, we would like to derive
the equivalence between time and frequency pictures. The following section
is mainly based on Bracewell [142].

**Definition 7.2.1.** Given the spectral amplitude equipped with the complex
phases, \( E(\omega) = A(\omega) \exp(i\phi(\omega)) \), consider its autocorrelation (convolution)
function \( E_2(\omega) \):

\[ E_2(\omega) = E(\omega) \ast E(\omega) = \int_{-\infty}^{\infty} E(\Omega) \cdot E(\omega - \Omega) d\Omega \]

We would like to show how this autocorrelation function in the frequency
domain is linked to the time domain:

**Theorem 7.2.2.** The autocorrelation function of the spectral amplitude,
\( E_2(\omega) \), is proportional to the Fourier transform of the squared time-dependent
electric field, i.e.:

\[ E_2(\omega) \propto \int_{-\infty}^{\infty} \tilde{E}(t)^2 \exp(-i\omega t) \, dt \quad (7.3) \]
7.2. Second Harmonic Generation

Proof.

\[
E_2(\omega) = \int_{-\infty}^{\infty} E(\Omega) \cdot E(\omega - \Omega) d\Omega = \\
= \int_{-\infty}^{\infty} \left[ \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{E}(t) \exp(-i\Omega t) dt \right] \cdot \left[ \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{E}(\tau) \exp(-i(\omega - \Omega)\tau) d\tau \right] d\Omega = \\
= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{E}(t) \tilde{E}(\tau) \exp(-i\Omega(t - \tau)) \cdot \exp(-i\omega\tau) d\Omega dt d\tau = \\
= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{E}(t) \tilde{E}(\tau) \delta(t - \tau) \exp(-i\omega\tau) d\tau dt = \\
= \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{E}(t)^2 \exp(-i\omega t) dt = \\
= \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{E}(t)^2 \exp(-i\omega t) dt
\]

where \(\delta(x - \bar{x})\) is the Dirac delta function.

Theorem 7.2.3. (Plancherel’s Theorem) Given \(f(x)\), which has the Fourier transform \(F(s)\), the integral over the squared modulus of \(f(x)\) is equal to the integral over the squared modulus of its spectrum \(F(s)\):

\[
\int_{-\infty}^{\infty} |f(x)|^2 dx = \int_{-\infty}^{\infty} |F(s)|^2 ds
\]

See [142]. Thus, we can conclude from Theorems 7.2.2 and 7.2.3 that

\[
\int_{-\infty}^{\infty} |E_2(\omega)|^2 d\omega = \int_{-\infty}^{\infty} |E(t)|^4 dt
\]

and, equivalently, in terms of the intensities

\[
S_t = \int_{-\infty}^{\infty} I_2(\omega) d\omega = \int_{-\infty}^{\infty} I(t)^2 dt
\]

(7.4)

where \(I_2(\omega) = |E_2(\omega)|^2\).

Global Maximum

Theorem 7.2.4. The Total-SHG signal is maximized by the phase being any linear function of frequency, and in particular by the constant phase:

\[
\arg \max_{\phi(\omega)} \{S_t(\phi(\omega))\} \equiv a \cdot \omega + b
\]
An important remark should be made concerning the existence of a single optimal solution for the SHG maximization problem: Due to the use of second-order perturbation theory, the constant phase is a point in the control space (the generalization to a linear phase stems from symmetry), i.e., the level-set collapses into a single point. In higher-order corrections for SHG the maximally attained yield can be obtained by various other phase profiles.

Figure 7.1 provides the reader with an illustration for the so-called frequency doubling effect - the contribution of two phase points around the central frequency $\omega_0$ at $E(\omega)$, $\phi(\omega_0 + \omega_1)$ and $\phi(\omega_0 + \omega_2)$, to the construction of $\tilde{E}(\omega)$ with $\phi(2 \cdot \omega_0 + \omega_1 + \omega_2)$. Note the shift in the central frequency, and the scaling of the Gaussian.

### 7.2.2 Filtered SHG

We consider another second-order quantum optical system, which could be considered as a filtered case of the SHG system. It corresponds to a two photon absorption (TPA) process, whose model describes, within the limits of second-order time-dependent perturbation theory, the probability of making a transition from a ground state $|g\rangle$ to an excited state $|e\rangle$, upon the activation of the laser field. Thus, a specific transition frequency is considered here, $\omega_{eg}$, which practically filters the signal,

$$SHG_f \equiv S_f(\omega_{eg}) = \int_{-\infty}^{\infty} \delta(\omega_{eg} - \omega) I_2(\omega) d\omega,$$

by means of the Dirac delta function $\delta(\Omega - \Omega')$. It explicitly reads

$$S_f(\omega_{eg}) = \left| \int_{-\infty}^{\infty} E(\omega) E(\omega_{eg} - \omega) d\omega \right|^2 \quad (7.5)$$
Figure 7.2: A spectral illustration for the total-SHG (left) versus the filtered-SHG (right) signals. Figure courtesy of Jonathan Roslund.

