A monotonic method for solving nonlinear optimal control problems with concave dependence on the state

Julien Salomon, Gabriel Turinici

A monotonic method for solving nonlinear optimal control problems with concave dependence on the state

Julien Salomon, Gabriel Turinici

CEREMADE, UMR CNRS 7534, Université Paris IX,

Place du Marechal de Lattre de Tassigny,

75775 Paris Cedex 16, France

November 9, 2010

Abstract

Initially introduced in the framework of quantum control, the so-called *monotonic algorithms* have demonstrated excellent numerical performance when dealing with bilinear optimal control problems. This paper presents a unified formulation that can be applied to more nonlinear settings compatible with the hypothesis detailed below. In this framework, we show that the well-posedness of the general algorithm is related to a nonlinear evolution equation. We prove the existence of the solution to this equation and give important properties of the optimal control functional. Finally we show how the algorithm works for selected models from the literature and compare it with the gradient algorithm.

1   Introduction

This paper presents a general unified formulation of several algorithms that were proposed in different areas of nonlinear control (see works in [34, 40, 10, 28]). Given a cost functional to minimize $J(v)$ depending on the control v, and a system described by a state function $X(t)$ solution of the evolution equation (4) below, these algorithms are iterative procedures that construct a sequence of solution candidates v^k with the important monotonic behavior $J(v^{k+1}) \leq J(v^k)$; the algorithms have been named after this property as "monotonic". A convenient advantage of these procedures is that the monotonicity does not requires any additional computational effort, but results from the definition of the procedure itself.

These algorithms have first been used in the field of quantum control where the dynamics is controlled by a laser field. In this framework the function that associates to a control v associates the final state $X(T)$ of the system is highly nonlinear. This induces poor performance of standard, gradient-based algorithms. The "monotonic schemes" introduced in [2, 34, 40] were found to perform excellently in this setting. These schemes were used in bi-linear situations i.e., when the operator $A(t, v(t))$ is linear in $v(t)$ and for a cost functional $J(v) = G(X(T)) + C(v)$ that is sum of a part $G(X(T))$ quadratic in the final state $X(T)$ and a part $C(v)$ quadratic in the control v. These were

soon followed by variants as those in [39, 41, 30, 32, 33] that included situations where $G(X)$ has negative semi-definite Hessian but $C(v)$ was still quadratic in the control and, most importantly, $A(t,v)$ was linear in v.

In works cited up to now, the function $G(X(T))$ depends only on the final state $X(T)$ but adaptations were proposed in [23, 22] to deal with the case where G depends on the whole dynamics of the control process $X(t)$ at intermediary times or when the dynamics involve bilinear integro-differential equations.

Similar procedures were also proposed in different control applications where the evolution equation is of parabolic type (see [28, 10]) or mixed hyperbolic-parabolic (see [19, 20]).

Up to this point all works presented above considered bilinear situations i.e., the evolution equation is linear in the state $X(t)$ and $A(t,v)$ is linear in the control v; only recently different cases were documented in the literature where $A(t,v)$ is polynomial in the control v up to power 3; in [29, 11] specific monotonic procedures were proposed that were showed to work in this setting.

A situation when $A(t,v)$ depends polynomially on a one-dimensional control $v(t) \in \mathbb{R}$ was proposed in [21]. A model where the system is a nonlinear Bose-Einstein condensate was given in [31].

This paper continues in this direction and treats situations with arbitrary nonlinear $A(t,v)$ but still keeps the requirement that the explicit dependence of J on X be concave (see the hypothesis detailed below). In all situations where monotonic algorithms were introduced the well-posedness of the algorithms (i.e., the existence of v^{k+1} given v^k) was proved by ad-hoc techniques although the algebraic computations share similar points. The purpose of this paper is to identify and exploit the similarities present in all these situations, and present a general setting that includes the "monotonic" algorithms. This allows to tackle a large class of non-linear situations that cannot be solved with techniques present today in the literature. We prove rigorously the existence and convergence of a procedure that from a control v^k constructs a control v^{k+1} such that the cost functional is monotonic. The question of whether such a procedure exists has never been asked before in the literature because up to now the authors considered only particular cost functionals J and particular evolution equations; in each case they proposed explicit analytic formulaes for v^{k+1} applicable to their situation. On the contrary, we show here that in all contexts covered by the theory (i.e., satisfy the hypothesis below) a control v^{k+1} can always be found to ensure $J(v^{k+1}) \leq J(v^k)$ and we also give a constructive procedure to compute it.

The paper is structured as follows: Section 2 defines the general framework in which where our procedure applies. The algorithm is presented in Section 3. At this point we show that the well-posedness of the algorithm is related to a nonlinear evolution equation and prove the existence of the solution to this equation. We also give important properties of the optimal control functional. Some examples of concrete realizations follow in Section 4 together with numerical results illustrating the application and the efficiency of the algorithm.

2 Problem formulation

Let E, $\mathcal{H}$ and $\mathbf{V}$ be Hilbert spaces with $\mathbf{V}$ densely included in $\mathcal{H}$. We denote by $\cdot_E$ and $\langle \cdot, \cdot \rangle_{\mathbf{V}}$ the scalar product associated with E and $\mathbf{V}$.

For any vector spaces $\mathcal{A}$ and $\mathcal{B}$, we denote by $\mathcal{L}(\mathcal{A}, \mathcal{B})$ the space of linear continuous operators between $\mathcal{A}$ and $\mathcal{B}$.

Given a real valued function φ, we denote by $\nabla_x \varphi$ its gradient with respect to the variable x.

We also denote by D_x and $D_{x,x}$ the first and the second derivative of vectorial functionals in the Fréchet sense.

Remark 1 *Recall that, given H_1 and H_2 two Banach spaces and $U \subset H_1$ an open subset of H_1, a function $f : U \to H_2$ is said to be Fréchet differentiable at $x \in U$ if there exists a continuous linear operator $A_x \in \mathcal{L}(H_1, H_2)$ such that*

$$\lim_{h \to 0} \frac{\|f(x+h) - f(x) - A_x(h)\|_{H_2}}{\|h\|_{H_1}} = 0.$$

The operator A_x is then called the Fréchet differential (or Fréchet derivative) of f at x and is denoted $D_x f \triangleq A_x$.

Let us also recall that given an open set $\Omega \subset \mathbb{R}^\gamma$ and a Hilbert space H_1, the set $L^\infty(\Omega; H_1)$ is the space of functions f from Ω with values in the Hilbert space H_1 such that for almost all $t \in \Omega$ the norm $\|f(t)\|_{H_1}$ is bounded by the same constant (the lowest of which is the $L^\infty(\Omega; H_1)$ norm of f). One can likewise define $L^2(\Omega; H_1)$:

$$L^2(\Omega; H_1) = \{f : \Omega \to H_1 \ such \ that \ \int_\Omega \|f(t)\|_{H_1}^2 dt < \infty\}. \tag{1}$$

When the derivatives of f are considered the Sobolev spaces $W^{1,\infty}$ have to be introduced; we refer to [1, 38] for further details.

