

Optimal laser control of molecular systems: methodology and results

A. AUGER, A. BEN HAJ YEDDER, E. CANCES, and C. LE BRIS

*CERMICS, École Nationale des Ponts et Chaussées
6 & 8, avenue Blaise Pascal, Cité Descartes,
Champs sur Marne, 77455 Marne-La-Vallée Cedex 2, FRANCE*

C. M. DION, A. KELLER and O. ATABEK

*Laboratoire de Photophysique Moléculaire
Laboratoire de Photophysique Moléculaire du CNRS,
Bâtiment 213, Campus d'Orsay, 91405 Orsay, FRANCE*

We report on some mathematical and numerical work related to the control of the evolution of molecular systems using laser fields. More precisely, the control of the orientation of molecules is our goal. We treat this as an optimal control problem and optimize the laser field to be used experimentally by using both deterministic and stochastic algorithms. Comparisons between the different strategies are drawn. In particular, when gradients of the cost functional are used, the different ways for their computation are compared and analyzed.

1. Introduction

We wish to report on theoretical and numerical work devoted to the modeling of the control of chemical reactions by laser fields. The laser control of chemical reactions is indeed a very active field of laser physics, at the crossroads between quantum chemistry, quantum mechanics, and theoretical and experimental femtophysics. Manipulation of molecular systems using laser fields is today an experimental reality [1], provided one restricts his aims to reasonable goals, as will be seen below. This leads to a mostly unexplored field for mathematical analysis and numerical simulation. Numerical simulations can indeed efficiently complement the experimental strategy, both by explaining the deep nature of the phenomena involved and by optimizing the parameters to be used experimentally.

We present here the contributions of our team, which is composed both of mathematicians and physicists. The emphasis is here on the mathematical aspects and

the numerical techniques. A companion article [2] focusing on the physical aspects appears elsewhere. The most striking result of our work is given in [17].

Before we discuss the technicalities, let us briefly state in a rather formal way the problem we shall deal with. All details will be given in Section 2, and for pedagogical purposes we prefer to only give a vague setting in this explanatory survey.

The evolution of a molecular system subjected to a laser field $\vec{\mathcal{E}}$ is modeled by the time-dependent Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = H_0 \psi + \vec{\mathcal{E}}(t) \cdot \vec{D}(\vec{\mathcal{E}}(t)) \psi, \quad (1.1)$$

complemented with the initial condition $\psi(t=0) = \psi_0$. In this equation, the wave function ψ is assumed to depend only on the coordinates of the various nuclei the molecular system is composed of. The presence of the electrons is accounted for through an effective potential acting on the nuclei, and contained in the Hamiltonian H_0 of the free system (when the laser is turned off). We denote by $\vec{D}(\vec{\mathcal{E}}(t))$ the dipole moment of the molecule in presence of an external electric field $\vec{\mathcal{E}}(t)$; at the first order perturbation theory, one can use the form $\vec{D}(\vec{\mathcal{E}}(t)) = \vec{\mu}_0 + \overline{\alpha} \vec{\mathcal{E}}$. More sophisticated models would feature higher order expansion of $\vec{D}(\vec{\mathcal{E}}(t))$ interactions, still in the perturbation setting, or even a true dependence of the wave function ψ and the Hamiltonian H with respect to the coordinates of *all nuclei and electrons* of the molecular system. To the present day, the latter model is out of reach of numerical treatment.

In order to state an optimal control problem, we need, in addition to the direct equation (1.1) modeling the evolution of the system, to define a cost function. Minimizing this cost function will give a formal sense to the physical target we want to reach. In our work, we consider a linear molecular system and intend to orient it in the direction of the linearly polarized laser field. The cost function we adopt will therefore reflect this wish. Just to fix the ideas, let us mention an example of cost function in the simplest case when the state of the system ψ (solution to equation 1.1) is a function of time t and of the angle θ between the axis of the system and the direction of the laser field:

$$J(\mathcal{E}) = \frac{1}{T} \int_{t=0}^{t=T} \int_{\theta=0}^{\theta=\pi} |\psi(t, \theta)|^2 \cos \theta \sin \theta d\theta dt. \quad (1.2)$$

The reason why we choose an orientation problem as our control problem, and consequently such an objective function will be made clear below. Other forms of the cost function will also be given later in this article.

The simple setting we have just indicated above suffices to now underline the peculiarities of the optimal control problem we have to tackle, with respect to other optimal control problems that the reader may have in mind and that come from more usual domains of the engineering world (aeronautics, ...). Let us now emphasize these peculiarities.

From the standpoint of the mathematical theory, this problem is *bilinear* (the control $\vec{\mathcal{E}}$ multiplies the state ψ) which at once puts the problem on a very high level of mathematical difficulty. Indeed, the mathematical theoretical results on bilinear control are very rare. In infinite dimension, i.e., for the PDE (1.1), since the celebrated work by Ball and Slemrod [4], no real progress has been made, to the best of our knowledge. For the finite dimensional approximation of (1.1), there exist some results that can also be extended to the infinite dimensional case but that are not very easy to exploit (so far). We refer to the work of G. Turinici et al. [32, 33, 34] for some recent progress on the theory of exact controllability for systems such as those we deal with here. For the optimal control problem, *some* minor things can be done. We refer in particular to [10] where some of us have proven the existence of an optimal field in a very academic and simplified setting. We shall not elaborate any longer on these theoretical aspects and now concentrate on more practical ones.

A noticeable peculiarity is the fact that, in most cases, the control $\vec{\mathcal{E}}$ is *distributed in time*, and not in space. It is not a crucial fact for the sequel (cases when $\vec{\mathcal{E}}$ depends both on time and space could be treated in the same fashion, however with slightly more tedious computations) but it is rather convenient and constitutes a very reasonable approximation in the case of small molecular systems such as atoms and small molecules. At the scale of such a system, the laser light is indeed seen as homogeneous in space. Such a distributed in time control is not that usual for a partial differential equation such as (1.1).

In addition, special attention must be paid to the fact that although our goal is to drive the system from one initial state to some other specific state through a controlled time-dependent evolution, the cost function we choose to formulate our mathematical problem is not a distance to a target state, but the mean value of an observable (a measure of the orientation of the molecular system with the field). We wish to comment a little bit further on this point. The ultimate goal of the manipulations we want to model is the control of chemical reactions. This means for instance making a system ABC split into $AB+C$ rather than into $A+BC$ (see [9] for an introduction to this problem). Succeeding in making a chemical reaction possible does not necessarily mean driving the initial state to the final one, but sometimes (and in fact most of the times) only succeeding in *preparing* the initial system in a good way so that afterwards the desired reaction spontaneously happens. In that respect, orienting a molecule in space is both a modest and sufficient goal. Once it is conveniently “geometrically” prepared, the goal is almost reached. Nature will do the rest of the job. In addition, there are today experimental evidences showing that aligning a molecule (orientation is one step forward alignment) with a laser *is* feasible, and constitutes a significant step that can be used to efficiently control reactions (see the groundbreaking work by H. Stapelfeldt [29, 26]). The problem of orientation is therefore a good problem to look at.

It is also enlightening to consider this problem from the practical standpoint. Let us first indicate some orders of magnitude. Typically, the space scale is that of

a molecule, namely a few angströms (10^{-10} m), and the time scale is that of the vibration of a molecular bond, namely ten femtoseconds (10^{-14} s). The total time of simulation for equation (1.1) is thus typically the picosecond (10^{-12} s). There are two main consequences of this time scale. First, the control needs to be an *open-loop* control, since it is clear that one cannot update the field in real time with electronic devices. In other words, the only system that can react as fast as the molecular system is precisely the system itself. The second consequence is that we must think of this problem in a completely different way from the way we think of usual control problems: we are here in a framework where we can do thousands of experiments within a minute (while we cannot launch a rocket thousands times). This ability to make many experiments has in turn two consequences. First, one can imagine, and it is indeed done, to couple the numerical search for the optimal field not with the numerical simulation of equation (1.1), but with the experiment itself [1, 24]. The experimental solution of (1.1) is indeed much faster than its resolution on a computer. There is here some matter of reflection for experts in scientific computing. Second, one of the major problems of this field is the tremendous amount of data that are at our disposal. A challenge is to find a way to exploit them in the optimization cycle. We shall not give in this article any definite answer to the questions and concerns raised above, but it is sound to keep in mind these points.

One must also know about the practical parameters for a laser field. One of us has presented in [6, 7] a rapid account of this point, and we refer to it, or to the comprehensive report [13] for more details^b. Let us only say that a trade-off has to be made between the power of the laser, its time resolution, its repetition frequency, and also its price and its size. The laser fields we shall make use of have intensities in the range $[10^{12}, 10^{13}]$ W/cm², are able to have a risetime of the order of 10^{-14} s, and the light they create has frequency around 10^{14} Hz. A very peculiar feature appears here again. One can ask the question whether it is better to optimize upon *only* the fields that are today experimentally feasible or to consider all fields without taking into account any contemporary technological constraint. Both approaches may be useful. In particular, the second one may help in designing the lasers physicists do need for the next generation. In the present article, we mostly choose the first approach, taking explicitly into account the technical requirements. We shall however also explore the second one (see more on this point below when we optimize with ten laser fields).

The stage is now set. Let us say a few words on the methodology we choose for the search of the optimal laser field.

First and foremost, we must emphasize that the present study is far from being the first attempt to find numerically the optimal laser field. There exist many theoretical studies based upon the construction of small systems of ODEs approximating (1.1) so that the optimal (or exact) control problem can be treated explicitly

^bThis report (in French !) presents a broad overview of the domain, indicating current approaches, both theoretical and experimental, and gives trends for the future. Another useful reference in the same spirit is [31].

“by hand”. The leading experts of this approach, fundamentally based upon a deep knowledge of (or intuition of) the main mechanisms are P. Brumer, P. Schapiro and coworkers [9, 8]. Other outstanding contributions, in particular on intense laser fields are due to A. Bandrauk [11]. On the other hand, the optimal control methodology in the sense applied mathematicians speak about it has already been thoroughly explored by physicists, in the first row of which stands H. Rabitz [37]. See also works by Fujimura [23], Sakai [25]. However, in all these contributions, the algorithms used for the numerical search for the optimized field are seen as black boxes, and not as topics for research. Our own approach aims at complementing the work of these leading researchers in physics by exploring the capabilities of the most recent optimization tools, by comparing them to one another on the present problem, by drawing conclusions on the best tools to be used, and also, when possible, by improving the physical conclusions.

On the present problem, we shall investigate mainly the following issues, which are of general interest, but whose response may differ from one problem to another:

- use on this specific case of deterministic algorithms (gradient-like algorithms), of stochastic algorithms (genetic algorithms and evolutionary strategies) and of algorithms mixing the two approaches, such as genetic algorithms accelerated by mutation by gradient
- comparison of the different ways to compute the gradient when needed: discretization of the adjoint equation, computation of the adjoint of the discrete equation, automatic differentiation
- impact of the choice of the cost function on the result, multicriteria approaches,...

The sequel of this article is organized as follows. In the next section, we give a detailed presentation of the problem under study, making more precise the quantities (Hamiltonian H , state ψ , electric field $\vec{\mathcal{E}}$, dipole moment \vec{D} , cost function J) we have described above in a somewhat vague way. Section 3 describes the different optimization methods we shall make use of. For some of them, we shall need to compute the gradient of the cost function. In Section 3.1, we therefore make a numerical analysis to determine which strategy is the best one to compute this gradient. In Section 3.2 we give a short description of stochastic algorithms we employed. Section 4 then gives the results obtained for our problem with deterministic algorithms and with stochastic ones. Finally, in Section 5, we shall summarize our main results and indicate the directions of our current and future research.

