



**HAL**  
open science

# A monotonic method for solving nonlinear optimal control problems

Julien Salomon, Gabriel Turinici

► **To cite this version:**

Julien Salomon, Gabriel Turinici. A monotonic method for solving nonlinear optimal control problems. 2010. hal-00335297v3

**HAL Id: hal-00335297**

**<https://hal.science/hal-00335297v3>**

Preprint submitted on 2 Jul 2010 (v3), last revised 9 Nov 2010 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A MONOTONIC METHOD FOR SOLVING NONLINEAR OPTIMAL CONTROL PROBLEMS WITH CONCAVE DEPENDENCE ON THE STATE

JULIEN SALOMON AND GABRIEL TURINICI

ABSTRACT. Initially introduced in the framework of quantum control, the so-called *monotonic algorithms* have demonstrated excellent numerical results when dealing with various bilinear optimal control problems. This paper presents a unified formulation that can be applied to more nonlinear settings. 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.

## 1. INTRODUCTION

This paper aims at presenting a general unified formulation of several algorithms that were proposed in different areas of nonlinear (bilinear) control. Given a cost functional  $J(v)$  depending on the control  $v$ , these algorithms are iterative procedures that construct a sequence of solution candidates  $v^k$  with the important "monotonic" behavior, i.e.  $J(v^{k+1}) \leq J(v^k)$ ; the algorithms have been named after this property as "monotonic". An convenient property of these procedures is that the monotonicity does not requires any additional computational effort, but results from the construction of the procedure itself.

These procedures have first been used in the field of quantum control, where quantum particles' dynamics is controlled by a laser field as described by the Schrödinger equation (cf. Section 4.1 for more details about the modeling of this problem). In this framework, the function that to each control associates the final state of the system is highly nonlinear. This induces poor performance of standard, gradient-based algorithms. The "monotonic schemes" introduced in [2, 49, 55] appeared in this context and were found to perform excellently in this very nonlinear setting. These schemes were used in bi-linear situations where the control multiplies the state. These were soon followed variants [54, 56] that generalized the cost functional to include situations more complicated that a distance to a given target.

At first the relationship between the procedures introduced in the cited works was not obvious but in [29] it was showed that there are all particular cases of a two-parameter class of algorithms.

Though the monotonic schemes are based on algebraic calculations, the specific setting induced by the Schrödinger equation enables in [43] to relate the monotonic schemes to trajectory tracking algorithm [30, 31]. At the numerical level, efficient discretizations of the procedure have been proposed in [28, 11] and a time parallelized version was introduced in [27]. Continuing interest in the monotonic schemes lead to the introduction in [14] of versions that ensure that the resulting field will fit in a given frequency window.

In previous works the objective was encoded through a criterion on the final state but adaptations were proposed in [36, 35] to deal with the case where the optimal control functional depends on the whole dynamics of the control process or when the dynamics involve integro-differential equations. Additional situations involving dissipation operators were proposed in [32] with a non-Markovian version in [33].

Further different examples consider the case where the system is described by a density matrix operator instead of a wave function: details on the computation and convergence proofs limited to this case were given in [44, 47, 48] and in [18] this is applied to create a quantum computer gate; further applications can be found in [37, 9].

At some point similar procedures were also proposed in other control applications ([8, 19]).

The convergence of the algorithm has been obtained in the case of quantum control (see Section 4.1) using Lojasiewicz-Simon inequality (see [7, 17, 26, 45] and the references therein) and also in discrete and continuous settings in [3, 41]. The structure of the proofs shows that when  $J$  is analytic and its gradient is Fredholm, convergence is guaranteed as soon as  $J$  contains a penalization term of the  $L^2$ -norm of the control  $v$ . Note also that another proof has been presented in the framework of semi-group theory [15] using compactness arguments. All these results are available for a bi-linear setting when the control multiplies the state and are specific to the Schrödinger equations.

Up to this point all works presented above considered bi-linear situations (in all cases the optimal control functional is a non-linear functional) ; only recently different cases were documented in the literature: in [42, 20] the procedures were tailored to tackle specific non-bilinear models in which the control field appears up to power 3. A situation when a unique control appears at arbitrary powers of polynomial was proposed in [34]. A model where the system is a nonlinear Bose-Einstein condensate was given in [46].