Within an optimal control formulation, the control of a system described by a state function $X(t)$ is encoded in the following optimization problem:

$$\min_v J(v), \tag{2}$$

where

$$J(v) \triangleq \int_0^T F\big(t, v(t), X(t)\big) dt + G\big(X(T)\big). \tag{3}$$

The functions $F : \mathbb{R} \times E \times \mathbf{V} \to \mathbb{R}$ and $G : \mathbf{V} \to \mathbb{R}$ are assumed to be differentiable and integral assumed to exist. The state function $X(t) \in \mathbf{V}$ satisfies the following evolution equation

$$\partial_t X + A(t, v(t))X \quad = \quad B(t, v(t)) \tag{4}$$
$$X(0) = X_0. \tag{5}$$

where $v : [0, T] \to E$ is the control. The possibly unbounded operator $A(t, v) : \mathbb{R} \times E \times \mathcal{H} \to \mathcal{H}$ is such that for almost all $t \in [0, T]$ the domain of $A(t, v)^{1/2}$ includes $\mathbf{V}$; furthermore we take $B(t, v)$ such that for almost all $t \in [0, T]$ and all $v \in E$ we have $B(t, v) \in \mathcal{L}(\mathcal{H}, \mathcal{H}) \cap \mathcal{L}(\mathbf{V}, \mathbf{V}^*)$[1]. We postpone to Section 3 (cf. Lemma 3.4, Theorem 3.5) the precise formulation of additional regularity assumptions to be imposed on A, B, F, G.

Remark 2 *Note that E is not necessarily of dimension one, not even finite dimensional, cf. Section 4.2. This means that the control can be a set of **several** time-dependent functions but also a distributed control depending on a spacial variable.*

---

[1]For a space $\mathbf{V}$ we denote by $\mathbf{V}^*$ its dual space.

Let us stress that although the equation is linear in X (for v fixed) the mapping $v \mapsto X$ is **not** linear ; the term $A(t, v(t))$ **multiplies** the state X and as such the mapping is highly nonlinear (of non-commuting exponential type).

Remark 3 *Most of the previous works considered a bilinear operator $A(t, v)$ i.e., $A(t, v)X = vX$; the only exceptions (cf. discussion in the Introduction) were of the polynomial type (of order at most 3 in [29, 11] and polynomial with $E = \mathbb{R}^1$ in [21]). The techniques present in the above papers cannot be used for general operators $A(t, v)$ and control sets E. On the contrary the results in this work include all the situations considered in the bibliography but also apply to nonlinearities in v compatible with the hypothesis of Lemma 3.4 and Thm. 3.5 below.*

The following concavity with respect to X will be assumed throughout the paper:

$$\forall X, X' \in \mathbf{V}, \ G(X') - G(X) \leq \langle \nabla_X G(X), X' - X \rangle_{\mathbf{V}}, \tag{6}$$

$$\forall t \in \mathbb{R}, \forall v \in E, \forall X, X' \in \mathbf{V},$$

$$F(t, v, X') - F(t, v, X) \leq \langle \nabla_X F(t, v, X), X' - X \rangle_{\mathbf{V}}. \tag{7}$$

Remark 4 *Unlike the more technical hypothesis that will be assumed latter, the properties* (6), (7) *and the linearity of* (4) *are crucial to the existence of the monotonic algorithms.*

3 Monotonic algorithms

We now present the structure of our optimization procedure together with the general algorithm.

3.1 Tools for monotonic algorithms

The monotonic algorithms exploit a specific factorization which is the consequence of the results described in this section. To ease the notations, we will make explicit the dependence of X on v, i.e. we will write X_v instead of X in Eqs. (4–5).

We define the adjoint state Y_v (see [8, 16]) which is the solution of the following evolution equation:

$$\partial_t Y_v - A^*\big(t, v(t)\big)Y_v + \nabla_X F\big(t, v(t), X_v(t)\big) = 0 \tag{8}$$

$$Y_v(T) = \nabla_X G\big(X_v(T)\big). \tag{9}$$

A first estimate about the variations in J can be obtained:

Lemma 3.1 *For any $v', v : [0, T] \to E$ denote*

$$\Upsilon\Big(t, X_v(t), v(t), v'(t), Y_v(t), X_{v'}(t)\Big) \triangleq$$

$$-\langle Y_v(t), \Big(A\big(t, v'(t)\big) - A\big(t, v(t)\big)\Big)X_{v'}(t)\rangle_{\mathbf{V}}$$

$$\langle Y_v(t), B\big(t, v'(t)\big) - B\big(t, v(t)\big)\rangle_{\mathbf{V}} \tag{10}$$

$$+F\big(t, v'(t), X_{v'}(t)\big) - F\big(t, v(t), X_{v'}(t)\big). \tag{11}$$

4

Then

$$J(v') - J(v) \le \int_0^T \Upsilon\Big(t, X_v(t), v(t), v'(t), Y_v(t), X_{v'}(t)\Big) dt. \tag{12}$$

Proof Using successively (6),(7), (4) and finally (9), we find that:

$$
\begin{aligned}
J(v') - J(v) &= \int_0^T F\big(t, v(t), X_{v'}(t)\big) - F\big(t, v(t), X_v(t)\big) \\
&\qquad + F\big(t, v'(t), X_{v'}(t)\big) - F\big(t, v(t), X_{v'}(t)\big) dt \\
&\quad + G\big(X_{v'}(T)\big) - G\big(X_v(T)\big) \\
&\le \int_0^T \langle \nabla_X F\big(t, v(t), X_v(t)\big), X_{v'}(t) - X_v(t)\rangle_{\mathbf{V}} \\
&\qquad + F\big(t, v'(t), X_{v'}(t)\big) - F\big(t, v(t), X_{v'}(t)\big) dt \\
&\quad + \langle Y_v(T), X_{v'}(T) - X_v(T)\rangle_{\mathbf{V}} \\
&\le \int_0^T \langle \frac{\partial}{\partial t} Y_v(t) - A\big(t, v(t)\big)^* Y_v(t) \\
&\qquad + \nabla_X F\big(t, v(t), X_v(t)\big), X_{v'}(t) - X_v(t)\rangle_{\mathbf{V}} \\
&\qquad - \langle Y_v(t), \Big(A\big(t, v'(t)\big) - A\big(t, v(t)\big)\Big) X_{v'}(t)\rangle_{\mathbf{V}} \\
&\qquad + \langle Y_v(t), B(t, v'(t)) - B(t, v(t))\rangle_{\mathbf{V}} \\
&\qquad + F\big(t, v'(t), X_{v'}(t)\big) - F\big(t, v(t), X_{v'}(t)\big) dt.
\end{aligned}
$$

Due to (8), the first term of the right-hand side of this last inequality cancels and the result follows.
$\square$

Remark 5 *The purpose of this result is not to obtain an estimation of the increment $J(v') - J(v)$ via the adjoint (which is well documented in optimal control theory, cf. [8, 16]); we rather emphasis that the evaluation of the integrand Υ at time t requires information on the control $v(s)$ for all $s \in [0, T]$ (in order to compute $X_v(T)$ then $Y_v(t)$) but on the second control $v'(s)$ only for $s \in [0, t]$ (because this is enough to compute $X_{v'}(t)$). This estimate can be useful to decide* **at time** *t if the current value of the control $v'(t)$ will imply an increase or decrease of $J(v')$. This localization property is a consequence of the concavity of F and G (in X) and bi-linearity induced by A. The purpose of the paper is to construct and theoretically support a general numerical algorithm that exploits this remark.*

Remark 6 *We can intuitively note that Υ has the factorized form:*

$$\Upsilon\Big(t, X_v(t), v(t), v'(t), Y_v(t), X_{v'}(t)\Big) = \Delta(v, v')(t) \cdot_E \big(v'(t) - v(t)\big), \tag{13}$$

with $\cdot_E$ the E scalar product. Thus v' can always be chosen so as to make it negative (in the worse case set it null by the choice $v' = v$). We will come back with a formal definition of $\Delta(v, v')(t)$ and a proof of the previous relation in Section 3.3.

