A MONOTONIC METHOD FOR SOLVING NONLINEAR OPTIMAL CONTROL PROBLEMS

JULIEN SALOMON, GABRIEL TURINICI*

Abstract. Initially introduced in the framework of quantum control, the so-called monotonic algorithms have shown excellent numerical results when dealing with various bilinear optimal control problems. This paper aims at presenting a unified formulation of such procedures and the intrinsic assumptions they require. In this framework, we prove the feasibility of the general algorithm. Finally, we explain how these assumptions can be relaxed.

Key words. Monotonic algorithms, nonlinear control, optimal control, non-convex optimization

AMS subject classifications. 49J20 , 80M50

1. Introduction. We document in this paper a general unified formulation for several algorithms that were proposed in different areas of nonlinear (bilinear) control.

Historically first to appear in quantum control (cf. Section 3.1), it was noted that the nonlinearity of the mapping control $\rightarrow$ state induces poor performance of the standard, gradient-based algorithms; as such new numerical procedures have been proposed [1, 47, 53] and were found to perform excellently in this very nonlinear setting. These were soon followed by scores of variants [18, 25, 44, 26, 27, 28, 29, 30, 31, 32, 35, 16, 40, 41, 42, 45, 46, 51, 52, 54, 9, 33, 12, 7]. At some point similar procedures were also proposed in other control or optimization settings ([39, 17], cf. Section 3.2).

Given a cost functional J, 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". It is interesting to note that the monotonicity does not requires any additional computational effort, but results from the construction of the procedure itself.

The purpose of this paper is to investigate what is the most general class to which "monotonic" algorithms apply and propose a general framing for procedures tailored to solve such classes of problems.

The paper is structured as follows: Section 2 provides the general framework where our procedure applies; some examples of concrete realization follow in Section 3. The algorithm itself is presented in Section 4. In Section 5 we briefly explain some extensions to more nonlinear settings and then we give details about the convergence of the procedure (in Section 6). A discussion on the the numerical implementation of the algorithm is given in Appendix.

2. Setting of the problem. Let E, $\mathcal{H}$ and $\mathbf{V}$ be Hilbert spaces with $\mathbf{V}$ densly included in $\mathcal{H}$. and denote by $\cdot_E$, $\langle\cdot,\cdot\rangle$ and $\langle\cdot,\cdot\rangle_{\mathbf{V}}$ their associated scalar products.

For any space $W = E, \mathcal{H}, \mathbf{V}$ by $\langle\cdot,\cdot\rangle_W$ the scalar product associated with it.

Given a real or complex 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 functions in the Fréchet sense.

* CEREMADE, UMR CNRS 7534, Université Paris IX, Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, France. salomon@ceremade.dauphine.fr, turinici@ceremade.dauphine.fr .
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)

We consider the unbounded operator $A(t,v) : \mathbf{R} \times E \times \mathcal{H} \to \mathcal{H}$ and suppose that for almost all $t \in [0,T]$ the domain of $A(t,v)^{1/2}$ includes $\mathbf{V}$; furthermore we take $B(t,v)$ that for almost all $t \in [0,T]$ and all $v \in E$ associates an element of $\mathcal{L}(\mathcal{H},\mathcal{H}) \cap \mathcal{L}(\mathbf{V},\mathbf{V}^*)$.

We can now introduce a system whose state $X(t) \in \mathbf{V}$ is governed by the evolution equation

$$\partial_t X + A(t,v(t))X = B(t,v(t)) \tag{2.1}$$

$$X(0) = X_0. \tag{2.2}$$

where $v : [0,T] \to E$ is the control. Note that E is not necessarily finite dimensional, cf. Section 3.2 where $E = W^{1,\infty}(0,1)$ for a probability measure m. 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).

Within an optimal control formulation, the desirable evolution of the system is encoded in the following optimization problem:

$$\min_v J(v), \tag{2.3}$$

where

$$J(v) := \int_0^T F\big(t, v(t), X(t)\big)dt + G\big(X(T)\big). \tag{2.4}$$

The functions $F : \mathbf{R} \times E \times \mathbf{V} \to \mathbf{R}$ and $G : \mathbf{V} \to \mathbf{R}$ are supposed to be differentiable and integral supposed to exist. We postpone to Section 4 (cf. Lemma 4.4, Theorem 1) the precise formulation of additional regularity assumptions to be imposed on A, B, F, G.