2. Statement of the control problem

2.1. *The system under study and the control problem*

The molecular system we study is the linear HCN molecule (hydrogen cyanide). This molecule has been chosen because it is linear in its ground state and should

stay so if the laser frequency is out of resonance with respect to the bending modes. Therefore it constitutes a perfect toy object for testing our methodology. We use the so-called Jacobi coordinates $(\mathbf{R} = (R, r), \theta, \varphi)$ to parameterize the state of the molecule (see Figure 1). The free Hamiltonian H_0 can be written as $H_0 = H_{\text{vib}}(\mathbf{R}) + H_{\text{rot}}(\mathbf{R}, \theta, \varphi) + V(\mathbf{R})$ and the dipole moment is written as $D(\mathcal{E}(t)) = -\mu_0(R, r) \cos \theta - \frac{\mathcal{E}(t)}{2} [\alpha_{\parallel}(R, r) \cos^2 \theta + \alpha_{\perp}(R, r) \sin^2 \theta]$. Then the general form for the Hamiltonian, given in [16], is

$$H(\mathbf{R}, \theta, \varphi, t) = H_{\text{vib}}(\mathbf{R}) + H_{\text{rot}}(\mathbf{R}, \theta, \varphi) + V(\mathbf{R}) + H_{\text{laser}}(\mathbf{R}, \theta, \varphi, t), \quad (2.3)$$

where $T_{\text{rot}} + H_{\text{rot}}$ denotes the kinetic energy operator with

$$H_{\text{vib}}(\mathbf{R}) = -\frac{\hbar^2}{2\mu_{\text{HCN}}} \frac{1}{R^2} \frac{\partial}{\partial R} \left(R^2 \frac{\partial}{\partial R} \right) - \frac{\hbar^2}{2\mu_{\text{CN}}} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right),$$

$$H_{\text{rot}}(\mathbf{R}, \theta, \varphi) = -\frac{\hbar^2}{2(\mu_{\text{HCN}}R^2 + \mu_{\text{CN}}r^2)} \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right],$$

where $V(\mathbf{R})$ denotes the effective potential resulting from the electrostatic interaction between nuclei and electrons (in their ground state), while

$$\begin{aligned} H_{\text{laser}}(\mathbf{R}, \theta, \varphi, t) &= \mathcal{E}(t) \cdot D(\mathcal{E}(t)) \\ &= -\mu_0(R, r) \mathcal{E}(t) \cos \theta - \frac{\mathcal{E}^2(t)}{2} [\alpha_{\parallel}(R, r) \cos^2 \theta + \alpha_{\perp}(R, r) \sin^2 \theta] \end{aligned}$$

denotes the interaction between the molecule and the laser field. In the former

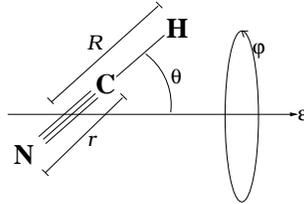


Fig. 1. Model for the HCN molecule.

formulas, μ_{CN} and μ_{HCN} represent the reduced masses:

$$\mu_{\text{CN}} = \frac{m_{\text{C}}m_{\text{N}}}{m_{\text{C}} + m_{\text{N}}}, \quad \mu_{\text{HCN}} = \frac{m_{\text{H}}(m_{\text{C}} + m_{\text{N}})}{m_{\text{H}} + m_{\text{C}} + m_{\text{N}}}$$

and μ_0 is the permanent dipole moment. The coefficients α_{\parallel} and α_{\perp} are respectively the parallel and the perpendicular components of the diagonal polarizability tensor $\bar{\bar{\alpha}}$ given by $\alpha_{\parallel} = \alpha_{zz}$ and $\alpha_{\perp} = \alpha_{xx} = \alpha_{yy}$ when (Oz) is the molecular axis.

As a first step toward the treatment of the sophisticated model (2.3), we consider in all the remainder of this article the case of a rigid rotor: the problem depends

only on the angular variables θ, ϕ . Furthermore, symmetry conservation around the laser polarization axis allows us to separate the motion in ϕ from the motion in θ , and consider only the latter in our calculations. The Hamiltonian (2.3) therefore reduces to

$$H = H(\theta, t) = H_{rot}(\theta) + H_{laser}(\theta, t), \quad (2.4)$$

with

$$H_{rot}(\theta) = -\frac{\hbar^2}{2(\mu_{HCN}R^2 + \mu_{CN}r^2)} \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right)$$

and

$$H_{laser}(\theta, t) = -\mu_0(R, r)\mathcal{E}(t) \cos \theta - \frac{\mathcal{E}^2(t)}{2} [\alpha_{\parallel}(R, r) \cos^2 \theta + \alpha_{\perp}(R, r) \sin^2 \theta],$$

where R and r are fixed at their equilibrium value. The objective function $J(\mathcal{E})$ we are optimizing will be detailed in Section 2.3 but let us now introduce the instantaneous criterion $j(t)$ used to compute $J(\mathcal{E})$ and which is the measure of the orientation at time t (see [21] for more details),

$$j(t) = \langle \cos \theta \rangle = \int_0^{\pi} \cos \theta \mathcal{P}(\theta, t) \sin \theta d\theta, \quad (2.5)$$

where $\mathcal{P}(\theta, t)$ is the angular distribution of the molecule. In the case of rigid rotor angular distribution is reduced to $\mathcal{P}(\theta, t) = \|\psi\|_{\mathcal{C}}^2$ where $\|\psi\|_{\mathcal{C}}^2$ denotes the squared norm of the complex ψ . The instantaneous criterion therefore becomes

$$j(t) = \int_0^{\pi} \cos \theta \|\psi\|_{\mathcal{C}}^2 \sin \theta d\theta. \quad (2.6)$$

The instantaneous criterion $j(t)$ takes its values in the range $[-1, 1]$, the values -1 and 1 corresponding respectively to a molecule pointing in the direction of the laser field polarization axis and in the opposite direction.

The Schrödinger equation

$$\begin{cases} i\hbar \frac{\partial \psi}{\partial t} = H \psi, \\ \psi(t=0) = \psi_0. \end{cases} \quad (2.7)$$

depending only on the variable θ is numerically solved with an operator splitting method [20] coupled with a FFT for the kinetic part as shown in [12, 28]. Table 1 summarizes the parameters of the HCN molecule for R and r fixed at their equilibrium value.

2.2. Choice of the set of electric fields

We now describe the set of laser fields we minimize upon. As said in the introduction, both strategies of restricting oneself to the experimental state of the art or of considering the most general laser fields are of some interest. We begin with the second one, by considering that the electric field $\mathcal{E}(t)$ we have at our disposal is the sum

Table 1. Parameters of the HCN molecule.

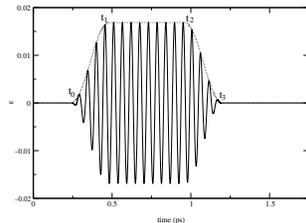
$B = \frac{\hbar^2}{2(\mu_{HCN}R^2 + \mu_{CN}r^2)}$ (a.u.)	μ_0 (a.u.)	α_{\parallel} (a.u.)	α_{\perp} (a.u.)
6.638×10^{-6}	1.141	20.05	8.638

of N (≤ 10) individual linearly-polarized pulses: $\mathcal{E}(t) = \sum_{n=1}^N \mathcal{E}_n(t) \sin(\omega_n t + \phi_n)$.

The envelope functions $\mathcal{E}_n(t)$ are of given sine-square form,

$$\mathcal{E}_n(t) = \begin{cases} 0 & \text{if } t \leq t_{0n} \\ \mathcal{E}_{0n} \sin^2 \left[\frac{\pi}{2} \left(\frac{t-t_{0n}}{t_{1n}-t_{0n}} \right) \right] & \text{if } t_{0n} \leq t \leq t_{1n} \\ \mathcal{E}_{0n} & \text{if } t_{1n} \leq t \leq t_{2n} \\ \mathcal{E}_{0n} \sin^2 \left[\frac{\pi}{2} \left(\frac{t_{3n}-t}{t_{3n}-t_{2n}} \right) \right] & \text{if } t_{2n} \leq t \leq t_{3n} \\ 0 & \text{if } t \geq t_{3n} \end{cases} \quad (2.8)$$

each pulse being characterized by a set of 7 adjustable parameters, namely its frequency ω_n , relative phase ϕ_n , maximum field amplitude \mathcal{E}_{0n} , together with 4 times determining its shape (origin t_{0n} , rise time $t_{1n} - t_{0n}$, plateau $t_{2n} - t_{1n}$, and extinction time $t_{3n} - t_{2n}$). All beams are polarized along the same axis. This makes

Fig. 2. A typical laser field $\mathcal{E}_i(t)$.

a total of $7 \times 10 = 70$ parameters. It should be once more emphasized that by considering such a superposition we do not have in mind to model a situation that is experimentally feasible, but only to generate a “generic” form of signal $\mathcal{E}(t)$.

As it will be seen below, using such a generic field has one main disadvantage (in addition to that obvious huge difficulty to minimize over \mathcal{R}^{70}): the optimized laser field that is obtained through minimization is likely to be too difficult to analyze! Indeed, as we have very pragmatic purposes, we aim at providing the experimenter with a well identified field to generate. Obviously, a typical field obtained by such a minimization and shown on Figure 10 cannot be easily analyzed. Therefore, the main part of our work will be along the first strategy: restrict ourselves to a superposition of two, or at most three, different lasers of the shape of Figure 2^c.

^ceven if the price to pay for this is to lose a little on the optimality

Apart from sticking to experimental reality (for instance a system of two lasers with the same pulsation but with two different phases is nothing else than the same laser with different optical paths), it greatly simplifies the post-treatment of results. In this view, one of our first results has been that when we use 3 lasers, i.e., when we allow for 3 different lasers in the minimization procedure, the algorithm ends up with an optimized field where the third laser has a very small amplitude (see Table 5 in Section 4). In other words, considering two lasers is enough. We shall therefore concentrate on this latter case.

2.3. Choice of the cost function

The cost function is the mathematical formulation of our physical goal. Its choice is so difficult in our context that it has not been done *a priori*, but has been the result of an “iterative process”. We have tested different ones and compared (on mathematical and physical bases) the results they produce. In this process, we have kept in mind the crucial following points: if a function produces (after minimization) a field which is too difficult to understand, it can be replaced however by another (possibly less) efficient that produces more understandable results. Most of the time we shall therefore handle many different cost functions, and not only one.

Basically, our physical goal is twofold:

- we want to have the molecule oriented with the field in a very good way at (at least) one time during the interval of time considered. The criterion for this purpose is:

$$J = \min_{t \in [0, T]} j(t), \quad (2.9)$$

- and/or we want this orientation to be kept as long as possible, even if it is not so perfect. Then the criterion to be used is:

$$J = \frac{1}{T} \int_0^T j(t) dt. \quad (2.10)$$

The latter criterion J is what we have written $J(\mathcal{E})$ in the equation (1.2). Unless otherwise mentioned, we shall deal henceforth with a criterion J that denotes either of the two criteria (2.9) or (2.10). In both formulas, let us recall that $j(t)$ is the quantity introduced in formula (2.6) of Section 2.1 and which the orientation at time t .

In the second setting, it should be made precise that “as long as possible” typically means relatively long compared to the rotation period of the molecule, namely 11 ps for HCN, which indeed is quite a long time in our context.

In the following, we shall call “narrow” (see Figure 3 (a)) a function $j(t)$ produced mainly by the optimization in the first setting and “wide” (see Figure 3 (b)) a function $j(t)$ produced mainly in the second one. Let us also mention that a multicriteria approach is possible and that it may possibly result in obtaining many different minima and/or the best one in some sense to be defined (see Section 4).

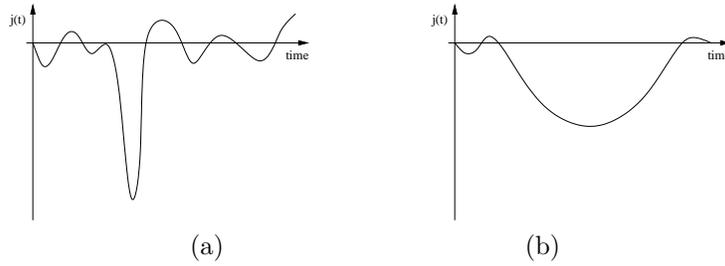


Fig. 3. Typical shape of an optimized $j(t)$ obtained with criterion J_1 (a) and criterion J_2 (b). They are respectively called a “narrow” and a “wide” in the text.