In all situations where monotonic algorithms were introduced the well-posedness of the algorithms were proved by ad-hoc techniques and the same for convergence, although the algebraic computations share similar points. The purpose of this paper is to identify the similarities present in all these situations, and present a general setting to which the "monotonic" algorithms belong ; we also propose corresponding formulas and procedures to solve such type of problems. This allows to tackle general non-linear situations the cannot be solved with techniques presented in the literature.

The paper is structured as follows: Section 2 defines the general framework where our procedure applies. The algorithm itself 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 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 two vector spaces  $\mathcal{A}$  and  $\mathcal{B}$  we denote by  $\mathcal{L}(\mathcal{A}, \mathcal{B})$  the space of linear continuous of operator between  $\mathcal{A}$  and  $\mathcal{B}$ .

Given a real or complex valued function  $\varphi$ , 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 functions 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  $f : U \rightarrow H_2$  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

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

If this is the case the operator  $A_x$  is called the Fréchet differential (or Fréchet derivative) of  $f$  at  $x$  and is denoted  $A_x = D_x f$ .

Let us also recall that for a open set  $\Omega \subset \mathbb{R}^\gamma$  and any Hilbert space  $H_1$ ,  $L^\infty(\Omega; H_1)$  is the space of functions  $f$  from  $\Omega$  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 will be the  $L^\infty(\Omega; H_1)$  norm of  $f$ ).

In the same way one can define  $L^2(\Omega; H_1)$

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

When the derivatives of functions of  $f$  are considered the Sobolev spaces  $W^{1,\infty}$  have to be introduced; we refer to [1, 53] for further details.

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

$$(2) \quad \min_v J(v),$$

where

$$(3) \quad J(v) := \int_0^T F(t, v(t), X(t)) dt + G(X(T)).$$

The functions  $F : \mathbb{R} \times E \times \mathbf{V} \rightarrow \mathbb{R}$  and  $G : \mathbf{V} \rightarrow \mathbb{R}$  are assumed to be differentiable and integral assumed to exist. The system is described by a state function  $X(t) \in \mathbf{V}$  being governed by the evolution equation

$$(4) \quad \partial_t X + A(t, v(t))X = B(t, v(t))$$

$$(5) \quad X(0) = X_0.$$

where  $v : [0, T] \rightarrow E$  is the control. The unbounded operator  $A(t, v) : \mathbb{R} \times E \times \mathcal{H} \rightarrow \mathcal{H}$  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  $B(t, v) \in \mathcal{L}(\mathcal{H}, \mathcal{H}) \cap \mathcal{L}(\mathbf{V}, \mathbf{V}^*)$ . We postpone to Section 3 (cf. Lemma 3.4, Theorem 1) the precise formulation of additional regularity assumptions to be imposed on  $A, B, F, G$ .

**Remark 2.** Finally, note that  $E$  is not necessarily a real number, neither finite dimensional, cf. Section 4.2. This means that the control can be a set of several time-dependent function but also a distributed control depending on time and also on a spacial variable.

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 [42, 20] and polynomial with  $E = \mathbb{R}^1$  in [34]). The techniques present in the above papers cannot be used for general operators  $A(t, v)$ . On the contrary the results in this paper include all the situations considered in the bibliography but also apply to all nonlinearities in  $v$  compatible with the hypothesis of Lemma 3.4 and Thm. 1 below.

Moreover, the following concavity with respect to  $X$  will be assumed throughout the paper:

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

$$(7) \quad \forall t \in \mathbb{R}, \forall v \in E, \forall X, X' \in \mathbf{V}, F(t, v, X') - F(t, v, X) \leq \langle \nabla_X F(t, v, X), X' - X \rangle_{\mathbf{V}}.$$

**Remark 4.** *Contrary to 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 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 [13, 25]) by:

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

$$(9) \quad Y_v(T) = \nabla_X G(X_v(T)).$$

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

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

$$(10) \quad \Upsilon\left(t, X_v(t), v(t), v'(t), Y_v(t), X_{v'}(t)\right) = -\langle Y_v(t), \left(A(t, v'(t)) - A(t, v(t))\right) X_{v'}(t) \rangle_{\mathbf{V}} \\ \langle Y_v(t), B(t, v'(t)) - B(t, v(t)) \rangle_{\mathbf{V}} + F(t, v'(t), X_{v'}(t)) - F(t, v(t), X_{v'}(t)).$$