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

$$\forall X, X' \in \mathbf{V},\ \ G(X') - G(X) \le \langle \nabla_X G(X), X' - X \rangle, \tag{2.5}$$

$$\forall t \in \mathbf{R}, \forall v \in E, \forall X, X' \in \mathbf{V},\ \ F(t,v,X') - F(t,v,X) \le \langle \nabla_X F(t,v,X), X' - X \rangle. \tag{2.6}$$

Remark 1. *Contrary to the more technical hypothesis that will be assumed latter, the properties* (2.5)*,* (2.6) *and the linearity of* (2.1) *are crucial to the existence of the monotonic algorithms. We will discuss in the Section 5 some possible ideas to relax the form* (2.1) *of the state evolution or the concavity.*

Remark 2. *The intrinsic nonlinear regime is manifest from the explicit* **concave** *dependence of the functionals F and G on the state ; certainly a linear $v \mapsto X$ minimization problem would only have trivial solutions for such functionals.*

3. Examples. Within the framework of control theory, nonlinear formulations prove useful nowadays in domains as diverse as the laser control of quantum phenomena [22, 34, 36, 37, 49, 50] or the modeling of a equilibrium (or again social beliefs, product prices, etc) within a game with infinite numbers of agents [19, 20, 21]. Yet other applications arise from modern formulations of the Monge-Kantorovich mass transfer problem [3, 4], see [39].

3.1. (I): Quantum control. 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 $H(t)$ is the Hamiltonian of the system and $z \in \mathbf{R}^\gamma$ the set of internal degrees of freedom. The Hamiltonian will be supposed to be an auto-adjoint operator over $L^2(\mathbf{R}^\gamma;\mathbf{C})$, i.e. $H(t)^* = H(t)$[1]. Note that this results in the following norm conservation property

$$\|X(t,\cdot)\|_{L^2(\mathbf{R}^\gamma;\mathbf{C})} = \|X_0\|_{L^2(\mathbf{R}^\gamma;\mathbf{C})},\ \forall t>0, \tag{3.1}$$

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

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 $v(t) \in \mathbf{R},\ t \geq 0$; 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:

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

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:

$$J(v) := \|X(T) - X_{target}\|^2_{L^2(\mathbf{R}^\gamma;\mathbf{C})} + \int_0^T \alpha(t)v^2(t)dt, \tag{3.2}$$

$$\widetilde{J}(v) := -\langle X(T), O(X(T))\rangle + \int_0^T \alpha(t)v^2(t)dt, \tag{3.3}$$

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 (3.1) the cost functional J is equal to

$$J(v) := 2 - 2Re\langle X(T), X_{target}\rangle_{L^2(\mathbf{R}^\gamma;\mathbf{C})} + \int_0^T \alpha(t)v^2(t)dt, \tag{3.4}$$

so that the functionals J and $\widetilde{J}$ satisfy assumptions (2.5) and (2.6).

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

3.2. (II) : Mean field games. Although the Nash equilibrium in game theory has been initially formulated for a finite number of players, modern results [19, 20, 21] 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 $\nu > 0$:

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

The control v is chosen to minimize the cost criterion (2.4). For reasons related to economic modeling interesting examples include situations where F, G are concave in X, e.g., in [17]

$$G = 0,\ F(t,z,X) = \int_Q p(t)(1-\beta z)X(t,z) + \frac{c_0 \cdot z \cdot X(t,z)}{c_1 + c_2 X(t,z)} dz + \frac{v^2(t)}{2}, \tag{3.5}$$

with positive constants β, c_0, c_1, c_2 and $p(t)$ a positive function. Another example is given in [39]:

$$G(X(T)) = \langle v, X(T) \rangle,\ F(t,z,X) = \int_Q X(t,z) v^2(t,z) dz, \tag{3.6}$$

The relevance of the monotonic algorithms to this setting has been established in several works [39, 17].

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

4.1. Tools for monotonic algorithms. The monotonic algorithms are mainly based on a special factorization obtained after algebraic manipulations built on the results presented 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. (2.1–2.2).

We define the adjoint state Y_v [11, 23] by:

$$\partial_t Y_v - A^*\big(t, v(t)\big) Y_v + \nabla_X F\big(t, v(t), X_v(t)\big) = 0 \tag{4.1}$$
$$Y_v(T) = \nabla_X G\big(X_v(T)\big). \tag{4.2}$$

Thanks to this auxiliary variable, a first estimate about the variations in J can be obtained.