A more general formulation can be obtained if we suppose that the backward propagation of the adjoint state is performed with the intermediate field $\widetilde{v}$ (cf. also [18]), i.e. according to the equation :

$$\frac{\partial}{\partial t} Y_{\widetilde{v}} - A^*\big(t, \widetilde{v}(t)\big) Y_{\widetilde{v}} + \nabla_X F\big(t, v(t), X_v(t)\big) = 0$$

$$Y_{\widetilde{v}}(T) = \nabla_X G\big(X_v(T)\big).$$

Note that because of its final condition, $Y_{\widetilde{v}}$ actually also depends on v. Nevertheless, for the sake of simplicity, we keep the previous notation. We then obtain the following lemma whose proof we let as exercise for the reader:

Lemma 3.2 *For any $v', \widetilde{v}, v : [0, T] \to E$,*

$$\begin{aligned}
J(v') - J(v) &\leq \int_0^T -\langle Y_{\widetilde{v}}(t), \big(A\big(t, v'(t)\big) - A\big(t, \widetilde{v}(t)\big)\big) X_{v'}(t)\rangle_{\mathbf{V}} \\
&\quad + \langle Y_{\widetilde{v}}(t), B\big(t, v'(t)\big) - B\big(t, \widetilde{v}(t)\big)\rangle_{\mathbf{V}} \\
&\quad + F\big(t, v'(t), X_{v'}(t)\big) - F\big(t, \widetilde{v}(t), X_{v'}(t)\big) dt \\
&\quad + \int_0^T -\langle Y_{\widetilde{v}}(t), \big(A\big(t, \widetilde{v}(t)\big) - A\big(t, v(t)\big)\big) X_v(t)\rangle_{\mathbf{V}} \\
&\quad + \langle Y_{\widetilde{v}(t)}(t), B\big(t, \widetilde{v}(t)\big) - B\big(t, v(t)\big)\rangle_{\mathbf{V}} \\
&\quad + F\big(t, \widetilde{v}(t), X_{\widetilde{v}}(t)\big) - F\big(t, v(t), X_{\widetilde{v}}(t)\big) dt.
\end{aligned}$$

In this lemma, the variation in the cost functional J is expressed as the sum of two terms, and can be considered as factorized with respect to $v' - \widetilde{v}$ and $\widetilde{v} - v$.

Remark 7 *Lemmas 3.1 and 3.2 are generalizations of previous results that were proved in the bilinear case. To the best of our knowledge, only specific corollaries requiring additional assumptions have appeared in the literature up to now.*

3.2 The algorithms

The factorization (13) that will be proved in Lemma 3.3 enables to design various ways to ensure that $J(v') \leq J(v)$, i.e. that guaranty the monotonicity resulting from the update $v' \leftarrow v$. This allows to present a general structure for our class of optimization algorithms. We focus on the one that results from Lemma 3.1.

Algorithm 1 *(Monotonic algorithm)*
Given an initial control v^0, the sequence $(v^k)_{k\in\mathbb{N}}$ is computed iteratively by:

1. *Compute the solution X_{v^k} of (4–5) with $v = v^k$.*

2. *Compute the solution Y_{v^k} of (8–9) with $v = v^k$ backward in time from*

$$Y_{v^k}(T) \triangleq \nabla_X G\big(X_{v^k}(T)\big).$$

3. *Define (as explained latter) v^{k+1} together with $X_{v^{k+1}}$ such that for all $t \leq T$ the following monotonicity condition be satisfied :*

$$\Delta(v^{k+1}, v^k)(t) \cdot_E \Big(v^{k+1}(t) - v^k(t)\Big) \leq 0. \tag{14}$$

Lemma 3.1 then guarantees that $J(v^{k+1}) \leq J(v^k)$. Several strategies can be used to ensure (14); we will present one below. Its importance stems from the fact that no further optimization is necessary once this condition is fulfilled. In order to guarantee (14), many authors (see [18, 34, 40]) consider an update formula of the form:

$$v^{k+1}(t) - v^k(t) = -\frac{1}{\theta}\Delta(v^{k+1}, v^k)(t), \tag{15}$$

where θ is a positive number, that can also depend on k and t. In what follows, we focus on the existence of solution of (15), and on practical methods to compute it. If v^{k+1} satisfies (15), the variations in J satisfy:

$$J(v^{k+1}) - J(v^k) \quad \leq \quad -\theta \int_0^T (v^{k+1}(t) - v^k(t))^2 dt.$$

Note that (15) reads as an update formula combining on the one hand a gradient method:

$$v^{k+1}(t) - v^k(t) = -\frac{1}{\theta}\Delta(v^k, v^k)(t),$$

and on the other hand the so-called Proximal Algorithm (as described in [5]) which prescribes:

$$v^{k+1}(t) - v^k(t) = -\frac{1}{\theta}\Delta(v^{k+1}, v^{k+1})(t).$$

Remark 8 *When $F = 0$ and A is independent of v i.e., linear control with final objective, (15) coincides with a gradient method.*

3.3 Well-posedness of the algorithm

In this section, we focus on the procedure obtained when using Algorithm 1 with the update formula (15). To the best of our knowledge no theoretical result exists in the literature to prove the existence of a solution to the Eq. (15) for general choices of $A(t, v)$ and general space of controls E because previous authors only dealt with particular choices of functionals F, G, operators A, B and managed

to obtain in each case an analytic solution; we provide here such a proof together with a convergent procedure to compute it. Since this procedure involves the resolution of an implicit equation, the proof is non-trivial and has been split in three parts: two preparatory Lemmas (3.3 and 3.4) and the final result in Theorem 3.5. As a by-product, we obtain a proof of the monotonicity of the algorithm.

Lemma 3.3 *Suppose that for any $t \in [0, T]$:*

- $\mathcal{A} : \mathbb{R} \times \mathbf{V} \times \mathbf{V} \times E \to \mathbb{R}$ defined by $\mathcal{A}(t, X, Y, v) = \langle Y, A(t, v)X \rangle_{\mathbf{V}}$ is of C^1 class with respect to v for any X, Y, v;

- $\mathcal{B} : \mathbb{R} \times \mathbf{V} \times E \to \mathbb{R}$ with $\mathcal{B}(t, Y, v) = \langle Y, B(t, v) \rangle_{\mathbf{V}}$ is of C^1 class with respect to v for any Y, v;

- F is of C^1 class with respect to $v \in E$ for any X, Y, v.