Then

$$(11) \quad J(v') - J(v) \leq \int_0^T \Upsilon\left(t, X_v(t), v(t), v'(t), Y_v(t), X_{v'}(t)\right) dt.$$

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

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

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

**Remark 5.** *The focus of the result is not on obtaining an estimation of the increment  $J(v') - J(v)$  via the adjoint (which is well documented in optimal control theory, cf. [13, 25]); we rather emphasize that the evaluation of the integrand  $\Upsilon(\dots)$  at time  $t$  requires information on the control  $v(s)$  for all  $s \in [0, T]$  (in order to compute  $X_v(T)$  then  $Y_{\tilde{v}}(t)$ ) 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 in deciding, **at time**  $t$  if the current value of the control  $v'(t)$  will result in 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  $\Upsilon(\dots)$  has the factorized form:*

$$(12) \quad \Upsilon\left(t, X_v(t), v(t), v'(t), Y_{\tilde{v}}(t), X_{v'}(t)\right) = \Delta(v, v')(t) \cdot_E (v'(t) - v(t)),$$

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 intermediate field  $\tilde{v}$  (cf. also [29]), i.e. according to the equation :

$$\begin{aligned} \frac{\partial}{\partial t} Y_{\tilde{v}} - A^*(t, \tilde{v}(t)) Y_{\tilde{v}} + \nabla_X F(t, v(t), X_v(t)) &= 0 \\ Y_{\tilde{v}}(T) &= \nabla_X G(X_v(T)). \end{aligned}$$

Note that because of its final condition,  $Y_{\tilde{v}}$  actually also depends on  $v$ . Nevertheless, for sake of simplicity, we keep the previous notation. We then obtain the following lemma.

**Lemma 3.2.** *For any  $v', \tilde{v}, v : [0, T] \rightarrow E$ ,*

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

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  $v' - \tilde{v}$  and  $\tilde{v} - v$ .

**Remark 7.** *Lemmas 3.1 and 3.2 are generalization of previous results that were proved in the bi-linear case. To the best of our knowledge, only specific corollaries requiring additional assumptions have been used in the literature up to now.*

**3.2. The algorithms.** The factorization obtained above 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  $v = v^k$ , starting from

$$Y_{v^k}(T) := \nabla_X G(X_{v^k}(T)).$$

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

$$(13) \quad \Delta(v^{k+1}, v^k)(t) \cdot_E (v^{k+1}(t) - v^k(t)) \leq 0.$$

Lemma 3.1 then guarantees that  $J(v^{k+1}) \leq J(v^k)$ . Several strategies can be used to ensure (13); 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 (13), many authors (see [29, 49, 55]) consider an update formula of the form:

$$(14) \quad v^{k+1}(t) - v^k(t) = -\frac{1}{\theta} \Delta(v^{k+1}, v^k)(t),$$

where  $\theta$  is a positive number, that can also depend on  $k$  and  $t$ . In what follows, we focus on the existence of solution of (14), and on practical methods to compute it. If  $v^{k+1}$  satisfies (14), the variations in  $J$  satisfy:

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

Note that (14) 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 (introduced by [6]), 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, (14) 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 (14). Since this procedure involves the resolution of an implicit equation, see Eq. (14), we prove the existence of a solution and present a convergent procedure to compute it. As a by-product, we obtain a proof of the monotonicity of the algorithm.

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

- $\mathcal{A} : \mathbb{R} \times \mathbf{V} \times \mathbf{V} \times E \rightarrow \mathbb{R}$  defined by  $\mathcal{A}(t, X, Y, v) = \langle Y, A(t, v)X \rangle_{\mathbf{V}}$  is Fréchet differentiable everywhere with respect to  $v$  for any  $X, Y, v$ .
- $\mathcal{B} : \mathbb{R} \times \mathbf{V} \times E \rightarrow \mathbb{R}$  with  $\mathcal{B}(t, Y, v) = \langle Y, B(t, v) \rangle_{\mathbf{V}}$  is Fréchet differentiable everywhere with respect to  $v$  for any  $Y, v$ .
- $F$  is Fréchet differentiable everywhere 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  $v, v' \in E$

$$(15) \quad \begin{aligned} \Delta(v', v; t, X, Y) \cdot_E (v' - v) &= - \left\langle Y, \left( A(t, v') - A(t, v) \right) X + B(t, v') - B(t, v) \right\rangle_{\mathbf{V}} \\ &\quad + F(t, v', X) - F(t, v, X). \end{aligned}$$