LEMMA 4.1. *For any $v', v : [0,T] \to E$,*

$$\begin{aligned} J(v') - J(v) \le \int_0^T & -\langle Y_v(t), \Big(A\big(t, v'(t)\big) - A\big(t, v(t)\big)\Big) X_{v'}(t) \rangle \\ & + \langle Y_v(t), B\big(t, v'(t)\big) - B\big(t, v(t)\big) \rangle \\ & + F\big(t, v'(t), X_{v'}(t)\big) - F\big(t, v(t), X_{v'}\big) dt. \end{aligned} \tag{4.3}$$

Proof. Using successively (2.5),(2.6), (2.1) and finally (4.2), we find that:

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

Thanks to (4.1), the first term of the left-hand side of this last inequality cancels and the results follows.□

Remark 3. *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.[11, 23]); we rather emphasis that $J(v') - J(v)$ is upper bounded by a quantity that only requires information on v' up to time t (in other words which is independent of v' on $]t, T]$); this estimate can be useful in deciding,* **at time** *t if the next contribution of the control $v'(t+\delta t)$ 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 4. *We can intuitively note that the right hand side term of* (4.3) *has the factorized form:*

$$\begin{aligned}
\Delta(v, v')(t) \cdot_E \big(v'(t) - v(t)\big) = & -\langle Y_v(t), \Big(A\big(t, v'(t)\big) - A\big(t, v(t)\big)\Big) X_{v'}(t)\rangle \\
& + \langle Y_v(t), B\big(t, v'(t)\big) - B\big(t, v(t)\big)\rangle \\
& + F\big(t, v'(t), X_{v'}(t)\big) - F\big(t, v(t), X_{v'}(t)\big), \qquad (4.4)
\end{aligned}$$

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 precise definition of $\Delta(v, v')(t)$ in Section 4.3.

A more general formulation can be obtained if we suppose that the backward propagation of the adjoint state is performed with intermediate field $\widetilde{v}$ [27], 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 sake of simplicity, we keep the previous notation. We then obtain the following lemma.

LEMMA 4.2. *For any $v', \widetilde{v}, v : [0,T] \to E$,*

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

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 $\widetilde{v} - v$.

4.2. The algorithms. The factorization obtained in the previous lemmas brings to light various arguments 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 4.1.

ALGORITHM 1. *(Monotonic algorithm)*
Given an initial control v^0, the sequence $(v^k)_{k\in\mathbf{N}}$ is computed iteratively by:
1. *Compute the solution X_{v^k} of (2.1–2.2) with $v = v^k$.*
2. *Compute the solution Y_{v^k} of (4.1–4.2) with $v = v^k$, starting from*

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

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

$$\Delta(v^{k+1}, v^k)(t) \cdot_E \Big(v^{k+1}(t) - v^k(t)\Big) \leq 0. \tag{4.5}$$

Lemma 4.1 then guarantees that $J(v^{k+1}) \leq J(v^k)$. Many strategies can be used to ensure (4.5). Its importance stems from the fact that no further optimization is necessary once this condition is fulfilled. In order to guarantee (4.5), many authors [27, 47, 53] consider an update formula of the form:

$$v^{k+1}(t) - v^k(t) = -\frac{1}{\theta}\Delta(v^{k+1}, v^k)(t), \tag{4.6}$$

where θ is a positive number, that can also depend on k and t. In this case, the variations in J are such that:

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

Note that (4.6) reads as an intermediate update formula between a gradient method:

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

and the proximal algorithm [5], which prescribes:

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

REMARK 5. *In the case $F = 0$ and A independent of v, i.e. linear control with final objective,* (4.6) *coincides with a gradient method.*

4.3. Wellposedness of the algorithm. In this section, we focus on the procedure obtained when using Algorithm 1 with the update formula (4.6).

LEMMA 4.3. *Suppose that for any $t \in [0, T]$:*

- $\mathcal{A} : \mathbf{R} \times \mathbf{V} \times \mathbf{V} \times E \to \mathbf{R}$, $\mathcal{A}(t, X, Y, v) = \langle Y, A(t, v)X\rangle$ is Fréchet differentiable everywhere with respect to v for any X, Y, v.

- $\mathcal{B} : \mathbf{R} \times \mathbf{V} \times E \to \mathbf{R}$, $\mathcal{B}(t, Y, v) = \langle Y, B(t, v)\rangle$ 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 $v, v' \in E$

$$\Delta(v', v; t, X, Y) \cdot_E \Big(v' - v\Big) = -\Big\langle Y, \Big(A(t, v') - A(t, v)\Big)X + B(t, v') - B(t, v)\Big\rangle \\ +F(t, v', X) - F(t, v, X). \tag{4.7}$$

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:

$$\Delta(v', v; t, X, Y) = \int_0^1 -\nabla_w\Big(\langle Y, A(t, w)X - B(t, w)\rangle\Big)\Big|_{w=v+\lambda(v'-v)} \\ + \nabla_v F(t, v + \lambda(v' - v), X) d\lambda. \tag{4.8}$$