Then there exists $\Delta(\cdot, \cdot; t, X, Y) \in C^0(E^2, E)$ such that, for all $v, v' \in E$

$$
\begin{aligned}
\Delta(v', v; t, X, Y) \cdot_E \left(v' - v \right) &= -\left\langle Y, \left(A(t, v') - A(t, v) \right) X + B(t, v') - B(t, v) \right\rangle_{\mathbf{V}} \\
&\quad + F(t, v', X) - F(t, v, X).
\end{aligned} \tag{16}
$$

Moreover, $\Delta(\cdot, \cdot; t, X, Y)$ can be defined through the explicit formula:

$$
\begin{aligned}
\Delta(v', v; t, X, Y) &= \int_0^1 -\nabla_w \left(\langle Y, A(t, w)X - B(t, w) \rangle_{\mathbf{V}} \right) \Big|_{w = v + \lambda(v' - v)} \\
&\quad + \nabla_v F(t, v + \lambda(v' - v), X) d\lambda.
\end{aligned} \tag{17}
$$

Proof We denote by $\| \cdot \|$ the norm associated with E. Since $\mathcal{A}, \mathcal{B}, F$ are Fréchet differentiable with respect to v the full expression in Eq. (16) is of the form $\Xi(v') - \Xi(v)$ with $\Xi(v) = -\mathcal{A}(t, X, Y, v) + \mathcal{B}(t, Y, v) - F(t, v, X)$ differentiable with respect to v; Eq. (17) is an application of the identity

$$
\Xi(v') - \Xi(v) = \int_0^1 \nabla_v \Xi(v + \lambda(v' - v)) d\lambda \cdot_E \left(v' - v \right).
$$

The continuity is obtained from that of $\nabla_v \Xi$. $\qquad \square$

Lemma 3.4 *Suppose that*

- $\mathcal{A}, \mathcal{B}, F$ are of (Fréchet) C^2 class with respect to v with $D_{vv}A$, $D_{vv}B$ uniformly bounded as soon as X, Y are in a bounded set;

- $\nabla_v F$ is of C^1 class with respect to X;

- $D_{vv}F(t, \cdot, X)$ is bounded by a positive, continuous, increasing function $X \mapsto k(\|X\|)$.

Then given $\varepsilon > 0$, $(t, v, X, Y) \in \mathbb{R} \times E \times \mathbf{V} \times \mathbf{V}$ and a bounded neighborhood W of (t, v, X, Y), there exists $\theta^\star > 0$ depending only on ε, W, $\|v\|$, $\|X\|$ and $\|Y\|$ such that, for any $\theta > \theta^\star$

1. *$\Delta(v', v; t, X, Y) = -\theta(v' - v)$ has a unique solution $v' = \mathcal{V}_\theta(t, v, X, Y) \in E$;*

2. *$\mathcal{V}_\theta(t, v, X, Y) = v$ implies*

$$
-\nabla_v \left(\langle Y, A(t, v)X \rangle_{\mathbf{V}} \right)(v) + \nabla_v \left(\langle Y, B(t, v) \rangle_{\mathbf{V}} \right)(v) + \nabla_v F(t, v, X) = 0, \tag{18}
$$

8

3. $\|\mathcal{V}_\theta(t,v,X,Y) - v\| \le \frac{\|X\|\|Y\|+\|Y\|+k(\|X\|)}{\theta}\{M_0(t)+M_1\|v\|\}$ *with $M_0(t)$ and M_1 independent of v, X, Y. If the dependence of A, B, F with respect to t is smooth then $M_0(t)$ is bounded on $[0, T]$;*

4. $\mathcal{V}_\theta(t, v, X, Y)$ *is continuous on W;*

5. *Let X belong to a bounded set; then $X \mapsto \mathcal{V}_\theta(t, v, X, Y)$ is Lipschitz with the Lipschitz constant smaller than ε.*

Proof

1. Denote $h = v' - v$ and define $\mathcal{G}_{t,v,X,Y}(h) = \frac{-\Delta(v+h,v;t,X,Y)}{\theta}$. When the dependence is clear we will write simply $\mathcal{G}(h)$ instead of $\mathcal{G}_{t,v,X,Y}(h)$. We look thus for a solution to the following fixed point problem: $\mathcal{G}(h) = h$. For θ large enough, the mapping $\mathcal{G}$ is a (strict) contraction and we obtain the conclusion by a Picard iteration. The uniqueness is a consequence of the contractivity of $\mathcal{G}$.

2. If $v' = v$ then $h = 0$ thus $\mathcal{G}(h) = 0$ which gives (18) after using (17).

3. For θ large enough, the mapping $\mathcal{G}$ is not only a contraction but has its Lipschitz constant less than, say, $1/2$. Because of the contractivity of $\mathcal{G}$, we have $\|h\| - \|\mathcal{G}(0)\| \le \|h - \mathcal{G}(0)\| = \|\mathcal{G}(h) - \mathcal{G}(0)\| \le \frac{1}{2}\|h\|$, which amounts to $\|h\| \le 2\|\mathcal{G}(0)\|$. Next, we note that

$$\|\mathcal{G}(0)\| \le \frac{\|\Delta(v,v,t,X,Y) - \Delta(0,0,t,X,Y)\| + \|\Delta(0,0,t,X,Y)\|}{\theta}$$
$$\le M_2\|v\| + M_3(t)$$

and the estimate follows.

4. Formula (17) shows that Δ depends continuously on t, v, v', X, Y. Consider converging sequences $t_n \to t$, $v_n \to v$, $X_n \to X$, $Y_n \to Y$ and define $h_n \triangleq \mathcal{V}_\theta(t_n, v_n, X_n, Y_n)$ and $h \triangleq \mathcal{V}_\theta(t, v, X, Y)$.
Given W and $\eta > 0$, consider a large enough value of θ such that:

- for any $(t', v', X', Y') \in W$, $\mathcal{G}_{t',v',X',Y'}$ is a contraction with Lipschitz constant less than $1/2$.

- for any $(t', v', X', Y'), (t'', v'', X'', Y'') \in W$,

$$\|\Delta(v''+h, v'', t'', X'', Y'') - \Delta(v'+h, v', t', X', Y')\| \le \eta.$$

This last property implies $\|\mathcal{G}_{t_n,v_n,X_n,Y_n}(h) - \mathcal{G}_{t,v,X,Y}(h)\| \le \frac{\eta}{\theta}$ for n large enough. On the other hand

$$
\begin{aligned}
\|h_n - h\| &= \|\mathcal{G}_{t_n,v_n,X_n,Y_n}(h_n) - \mathcal{G}_{t,v,X,Y}(h)\| \\
&\le \|\mathcal{G}_{t_n,v_n,X_n,Y_n}(h_n) - \mathcal{G}_{t_n,v_n,X_n,Y_n}(h)\| \\
&\quad + \|\mathcal{G}_{t_n,v_n,X_n,Y_n}(h) - \mathcal{G}_{t,v,X,Y}(h)\| \\
&\le \frac{1}{2}\|h_n - h\| + \frac{\eta}{\theta}.
\end{aligned}
$$

We have thus obtained that for n large enough : $\frac{1}{2}\|h_n - h\| \le \frac{\eta}{\theta}$ and the continuity follows.

5. Subtracting the two equalities

$$\Delta(V_1, v; t, X_1, Y) = -\theta(V_1 - v), \ \ \Delta(V_2, v; t, X_2, Y) = -\theta(V_2 - v)$$

and using that $\Delta(V, v; t, X, Y)$ is C^1 in X and v gives to first order

$$\Delta_V(...)(V_1 - V_2) + \Delta_X(...)(X_1 - X_2) = -\theta(V_1 - V_2).$$

For θ large enough the operator $\Delta_V(...) + \theta \cdot Id$ is invertible and the conclusion follows.