Moreover, if  $\mathcal{A}, \mathcal{B}, F$  are of  $C^1$  class in  $v$  then  $\Delta(\cdot, \cdot; t, X, Y)$  can be defined through the explicit formula:

$$(16) \quad \begin{aligned} \Delta(v', v; t, X, Y) &= \int_0^1 -\nabla_w \left( \langle Y, A(t, w) X - B(t, w) \rangle_{\mathbf{V}} \right) \Big|_{w=v+\lambda(v'-v)} \\ &\quad + \nabla_v F(t, v + \lambda(v' - v), X) d\lambda. \end{aligned}$$

*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. (15) 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 in  $v$ ; we introduce

$$(17) \quad \Delta_{\Xi}(v', v) := \frac{\Xi(v') - \Xi(v)}{\|v' - v\|^2} (v' - v) \in E.$$

Since  $\Xi$  is differentiable, we obtain the continuity of  $\Delta_{\Xi}(v', v)$  for all points  $v' = v$  and  $\Delta_{\Xi}(v, v) = \nabla_v \Xi(v)$  (the continuity is obvious everywhere else) hence the conclusion.

Finally, Eq. (16) is an application of the identity

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

□

**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 in  $X$ ;

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

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  $\theta^* > 0$  depending only on  $\varepsilon, W, \|v\|, \|X\|$  and  $\|Y\|$  such that, for any  $\theta > \theta^*$

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

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

- (3)  $\|\mathcal{V}_{\theta}(t, v, X, Y) - v\| \leq \frac{\|X\|\|Y\| + \|Y\| + k(\|X\|)}{\theta} \{M_0(t) + M_1\|v\|\}$  with  $M_0(t)$  and  $M_1$  independent of  $v, X, Y$ . If the dependence of  $A, B, F$  on  $t$  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  $\varepsilon$ .

*Proof.*

- (1) Denote  $h = v' - v$  and  $\mathcal{G}_{t,v,X,Y}(h) = \frac{-\Delta(v+h,v;t,X,Y)}{\theta}$ . 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  $\theta$  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 (16).  
(3) For  $\theta$  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  $\|h\| - \|\mathcal{G}(0)\| \leq \|h - \mathcal{G}(0)\| = \|\mathcal{G}(h) - \mathcal{G}(0)\| \leq \frac{1}{2}\|h\|$ , which amounts to  $\|h\| \leq 2\|\mathcal{G}(0)\|$ . Next, we note that

$$\begin{aligned} \|\mathcal{G}(0)\| &\leq \frac{\|\Delta(v, v, t, X, Y) - \Delta(0, 0, t, X, Y)\| + \|\Delta(0, 0, t, X, Y)\|}{\theta} \\ &\leq M_2\|v\| + M_3(t) \end{aligned}$$

and the estimates follows.

- (4) Formula (16) shows that  $\Delta$  depends continuously on  $t, v, v', X, Y$ . Consider converging sequences  $t_n \rightarrow t, v_n \rightarrow v, X_n \rightarrow X, Y_n \rightarrow Y$  and define  $h_n := \mathcal{V}_\theta(t_n, v_n, X_n, Y_n)$  and  $h := \mathcal{V}_\theta(t, v, X, Y)$ .

Given  $W$  and  $\eta > 0$ , consider large value of  $\theta$  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  $(t', v', X', Y'), (t'', v'', X'', Y'') \in W$ ,

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

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

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

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

- (5) Subtracting the two equalities

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

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

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

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

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

We are thus able to give an example of a setting where the existence of  $v^{k+1}(t)$  satisfying (13) is guaranteed.