2.4. Identification and classification of the fields obtained

Of primary interest is the need to understand the fields produced by the optimization algorithm. It will allow one to identify the underlying main mechanisms, to imagine scenarii, and to further simplify the electric field to suggest the most simple field to be experimentally generated.

The huge number of optimization processes we have run, with different sets of parameters, with different ranges of values of these parameters, and with different criteria, has resulted in an enormous data set of optimized fields $\mathcal{E}(t)$. We believe that a good way to classify them is:

- fields of the form of a *kick* (see Figure 4), which is an initial sudden (of approximately 0.25 ps, *i.e.*, much shorter than the rotational period of 11 ps) and asymmetric (with respect to its sign) pulse.

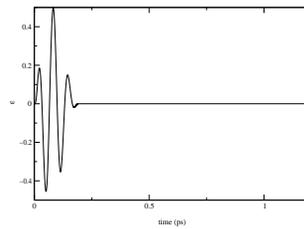


Fig. 4. Example of a “kick” field.

- fields of the form $(\omega, 2\omega)$ (see Figure 5), which are a superposition of two laser fields with the pulsation of one being twice the pulsation of the other one.
- succession in time of two fields with a short overlay time (see Figure 6)
- other types of fields, apparently too complicated to be easily described.

3. Methodology

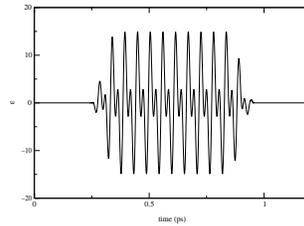
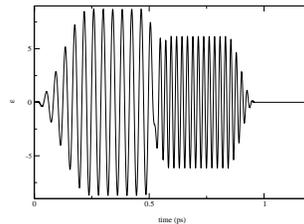
Fig. 5. Example of an $(\omega, 2\omega)$ field.

Fig. 6. Example of a succession of two laser fields.

The way we have tackled the optimization of the orientation problem is based on two different classes of algorithms: first the gradient like algorithms, and second the evolutionary algorithms (EAs). The former ones are purely deterministic and are known to be from far the more rapidly convergent ones but present the drawback from running the risk of remaining trapped in a local minima. The latter ones are stochastic algorithms based on Artificial Darwinism. They are less sensitive to the number of local minima but as they are zero-order methods, the convergence is slower. A way to exploit the forces of both deterministic and stochastic algorithms is to use hybrid methods. We have explored one of these methods with an evolutionary algorithm using a gradient mutation operator.

This section presents in a first part different ways of computing the gradient of the criterion to optimize and compares the different methods. In a second part, this section briefly explains the basic steps of EAs, and next mention which purely stochastic and hybrid EAs have been used for the results presented in Section 4.

3.1. Gradient like algorithms

In this part, we present different ways to compute the gradient of the differentiable cost function $J(\mathcal{E})$, defined by (2.10) (the criterion (2.9) is not differentiable) needed for the gradient-like algorithms. We use two gradient like algorithms: the Polak-Ribière non linear conjugated gradient algorithm with Wolfe or Goldstein-Price line-

search (hereafter abbreviate as PRLS) and the BFGS algorithm. For a complete presentation of these algorithms see [5]. The most natural and the most easiest way to compute the gradient is the finite differences method, which is unfortunately very time consuming. So the need is to find another, less time consuming, way to compute the gradient. The well known adjoint method may be implemented in (at least) two ways: one can either discretize the continuous adjoint equation or one can do the adjoint calculus on the discretized form of the direct equation. It is not clear at all (at least to us) whether there is a general recipe claiming which of the two approaches is the best one. Therefore we shall test both approaches on our specific situation. In fact, the second approach (adjoint calculus on the discretized form) can be itself subdivided into two approaches: the semi-discrete approach, and the fully discrete one (see below). In addition, we shall also compare these methods with that of automatic differentiation (which in principle amounts to doing calculus on the fully discretized form of the equation, but which, in fact differs from this strategy because of implementation details). The tool we use in this latter approach is *Odyssée* [19].

We begin in Section 3.1.1 by presenting the continuous approach which consists in discretizing the continuous adjoint equation. Next Section 3.1.2 details an intermediate approach where one does the adjoint calculus on the semi-discretized equations (which means equations only discretized in time) and next discretizes in space (θ) the so-obtained adjoint problem. In Section 3.1.3 we then compare this approach to the continuous one on a simplified example. In Section 3.1.4, we present the approach (called the discrete approach) consisting in doing adjoint calculus on the fully discretized equations (both in time and space). Finally, in Section 3.1.5, we present the automatic differentiation approach. The numerical results are presented in Section 3.1.6.

3.1.1. *Discretization of the adjoint of the continuous problem*

To find the equations satisfied by the adjoint state p , let us see the control problem as a minimization problem under the constraint $i\hbar\frac{\partial\psi}{\partial t} = H\psi$ and $\psi(t=0) = \psi_0$. We emphasize that this is only a formal method to determine the adjoint problem and to compute the gradient. We shall skip the rigorous verification that the adjoint problem we find is indeed the correct one and that it yields the correct gradient. Using definitions given by Equation (2.4), we will write in this section the Hamiltonian H in the form: $H = H_{rot} + H_{laser}$. We recall that only H_{laser} depends on \mathcal{E} .

Let us first introduce some definitions and notations that we will use throughout this section. For $\mathcal{E}_1, \mathcal{E}_2 \in V_t = L^2([0, T], \mathcal{R})$ we define the scalar product

$$\langle \mathcal{E}_1 | \mathcal{E}_2 \rangle_{t,C} = \int_0^T \mathcal{E}_1(t) \mathcal{E}_2(t) dt,$$

and for $\phi_1, \phi_2 \in V_\theta = L^2([0, 2\pi], \mathcal{C})$ the scalar product

$$\langle \phi_1 | \phi_2 \rangle_{\theta, \mathcal{C}} = \int_0^\pi \Re \left(\phi_1(\theta) \overline{\phi_2(\theta)} \right) \sin \theta d\theta.$$

We also define for $\psi_1, \psi_2 \in V = L^2([0, T] \times [0, 2\pi], \mathcal{C})$ the scalar product

$$\langle \psi_1 | \psi_2 \rangle_{t, \theta, \mathcal{C}} = \int_0^T \int_0^\pi \Re \left(\psi_1(\theta, t) \overline{\psi_2(\theta, t)} \right) \sin \theta d\theta dt.$$

The subscript \mathcal{C} aims at recalling the ‘‘continuous’’ nature of the scalar product, in comparison with the semi-discrete or the discrete ones which will be used later on. We emphasize that when differentiating functions with complex variables we consider these functions as two-variable functions and more precisely the complex variable is taken as an element of \mathcal{R}^2 . For a given laser field \mathcal{E} , we denote by $\psi_\mathcal{E}$ the solution of Equation (1.1). Therefore, we define \tilde{J} using the criterion J as: $J(\mathcal{E}) = \tilde{J}(\psi_\mathcal{E})$. Thus for $\mathcal{E} \in V_t$ and $(\psi, p) \in V^2$ we write the Lagrangian \mathcal{L}^C of the continuous problem as follows:

$$\begin{aligned} \mathcal{L}^C(\mathcal{E}, \psi, p) &= \tilde{J}(\psi) + \left\langle \left(i\hbar \frac{\partial}{\partial t} - H_{rot} - H_{laser} \right) \psi \middle| p \right\rangle_{t, \theta, \mathcal{C}} \\ &\quad + \langle \psi(\cdot, t=0) - \psi^0 | p(\cdot, t=0) \rangle_{\theta, \mathcal{C}}. \end{aligned} \quad (3.11)$$

With standard, but tedious, calculations mainly based upon the linearity of the scalar product and that of the operators H_{rot} and H_{laser} , with an integration by part and with

$$\tilde{J}'(\psi) \cdot \delta\psi = \frac{1}{T} \int_0^T \int_0^\pi \Re [2\bar{\psi} \cos \theta \delta\psi] \sin \theta d\theta dt, \quad (3.12)$$

we obtain $\frac{\partial \mathcal{L}^C}{\partial \psi}(\mathcal{E}, \psi_\mathcal{E}, p) \cdot \delta\psi$, which when set to zero gives the adjoint problem

$$\begin{cases} i\hbar \frac{\partial p}{\partial t} = H_{rot} p + H_{laser} p - \frac{2}{T} \psi_\mathcal{E} \cos \theta, \\ p(T) = 0. \end{cases} \quad (3.13)$$

We next formally compute the gradient $\nabla^C J$ using the Lagrangian \mathcal{L}^C . When using $\psi = \psi_\mathcal{E}$ the expression of the Lagrangian is $\mathcal{L}^C(\mathcal{E}, \psi_\mathcal{E}, p) = \tilde{J}(\psi_\mathcal{E}) = J(\mathcal{E})$, thus we get

$$J'(\mathcal{E}) \cdot \delta\mathcal{E} = \frac{\partial \mathcal{L}^C}{\partial \psi}(\mathcal{E}, \psi_\mathcal{E}, p) \cdot \frac{\partial \psi_\mathcal{E}}{\partial \mathcal{E}} \cdot \delta\mathcal{E} + \frac{\partial \mathcal{L}^C}{\partial \mathcal{E}}(\mathcal{E}, \psi_\mathcal{E}, p) \cdot \delta\mathcal{E},$$

which is simplified into $J'(\mathcal{E}) \cdot \delta\mathcal{E} = \frac{\partial \mathcal{L}^C}{\partial \mathcal{E}}(\mathcal{E}, \psi_\mathcal{E}, p) \cdot \delta\mathcal{E}$ when p is the adjoint state $p_\mathcal{E}$. Therefore, the gradient $\nabla^C J = \frac{dJ}{d\mathcal{E}}$ is obtained by

$$\langle \nabla^C J | \delta\mathcal{E} \rangle_{t, \mathcal{C}} = \frac{\partial \mathcal{L}^C}{\partial \mathcal{E}}(\mathcal{E}, \psi_\mathcal{E}, p_\mathcal{E}) \cdot \delta\mathcal{E}$$

$$\begin{aligned}
&= \left\langle - \left(\frac{\partial H_{laser}}{\partial \mathcal{E}}(\mathcal{E}) \cdot \delta \mathcal{E} \right) \psi_{\mathcal{E}} \middle| p_{\mathcal{E}} \right\rangle_{t, \theta, C} \\
&= \int_0^T \int_0^\pi \Re \left[(\mu_0 \cos \theta + \mathcal{E} [\alpha_{\parallel} \cos^2 \theta + \alpha_{\perp} \sin^2 \theta]) \psi_{\mathcal{E}} \overline{p_{\mathcal{E}}} \delta \mathcal{E} \right] \sin \theta d\theta dt.
\end{aligned}$$

The discretization of Equation (2.7) is done with an operator splitting method,

$$\begin{cases} \psi^0, \\ \psi^{n+1} = e^{-\frac{i}{\hbar} \frac{\Delta t}{2} H_{laser}^n} e^{-\frac{i}{\hbar} \Delta t H_{rot}} e^{-\frac{i}{\hbar} \frac{\Delta t}{2} H_{laser}^n} \psi^n, \end{cases} \quad (3.14)$$

where H_{laser}^n is the time-dependent operator taken at time step t^n . Using this scheme to discretize the linear part of Equation (3.13) we obtain

$$\begin{cases} p^N = 0, \\ p^{n-1} = e^{\frac{i}{\hbar} \frac{\Delta t}{2} H_{laser}^n} e^{\frac{i}{\hbar} \Delta t H_{rot}} e^{\frac{i}{\hbar} \frac{\Delta t}{2} H_{laser}^n} p^n + \frac{2}{T} \psi_{\mathcal{E}}^n \cos \theta \frac{\Delta t}{i\hbar}. \end{cases} \quad (3.15)$$