$\square$

Remark 9 *Note that $\theta^\star$ is proportional to $(\|X\|_\mathbf{V}\|Y\|_\mathbf{V} + \|Y\|_\mathbf{V} + k(\|X\|_\mathbf{V}))$.*

We are now able to construct a procedure such that the existence of $v^{k+1}(t)$ satisfying (14) is guaranteed.

Theorem 3.5 *Suppose that A, B, F satisfy the hypothesis of the Lemma 3.4. Also suppose that the operators A, B are such that Eqs. (4–5) and (8-9) have solutions for any $v \in L^\infty(0, T; E)$ with $v \mapsto X$, $v \mapsto Y$ locally Lipschitz. Then:*

1. *For any $v \in L^\infty(0, T; E)$, there exists $\theta^\star > 0$ such that for any $\theta > \theta^\star$, the (nonlinear) evolution system*

$$\partial_t X_{v'}(t) + A(t, v') X_{v'}(t) = B(t, v') \tag{19}$$
$$v'(t) = \mathcal{V}_\theta(t, v(t), X_{v'}(t), Y_v(t)) \tag{20}$$
$$X_{v'}(0) = X_0 \tag{21}$$

   has a solution. Here Y_v is the adjoint state defined by (8–9) and corresponding to control v.

2. *There exists a sequence $(\theta_k)_{k \in \mathbb{N}}$ such that the algorithm 1 (cf. Section 3.2)*
   a/ initialization $v^0 \in L^\infty(0, T; E)$,
   b/ $v^{k+1}(t) = \mathcal{V}_{\theta_k}(t, v^k(t), X_{v^{k+1}}(t), Y_{v^k}(t))$
   is monotonic and satisfies

$$J(v^{k+1}) - J(v^k) \leq -\theta_k \|v^{k+1} - v^k\|^2_{L^2([0,T])}.$$

3. *With the notations above, if for all $t \in [0, T]$ $v^{k+1}(t) = v^k(t)$ (i.e. algorithm stops) then v^k is a critical point of J: $\nabla_v J(v^k) = 0$.*

Proof Some of the proof is contained in the previous lemmas. The part that still has to be proved is the existence of a solution to (19)-(21).

Given $v \in L^\infty(0, T; E)$, consider the following iterative procedure :

$$v_0 = v, \ \ v_{l+1}(t) = \mathcal{V}_\theta(t, v(t), X_{v_l}(t), Y_v(t)).$$

We take a spherical neighborhood $B_v(R)$ of v of radius R and suppose that $\forall k \leq l$, $v_k \in B_v(R)$. Since the correspondence $v \mapsto X_v$ is continuous, it follows that the set of solutions $S_{v,R} \triangleq \{X_w; w \in$

$B_v(R)\}$ of (4) is bounded. In particular for $w = v_l$ by the item 3 of Lemma 3.4 the quantity $\|\mathcal{V}_\theta(t, v(t), X_{v_l}(t), Y_v(t)) - v\|$ will be bounded by R for θ large enough (depending on R, independent of l), i.e. $v_{l+1} \in B_v(R)$. Thus $v_l \in B_v(R)$ for all $l \geq 1$.

Since $S_{v,R}$ is bounded, recall that by item 5 of Lemma 3.4 the mapping $X \mapsto \mathcal{V}_\theta(t, v(t), X, Y_v(t))$ has on $S_{v,R}$ a Lipschitz constant as small as desired. Since the mapping $w \mapsto X_w$ is Lipschitz, for θ large enough, $w \in B_v(R) \mapsto \mathcal{V}_\theta(t, v(t), X_w, Y_v(t))$ is a contraction. By a Picard argument the sequence v_l is converging. The limit will be the solution of (19–20). $\square$

4 Examples

We now present three examples that fit into the setting of Theorem 3.5. The space does not allow to treat different variants (cf. references in Introduction) so we leave them as an exercise to the reader.

Within the framework of control theory, nonlinear formulations prove useful nowadays in domains as diverse as the laser control of quantum phenomena (see [15, 25, 26, 27, 36, 37]) or the modeling of a equilibrium (or again social beliefs, product prices, etc) of a game with an infinite numbers of agents (see [12, 13, 14]). Yet other applications arise from modern formulations of the Monge-Kantorovich mass transfer problem (see [4, 3, 28]).

In the following, we present some examples coming from these fields of application and present the corresponding monotonic algorithm resulting from Theorem 3.5.

4.1 (I): Quantum control

4.1.1 Setting

The evolution of a quantum system is described by the Schrödinger equation

$$\partial_t X + iH(t)X = 0$$
$$X(0, z) = X_0(z),$$

where $i = \sqrt{-1}$, $H(t)$ is the Hamiltonian of the system and $z \in \mathbb{R}^\gamma$ the set of internal degrees of freedom. We assume that the Hamiltonian is a self-adjoint operator over $L^2(\mathbb{R}^\gamma; \mathbb{C})$, i.e. $H(t)^* = H(t)$[2]. Note that this implies the following norm conservation property

$$\|X(t, \cdot)\|_{L^2(\mathbb{R}^\gamma; \mathbb{C})} = \|X_0\|_{L^2(\mathbb{R}^\gamma; \mathbb{C})}, \ \forall t > 0, \tag{22}$$

so that the state (also called wave-) function $X(t, z)$, evolves on the (complex) unit sphere $S \triangleq \left\{ X \in L^2(\mathbb{R}^\gamma; \mathbb{C}) : \|X\|_{L^2(\mathbb{R}^\gamma; \mathbb{C})} = 1 \right\}$.

The Hamiltonian is composed of two parts: a free evolution Hamiltonian H_0 and a part that describes the coupling of the system with an external laser source of intensity $v(t) \in \mathbb{R}$, $t \geq 0$; a first order approximation leads to adding a time-independent dipole moment operator $\mu(z)$ resulting in the formula $H(t) = H_0 - v(t)\mu$ and the dynamics:

$$\partial_t X + i \left(H_0 - v(t)\mu \right) X = 0$$
$$X(0) = X_0.$$

---

[2]For any operator M, we denote by M^* its adjoint.

The purpose of control may be formulated as to drive the system from its initial state X_0 to a final state X_{target} compatible with predefined requirements. Here, the control is the laser intensity $v(t)$. Because the control is multiplying the state, this formulation is called "bilinear" control. The dependence $v \mapsto X(T)$ is of course not linear.

The optimal control approach can be implemented by introducing a cost functional. The following functionals are often considered:

$$J(v) \triangleq \|X(T) - X_{target}\|^2_{L^2(\mathbb{R}^\gamma;\mathbb{C})} + \int_0^T \alpha(t)v^2(t)dt, \tag{23}$$

$$\widetilde{J}(v) \triangleq -\langle X(T), OX(T)\rangle_{L^2(\mathbb{R}^\gamma;\mathbb{C})} + \int_0^T \alpha(t)v^2(t)dt, \tag{24}$$

where O is a positive linear operator defined on $\mathcal{H}$, characterizing an observable quantity and $\alpha(t) > 0$ is a parameter that penalizes large (in the L^2 sense) controls. The goal is to minimize these functionals with respect to v. According to (22) the cost functional J is equal to

$$J(v) \triangleq 2 - 2Re\langle X(T), X_{target}\rangle_{L^2(\mathbb{R}^\gamma;\mathbb{C})} + \int_0^T \alpha(t)v^2(t)dt, \tag{25}$$

so that the functionals J and $\widetilde{J}$ satisfy assumptions (6) and (7).