**Theorem 1.** Suppose that  $A, B, F$  satisfy hypothesis of 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.

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

$$(19) \quad \partial_t X_{v'}(t) + A(t, v')X_{v'}(t) = B(t, v')$$

$$(20) \quad v'(t) = \mathcal{V}_\theta(t, v(t), X_{v'}(t), Y_v(t))$$

$$(21) \quad X_{v'}(0) = X_0$$

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 (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

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

(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.* Most of the proof is already 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 :

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

We take a spherical neighborhood  $B_v(R)$  of  $v$  of radius  $R$  and suppose

$$\forall k \leq l, \quad v_k \in B_v(R).$$

Since the correspondence  $v \mapsto X_v$  is continuous, it follows that the set of solutions  $S_{v, R} := \{X_w; w \in B_v(R)\}$  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  $\theta$  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  $\theta$  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 a solution of (19–20).  $\square$

#### 4. EXAMPLES

We now present two examples that fit into the setting of Theorem 1. The space does not allow to treat all other 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 [24, 38, 39, 40, 51, 52]) or the modeling of an equilibrium (or again social beliefs, product prices, etc) within a game with infinite numbers of agents (see [21, 22, 23]). Yet other applications arise from modern formulations of the Monge-Kantorovich mass transfer problem (see [5, 4, 8]).

In the following, we present some examples coming from these fields of application and present the corresponding monotonic algorithm resulting from Theorem 1.

##### 4.1. (I): Quantum control.

4.1.1. *Setting.* The evolution of a quantum system is described by the Schrödinger equation

$$\begin{aligned}\partial_t X + iH(t)X &= 0 \\ X(0, z) &= X_0(z),\end{aligned}$$

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 an auto-adjoint operator over  $L^2(\mathbb{R}^\gamma; \mathbb{C})$ , i.e.  $H(t)^* = H(t)$ <sup>1</sup>. Note that this results in the following norm conservation property

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

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

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  $\mu(x)$  resulting in the formula  $H(t) = H_0 - v(t)\mu$  and the dynamics:

$$\begin{aligned}\partial_t X + i(H_0 - v(t)\mu)X &= 0 \\ X(0) &= X_0.\end{aligned}$$

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:

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

$$(24) \quad \tilde{J}(v) := -\langle X(T), O(X(T)) \rangle_{L^2(\mathbb{R}^\gamma; \mathbb{C})} + \int_0^T \alpha(t)v^2(t)dt,$$

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 thus to minimize these functionals with respect to  $v$ . According to (22) the cost functional  $J$  is equal to

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

so that the functionals  $J$  and  $\tilde{J}$  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  $\mu$ . The dependence of  $A$  on  $v$  is smooth (linear) and therefore all hypotheses on  $A$  are satisfied.

- $E = \mathbb{R}$ ,  $\mathcal{H} = L^2(\mathbb{R}^d; \mathbb{C})$ ,  $\mathbf{V} = \operatorname{dom}(H_0^{1/2})$  (over  $\mathbb{C}$ ), or their realifications  $\mathcal{H} = L^2 \times L^2$ ,  $\mathbf{V} = \operatorname{dom}(H_0^{1/2}) \times \operatorname{dom}(H_0^{1/2})$  (over  $\mathbb{R}$ ) see [16];

- $B(t, v) = 0$ .

- $F(t, v, X) = \alpha(t)v(t)^2$  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.

---

<sup>1</sup>For any operator  $M$ , we denote by  $M^*$  its adjoint.

- $G$  is either (see, e.g., [27, 29])  $2 - 2\text{Re}\langle X_{target}, X(T) \rangle_{\mathbf{V}}$  or  $-\langle X(T), OX(T) \rangle_{\mathbf{V}}$  where  $O$  is a positive semi-definite operator; both are concave in  $X$ .
- Here

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

and the equation in  $v'$   $\Delta(v', v; t, X, Y) = -\theta(v' - v)$  has, for  $\theta$  large enough, an unique solution  $v' = \mathcal{V}_\theta(t, v, X, Y) := \frac{(\theta - \alpha(t))v + \text{Re}\langle Y, i\mu X \rangle_{\mathbf{V}}}{\theta + \alpha(t)}$ .