Global Maximum

Theorem 7.2.5. The filtered-SHG signal is maximized by the phase being any odd function of frequency antisymmetric about $\frac{\omega}{2}$, i.e., spectral phases of the form $\phi(\frac{\omega}{2} - \omega) = -\phi(\frac{\omega}{2} + \omega)$.

See [144, 145]. Figure 7.2 provides an illustrative comparison between the two SHG variants considered here.

Problem Difficulty: Numerical Assessment

In order to assess the optimization difficulty of the Second Harmonic Generation maximization problems, we considered numerical simulations of the two SHG problem variants and conducted the following simple statistical test. We considered phase functions pixelized by $n = 64$ function values, which are randomly initialized in the interval $[0, 2\pi]$. We then gradually transformed the given random phases into a zero-phase in two different routines: (1) Setting function values to zero when consistently indexing from right to left, or (2) Setting function values to zero in random permutation of indices, with no repetition. Both routines eventually obtain zero-phases, which attain the maximal yield of 1 for both SHG problem variants.

Figure 7.3 presents typical runs for the two routines when applied to both SHG problem variants. It is observed in these plots that approximately 50% of the function values must be set to zero in order to enhance the yield value, for all cases. Once this threshold is exceeded, the yield value increases consistently until it reaches the value of 1. The actual profiles of routine (1) versus routine (2) differ, for both SHG variants. More variables are required to be set to zero in the random indexing routine, in comparison to
Figure 7.3: Transforming randomly-initialized phases into a zero-phase, pixel-by-pixel, either by (1) Consistently indexing the phase function from right to left, or by (2) Randomly selecting phase function indices, without repetition. The attained yield per index-step is recorded for each test-case. Typical runs are presented for the two routines applied to the SHG problem variants. Left: Filtered-SHG system; Right: Total-SHG system.

the consistent indexing. This is due to the shape of the weighting function (i.e., a Gaussian), which limits the contribution to the yield value from pixels which are not in the proximity of the central frequency.

This statistical test reveals that the SHG problems under investigation are non-separable upon following the formal definition.

7.3 Numerical Simulations

We present here results of the four derandomized ES comma-variants when applied to numerical simulations of second-order photon processes: The maximization of the Total-SHG as well as the Filtered-SHG signals.

7.3.1 Preliminary ES Failure: Stretched Phases

When applied to both SHG simulations, the derandomized ES variants suffered from pre-mature convergence to sub-optimal solutions of low yield. Upon examination of the attained optimized phases in the decision space, they were always observed to be highly steep linear phases. We offer the following explanation for that.

The ES is not subject to any restrictions concerning its decision parameters, in particular in the context of the periodic nature of the phase. It seems that an unrestricted search, as employed by the ES variants in hand, is likely to stretch the candidate phases, with no way to reverse it. It suffers accordingly from convergence to highly steep linear phases with sub-optimal
7.3. Numerical Simulations

Table 7.1: Derandomized Evolution Strategies optimizing the Total-SHG simulation: Mean and standard-deviation of attained yield over 100 runs for the three procedures – unrestricted, wrapped and bounded.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Unrestricted</th>
<th>Wrapped</th>
<th>Bounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>DR1</td>
<td>$0.208 \pm 0.072$</td>
<td>$0.873 \pm 0.187$</td>
<td>$0.574 \pm 0.189$</td>
</tr>
<tr>
<td>DR2</td>
<td>$0.181 \pm 0.064$</td>
<td>$0.967 \pm 0.019$</td>
<td>$0.725 \pm 0.185$</td>
</tr>
<tr>
<td>DR3</td>
<td>$0.457 \pm 0.198$</td>
<td>$0.718 \pm 0.274$</td>
<td>$0.529 \pm 0.278$</td>
</tr>
<tr>
<td>CMA</td>
<td>$0.581 \pm 0.136$</td>
<td>$1 \pm 0$</td>
<td>$0.997 \pm 0.002$</td>
</tr>
</tbody>
</table>