4.1.2 Mathematical formulation

We have
- $A(t,v) = H_0 + v(t)\mu$ with (possibly) unbounded v-independent operator H_0 (but which generates a C^0 semi group) and bounded operator μ. The dependence of A on v is smooth (linear) and therefore all hypotheses on A are satisfied.
- $E = \mathbb{R}$, $\mathcal{H} = L^2(\mathbb{R}^d;\mathbb{C})$, $\mathbf{V} = dom(H_0^{1/2})$ (over $\mathbb{C}$), or their realifications $\mathcal{H} = L^2 \times L^2$, $\mathbf{V} = dom(H_0^{1/2}) \times dom(H_0^{1/2})$ (over $\mathbb{R}$) as explained in [9];
- $B(t,v) = 0$.
- $F(t,v,X) = \alpha(t)v(t)^2$ with $\alpha(t) \in L^\infty(\mathbb{R})$; here the second derivative $D_{vv}F$ is obviously bounded. Since it is independent of X it will be trivially concave.
- G is either (see e.g., [17, 18]) $2 - 2Re\langle X_{target}, X(T)\rangle_{\mathbf{V}}$ or $-\langle X(T), OX(T)\rangle_{\mathbf{V}}$ where O is a positive semi-definite operator; both are concave in X.
- Here

$$\Delta(v', v; t, X, Y) = -Re\langle Y, i\mu X\rangle_{\mathbf{V}} + \alpha(t)(v' + v) \tag{26}$$

and the equation in v' is: $\Delta(v', v; t, X, Y) = -\theta(v' - v)$ and has for θ large enough a unique solution $v' = \mathcal{V}_\theta(t, v, X, Y) \triangleq \frac{(\theta - \alpha(t))v + Re\langle Y, i\mu X\rangle_{\mathbf{V}}}{\theta + \alpha(t)}$.
- at the $k+1$-th iteration, Theorem 3.5 guarantees the existence of the solution X^{k+1} of the following nonlinear evolution equation:

$$i\partial_t X^{k+1}(t) = \left(H_0 + \frac{(\theta - \alpha(t))v^k + Re\langle Y_{v^k}, i\mu X^{k+1}\rangle_{\mathbf{V}}}{\theta + \alpha(t)}\mu\right) X^{k+1}(t) \tag{27}$$

Then

$$v^{k+1} = \frac{(\theta - \alpha(t))v^k + Re\langle Y_{v^k}, i\mu X^{k+1}\rangle_{\mathbf{V}}}{\theta + \alpha(t)}, \ X_{v^{k+1}} = X^{k+1}. \tag{28}$$

12

4.1.3 Numerical test

In order to test the performance of the algorithm we have chosen a case already treated in the literature (see [40]). The system under consideration is the $O - H$ bond that vibrates in a Morse type potential $V(z) = D_0(\exp(-\beta(z - z')) - 1)^2 - D_0$ and $H_0 = -m\frac{\partial^2}{\partial z^2} + V(z)$. The dipole moment operator of this system is modeled by $\mu(z) = \mu_0.ze^{-\frac{z}{z^\star}}$. The objective is to localize the wavefunction at time $T = 131000$ at a given location z_0 ; this is expressed through the requirement that the functional $\widetilde{J}$ is maximized, with the observable O defined by $O(z) = \frac{\gamma_0}{\sqrt{\pi}}e^{-\gamma_0^2(z-z_0)^2}$. The numerical values we use are given below:

D_0	β	z'	$z^\star$	z_0	γ_0	μ_0	m
0.1994	1.189	1.821	0.6	2.5	25	3.088	$2.8694.10^{-4}$

We consider a constant penalization parameter $\alpha = 1$ and optimization parameter $\theta = 10^{-2}$. To compare this procedure with a standard algorithm, we have also minimized $J(v)$ with an optimal step gradient method. The line search is achieved through a golden section search cf. [24]. Results are presented in Fig. 1. This test shows that the gradient method fails in efficiently solving the
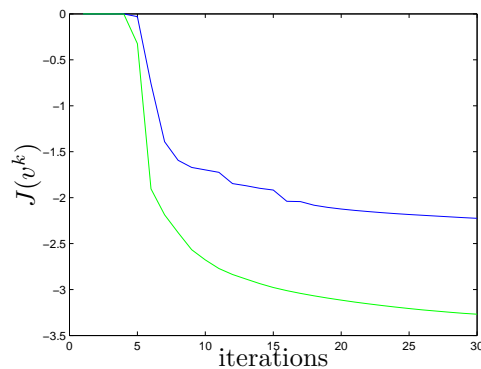

Figure 1: Numerical resolution of the example of Section 4.1. The cost functional $J(v)$ is optimized using the monotonic algorithm (28) (green line) and the optimal step gradient algorithm (blue line).

problem, whereas the monotonic procedure ensures that the cost functional values rapidly decrease. Note that the non-convexity of the problem renders difficult the convergence of the gradient method. On the other hand, the monotonic scheme fully exploits the concavity of the cost functional with respect to the state variable. In our implementation the time of computation is about two times larger for the gradient method as the line search requires about 3 evaluations of the cost functional per iteration.

4.2 (II) : Mean field games

4.2.1 Setting

Although the Nash equilibrium in game theory has been initially formulated for a finite number of players, modern results (see [12, 13, 14]) indicate that it is possible to extend it to an infinite number of players and obtain the equations that describe this equilibrium; applications have already been proposed in economic theory and other are expected in the behavior of multi-agents ensembles and decision theory.

The equations describe evolution of the density $X(t,z)$ of players at time t and position $z \in Q = [0,1]$ in terms of a control $v(t,z)$ and a fixed parameter $\nu > 0$:

$$\partial_t X - \nu \Delta X + div(v(t,z)X) = 0,$$
$$X(0) = X_0.$$

The control v is chosen to minimize the cost criterion (3). For reasons related to economic modeling interesting examples include situations where F, G are concave in X, e.g., as in [10]

$$G = 0, \ F(t,z,X) = \int_Q p(t)(1-\beta z)X(t,z) + \frac{c_0 \cdot z \cdot X(t,z)}{c_1 + c_2 X(t,z)} + \frac{v^2(t)}{2}X(t,z)dz, \qquad (29)$$

with positive constants β, c_0, c_1, c_2 and $p(t)$ a positive function. Another example is given in [28]:

$$G(X(T)) = \int_Q V(z)X(T,z)dz, \ F(t,z,X) = \int_Q X(t,z)v^2(t,z)dz, \qquad (30)$$

where V encodes a potential. The interpretation of this terminal cost is that the crowd aims at reaching zones of low potential V at the terminal time T while minimizing the cost of changing state.

The numerical relevance of the monotonic algorithms to this setting has been established in several works, see [28, 10].