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

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

Then

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

4.1.3. *Numerical test.* In order to test the performance of the algorithm we have chosen a case already treated in the literature. The system under consideration is the  $O - H$  bond that vibrates in a Morse type potential  $V(x) = D_0(\exp(-\beta(x - x')) - 1)^2 - D_0$ . The dipole moment operator of this system is modeled by  $\mu(x) = \mu_0.xe^{-\frac{x}{x^*}}$ . We refer to [55] for more details concerning this system. The objective is to localize the wavefunction in a time  $T = 131000$  at a given location  $x_0$ ; this is expressed through the requirement that the functional  $\tilde{J}$  is maximized, where the observable  $O$  is defined by  $O(x) = \frac{\gamma_0}{\sqrt{\pi}}e^{-\gamma_0^2(x-x_0)^2}$ . We consider a constant penalization parameter  $\alpha = 1$  and optimization parameter  $\theta = 10^{-2}$ . The numerical values we use are give in the next table.

$D_0$	$\beta$	$x'$	$x^*$	$x_0$	$\gamma_0$	$\mu_0$
0.1994	1.189	1.821	0.6	2.5	25	3.088

Results are presented in Fig. 1.

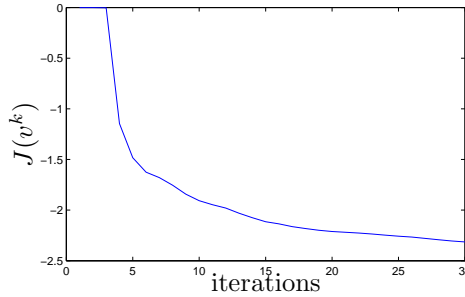


FIGURE 1. For the example in Section 4.1 we plot the evolution of the cost functional values  $J(v^k)$  as function of the iteration number  $k$ . Monotonic decrease is observed as expected by the theoretical arguments.

#### 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 [21, 22, 23]) indicate that it is possible to extend it to a 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  $\nu > 0$ :

$$\begin{aligned}\partial_t X - \nu \Delta X + \operatorname{div}(v(t, z)X) &= 0, \\ X(0) &= X_0.\end{aligned}$$

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 [19]

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

with positive constants  $\beta, c_0, c_1, c_2$  and  $p(t)$  a positive function. Another example is given in [8]:

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

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  $T$ .

The relevance of the monotonic algorithms to this setting has been established in several works [8, 19].

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 [19] and [10] (Chap XVIII §4.4)
- $A(t, v) = -\nu \Delta \cdot + \operatorname{div}(v \cdot)$ . The dependence of  $A$  on  $v$  is smooth (linear) and therefore all hypotheses on  $A$  are satisfied ( $D_{vv}A = 0, \dots$ ).
- $B(t, v) = 0$ .
- 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

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

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

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

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

Then

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

4.2.3. *Numerical test.* The algorithm is test on the time interval  $[0, 1]$  with  $p(t) = 1$  and the numerical values  $\beta = 0.8$ ,  $c_0 = c_2 = 1$ ,  $c_1 = 0.1$ . Results are presented in Fig. 2.

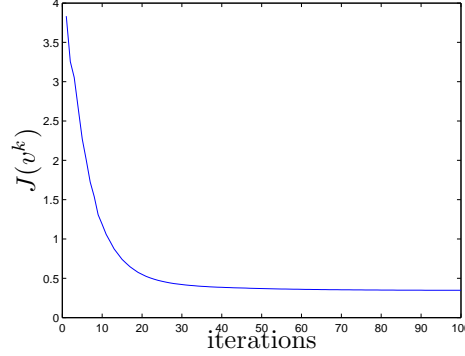


FIGURE 2. For the example in Section 4.2 we plot the evolution of the cost functional values  $J(v^k)$  as function of the iteration number  $k$ . Monotonic decrease is observed as expected by the theoretical arguments.

**4.3. Additional application.** As a third example we consider a nonlinear vectorial case from [12, 50] 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  $A(t, v) = i[H_0 + (v_1(t)^2 + v_2(t)^2)\mu_1 + v_1(t)^2v_2(t)\mu_2]$ . Here, denoting  $\xi_1 = -\text{Re}\langle Y, i\mu_1 X \rangle_{\mathbf{V}} + \alpha(t)$ ,  $\xi_2 = -\text{Re}\langle Y, i\mu_2 X \rangle_{\mathbf{V}}$  we obtain

$$(34) \quad \Delta(v', v; t, X, Y) = \xi_1 \begin{pmatrix} v_1 + v'_1 \\ v_2 + v'_2 \end{pmatrix} + \xi_2 \begin{pmatrix} (v_1 + v'_1)v'_2 \\ (v_1)^2 \end{pmatrix}$$

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

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

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  $\gamma$  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* [12, 50], 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. In this way 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 [12, 50]:

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

The results are presented in Fig. 3. As in the previous examples, a rapid convergence is obtained, since about 100 iterations are necessary to reach the numerical convergence.