In addition, we use the same schemes for the time and space discretizations as the ones used for computing $J(\mathcal{E})$. More precisely, we use for time discretization a simple Riemann rule integration scheme. For the integration in θ , the method used is the Simpson rule

$$\begin{aligned}
\int_0^\pi g(\theta) d\theta &= \frac{\pi}{2N} \sum_{k=1}^{2N} \alpha_k g(\theta_k) \\
&= \frac{\Delta \theta}{3} \left[g(\theta_0) + 4 \sum_{k=1}^{N-1} g(\theta_{2k+1}) + 2 \sum_{k=1}^{N-2} g(\theta_{2k+2}) + g(\theta_N) \right] \quad (3.16)
\end{aligned}$$

where $\Delta \theta = \frac{\pi}{2N}$ and where $(\theta_k)_{k=0, 2N}$ are the equally-spaced integration points. Therefore the discretization of the gradient (3.14) reads, with an approximation in $(\Delta t)^2$ and in $(\Delta \theta)^4$,

$$\langle \nabla^C J | \delta \mathcal{E} \rangle_{t, C} =$$

$$\sum_{n=0}^{N-1} \sum_{k=0}^{2M} \Re \left[\mu_0 \cos \theta_k + \mathcal{E}^n [\alpha_{\parallel} \cos^2 \theta_k + \alpha_{\perp} \sin^2 \theta_k] \psi_k^n \overline{p_k^n} \right] \delta \mathcal{E}^n \Delta t \alpha_k \sin \theta_k \Delta \theta. \quad (3.17)$$

3.1.2. Adjoint calculus on the semi-discretized equations

The discretization of the time-dependent Schrödinger Equation (2.7) is given by (3.14) while the discretization of the criterion (again by the Riemann rule integration scheme) yields the semi-discrete Lagrangian

$$\begin{aligned}
\mathcal{L}^{SD}(\mathcal{E}, \Psi, P) &= \frac{1}{T} \sum_{n=0}^{N-1} \int_0^\pi \|\psi^n\|_C^2 \cos \theta \sin \theta d\theta \Delta t + \langle \psi^0 - \psi_0 | p^0 \rangle_{\theta, C} \\
&\quad + \left\langle \Psi^S - e^{-\frac{i}{\hbar} \frac{\Delta t}{2} H_{laser}} e^{-\frac{i}{\hbar} \Delta t H_{rot}} e^{-\frac{i}{\hbar} \frac{\Delta t}{2} H_{laser}} \Psi \middle| P \right\rangle_{t, \theta, SD} \quad (3.18)
\end{aligned}$$

where $\Psi^S = (\psi^1, \dots, \psi^N)$, $\Psi = (\psi^0, \dots, \psi^{N-1})$ and $P = (p^0, \dots, p^{N-1})$ are elements of $(V_\theta)^N$ and where $\mathcal{E} = (\mathcal{E}^0, \dots, \mathcal{E}^{N-1})$ is an element of \mathcal{R}^N . The scalar product $\langle \cdot | \cdot \rangle_{\theta, C}$ is the one given in the previous section and the scalar product

$\langle \cdot | \cdot \rangle_{t, \theta, SD}$ is given by $\langle \Psi_1 | \Psi_2 \rangle_{t, \theta, SD} = \sum_{n=1}^N \langle \psi_1^n | \psi_2^n \rangle_{\theta, C} \Delta t$ with $\Psi_1, \Psi_2 \in (V_\theta)^N$. We

also define for $\mathcal{E}_1, \mathcal{E}_2 \in \mathcal{R}^n$ the scalar product $\langle \mathcal{E}_1 | \mathcal{E}_2 \rangle_{t, SD} = \sum_{n=1}^N \mathcal{E}_1^n \mathcal{E}_2^n$. For a given laser field \mathcal{E} we denote $\Psi_\mathcal{E}$ the solution of Equation (3.14). As in the previous section, by computing $\frac{\partial \mathcal{L}^{SD}}{\partial \Psi}(\mathcal{E}, \Psi_\mathcal{E}, P) \cdot \delta \Psi$ and then by setting it to zero we get the following discrete adjoint problem:

$$\begin{cases} p^{N-1} = 0, \\ p^{n-1} = e^{\frac{i}{\hbar} \frac{\Delta t}{2} H_{laser}^n} e^{\frac{i}{\hbar} \Delta t H_{rot}} e^{\frac{i}{\hbar} \frac{\Delta t}{2} H_{laser}^n} p^n - \frac{2}{T} \psi_\mathcal{E}^n \cos \theta. \end{cases} \quad (3.19)$$

And the gradient is obtained by:

$$\begin{aligned} & \frac{\partial \mathcal{L}^{SD}}{\partial \mathcal{E}}(\mathcal{E}, \Psi_\mathcal{E}, P_\mathcal{E}) \cdot \delta \mathcal{E} \\ &= \sum_{n=0}^{N-1} \left\langle -\frac{i \Delta t}{\hbar} \frac{\partial H_{laser}^n}{\partial \mathcal{E}_n} \delta \mathcal{E}^n e^{-\frac{i}{\hbar} \frac{\Delta t}{2} H_{laser}^n} e^{-\frac{i}{\hbar} \Delta t H_{rot}} e^{-\frac{i}{\hbar} \frac{\Delta t}{2} H_{laser}^n} \psi^n \middle| p^n \right\rangle_{\theta, C} \\ &= \sum_{n=0}^{N-1} \left\langle -\frac{i \Delta t}{\hbar} \frac{\partial H_{laser}^n}{\partial \mathcal{E}_n} \delta \mathcal{E}^n e^{-\frac{i}{\hbar} \frac{\Delta t}{2} H_{laser}^n} \psi^{n+1} \middle| p^n \right\rangle_{\theta, C}. \end{aligned} \quad (3.20)$$

Thus, with an approximation at the order $(\Delta \theta)^4$, we obtain

$$\begin{aligned} & \langle \nabla^{SD} J | \delta \mathcal{E} \rangle_{t, SD} = \\ & - \sum_{n=0}^{N-1} \sum_{k=0}^{2M} \Re \left[\frac{i}{\hbar} \psi_k^{n+1} \bar{p}_k^n (\mu_0 \cos \theta_k + \mathcal{E}^n [\alpha_{\parallel} \cos^2 \theta_k + \alpha_{\perp} \sin^2 \theta_k]) \delta \mathcal{E}^n \right] (\Delta t)^2 \sin \theta_k \alpha_k \Delta \theta. \end{aligned} \quad (3.21)$$

3.1.3. Comparison of the continuous and the semi-discretized approaches

In order to understand which of the formulae (3.17) or (3.21) is more accurate, we give below some illustrative example. Although very basic, this example allows one to understand the fundamental difference between formulae (3.17) and (3.21). Let us argue on the following Schrödinger equation:

$$\begin{cases} i \hbar \frac{\partial \psi}{\partial t} = \psi \mathcal{E}(t) \cos \theta, \\ \psi(0) = \psi^0, \end{cases} \quad (3.22)$$

(obtained by simply setting H_0 to zero in (2.7)) with the criterion written in the form

$$J(\mathcal{E}) = \tilde{J}(\psi) = \frac{1}{T} \int_0^T \int_0^\pi f(\psi) \sin \theta d\theta dt.$$

The first way to proceed is the one we have followed in Section 3.1.1, namely by discretizing the adjoint equation. For Equation (3.22), basic calculus shows that the adjoint equation is given by

$$\begin{cases} i\hbar \frac{\partial p}{\partial t} = \mathcal{E}(t)p \cos \theta - \overline{f'}(\psi_{\mathcal{E}}), \\ p(T) = 0, \end{cases} \quad (3.23)$$

which once discretized with the same scheme as the one used for the direct equation, yields

$$\begin{cases} p^N = 0, \\ p^{n-1} = e^{\frac{i}{\hbar} \frac{\Delta t}{2} H_{laser}^n} e^{\frac{i}{\hbar} \Delta t H_{rot}} e^{\frac{i}{\hbar} \frac{\Delta t}{2} H_{laser}^n} p^n - \overline{f'}(\psi_{\mathcal{E}}^n) \frac{\Delta t}{i\hbar}. \end{cases} \quad (3.24)$$

We compute the gradient of the criterion,

$$\nabla J(\mathcal{E}) = \frac{1}{T} \int_0^T \int_0^\pi f'(\psi_{\mathcal{E}}) \frac{\partial \psi}{\partial \mathcal{E}} \sin \theta d\theta dt,$$

where $\frac{\partial \psi}{\partial \mathcal{E}}$ solves

$$\begin{cases} i\hbar \frac{\partial}{\partial t} \left(\frac{\partial \psi}{\partial \mathcal{E}} \right) = \psi \cos \theta + \mathcal{E}(t) \cos \theta \frac{\partial \psi}{\partial \mathcal{E}}, \\ \left. \frac{\partial \psi}{\partial \mathcal{E}} \right|_{t=0} = 0. \end{cases} \quad (3.25)$$

Thus, by using Equation (3.23) and integration by part we obtain

$$\langle \nabla J | \delta \mathcal{E} \rangle_{t,C} = - \int_0^T \left(\int_0^\pi \psi \overline{p} \cos \theta \sin \theta d\theta \right) \delta \mathcal{E} dt. \quad (3.26)$$

We now discretize this integral by the Riemann scheme which yields

$$\int_0^T \left(\int_0^\pi \psi \overline{p} \cos \theta \sin \theta d\theta \right) \delta \mathcal{E} dt = \sum_{n=0}^{N-1} \left(\int_0^\pi \psi^n \overline{p}^n \cos \theta \sin \theta d\theta \right) \delta \mathcal{E}^n \Delta t,$$

and thus the following approximation of the gradient:

$$\langle \nabla J(\psi) | \delta \mathcal{E} \rangle_{t,C} = - \sum_{n=0}^{N-1} \left(\int_0^\pi \psi^n \overline{p}^n \cos \theta \sin \theta d\theta \right) \delta \mathcal{E}^n \Delta t, \quad (3.27)$$

is the exact analogous of formula (3.17). Using this Riemann discretization scheme, the numerical error is controlled by the following estimate:

$$\left| \int_0^T g(t) dt - \sum_{n=0}^{N-1} g^n \Delta t \right| \leq T \Delta t \|g'\|_{L^\infty}. \quad (3.28)$$

Applying this result to $g = \left(\int_0^\pi \psi \bar{p} \cos \theta \sin \theta d\theta \right) \delta \mathcal{E}$, we obtain the control of the numerical error of the approximation (3.27) of the gradient

$$\begin{aligned} |\varepsilon_{\Delta t}^C| &\leq T \Delta t \left\| \frac{\partial}{\partial t} \left[\left(\int_0^\pi \psi \bar{p} \cos \theta \sin \theta d\theta \right) \delta \mathcal{E} \right] \right\|_{L^\infty} \\ &\leq T \Delta t \left\| \left[\frac{\partial}{\partial t} \left(\int_0^\pi \psi \bar{p} \cos \theta \sin \theta d\theta \right) \right] \delta \mathcal{E} + \left(\int_0^\pi \psi p \cos \theta \sin \theta d\theta \right) \frac{\partial}{\partial t} (\delta \mathcal{E}) \right\|_{L^\infty} \\ &\leq CT \Delta t \left(\|\delta \mathcal{E}\|_{L^\infty} + \left\| \frac{\partial}{\partial t} (\delta \mathcal{E}) \right\|_{L^\infty} \right), \end{aligned} \quad (3.29)$$

where the constant C depends on norms of $\psi|_{t=0}$ and \mathcal{E} but not on $\delta \mathcal{E}$.