4.2.2 Mathematical formulation

We have
- $E = W^{1,\infty}(0,1)$, $\mathcal{H} = L^2(0,1)$, $\mathbf{V} = H^1(0,1)$ see [10] and [6] (Chap XVIII §4.4)
- $A(t,v) = -\nu\Delta \cdot + div(v\cdot)$. The dependence of A on v is smooth (linear) and therefore all hypotheses on A are satisfied ($D_{vv}A = 0, ...$).
- $B(t,v) = 0$.
- with definitions in (29) $F(t,v,X) = \int_Q p(t)(1-\beta z)X(t,z) + \frac{c_0 \cdot z \cdot X(t,z)}{c_1 + c_2 X(t,z)} + \frac{v(t,z)^2}{2}X(t,z)dz$; F is concave in X; the second differential $D_{vv}F$ has all required properties.
- $G = 0$ (algorithm will apply in general when G is concave with respect to X).
- Here

$$\Delta(v',v;t,X,Y) = \nabla Y + \frac{v'+v}{2} \qquad (31)$$

and the equation in v' is: $\Delta(v',v;t,X,Y) = -\theta(v'-v)$ and has for all $\theta > 0$ a unique solution $v' = \mathcal{V}_\theta(t,v,X,Y) \triangleq \frac{(\theta-1/2)v-\nabla Y}{\theta+1/2}$.

14

• at the $k + 1$-th iteration, Theorem 3.5 guarantees the existence of the solution X^{k+1} of the following nonlinear evolution equation:

$$\partial_t X^{k+1}(t) - \nu \Delta X^{k+1} + div(\frac{(\theta - 1/2)v^k - \nabla Y_{v^k}}{\theta + 1/2} X^{k+1}) = 0. \tag{32}$$

Then

$$v^{k+1} = \frac{(\theta - 1/2)v^k - \nabla Y_{v^k}}{\theta + 1/2}, \ X_{v^{k+1}} = X^{k+1}. \tag{33}$$

4.2.3   Numerical test

The algorithm is tested on the time interval $[0, 1]$ with $p(t) = 1$ and the numerical values $\beta = 0.8$, $c_0 = c_2 = 1$, $c_1 = 0.1$. The same gradient method as in Section 4.1.3 is also tested. Results are presented in Fig. 2. In this example, the gradient method gives better results in the first iterations.
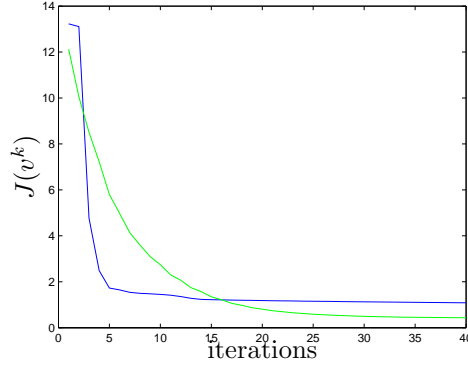

Figure 2: Numerical resolution of the example of Section 4.2. The cost functional $J(v)$ is optimized using the monotonic algorithm (33) (green line) and the optimal step gradient algorithm (blue line).

However, the monotonic algorithm converges asymptotically faster.

4.3   Additional application

As a third example we consider a nonlinear vectorial case from [7, 35] which differs from that of Section 4.1 in that $v(t) = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in E = \mathbb{R}^2$ and $A(t, v) = i[H_0 + (v_1(t)^2 + v_2(t)^2)\mu_1 + v_1(t)^2 v_2(t)\mu_2]$. Here, denoting $\xi_1 = -Re\langle Y, i\mu_1 X\rangle_{\mathbf{V}} + \alpha(t)$, $\xi_2 = -Re\langle Y, i\mu_2 X\rangle_{\mathbf{V}}$ we obtain

$$\Delta(v', v; t, X, Y) = \xi_1 \begin{pmatrix} v_1 + v_1' \\ v_2 + v_2' \end{pmatrix} + \xi_2 \begin{pmatrix} (v_1 + v_1')v_2' \\ (v_1)^2 \end{pmatrix} \tag{34}$$

15

and the equation in v' is: $\Delta(v', v; t, X, Y) = -\theta(v' - v)$ and has for θ large enough a unique solution

$$v' = \mathcal{V}_\theta(t, v, X, Y) = \begin{pmatrix} \frac{(\theta - \xi_1)v_2 - \xi_2 v_1^2}{\theta + \xi_1} \\ -\frac{\theta - \xi_1 + \xi_2 \frac{(\theta - \xi_1)v_2 - \xi_2 v_1^2}{\theta + \xi_1}}{\theta + \xi_1 + \xi_2 \frac{(\theta - \xi_1)v_2 - \xi_2 v_1^2}{\theta + \xi_1}} v_1 \end{pmatrix}.$$ We leave as an exercise to the reader the writing

of the equation for X^{k+1} and the formula for v^{k+1}.

This model corresponds to the problem of controlling the orientation γ of a molecule, considered as rigid rotator.

4.3.1 Numerical test

To test our approach we have used the parameters of the molecule CO (see [7, 35]), namely $H_0 = BJ^2$, where B is the rotational constant and J is the angular momentum. We consider the basis given by the spherical harmonics ; the corresponding matrix is diagonal with diagonal coefficients given by $(H_0)_{k,k} = k(k+1)$. The controlled is performed over an interval of length $T = 20T_{per} = 20\frac{\pi}{B}$. We consider constant penalization factor $\alpha = 10^{-1}$ and optimization parameter $\theta = 10^3$.

The other parameters correspond to the polarizability and the hyperpolarizability components of the molecule. We have $\mu_1 = -\frac{1}{2}\lambda$, and $\mu_2 = -\frac{3}{4}\beta$, with $\lambda = \frac{1}{2}(\lambda_\parallel \cos^2\gamma + \lambda_\perp \sin^2\gamma)$, $\beta = \frac{1}{6}((\beta_\parallel - 3\beta_\perp)\cos^3\gamma + 3\beta_\perp \cos\gamma)$. The matrix $\cos\gamma$ is tridiagonal, with:

$$(\cos\gamma)_{k,k} = 0, \ (\cos\gamma)_{k,k+1} = (\cos\gamma)_{k+1,k} = \frac{k+1}{\sqrt{(2k+1)(2k+3)}}.$$

We use the numerical values given in [7, 35]:

B	$\lambda_\perp$	$\lambda_\parallel$	$\beta_\parallel$	$\beta_\perp$
1.93	11.73	15.65	28.35	6.64

A gradient method similar to the one that is used in Section 4.1.3 is also performed. The results are presented in Fig. 3. The monotonic algorithm shows a fast convergence whereas the gradient method does not optimizes efficiently the cost functional values.

5 Conclusion

Motivated by a set of control algorithms that were initially introduced in the specific context of quantum control we have presented an abstract formulation that includes them all. It is seen that the algorithm involves at each step the resolution of a highly nonlinear evolution equation. We identified the theoretical assumptions that ensure that the evolution equation is well posed and has a solution. The proof being constructive it serves as basis for numerical approximations of the solution. We also proved several properties concerning the algorithms and more specifically concerning its convergence. Examples are provided to indicate how the proposed procedure solves cases from the literature and also new situations that were not previously considered. Numerical simulations indicate that the procedures have indeed the expected behavior.
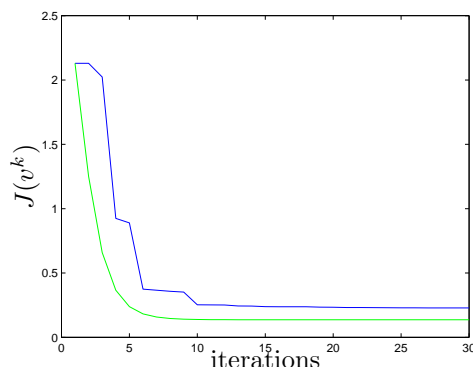
Figure 3: Numerical resolution of the example of Section 4.3. The cost functional $J(v)$ is optimized using the monotonic algorithm (34) (green line) and a gradient method (blue line).