Table 7.2: Derandomized Evolution Strategies optimizing the Filtered-SHG simulation: Mean and standard-deviation of attained yield over 100 runs for the three procedures – unrestricted, wrapped and bounded.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Unrestricted</th>
<th>Wrapped</th>
<th>Bounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>DR1</td>
<td>$0.257 \pm 0.087$</td>
<td>$0.666 \pm 0.247$</td>
<td>$0.713 \pm 0.152$</td>
</tr>
<tr>
<td>DR2</td>
<td>$0.248 \pm 0.091$</td>
<td>$0.804 \pm 0.195$</td>
<td>$0.908 \pm 0.125$</td>
</tr>
<tr>
<td>DR3</td>
<td>$0.539 \pm 0.162$</td>
<td>$0.762 \pm 0.209$</td>
<td>$0.554 \pm 0.173$</td>
</tr>
<tr>
<td>CMA</td>
<td>$0.487 \pm 0.134$</td>
<td>$0.990 \pm 0.008$</td>
<td>$0.964 \pm 0.052$</td>
</tr>
</tbody>
</table>

yield values, as outlined earlier in Section 6.2.2. By implementing periodic boundary conditions into the ES algorithms, by means of coupling the wrapping operator (Eq. 6.35) to the mutation operator, this problem was solved. This procedure will be referred to as the wrapped procedure.

As a third procedure, we also considered the application of a boundary operator that fixes an exceeded value to the lower or upper bounds. Given $\varepsilon > 0$, it reads:

$$
\phi_i = 2\pi + \varepsilon \quad \rightarrow \quad \tilde{\phi}_i := 2\pi \\
\phi_j = -\varepsilon \quad \rightarrow \quad \tilde{\phi}_j := 0
$$

(7.6)

It is referred to as the bounded procedure.

7.3.2 Numerical Observation

Tables 7.1 and 7.2 summarize the numerical results of the application of the four derandomized ES comma-variants to the total-SHG and filtered-SHG simulation problems, respectively, subject to the three specified procedures, with $n = 64$ decision parameters. There are two clear observations from the given calculations:

1. The wrapping operator seems to be an essential component for the unrestricted ES optimization, and should be implemented into ES when
optimizing "phase" variables on a QC landscapes. This is an expected conclusion, given the nature of the search space. However, it is interesting to note the relatively high standard deviations for the results obtained subject to wrapping for the filtered-SHG case for the first three DES variants. Also, it is observed that the bounded approach works better for the DR2 on the filtered-SHG landscape.

2. The CMA outperformed the other algorithms on these two landscapes, with consistent winning performance. The DR2 was second-best, and it performed in a highly satisfactory manner. We thus hold two DES variants, each representing first- or second-order information approach, respectively, which performed well on these QC landscapes.

Intermediate Discussion

We found that employing the ES variants with default settings unrestrictively on the given QC landscapes resulted in pre-mature convergence to sub-optimal phases with highly sloped linear profiles. We analyzed this effect, and introduced the wrapping operator into the ES framework. The latter solved the observed problem.

7.4 Laboratory Experiments

We report here on laboratory experiments where we aimed at optimizing the two quantum control systems described in Section 7.2. Due to the tremendous effort and time which are required for a reliable experiment, we had no choice but to restrict ourselves to a limited number of experiments as well as optimization routines.

We chose to employ three optimization routines in the laboratory:

- DR2: First-order DES.
- CMA: Second-order DES.
- GA: Laboratory reference.

Concerning the technical details, for total-SHG signal, \( S_0 \), the amplified pulses are delivered to a 100 \( \mu m \) type-I BBO crystal, and the time integrated SHG signal is recorded with a photodiode and boxcar integrator. For the filtered-SHG signal, \( S_f \), unamplified seed pulses are focused onto a 100 \( \mu m \) type-I BBO crystal, and the resultant up-converted light is analyzed with a spectrometer. Regarding the actual yield values recorded by us, we choose to normalize the FTL signal as yield 1.0 for both systems.

It should be noted that the SHG optimization problems have been widely investigated at several levels, including at laboratory experiments [146], where it was shown to have a highly complex landscape.
Table 7.3: Laboratory SHG Optimization: Performance Evaluation. The experimental results of the two SHG systems, averaged over 10 experiments. The final yield (averaged over the last 50 iterations) and the number of evaluations required to cross a yield threshold of 0.90 are considered here.