On the other hand, if we now discretize the equation and the criterion, we obtain as in (3.20)

$$\langle \nabla J | \delta \mathcal{E} \rangle_{t,SD} = \sum_{n=0}^{N-1} \left(\int_0^\pi \psi^{n+1} p^n \cos \theta \sin \theta d\theta \right) \delta \mathcal{E}^n \Delta t. \quad (3.30)$$

Applying the same numerical analysis, we see that the error in the approximation of the gradient is now obtained by setting $g = \int_0^\pi f'(\psi) \delta \psi \sin \theta d\theta$ in (3.28), which yields

$$|\varepsilon_{\Delta t}^{SD}| \leq T \Delta t \left\| \int_0^\pi \left(\frac{\partial}{\partial t} [f'(\psi) \delta \psi] \right) \sin \theta d\theta \right\|_{L^\infty}. \quad (3.31)$$

Now

$$\frac{\partial}{\partial t} (f'(\psi) \delta \psi) = f''(\psi) \frac{\partial \psi}{\partial t} \delta \psi + f'(\psi) \frac{\partial (\delta \psi)}{\partial t},$$

where $\frac{\partial \psi}{\partial t} = \frac{1}{i\hbar} (\mathcal{E} \psi \cos \theta)$ and $\frac{\partial (\delta \psi)}{\partial t} = \frac{1}{i\hbar} ((\delta \mathcal{E}) \psi \cos \theta + \mathcal{E} x (\delta \psi))$. It follows that

$$\left\| \int_0^\pi \left(\frac{\partial}{\partial t} [f'(\psi) \delta \psi] \right) \sin \theta d\theta \right\|_{L^\infty} \leq C \|\delta \mathcal{E}\|_{L^\infty},$$

where C only depends on norms on $\psi|_{t=0}$ and \mathcal{E} . Therefore

$$|\varepsilon_{\Delta t}^{SD}| \leq CT \Delta t \|\delta \mathcal{E}\|_{L^\infty}. \quad (3.32)$$

Comparing this estimate to (3.29), we see that the control in (3.32) is better, in particular for variations $\delta \mathcal{E}$ of \mathcal{E} that have large variations in time, which will precisely be the case for us (oscillatory laser fields). It is therefore expected that *in our case* the adjoint calculus on the discrete equation will yield a better accuracy for the computation of the gradient than the approach consisting in discretizing the continuous adjoint equation. Let us emphasize that the main difference between the two approaches is the following formal (non rigorous) integration by parts:

$$\int_0^T f'(\psi) \delta \psi \approx \int_0^T \left(\frac{\partial p}{\partial t} \right) \delta \psi \approx \int_0^T p \left(\frac{\partial \delta \psi}{\partial t} \right) \approx \int_0^T p \psi \delta \mathcal{E},$$

which is done before or after discretization and thus allows one to have the control of the error basically either by

$$\frac{\partial}{\partial t} (f'(\psi)\delta\psi) \approx f'(\psi)\delta\mathcal{E} \quad (3.33)$$

or by

$$\frac{\partial}{\partial t} (p\psi\delta\mathcal{E}) \approx p\psi\frac{\partial}{\partial t} (\delta\mathcal{E}). \quad (3.34)$$

In the case (3.33) the numerical error of integration is reported on ψ and $\delta\psi$ while in the case (3.34) the numerical error of integration is directly reported on $\delta\mathcal{E}$. Figure 7 summarizes the main ideas presented here.

3.1.4. Adjoint calculus on the fully discretized equations

In this section we begin by discretizing Equation (2.7) both in time and in θ -space and then do the adjoint calculus. The numerical propagation of the θ operator H_θ and the laser operator H_{laser} can be written in the matrix form

$$\Psi^{n+1} = A_\theta^n B A_\theta^n \Psi^n,$$

where Ψ^n is the vector (ψ_k^n) , where A_θ^n is the diagonal matrix of the laser operator propagation, and where B is the matrix corresponding to the θ operator propagation. Only the matrix A_θ^n depends on the laser field \mathcal{E} .

We write the discrete Lagrangian as follows:

$$\begin{aligned} \mathcal{L}^D(\mathcal{E}, \underline{\Psi}, \underline{P}) &= \sum_{n=0}^{N-1} \sum_{k=0}^{2M} \frac{1}{T} \|\psi_k^n\|_{\mathcal{C}}^2 \cos \theta \sin \theta \alpha_k \Delta \theta \Delta t \\ &+ \langle \underline{\Psi}^S - \underline{\Psi}^M | \underline{P} \rangle_{t,D} + \langle \Psi^0 - \Psi_0 | P^0 \rangle_{\theta,D}, \end{aligned} \quad (3.35)$$

where

$$\underline{\Psi}^S = (\Psi^1, \dots, \Psi^N),$$

$$\underline{\Psi}^M = (A_\theta^0 B A_\theta^0 \Psi^0, \dots, A_\theta^{N-1} B A_\theta^{N-1} \Psi^{N-1}),$$

and

$$\underline{P} = (P^0, \dots, P^{N-1})$$

are elements of $(\mathcal{R}^{2M})^N$ and where $\mathcal{E} = (\mathcal{E}^0, \dots, \mathcal{E}^{N-1})$ is an element of \mathcal{R}^N .

The scalar product $\langle \cdot | \cdot \rangle_{t,D}$ is given for $\underline{\Psi}_1, \underline{\Psi}_2 \in (\mathcal{R}^{2M})^N$ by

$$\langle \underline{\Psi}_1 | \underline{\Psi}_2 \rangle_{t,D} = \sum_{n=1}^N \langle \Psi_1^n | \Psi_2^n \rangle_{\theta,D} \Delta t$$

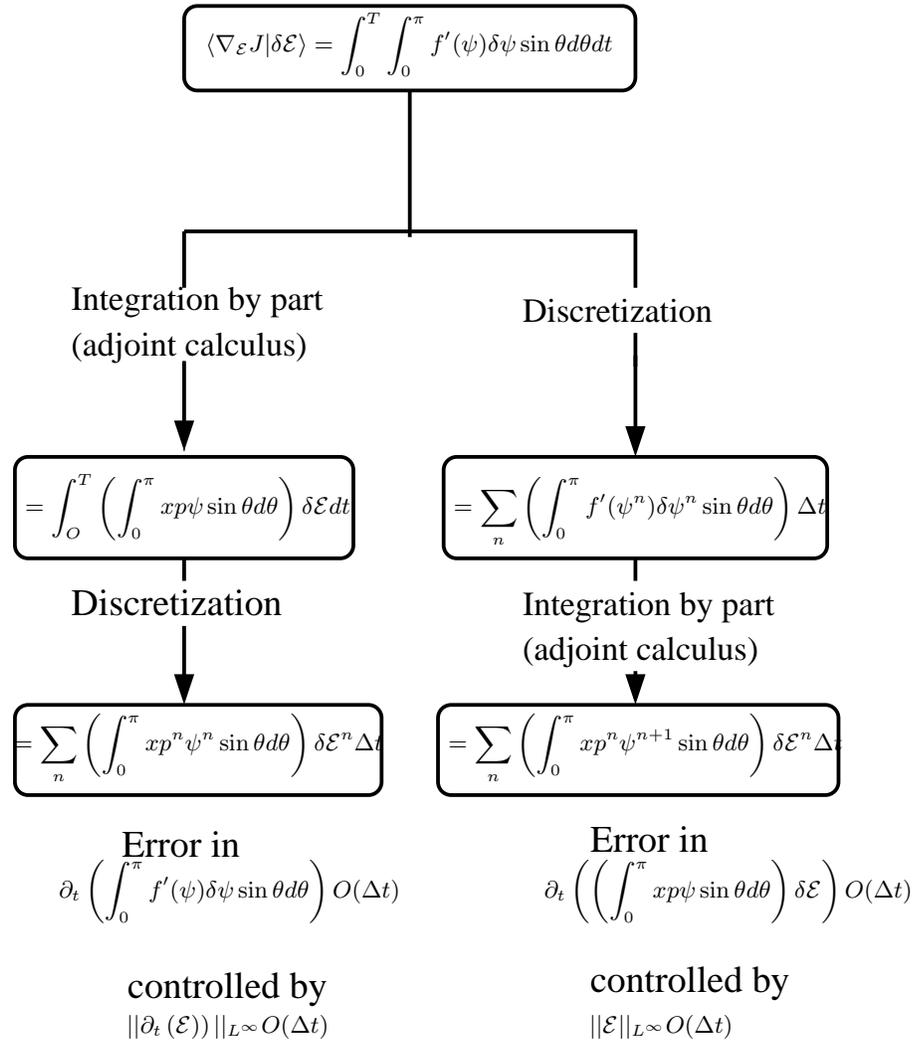


Fig. 7. Comparison of the two approaches of Section 3.1.1 and 3.1.2 to compute the gradient.

with $\langle \Psi_1 | \Psi_2 \rangle_{\theta, D} = \sum_{n=0}^{2M} \Re(\psi_{1k} \overline{\psi_{2k}}) \alpha_k \sin \theta \Delta \theta$ when $\Psi_1, \Psi_2 \in \mathcal{R}^{2M}$.

Therefore, equation $\frac{\partial \mathcal{L}^D}{\partial \underline{\Psi}}(\mathcal{E}, \underline{\Psi}, \underline{P}) \cdot \delta \underline{\Psi} = 0$, we get

$$\begin{cases} P^{N-1} = 0 \\ p_k^{n-1} = \left[(A_\theta^n \overline{B} A_\theta^n)^T P^n \right]_k - \alpha_k \frac{2}{T} \psi_k^n \cos \theta. \end{cases} \quad (3.36)$$

The gradient is obtained by

$$\begin{aligned} & \frac{\partial \mathcal{L}^D}{\partial \mathcal{E}}(\mathcal{E}, \underline{\Psi}, \underline{P}) \cdot \delta \mathcal{E} \\ &= - \left\langle \left(\frac{\partial \underline{\Psi}^M}{\partial \mathcal{E}} \right) \cdot \delta \mathcal{E} \middle| \underline{P} \right\rangle_{t, D} \\ &= - \sum_{n=0}^{N-1} \sum_{k=0}^{2M} \Re \left(\left[\left(\frac{\partial A_\theta^n}{\partial \mathcal{E}^n} B A_\theta^n + A_\theta^n B \frac{\partial A_\theta^n}{\partial \mathcal{E}^n} \right) \Psi^n \right]_k \overline{p_k^n} \delta \mathcal{E}^n \right) \alpha_k \sin \theta \Delta \theta \Delta t, \end{aligned}$$

where $\frac{\partial A_\theta^n}{\partial \mathcal{E}^n}$ is the matrix obtained by differentiating the matrix A_θ^n . We obtain for the gradient formula

$$\langle \nabla^D J | \delta \mathcal{E} \rangle_{t, D} =$$

$$\sum_{n=0}^{N-1} \sum_{k=0}^{2M} \Re \left[(\mu_0 \cos \theta_k + \mathcal{E}^n [\alpha_{\parallel} \cos^2 \theta_k + \alpha_{\perp} \sin^2 \theta_k]) \widetilde{\psi}_k^n \overline{p_k^n} \delta \mathcal{E}^n \right] \alpha_k \sin \theta \Delta \theta \Delta t, \quad (3.37)$$

where $\widetilde{\Psi}^n = \left(\frac{\partial A_\theta^n}{\partial \mathcal{E}^n} B A_\theta^n + A_\theta^n F^{-1} A B F \frac{\partial A_\theta^n}{\partial \mathcal{E}^n} \right) \Psi^n$. This formula is to be compared with (3.17) and (3.21).

3.1.5. Computing the gradient using Automatic Differentiation tools

In this section we briefly present another method to compute the gradient, which uses the Automatic Differentiation tool *Odyssée* [36, 19]. Automatic Differentiation tools can be seen as black boxes taking as input a program computing a cost function $f : \mathcal{R}^n \rightarrow \mathcal{R}^m$ and giving as output another program computing the gradient $\frac{\partial f(x)}{\partial x}$.