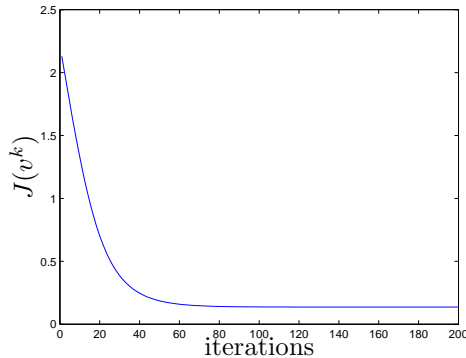


FIGURE 3. For the example in Section 4.3 we plot the evolution of the cost functional values  $J(v^k)$  as function of the iteration number  $k$ . Monotonic decrease is observed as expected by the theoretical arguments.

## 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 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 procedure proposed solves previous cases from the literature and also new situations that were not previously considered. Numerical simulations indicate that the procedures have indeed the expected behavior.

## ACKNOWLEDGEMENTS

This work is partially supported by the French ANR programs OTARIE (ANR-07-BLAN-0235 OTARIE) and C-Quid, and by a CNRS-NFS PICS grant. G.T. acknowledges partial support by INRIA Rocquencourt (MicMac and OMQP).

## 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] Lucie Baudouin and J. Salomon. Constructive solution of a bilinear control problem. *Syst. Cont. Lett.*, 57(6):453–464, 2008.

- [4] 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.
- [5] 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.
- [6] 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.
- [7] J. Bolte, A. Daniilidis, O. Ley, and L. Mazet. Characterization of Lojasiewicz inequalities and applications. *arXiv:0802.0826*, 2008.
- [8] Guillaume Carlier and Julien Salomon. A monotonic algorithm for the optimal control of the fokker-planck equation. december 2008.
- [9] Alberto Castro and E. K. U. Gross. Acceleration of quantum optimal control theory algorithms with mixing strategies. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 79(5):056704, 2009.
- [10] 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.
- [11] P. Ditz and A. Borzi. A cascadic monotonic time-discretized algorithm for finite-level quantum control computation. *Computer Physics Communications*, 178(5):393–399, 2008.
- [12] Bretislav Friedrich and Dudley Herschbach. Alignment and trapping of molecules in intense laser fields. *Phys. Rev. Lett.*, 74(23):4623–4626, Jun 1995.
- [13] 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.
- [14] C. Gollub, M. Kowalewski, and R. de Vivie-Riedle. Monotonic convergent optimal control theory with strict limitations on the spectrum of optimized laser fields. *Physical Review Letters*, 101(7):073002, 2008.
- [15] Kazufumi Ito and Karl Kunisch. Optimal bilinear control of an abstract Schrödinger equation. *SIAM Journal on Control and Optimization*, 46(1):274–287, 2007.
- [16] Kazufumi Ito and Karl Kunisch. Optimal bilinear control of an abstract Schrödinger equation. *SIAM J. Control Optim.*, 46(1):274–287 (electronic), 2007.
- [17] M. A. Jendoubi. A simple unified approach to some convergence theorems of L. Simon. *J. Func. Anal.*, 153:187–202, 1998.
- [18] Christiane P. Koch, José P. Palao, Ronnie Kosloff, and Françoise Masnou-Seeuws. Stabilization of ultracold molecules using optimal control theory. *Physical Review A (Atomic, Molecular, and Optical Physics)*, 70(1):013402, 2004.
- [19] A. Lachapelle, J. Salomon, and G. Turinici. A monotonic algorithm for a mean field games model in economics. *M3AS*, page accepted, 2009.
- [20] 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.
- [21] 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.
- [22] 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.
- [23] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.
- [24] 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.
- [25] J.-L. Lions. *Optimal control of systems governed by partial differential equations*. Springer-Verlag, New York, 1971.
- [26] S. Lojasiewicz. Sur la géométrie semi- et sous-analytique. *Ann. Inst. Fourier*, 43:1575–1595, 1993.
- [27] Yvon Maday, J. Salomon, and Gabriel Turinici. Monotonic parareal control for quantum systems. *SIAM Journal on Numerical Analysis*, 45(6):2468–2482, 2007.
- [28] Yvon Maday, Julien Salomon, and Gabriel Turinici. Monotonic time-discretized schemes in quantum control. *Numerische Mathematik*, 103(2):323–338, 2006.
- [29] Yvon Maday and Gabriel Turinici. New formulations of monotonically convergent quantum control algorithms. *J. Chem. Phys.*, 118(18), 2003.
- [30] M. Mirrahimi, G. Turinici, and P. Rouchon. Lyapunov control of bilinear Schrödinger equations. *Automatica*, 41:1987–1994, 2005.
- [31] M. Mirrahimi, G. Turinici, and P. Rouchon. Reference trajectory tracking for locally designed coherent quantum controls. *J. Phys. Chem. A*, 109(11):2631–2637, 2005.
- [32] Y. Ohtsuki, W. Zhu, and H. Rabitz. Monotonically convergent algorithm for quantum optimal control with dissipation. *J. Chem. Phys.*, 110:9825–9832, 1999.