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. (4.7) 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

$$\Delta_\Xi(v', v) := \frac{\Xi(v') - \Xi(v)}{\|v' - v\|^2}(v' - v) \in E. \tag{4.9}$$

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

Finally, Eq. (4.8) is an application of the identity

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

$\square$

LEMMA 4.4. *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 $\varepsilon > 0$, $(t, v, X, Y) \in \mathbf{R} \times E \times \mathbf{V} \times \mathbf{V}$ *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 an unique solution* $v' = \mathcal{V}_\theta(t, v, X, Y) \in E$.
2. $\mathcal{V}_\theta(t, v, X, Y) = v$ *implies*

$$-\nabla_v \Big(\langle Y, A(t, v)X \rangle \Big)(v) + \nabla_v \Big(\langle Y, B(t, v) \rangle \Big)(v) + \nabla_v F(t, v, X) = 0. \quad (4.10)$$

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* ε.

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 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 contrativity of $\mathcal{G}$.
2. If $v' = v$ then $h = 0$ thus $\mathcal{G}(h) = 0$ which gives (4.10) after using (4.8).
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 $\|h\| - \|\mathcal{G}(0)\| \leq \|h - \mathcal{G}(0)\| = \|\mathcal{G}(h) - \mathcal{G}(0)\| \leq \frac{1}{2}\|h\|$, which amounts to $\|h\| \leq 2\|\mathcal{G}(0)\|$. Next, we note that

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

   and the estimates follows.
4. Formula (4.8) 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 $h_n := \mathcal{V}_\theta(t_n, v_n, X_n, Y_n)$ and $h := \mathcal{V}_\theta(t, v, X, Y)$.

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

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

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

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

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 6. *Note that $\theta^\star$ 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 (4.5) is guaranteed.

Theorem 1. *Suppose that A, B, F satisfy hypothesis of Lemma 4.4. Also suppose that the operators A, B are such that Eqs. (2.1–2.2) and (4.1-4.2) 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 $\theta^\star > 0$ such that for any $\theta > \theta^\star$, the (nonlinear) equation*

$$\partial_t X_{v'}(t) + A(t, v')X_{v'}(t) = B(t, v') \tag{4.11}$$
$$v'(t) = \mathcal{V}_\theta(t, v(t), X_{v'}(t), Y_v(t)) \tag{4.12}$$
$$X_{v'}(0) = X_0 \tag{4.13}$$

has a solution. Here Y_v is the adjoint state defined by (4.1–4.2) and corresponding to control v.

2. *There exists a sequence $(\theta_k)_{k\in\mathbf{N}}$ such that the algorithm (cf Section 4.2)*
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. Most of the proof is already contained in the previous lemmas. The part that still has to be proven is the existence of a solution to (4.11)-(4.13).

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

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

Since the correspondence $v \mapsto X_v$ is continuous, it follows that the set of solutions $S_{v,R} := \{X_w; w \in B_v(R)\}$ of (2.1) is bounded. In particular for $w = v_l$ by the item 3 of Lemma 4.4 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 $l \ge 1$.

Since $S_{v,R}$ is bounded, recall that by item 5 of Lemma 4.4 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 a solution of (4.11–4.12).□

4.4. Applications. We illustrate here how the examples in Sections 3.1 and 3.2 fit into the setting of the 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.

4.4.1. Example 3.1. 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.

• $E = \mathbf{R}$, $\mathcal{H} = L^2(\mathbf{R}^d; \mathbf{C})$, $\mathbf{V} = dom(H_0^{1/2})$ (over $\mathbf{C}$), or their realifications $\mathcal{H} = L^2 \times L^2$, $\mathbf{V} = dom(H_0^{1/2}) \times dom(H_0^{1/2})$ (over $\mathbf{R}$) see [13];

• $B(t,v) = 0$.

• $F(t,v,X) = \alpha(t)v(t)^2$ with $\alpha(t) \in L^\infty(\mathbf{R})$; 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., [26, 27]) $2 - 2Re\langle X_{target}, X(T)\rangle$ or $-\langle X(T), OX(T)\rangle$ where O is a positive semi-definite operator; both are concave in X.