Odyssée is able to use two modes: the tangent mode and the adjoint mode, which is similar to the adjoint method. We emphasize that the cost of the gradient computation is proportional to n with the tangent mode (as with finite differences) and it is proportional to m with the adjoint mode. Thus, the tangent mode has to be used when $n \ll m$ and the adjoint mode has to be used when $n \gg m$. For our problem we have $m = 1$ and $1 \ll n \leq 70$, so we use only the adjoint mode. As in

the direct program we have 50 000 iterations, the adjoint program needs a lot of memory to run. In order to reduce the size of memory needed, we have modified the adjoint program by deleting the temporary variables in the linear parts of the program. Table 2 gives an idea of the size of the direct code and that of the adjoint code. The calculation times refer to a Pentium II, 466 Mhz Celeron with 128 Mb RAM running with Linux.

Table 2. Technical requirements with *Odyssée*.

	direct code	standard adjoint code	post-processed adjoint code
Size (lines)	433	2075	1190
Memory needed	12 Ko	520 Mo	103 Mo
Time (CPU)	60 s	—	141 s

3.1.6. Numerical results

The purpose of this section is to compare numerically the different methods presented above for computing the gradient. For the numerical tests we use one laser field of the form $\mathcal{E}(t) = E \sin(\omega t + \phi)$. The gradient with respect to the parameters E , ω and ϕ is denoted as

$$\nabla J = \begin{pmatrix} \nabla_E J \\ \nabla_\omega J \\ \nabla_\phi J \end{pmatrix}.$$

We have computed the gradient using the methods presented in the previous sections, more precisely the continuous approach (C), the semi-discrete approach (SD), the discrete approach (D), and *Odyssée* (AD). We have also computed the gradient using the finite differences approach (FD), where for each variable x ($x = E, \omega, \phi$)

$$\nabla_x^{FD} J = \lim_{\delta x \rightarrow 0} \frac{J(x + \delta x) - J(x)}{\delta x}.$$

The gradient given by FD has been computed with different values of δx to make sure that we have reached the $\delta x \rightarrow 0$ limit. Next we compare the gradient obtained using the different approaches with the gradient obtained using the finite differences approach, which is therefore taken as a reference value. For each method we will compute the relative error

$$e_{E,\omega,\phi} = \left| \frac{\nabla_{E,\omega,\phi} J - \nabla_{E,\omega,\phi}^{FD} J}{\nabla_{E,\omega,\phi}^{FD} J} \right|.$$

The comparison is done for both low and high frequencies, using two different representative points $(E, \omega, \phi) = (10^{11} \text{ W/cm}^2, 500 \text{ cm}^{-1}, 0)$ and $(E, \omega, \phi) = (10^{11} \text{ W/cm}^2, 4000 \text{ cm}^{-1}, 0)$, respectively.

Table 3 shows that all the methods we have presented in this section give good

results compared to the finite differences method. We can also see on this table that the best results are obtained using *Odyssée*, where we have a better precision than with the other methods. In general the precision is increased by at least one order.

We can also see on this table that the results agree with the comparison we made in Section 3.1.3 between the continuous approach and the semi-discrete approach, except for the ∇_E component.

Let us now take the results given by automatic differentiation as a reference and make the same comparison with the other methods as we have done with the finite differences approach. On Table 4 we see that, compared to the AD approach, the best results are those given by the discrete approach (again except for the component ∇_E). We recall that with the discrete approach we make the adjoint calculus on the fully discretized equation and that the automatic differentiation tools make also adjoint calculus on the fully discretized equation with some implementation differences. We also emphasize that for the component ∇_E we obtain results which are different from the results we obtain with the components ∇_ω and ∇_ϕ . We still unable to explain such a difference.

In practice, the size of the parameter vector for our problem can go up to 70, so we can only use an adjoint based method and not the finite differences one. Indeed the CPU time needed to compute the gradient depends on the parameter vector size for the finite differences method and is independent of this size for the other methods presented below. More precisely, the CPU time needed for these other methods is about 3 times the one needed to compute the criterion. For implementation, the continuous approach and the semi-discrete approach are easiest to implement than the discrete approach. Finally, for *Odyssée*, let us recall that even if it gives automatically the gradient, some post-processing of the adjoint code is needed before running it.

Table 3. Relative error, with respect to the FD, of the gradient.

(a) : gradient computed at $(E, \omega, \phi) = (10^{11}, 500, 0)$					
	FD	AD	C	SD	D
e_E	0.	$25. \times 10^{-8}$	$89. \times 10^{-6}$	$20. \times 10^{-3}$	$41. \times 10^{-3}$
e_ω	0.	$92. \times 10^{-6}$	$13. \times 10^{-4}$	$47. \times 10^{-5}$	$32. \times 10^{-5}$
e_ϕ	0.	$18. \times 10^{-6}$	$81. \times 10^{-5}$	$25. \times 10^{-5}$	$25. \times 10^{-5}$
(b) : gradient computed at $(E, \omega, \phi) = (10^{11}, 4000, 0)$					
	FD	AD	C	SD	D
e_E	0.	$75. \times 10^{-5}$	$68. \times 10^{-5}$	$20. \times 10^{-3}$	$40. \times 10^{-3}$
e_ω	0.	$11. \times 10^{-5}$	$76. \times 10^{-4}$	$51. \times 10^{-4}$	$25. \times 10^{-4}$
e_ϕ	0.	$22. \times 10^{-3}$	$41. \times 10^{-3}$	$33. \times 10^{-3}$	$26. \times 10^{-3}$

3.2. Evolutionary Algorithms

This section presents the stochastic algorithms that we have used for the orientation problem which belong to the family of Evolutionary Algorithms (EAs). Their

Table 4. Relative error, with respect to the AD, of the gradient.

(a) : gradient computed at $(E, \omega, \phi) = (10^{11}, 500, 0)$					
	FD	AD	C	SD	D
e_E	$25. \times 10^{-8}$	0.	$89. \times 10^{-6}$	$20. \times 10^{-3}$	$41. \times 10^{-3}$
e_ω	$92. \times 10^{-6}$	0.	$14. \times 10^{-4}$	$57. \times 10^{-5}$	$22. \times 10^{-5}$
e_ϕ	$18. \times 10^{-6}$	0.	$79. \times 10^{-5}$	$23. \times 10^{-5}$	$27. \times 10^{-5}$
(b) : gradient computed at $(E, \omega, \phi) = (10^{11}, 4000, 0)$					
	FD	AD	C	SD	D
e_E	$75. \times 10^{-5}$	0.	$73. \times 10^{-6}$	$20. \times 10^{-3}$	$41. \times 10^{-3}$
e_ω	$11. \times 10^{-5}$	0.	$75. \times 10^{-4}$	$50. \times 10^{-4}$	$24. \times 10^{-4}$
e_ϕ	$22. \times 10^{-3}$	0.	$18. \times 10^{-3}$	$10. \times 10^{-3}$	$36. \times 10^{-4}$

common feature is to imitate the principle of natural evolution. This section is organized as follows: Section 3.2.1 briefly introduces EAs and their basic terminology and gives also a short state of the art in EAs while Section 3.2.2 presents more precisely the EAs that we have implemented for the orientation problem.

3.2.1. Introduction to Evolutionary Algorithms

This section briefly explains the basic steps of an EA. The problem is to optimize a given *objective function* f over a given search space. A population of individuals (i.e., a P-uple of points in the search space) undergoes some artificial Darwinian evolution based on the *fitness* F of each individual. The fitness of an individual is directly related to the value of the objective function of this individual (a typical example of a fitness function is the objective function itself, denoted by J in our work). The evolution operators applied to the individuals are defined upon the so-called *genotype space* noted E . It may be different from the definition space of the fitness called the *phenotype space*. The choice of this genotype space is the *representation*.

Figure 8 illustrates the framework of an EA: after an initialization of the population (generally a uniform random initialization) the fitness of each individual is computed. This is the *evaluation step*. Then, the loop of the algorithm called a *generation* is made up of the following steps:

- *Stopping criterion*: a basic stopping criterion is when the maximum number of generations fixed by the user is reached.
- *Selection*: the selection operator selects among the parents those who will generate offsprings, the genitors. There exists several selection operators, either of deterministic or of stochastic type. All of them are based on the fitness of the individuals and implement the first phase of Artificial Darwinism: the fittest allowed to reproduce.
- *Creation of new individuals*: there are basically two ways to create new individuals in the population from the genitors, namely the *crossover* and the *mutation*. These variation operators are stochastic operators: the crossover is a stochastic operator from E^k into E (typically $k = 2$), it is a recombination

of k parents, and the mutation operator is a stochastic operator from E into E .

- *Evaluation*: for each offspring the fitness is computed.
- *Replacement*: this operator discriminates among the individuals of the current population those who will be the parents for the next generation. This operator, like the selection operator, is based on the fitness of the individuals and implements the second step of Darwin's theory: survival of the fittest.

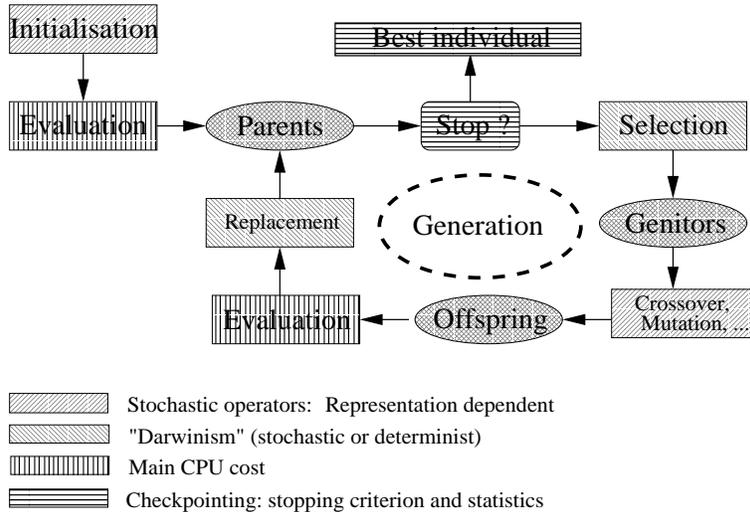


Fig. 8. General EA scheme.

Despite the common features of all EAs, several trends can be discriminated, mainly due to historical differences. We will only detail here the instances of EAs we have been using, referring to [3] and references therein for a complete description. The four main branches are (in alphabetical order):

- The evolutionary programming (EP), originally developed in California to evolve finite state machines.
- The evolution strategies (ESs) developed in Germany to solve numerical optimization problems for real search spaces. The genotype space is the phenotype space, namely a subset of \mathcal{R}^N . A precise description is given below.
- The genetic algorithms (GAs) developed in Michigan to study some adaptation mechanisms of populations for biology. These algorithms have later been used for optimization problems. More precision are given below.
- The genetic programming (GP), which has appeared more recently, consists in evolving tree structures.

The Canonical GA:

The genotype space is $\{0, 1\}^n$, the selection operator is the so-called roulette wheel, where the probability P_{X_p} to select the individual X_p is proportional to the fitness $F(X_p)$:

$$P_{X_p} = \frac{F(X_p)}{\sum_{i \in Population} F(X_i)}.$$

The crossover operator replaces some bits in the first parent string by the corresponding bits from the second parent, and the mutation operator randomly flips a bit of the parent. The replacement is *generational*: the offsprings at the generation n become the parents of the generation $n + 1$. Modern GAs are commonly used with any kind of representation as long as crossover and mutation can be defined.