- [33] Yukiyoishi Ohtsuki. Non-markovian effects on quantum optimal control of dissipative wave packet dynamics. *The Journal of Chemical Physics*, 119(2):661–671, 2003.
- [34] Yukiyoishi 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.
- [35] Yukiyoishi 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.
- [36] Yukiyoishi 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.
- [37] J. P. Palao and R. Kosloff. Optimal control theory for unitary transformations. *Phys. Rev. A*, 68(6):062308, Dec 2003.
- [38] H. Rabitz, R. de Vivie-Riedle, M. Motzkus, and K. Kompa. Wither the future of controlling quantum phenomena? *Science*, 288:824–828, 2000.
- [39] H Rabitz, M. R. Hsieh, and C. Rosenthal. Quantum optimally controlled transition landscapes. *Science*, 303:1998, 2004.
- [40] S. Rice and M. Zhao. *Optical Control of Quantum Dynamics*. Wiley, 2000.
- [41] J. Salomon. Convergence of the time-discretized monotonic schemes. *M2AN*, 41(1):77–93, 2007.
- [42] 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.
- [43] J. Salomon and Gabriel Turinici. On the relationship between the local tracking procedures and monotonic schemes in quantum optimal control. *J. Chem. Phys.*, 124:074102, 2006.
- [44] S. Schirmer, M. Girardeau, and J. Leahy. Efficient algorithm for optimal control of mixed-state quantum systems. *Phys. Rev. A*, 61:012101, 2000.
- [45] L. Simon. Asymptotics for a class of non-linear evolution equations, with applications to geometric problems. *Ann. of Math.*, 118:525–571, 1983.
- [46] 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.
- [47] M. Sugawara. General formulation of locally designed coherent control theory for quantum systems. *J. Chem. Phys.*, 118(15):6784–6800, 2003.
- [48] 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.
- [49] 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.
- [50] R. Tehini and D. Sugny. Field-free molecular orientation by non-resonant and quasi-resonant two-color laser pulses. *Phys. Rev. A*, 77:023407, 2008.
- [51] W.S. Warren, H. Rabitz, and M. Dahleh. Coherent control of quantum dynamics: The dream is alive. *Science*, 259:1581–1589, 1993.
- [52] T.C. Weinacht, J. Ahn, and P.H. Bucksbaum. Controlling the shape of a quantum wavefunction. *Nature*, 397:233–235, 1999.
- [53] K. Yosida. *Functional analysis*. Springer-Verlag, Berlin, 1995.
- [54] W. Zhu, J. Botina, and H. Rabitz. Rapidly convergent iteration methods for quantum optimal control of population. *J. Chem. Phys.*, 108:1953–1963, 1998.
- [55] 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.
- [56] 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.

UNIVERSITÉ PARIS-DAUPHINE/CEREMADE

*E-mail address:* `salomon@ceremade.dauphine.fr`

UNIVERSITÉ PARIS-DAUPHINE/CEREMADE

*E-mail address:* `turinici@ceremade.dauphine.fr`