<table>
<thead>
<tr>
<th>Routine</th>
<th>Filtered-SHG</th>
<th>Total-SHG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Avg. Yield</td>
<td>0.9 Eval</td>
</tr>
<tr>
<td>GA</td>
<td>0.95</td>
<td>4665</td>
</tr>
<tr>
<td>DR2</td>
<td>0.93</td>
<td>2159</td>
</tr>
<tr>
<td>CMA</td>
<td>0.95</td>
<td>841</td>
</tr>
</tbody>
</table>

**ES Failure Revisited: Stretched Phases** When applied to the experimental setup, the derandomized ES variants initially suffered from premature convergence to sub-optimal solutions of yield \(\approx 0.75\), where the maximum value is 1.0. Upon examination of the attained optimized phases in the decision space, the stretching effect as reported in Section 7.3.1 was observed. Thus, we used the *wrapping operator* in the two DES variants in all the reported experiments. The GA, on the other hand, did not typically locate highly-steep linear phases since the \([0, 2\pi]\) bounds are implicitly implemented by means of the *phenotypic mapping* (see, e.g., [22]).

### 7.4.1 Performance Evaluations

Table 7.3 presents the results of the two reported systems, averaged over 10 experiments. We consider the final yield (averaged over the last 50 iterations), as well as the number of evaluations required to cross a yield threshold of 0.90, as the performance criteria per experiment. Figure 7.4 presents averaging of the runs, with attained yield as a function of the required number of function evaluations. Note that this averaging procedure takes into account all 10 runs, whereas the convergence data shown in Table 7.3 considers only the relevant runs that exceeded the 0.90 yield threshold. Figure 7.5 presents histograms for the different algorithms with final yield versus the number of runs.

As reflected from the experimental results, the CMA performed best on the given experimental systems, both in terms of final yield as well as convergence speed. We would like to emphasize the extraordinary boost of convergence speed provided by the CMA relative to the GA, which is significant in the laboratory. Moreover, the CMA has a sharp and rapid convergence profile, in contrast to the inefficient hill-climbing capability of the GA. This profile is easy to identify as there is no ambiguity about convergence, and thus it is another attractive feature for the laboratory user.

Next, we discuss the experimental results and the algorithmic behavior.
Figure 7.4: Averaged runs of the algorithms over 10 runs. Left: Filtered-SHG system; Right: Total-SHG system.

Figure 7.5: Success-rate (yield) histograms. Left: Filtered-SHG system; Right: Total-SHG system.

Diversity of Solutions

As mentioned earlier in Section 7.2.2, the filtered SHG system possesses a family of nontrivial phases that correspond to global maxima. Interestingly, each run for the filtered SHG case converged to a distinct antisymmetric phase. This collection of different solutions provided a practical perspective concerning the richness of QC landscapes and their underlying level sets.

Sensitivity to Noise

The CMA-ES and the GA performed in a satisfactory manner on the given control problems and did not seem to be significantly impaired by the existence of noise in the experimental system. The DR2, on the other hand, suffered from high-sensitivity to the initial step-size. Its performance was disappointing, in particular in comparison to noise-free calculations that were
7.4. Laboratory Experiments

reported in the past [147, 148]. A proposed explanation for this behavior could be the lack of recombination, which has been shown to be a crucial ES component in noisy environments (see, e.g., [149]).

Covariance Learning

Recording the CMA data during the optimizations allows an analysis of the evolutionary search process. Upon examination of the data, it is found that the covariance matrix remains diagonal during the search (Eq. 1.41), or equivalently, the CMA does not utilize its second-order mechanism (i.e., rotations) when climbing up the landscape. This is not a surprising result, but rather an important piece of experimental evidence toward the corroboration of the OCT landscape analysis as outlined in Corollary 6.1.2.

Figure 7.6 presents a typical CMA run for the optimization of total-SHG in the laboratory and shows the yield and step-size upon function evaluations. Figure 7.7 presents the square-roots of the covariance matrix eigenvalues as a function of the number of experiments as well as the Euclidean distances between the best phase variables of successive iterations, i.e.,

\[ d^{(g+1)} = \| \vec{\phi}^{(g+1)}(\omega) - \vec{\phi}^{(g)}(\omega) \|, \]  

(7.7)

where \( \vec{\phi}^{\text{best}}(\omega) \) is as in Eq. 6.33.