The ES:

The ES [30] have been designed to optimize real functions, thus the natural search space is \mathcal{R}^N . The individuals undergo Gaussian mutations, namely addition of zero-mean Gaussian variables of standard deviation σ . The particularity of ES is that the parameter σ is a part of the genetic information. For a so-called isotropic ES, an individual is of the form $I = (x_1, \dots, x_N, \sigma)$ and, for a non isotropic ES, $I = (x_1, \dots, x_N, \sigma_1, \dots, \sigma_N)$ (there also exists a third type of ES not discussed here, the correlated ES). Consequently the mutation parameters are subjected to recombination and mutation as well. More precisely, the *adaptive mutation* takes place in two steps, first a mutation of the mutation parameters, second a mutation of object variables x_i . For an isotropic ES the two steps are

$$\begin{aligned} \sigma^{(t+1)} &= \sigma^{(t)} \exp(\tau_0 N(0, 1)), \\ x_i^{(t+1)} &= x_i^{(t)} + N_i(0, \sigma^{(t+1)}), \end{aligned}$$

and, for a non isotropic ES,

$$\begin{aligned} \sigma_i^{(t+1)} &= \sigma_i^{(t)} \exp(\tau_0 N(0, 1) + \tau N_i(0, 1)), \\ x_i^{(t+1)} &= x_i^{(t)} + N_i(0, \sigma_i^{(t+1)}), \end{aligned}$$

where $N(0, 1)$ stands for a Gaussian random variable. The crossover operator selects randomly two parents, $(x_1^1, \dots, x_N^1, \sigma_1^1, \dots, \sigma_N^1)$ and $(x_1^2, \dots, x_N^2, \sigma_1^2, \dots, \sigma_N^2)$, to produce an offspring $(x_1^{q_1}, \dots, x_N^{q_N}, \sigma_1^{q_1}, \dots, \sigma_N^{q_N})$ where $q_i = 1$ or $q_i = 2$ with equal probability. This crossover operator can also involve all individuals in the population, this is a *global crossover*. The replacement operator is strictly deterministic, based on the rank. For example, if μ (respectively λ) is the number of parents (respectively offsprings), $(\mu, \lambda) - ES$ selects the parents for next generation by taking the μ best offsprings and $(\mu + \lambda) - ES$ selects the parents for next generation by taking the μ best among the λ offsprings and μ parents.

It is now commonly accepted that the incorporation of specific knowledge, of the problem to optimize, by means of representation and specific operators, is the best way and the only way to enhance the performances of an EA. But, when using

an EA without introducing some specificities of the problem, ES is generally the most efficient EA for parametric optimization. ES, like GA, are implemented on the EOLib class library available from [18].

3.2.2. *The algorithms used*

The orientation problem is a minimization problem on a real space of size $7N$, where N is the number of laser fields to superpose. We use two kinds of EAs: the first one is based on a classical GA with a real representation (roulette wheel selection and *barycentric* or *multi-point* crossover [27]) and the second one is the ES described above and taken from EOLib [18].

The first algorithm is an improved GA, adding some specific operators and some specific features, which are known to improve the performances of GAs. We will name this algorithm in the sequel EGA (for Enhanced GA). *Niching* and *Rescaling* are two specific features of this algorithm. Rescaling is a way to avoid some bias in the roulette wheel selection; niching is to avoid that all the population concentrates on a region of the search space (see [27] for more precisions). Then, the mutation strength on EGA decreases with the number of generations. A specific gradient mutation operator is also used (EGA-CG), replacing the parent by the result of a few iterations of a conjugated gradient algorithm using the parent as initial value. The purpose of such an operator is to accelerate the convergence by taking advantages of a gradient algorithm.

We have tested EGA, EGA-CG and ES on several test functions taken from the literature (Sphere, Rosenbrock and Shekel functions). We refer to [35] for the details. We present here shortly some conclusions of these tests. First, for all the functions tested, a comparison with a classical GA has shown that EGA, EGA-CG, and ES converge more often, and faster than GA. Moreover, they are able to improve continuously their precision whereas GA stops at some non-zero distance of the solution. Second, the tests have confirmed that the gradient mutation operator accelerates the convergence, except for too “chaotic” functions. Third, the test cases have helped us for a crucial point of EAs, namely the setting of the parameters, which is specific to each function. Several trends can be discriminated for the setting, mainly taken from the literature and confirmed with test cases. As we have built our own EGA, it is difficult to give succinctly the parameters to set. With respect to ES, three important steps are given: First, the probability to mutate an individual is greater than the probability to cross two individuals (typically $p_{mut} = 0.8$ and $p_{cross} = 0.2$). Second, the size of the population is typically (7, 49) – ES and the number of parents should be increased if the number of local minima increases. Third, the initial mutation strength σ should also be increased when the number of local minima increases.

4. Results for the orientation problem

A preliminary study of the orientation problem with the purely deterministic PRLS

and BFGS algorithms (see Section 3.1), for the differentiable criterion (2.10), showed the need to use stochastic methods. Indeed, these algorithms converge after a few iterations towards a local minimum close to the initial guess: the cost function presents numerous local minima. We know from the literature and from test cases that for such functions, ES, EGA, and EGA-CG perform better. As far as EGA-CG is concerned, using it to minimize the criterion (2.10) does not improve the results in a significant way. More precisely, our best results have been obtained without the gradient mutation operator. However, using the gradient algorithm after EGA can improve the result of the optimization, as we will see in Section 4.1.

We present in this section the main results obtained with our algorithms on the orientation problem. The sequel of this section is organized as follows: in Section 4.1, we give the fields we have obtained by minimizing criteria (2.9) and (2.10). For both criteria we give the best results. Next, we explain how the addition of the CG at the end of the EGA improves the optimization of the criterion (2.10). In Section 4.2, we introduce a new hybrid criterion in order to approach both goals of Section 2.3 : obtaining at some given time a good orientation and keeping it as long as possible. Then, in Section 4.3, we present results obtained by a different form of laser fields. This form of laser field, named a *train of kicks* is a succession of fields of *kick* form presented in Section 4.1. As we will see, these fields really improve the results of Section 4.1 on both criteria (2.9) and (2.10).

4.1. Optimized fields for (2.9) and (2.10)

All the results presented in the sequel have been obtained by optimizing upon a superposition of two or three lasers in order to better understand the physical meaning of the results. Indeed, our trials for optimizing (2.10) on a superposition of ten laser fields have given results shown on Figure 10, which are not sufficiently easy to understand and interpret. We have therefore left this strategy aside.

Figures 11 and 12 show the optimized fields and their instantaneous criterion $j(t)$ obtained respectively with criteria (2.9) and (2.10). They have been obtained with a non-isotropic ES and with EGA, respectively. However, let us emphasize that the two algorithms give similar results. Indeed, EGA has given fields and instantaneous criterion of the same form as the ones shown on Figure 11 and ES has also given results of the form shown on Figure 12.

As it may be noticed on Figure 11, the minimum value of $j(t)$, namely -0.46 , is less than that on Figure 12 but the orientation does not last as long, which is expected in view of the criterion chosen. The first instantaneous criterion (Figure 11) is what we call a *narrow* $j(t)$ (see Section 2.3) and the second one (Figure 12) is what we call a *wide* $j(t)$. As for the fields, the first field is what we call in Section 2.4 a $(\omega, 2\omega)$ field and the second one is what we call a *kick* field. Table 5 shows the parameters of this latter field. The fact that a field of the form of a *kick* is a very efficient field for optimizing the criterion 2.10 is one of our most striking result from a physical viewpoint. It is reported and commented on in [17]. In the

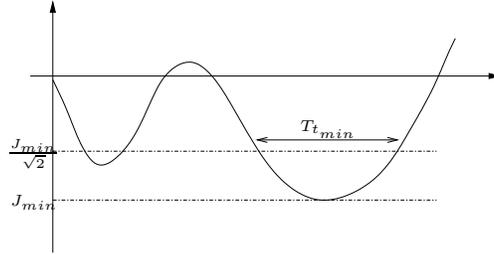


Figure 9: Construction of the hybrid criterion.

latter reference, the $(\omega, 2\omega)$ field is also analyzed. As explained above, using the

Table 5. Parameters of the optimized pulse with 3 laser fields.

n	\mathcal{E} (W/cm ²)	ω (cm ⁻¹)	ϕ (π rad)	t_0 (ps)	t_1 (ps)	t_2 (ps)	t_3 (ps)
1	1.01364×10^{08}	1389.541	1.98066	0.	0.312024	0.613023	1.193727
2	2.99976×10^{12}	500.051	1.82249	0.075077	0.270294	0.838110	1.562814
3	2.99989×10^{12}	500.000	0.82337	0.109518	0.235767	0.808280	1.080066

EGA-CG does not improve the results. However, CG is useful for a local search and we have tested how it could improve the result when used only at the end of a stochastic search. For this purpose we have first made an optimization on criterion (2.10) using EGA (the result is presented on Figure 13 with dotted lines) and then, we have applied the BFGS algorithm (the gradient has been computed with *Odyssée*) using the laser field so obtained as an initial guess for the conjugate gradient algorithm. After 100 CG iterations, the criterion is improved as may be seen on Figure 13 with solid lines. Such a result reconfirms that CG is useful for the local improvement search.

4.2. Results for the hybrid criterion

In view of the results of the previous section, it is a natural idea to introduce a new criterion aimed at approaching two goals together: obtaining at some given time a good orientation and keeping it as long as possible. Thus, we basically define a new criterion

$$J = J_{min} - J_{kept} + |J_{min} + J_{kept}|, \quad (4.38)$$

where $J_{min} = \min_{t \in [0, T]} j(t)$ and $J_{kept} = \frac{T_{t_{min}}}{T}$ where $T_{t_{min}}$ is the length the connex component of $\{t \in [0, T] \mid J_{min} \leq j(t) \leq \frac{J_{min}}{\sqrt{2}}\}$ including $t_{min} = \sup\{t \mid J(t) = J_{min}\}$ (see Figure 9). This criterion is a sum of three terms. The first one, J_{min} , measures the way the molecule is oriented. The second one, J_{kept} , measures how long the orientation is kept. The third part, $|J_{min} + J_{kept}|$, is a penalty term to ensure that J_{min} and $-J_{kept}$ are simultaneously minimized.

On Figure 14, we show a field obtained with this criterion and which is a succession in time of two fields with a short overlay time (see Section 2.4). For the physical

meaning of such a result, we refer to [2].

4.3. Results for the train of kicks

An other idea consists of starting with a field previously classified as a kick shape and using a succession of such fields in order to orient the molecule. The purpose of the optimization is thus to find the good delay between two successive *kicks*. Indeed we hope that by kicking several times the molecule we can lower the instantaneous criterion. The results are quite interesting: Figure 15 (a), is the result of an optimization of the criterion (2.10) with ES and it clearly illustrates the idea of kicking several times the molecule. This result is also interesting because the instantaneous criterion remains for a long time under the value -0.2 . Figure 15 (b) is the result of the optimization with the criterion (2.9). The criterion value (-0.82) is the best value we have ever had. However, the production of such fields remains an experimental challenge.

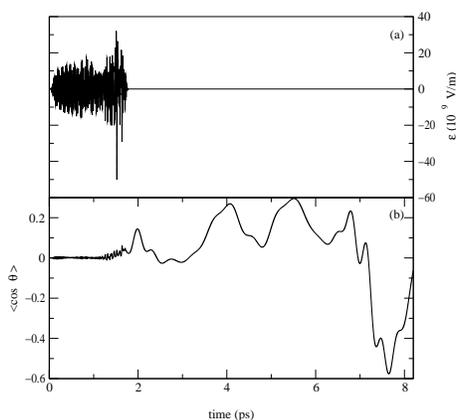


Fig. 10. Results obtained by optimizing upon 10 laser fields. In this figure and the following ones (11 to 14), the electric field is shown on top while $\langle \cos \theta \rangle$ which measures the instantaneous orientation of the molecule is shown on bottom. Time evolves from left to right at the same scale.

5. Conclusion and future directions

We have implemented and tested various strategies for the optimization of the laser field to be used for the orientation of the HCN molecule.

The best results have been obtained using evolutionary algorithms rather than purely deterministic algorithms such as gradient-like algorithms. However, in the case where the criterion is differentiable, we have shown that gradient like algorithms can efficiently complement the EA, not necessary when being used throughout the generations as mutations operators (the genetic algorithms with mutation by gradient have not yielded a real benefit in our specific case), but when being used as a final step in the optimization, once the population has been optimized by EA.