Acknowledgements

References

[1] R.A. Adams and J.J.F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003.

[2] A. Bartana and R. Kosloff. Laser cooling and internal degrees of freedom. II. *J. Chem. Phys.*, 106:1435–1448, 1997.

[3] J. D. Benamou and Y. Brenier. Mixed L^2-Wasserstein optimal mapping between prescribed density functions. *J. Optim. Theory Appl.*, 111(2):255–271, 2001.

[4] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.

[5] J. Bolte and E. Attouch. On the convergence of the proximal algorithm for nonsmooth functions involving analytic features. *Mathematical Programming*, DOI 10.1007/s10107-007-0133-5, 2007.

[6] Robert Dautray and Jacques-Louis Lions. *Mathematical analysis and numerical methods for science and technology. Vol. 5.* Springer-Verlag, Berlin, 1992. Evolution problems. I, With the collaboration of Michel Artola, Michel Cessenat and Hélène Lanchon, Translated from the French by Alan Craig.

[7] B. Friedrich and D. Herschbach. *Phys. Rev. Lett.*, 74:4623, 1995.

[8] R. Gabasov, F. M. Kirillova, and S. V. Prischepova. *Optimal feedback control*, volume 207 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag London Ltd., London, 1995.

[9] Kazufumi Ito and Karl Kunisch. Optimal bilinear control of an abstract Schrödinger equation. *SIAM J. Control Optim.*, 46(1):274–287 (electronic), 2007.

[10] A. Lachapelle, J. Salomon, and G. Turinici. A monotonic algorithm for a mean field games model in economics. *M3AS*, page accepted, 2009.

[11] M. Lapert, R. Tehini, G. Turinici, and D. Sugny. Monotonically convergent optimal control theory of quantum systems under a nonlinear interaction with the control field. *Phys. Rev A*, 2008.

[12] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343(9):619–625, 2006.

[13] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris*, 343(10):679–684, 2006.

[14] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.

[15] R. J. Levis, G.M. Menkir, and H. Rabitz. Selective bond dissociation and rearrangement with optimally tailored, strong-field laser pulses. *Science*, 292:709–713, 2001.

[16] J.-L. Lions. *Optimal control of systems governed by partial differential equations*. Springer-Verlag, New York, 1971.

[17] Yvon Maday, J. Salomon, and Gabriel Turinici. Monotonic parareal control for quantum systems. *SIAM Journal on Numerical Analysis*, 45(6):2468–2482, 2007.

[18] Yvon Maday and Gabriel Turinici. New formulations of monotonically convergent quantum control algorithms. *J. Chem. Phys.*, 118(18), 2003.

[19] Y. Ohtsuki, W. Zhu, and H. Rabitz. Monotonically convergent algorithm for quantum optimal control with dissipation. *J. Chem. Phys.*, 110:9825–9832, 1999.

[20] Yukiyoshi Ohtsuki. Non-markovian effects on quantum optimal control of dissipative wave packet dynamics. *The Journal of Chemical Physics*, 119(2):661–671, 2003.

[21] Yukiyoshi Ohtsuki and Kazuyuki Nakagami. Monotonically convergent algorithms for solving quantum optimal control problems of a dynamical system nonlinearly interacting with a control. *Phys. Rev. A*, 77(3):033414, Mar 2008.

[22] Yukiyoshi Ohtsuki, Yoshiaki Teranishi, Gabriel Turinici, and Herschel Rabitz. Monotonically convergent algorithms for solving quantum optimal control problems described by an integro-differential equation of motion. *PRA*, 75:033407, 2007.

[23] Yukiyoshi Ohtsuki, Gabriel Turinici, and Herschel Rabitz. Generalized monotonically convergent algorithms for solving quantum optimal control problems. *J. Chem. Phys.*, 120(12):5509–5517, 2004. also published in the Virtual Journal of Ultrafast Science.

[24] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical recipes in C++*. Cambridge University Press, Cambridge, 2002. The art of scientific computing, Second edition, updated for C++.

[25] H. Rabitz, R. de Vivie-Riedle, M. Motzkus, and K. Kompa. Wither the future of controlling quantum phenomena? *Science*, 288:824–828, 2000.

[26] H Rabitz, M. R. Hsieh, and C. Rosenthal. Quantum optimally controlled transition landscapes. *Science*, 303:1998, 2004.

[27] S. Rice and M. Zhao. *Optical Control of Quantum Dynamics*. Wiley, 2000.

[28] J. Salomon and Guillaume Carlier. A monotonic algorithm for the optimal control of the fokker-planck equation. In *Proceedings of the 47th IEEE Conference on Decision and Control, Cancun, Mexico*, december 2008.

[29] J. Salomon, C.M. Dion, and G. Turinici. Optimal molecular alignment and orientation through rotational ladder climbing. *J. Chem. Phys.*, 123(14):144310, October 2005.

[30] S. Schirmer, M. Girardeau, and J. Leahy. Efficient algorithm for optimal control of mixed-state quantum systems. *Phys. Rev. A*, 61:012101, 2000.

[31] Shlomo E. Sklarz and David J. Tannor. Loading a bose-einstein condensate onto an optical lattice: An application of optimal control theory to the nonlinear Schrödinger equation. *Phys. Rev. A*, 66(5):053619, Nov 2002.

[32] M. Sugawara. General formulation of locally designed coherent control theory for quantum systems. *J. Chem. Phys.*, 118(15):6784–6800, 2003.

[33] M. Sugawara and Y. Fujimura. Control of quantum dynamics by a locally optimized laser field. application to ring puckering isomerization. *J. Chem. Phys.*, 100:5646–5655, 1993.

[34] D. Tannor, V. Kazakov, and V. Orlov. Control of photochemical branching: Novel procedures for finding optimal pulses and global upper bounds. In Broeckhove J. and Lathouwers L., editors, *Time Dependent Quantum Molecular Dynamics*, pages 347–360. Plenum, 1992.

[35] R. Tehini and D. Sugny. *Phys. Rev. A*, 77:023407, 2008.

[36] W.S. Warren, H. Rabitz, and M. Dahleh. Coherent control of quantum dynamics: The dream is alive. *Science*, 259:1581–1589, 1993.

[37] T.C. Weinacht, J. Ahn, and P.H. Bucksbaum. Controlling the shape of a quantum wavefunction. *Nature*, 397:233–235, 1999.

[38] K. Yosida. *Functional analysis*. Springer-Verlag, Berlin, 1995.

[39] W. Zhu, J. Botina, and H. Rabitz. Rapidly convergent iteration methods for quantum optimal control of population. *J. Chem. Phys.*, 108:1953–1963, 1998.

[40] 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.

[41] W. Zhu and H. Rabitz. Uniform rapidly convergent algorithm for quantum optimal control of objectives with a positive semi-definite Hessian matrix. *Phys. Rev. A.*, 58:4741–4748, 1998.