We conducted an equivalent test in a noise-free simulator for the total-SHG problem\(^1\). Figure 7.8 presents a typical CMA run on the simulator. The convergence profile on the simulator is observed to be similar to the laboratory experiment, i.e., rapid climbing-up of the landscape without utilizing the second-order mechanism. However, upon approaching the top of the landscape, one of the covariance matrix eigenvalues dramatically grows, as shown in Figure 7.9. This behavior was observed to be typical in all runs. The corresponding eigenvector is always a flat phase, suggesting that the CMA discovers the invariance of a constant phase on the total-SHG signal. The phase Euclidean trajectories are plotted as well in Figure 7.9, showing some minor activity during this growth stage, corresponding to super-fine tuning of the spectral phase. The yield values, nonetheless, do not seem to be further improved during this process, at least in the precision available. In practice, the parameter adaptation during this fine-tuning stage produces fitness variations below that of the system noise in the laboratory, which explains its absence in laboratory optimizations.

Simulations: Zeroth-Order CMA

Given the experimental observation reported in the previous section, we were interested in testing the CMA while removing its covariance learning

\(^1\)The simulator was implemented in LabView with the Lab2 package.
components. In essence, we leave the CMA only with the step-size as a strategy parameter, and fix the covariance matrix as an identity matrix. This is a zeroth-order ES with normal mutations subject to hyperspheres as the equidensity probability surfaces. In order to assess the zeroth-order CMA behavior on the given QC systems, we conducted additional simulations with two variants of the algorithm:

- $(\mu_W, \lambda)$-CMA with $C = I$
- $(1, \lambda)$-CMA with $C = I$

The simulations were conducted for both systems - total-SHG as well as filtered-SHG - both with a noise-free simulator and a simulator with noise.
Figure 7.8: CMA optimization of the Total-SHG on a **noise-free simulator**. Yield (solid line, left axis) and step-size (dashed line, right log-scaled axis), versus function evaluations.

Figure 7.9: CMA optimization of the Total-SHG on a **noise-free simulator**. Square-root of the 64 eigenvalues of the covariance matrix (solid thin lines, left log-scaled axis), and phase Euclidean trajectories (bold points, right log-scaled axis), versus function evaluations. Missing trajectory points correspond to zero values. The single exploding eigenvalue can easily be identified in this scale.

The results of the simulations show that the CMA performance is not hampered at all on both systems when removing its covariance learning components: the \((\mu_W, \lambda)\)-CMA with \(C = I\) performs as well as the original CMA, in terms of final attained yield and convergence speed. This observation is valid for noise-free as well as for noisy simulations. However, when the weighted recombination operator was removed, the \((1, \lambda)\)-CMA with \(C = I\) did not converge, nor did it even climb-up from the initial yield at the bottom of the landscape. We thus conclude that it is possible to optimize the given simulated QC landscapes by a zeroth-order ES, as long as the weighted-recombination operator is kept.
7.4.2 Discussion

We presented a survey of derandomized Evolution Strategies and a Genetic Algorithm to a set of Quantum Control systems in the laboratory. As far as we know, this was one of the first applications of derandomized ES to experimental QC in general, and the first study to conduct a comparison between ES to GA as well as to explore the evolutionary path of the CMA, in particular. We would like to mention, however, two studies [150, 151] that applied Evolution Strategies to OCE, and explored a specific QC system both in experiments and simulations. The latter studies concluded that the employed Evolution Strategies were promising optimization routines.

While the QC systems examined here possess easily understood global optima, the search is conducted over a highly complex, curvilinear control landscape, which provides a good testbed for optimization algorithms. From the practical point of view, these systems are relatively easy for implementation in the laboratory.

We found that employing the ES variants with default settings unrestriccredly on the given QC landscapes resulted in pre-mature convergence to sub-optimal phases with highly sloped linear profiles. We analyzed this effect, and introduced the wrapping operator into the ES framework. The latter solved the observed problem.

The CMA-ES outperformed the other algorithms in terms of final yield as well as in convergence speed. It introduced a significant increase in convergence speed to the typical performance of the GA in the laboratory and is a promising tool for future laboratory experiments. While analyzing its behavior, it was experimentally confirmed that its second-order mechanism was not utilized when climbing-up the landscape. This may be considered as an experimental corroboration of the OCT landscape analysis.

We also conducted noise-free simulations of the CMA-ES applied to the systems. The latter calculations revealed interesting behavior of the covariance matrix, upon approaching the top of the landscape. A single eigenvalue consistently explodes with a corresponding eigenvector of a flat phase. We suggest that this is due to the fact that the CMA successfully learned the invariance of a constant phase in these problems. Furthermore, we considered zeroth-order versions of the CMA in simulations, where the covariance learning component was removed. The latter performed extremely well, as long as the weighted-recombination operator was kept.