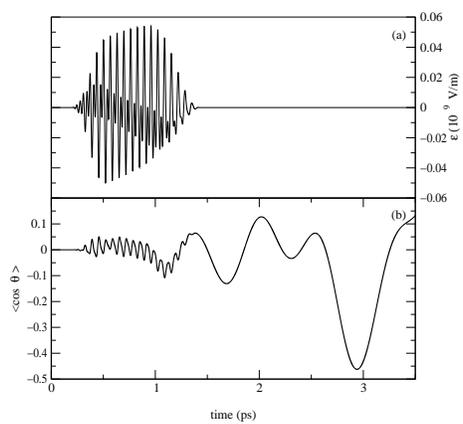


Fig. 11. Best result for $J = \min_{t \in [0, T]} j(t)$.

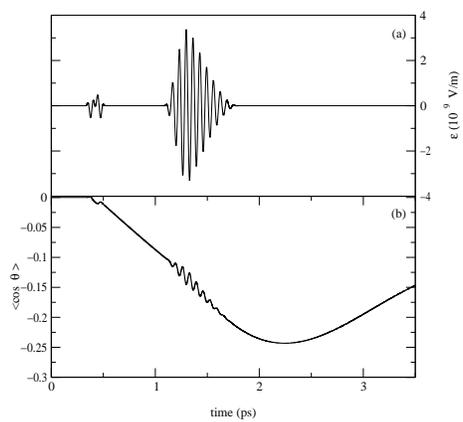


Fig. 12. Best result for $J = \frac{1}{T} \int j(t) dt$.

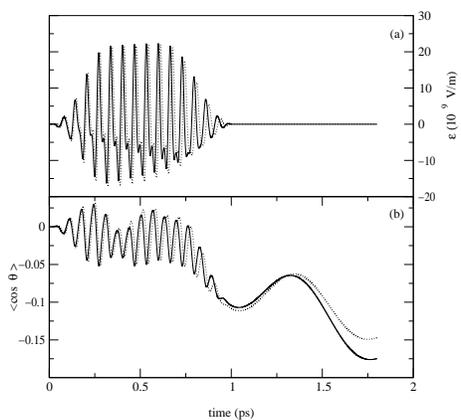


Fig. 13. Optimization by CG after optimization by GA.

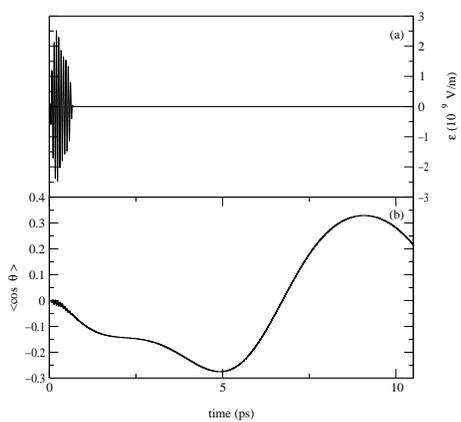


Fig. 14. Best result for the hybrid criterion given by equation 4.38.

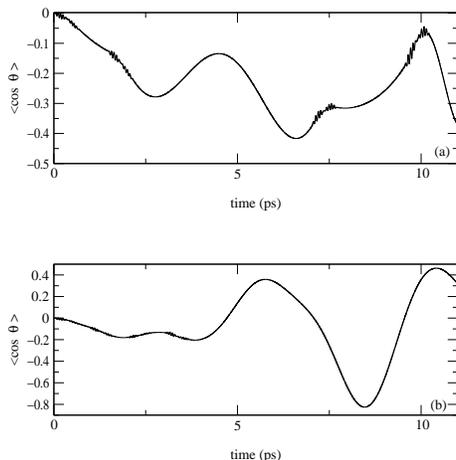


Fig. 15. (a) : Best result for $J = \frac{1}{T} \int j(t) dt$ with the train of kicks. (b) : Best result for $J = \frac{1}{T} \int j(t) dt$ with the train of kicks.

In order to understand how to compute the gradient of the criterion when needed, we have performed many tests, together with a numerical analysis on a toy equation related to our case of interest. They both show that the most efficient strategy (amenable in any case) is to compute the gradient by adjoint calculus on the discretized form of the equation or, if one does not fear a tedious post-processing work, to compute the gradient with an automatic differentiation tool.

As far as the choice of the criterion is concerned, we have tested many criteria, depending upon our physical aims. A multicriteria approach has also been implemented.

From the physical standpoint, our results have allowed us to identify two specific forms of laser fields that are most promising for the future: the $(\omega, 2\omega)$ field [14, 25] and the *kick* field [15, 22]. Definite conclusions about the efficiency of these fields are yet to be obtained and will be the purpose of some of our work in the future. It is anyway to be emphasized that such physically relevant fields have been obtained through our optimization methodology used as a blind tool, *i.e.*, without any specification of this form of fields. This is sufficient to give us some hope and confidence both in the physical and in the mathematical validity of our methodology.

Acknowledgements

The financial support of the *Action Concertée Incitative Jeunes Chercheurs* from the French Ministry of Research and Technology is gratefully acknowledged. CLB would like to thank Herschel Rabitz and André Bandrauk for enlightening discussions. The expertise of Marc Schöenauer on evolutionary strategies has been instrumental. Finally, Gabriel Turinici and Yvon Madaï are to be thanked for constant stimulating

interactions.

References

1. A. Assion, T. Baumer, M. Bergt, T. Brixner and B. Kiefer, V. Seyfried, M. Strehle, and G. Gerber. Control of chemical reactions by feedback-optimized phase-shaped femtosecond laser pulses. *Science*, 282:919–922, 1998.
2. A. Auger, C. M. Dion, A. Ben Haj Yedder, E. Cancès, A. Keller, C. Le Bris, and O. Atabek. Numerical optimization of laser fields to control molecular orientation. submitted.
3. Th. Bäck, D.B. Fogel, and Z. Michalewicz, editors. *Handbook of Evolutionary Computation*. Oxford University Press, Institute of Physics Publishing and Oxford University Press: Bristol and New York, 1997.
4. J.M. Ball, J.E. Marsden, and M. Slemrod. Controllability for distributed bilinear systems. *SIAM J. Control Optim.*, 20:575–597, 1982.
5. J. F. Bonnans, J. C. Gilbert, C. Lemaréchal, and C. Sagastizabal. *Optimisation Numérique : aspects théoriques et pratiques*. Springer, Berlin, 1997.
6. C. Le Bris. Control theory applied to quantum chemistry: Some tracks. In *International Conference on Control of Systems Governed by PDEs (Nancy, March 1999)*, volume 8, pages 77–94. ESAIM PROC, 2000.
7. C. Le Bris. Problématiques numériques pour la simulation moléculaire. In *Actes du 32ème Congrès national d'analyse numérique*, volume 9. ESAIM : Proceedings, September 2000.
8. P. Brumer and M. Shapiro. Control of unimolecular reactions using coherent light. *Chem. Phys. Lett.*, 126:541–546, 1986.
9. P. Brumer and M. Shapiro. Laser control of chemical reactions. *Scientific American*, pages 34–39, March 1995.
10. E. Cancès, C. Le Bris, and M. Pilot. Contrôle optimal bilinéaire sur une équation de Schrödinger. *Note aux Comptes Rendu de l'Académie des Sciences*, pages 567–571, 2000.
11. S. Chelkowski, M. Zamojski, and A. D. Bandrauk. Laser-phase directional control of photofragments in dissociative ionization of H_2^+ using two-color intense laser pulses. *Phys. Rev. A*, 63:023409, 2001.
12. C. E. Dateo and H. Metiu. Numerical solution of the time-dependent schrodinger equation in sperical coordinates by fourier-transformation methods. *J. Chem. Phys*, 95:7392–7400, 1991.
13. Académie des Sciences. Sciences aux temps ultracours ; de l'atoseconde aux petawatts. *Rapport sur la science et la technologie*, 2000, September.
14. C. M. Dion, A. D. Bandrauk, O. Atabek, A. Keller, H. Umeda, and Y. Fujimura. Two-frequency ir laser orientation of polar molecules. numerical simulations for hcn. *Chem. Phys. Lett.*, 302:215–223, 1999.
15. C. M. Dion, A. Keller, and O. Atabek. Orienting molecules using half-cycle pulses. *Eur. Phys. J. D*, 14:249–255, 2001.
16. C. M. Dion, A. Keller, O. Atabek, and A. D. Bandrauk. Laser-induced alignment dynamics of HCN : Roles of the permanent dipole moment and the polarizability. *Phys. Rev. A*, 59:1382–1391, 1999.
17. C. M. Dion, A. Ben Haj Yedder, E. Cancès, A. Keller, C. Le Bris, and O. Atabek. Optimal laser control of orientation: The kicked molecule. submitted.
18. EO. C++ class library, <http://eodev.sourceforge.net/>.
19. C. Faure and Y. Papegay. *Odyssée User's Guide Version 1.7. Rapport Technique*

- INRIA RT-0224*, 1998.
20. M.D. Feit, J.A. Fleck, and A. Steiger. Solution of a Schrodinger equation by a spectral method. *J. Comput. Phys.*, 47:412–433, 1982.
 21. B. Freidrich and D. R. Herschbach. On the possibility of orienting rotationally cooled polar molecules in an electric field. *Z. Phys. D*, 18:153–161, 1991.
 22. N. E. Henriksen. Molecular alignment and orientation in short pulse laser fields. *Chem. Phys. Lett.*, 312:196–202, 1999.
 23. K. Hoki and Y. Fujimura. Quantum control of alignment and orientation of molecules by optimized laser pulses. *Chem. Phys.*, 267:187–193, 2001.
 24. R. S. Judson and H. Rabitz. Teaching lasers to control molecules. *Phys. Rev. Lett.*, 68:1500–1503, 1992.
 25. T. Kanai and H. Sakai. Numerical simulation of molecular orientation using strong, nonresonant, two-color laser fields. *J. Chem. Phys.*, 115:5492–5497, 2001.
 26. J. J. Larsen, I. Wendt-Larsen, and H. Stapelfeldt. Controlling the branching ratio of photodissociation using aligned molecules. *Phys. Rev. Lett.*, 83:1123–1126, 1999.
 27. Z. Michalewicz. *Genetic algorithms + data structure = evolution programs*. Springer, 1999.
 28. R. Numico, A. Keller, and O. Atabek. Laser-induced molecular alignment in dissociation dynamics. *Phys. Rev. A*, 52:1298–1309, 1995.
 29. H. Sakai, C. P. Safvan, J. J. Larsen, K. M. Hilligsoe, K. Hald, and H. Stapelfeldt. Controlling the alignment of neutral molecules by a strong laser field. *J. Chem. Phys.*, 110:10235–10238, 1999.
 30. H. P. Schwefel. *Numerical Optimization of Computer Models*. John Wiley & Sons, New-York, 1981. 1995 – 2nd edition.
 31. P. Tournois, D. Hulin, F. Amiranoff, G. Mourou, and D. Kaplan. Les implusions lasers ultra-brèves. *Compte rendu des 3^{es} Entretiens de la physiques*, 1998.
 32. G. Turinici. Controlabilité exacte de la population des états propres dans les systèmes quantiques bilinéaires. *Note aux Compte Rendu de l'Académie des Sciences*, pages 327–332, 2000.
 33. G. Turinici and H. Rabitz. Quantum wave function controllability. *Chem. Phys.*, 267:1–9, 2001.
 34. G. Turinici and H. Rabitz. Wavefunction controllability in quantum systems. *Preprint*, 2001.
 35. A. Ben Haj Yedder. PhD thesis, Ecole Nationale des Ponts et Chaussées, in preparation.
 36. A. Ben Haj Yedder, E. Cancès, and C. Le Bris. Optimal laser control of chemical reactions using automatic differentiation. In George Corliss, Christèle Faure, Andreas Griewank, Laurent Hascoët, and Uwe Naumann (eds.), editors, *Proceedings of Automatic Differentiation 2000: From Simulation to Optimization*, pages 203–213, New York, 2001. Springer-Verlag.
 37. W. Zhu and H. Rabitz. A rapid monotonically convergent iteration algorithm for quantum optimal control over the expectation value of a positive definite operator. *J. Chem. Phys.*, 109:385–391, 1998.