• Here

$$\Delta(v', v; t, X, Y) = -Re\langle Y, i\mu X\rangle + \alpha(t)(v' + v) \tag{4.14}$$

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

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

$$i\partial_t X^{k+1}(t) = \left(H_0 + \frac{(\theta - \alpha(t))v^k + Re\langle Y_{v^k}, i\mu X^{k+1}\rangle}{\theta + \alpha(t)}\mu \right) X^{k+1}(t) \tag{4.15}$$

Then

$$v^{k+1} = \frac{(\theta - \alpha(t))v^k + Re\langle Y_{v^k}, i\mu X^{k+1}\rangle}{\theta + \alpha(t)},\ \ X_{v^{k+1}} = X^{k+1}. \tag{4.16}$$

4.4.2. Example 3.2. We have

- $E = W^{1,\infty}(0,1)$, $\mathcal{H} = L^2(0,1)$, $\mathbf{V} = H^1(0,1)$ see [17] and [8] (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 (3.5) $F(t,v,X) = \int_Q p(t)(1-\beta z)X(t,z) + \frac{c_0 \cdot z \cdot X(t,z)}{c_1 + c_2 X(t,z)} + \frac{v(t,z)^2}{2} X(t,z)dz$; 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

$$\Delta(v', v; t, X, Y) = \nabla Y + \frac{v' + v}{2} \tag{4.17}$$

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 iteraction, Theorem 1 guarantees the existence of the solution X^{k+1} of the following nonlinear evolution equation:

$$\partial_t X^{k+1}(t) - \nu\Delta X^{k+1} + div\Big(\frac{(\theta - 1/2)v^k - \nabla Y_{v^k}}{\theta + 1/2} X^{k+1}\Big) = 0. \tag{4.18}$$

Then

$$v^{k+1} = \frac{(\theta - 1/2)v^k - \nabla Y_{v^k}}{\theta + 1/2}, \ X_{v^{k+1}} = X^{k+1}. \tag{4.19}$$

4.4.3. Additional application. To illustrate the application of the methodology for a more nonlinear vectorial case, as a third example we consider a situation from [10, 48] which differs from that of Section 3.1 in that $v(t) = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in E = \mathbf{R}^2$ 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 $\xi_1 = -Re\langle Y, i\mu X\rangle + \alpha(t)$, $\xi_2 = -Re\langle Y, i\mu_2 X\rangle$ we obtain

$$\Delta(v', v; t, X, Y) = \xi_1 \begin{pmatrix} v_1 + v_1' \\ v_2 + v_2' \end{pmatrix} + \xi_2 \begin{pmatrix} (v_1 + v_1')v_2' \\ (v_1)^2 \end{pmatrix} \tag{4.20}$$

and the equation in v': $\Delta(v', v; t, X, Y) = -\theta(v' - v)$ has, for θ large enough, an unique solution $v' = \mathcal{V}_\theta(t, v, X, Y) = \begin{pmatrix} \frac{(\theta - \xi_1)v_2 - \xi_2 v_1^2}{\theta + \xi_1} \\ -\frac{\theta - \xi_1 + \xi_2 \frac{(\theta - \xi_1)v_2 - \xi_2 v_1^2}{\theta + \xi_1}}{\theta + \xi_1 + \xi_2 \frac{(\theta - \xi_1)v_2 - \xi_2 v_1^2}{\theta + \xi_1}} v_1 \end{pmatrix}$. We leave as an exercice to the reader the writing of the equation for X^{k+1} and the formula for v^{k+1}.

5. Extension of the monotonic algorithms. In this section, we discuss the relaxation of some of the assumptions concerning either the concavity of (parts of) J or the linearity in Eq. (2.1) (cf. Remark 1).

5.1. Relaxation of concavity assumptions for norm preserving evolution. In some cases, Eq. 2.1 is endowed with additional properties that enable to relax the hypothesis of concavity of the cost functional J. For instance, in Section 3.1 the L^2 norm of X is preserved. Thus, for any G whose second differential with respect to $X \in L^2$ is bounded (e.g. by M), our algorithm applies : in this case use $G - M \cdot Id$ instead of G (see e.g., [40]). The same conclusions also hold for F.

5.2. General evolution equation. We consider a general form of the semi-group generator

$$\partial_t X_v + L(t, v(t), X_v(t)) = 0$$
$$X_v(0) = X_0.$$

For a given v the corresponding adjoint state Y_v is:

$$\partial_t Y_v - D_X L^*(t, v(t), X_v(t))Y_v + \nabla_X F(t, v(t), X_v) = 0$$
$$Y_v(T) = \nabla_X G\big(X_v(T)\big).$$

In the case of the cost functional J defined in (2.4), the arguments of the proof of Lemma 4.1 apply, and we obtain the following result.

LEMMA 5.1. *For any $v', v : [0, T] \to E$,*

$$J(v') - J(v) \leq \int_0^T \mathcal{D}(v, v', t)dt$$

where

$$\mathcal{D}(v, v', t) = F(t, v'(t), X_{v'}) - F(t, v(t), X_{v'})$$
$$+\langle Y_v(t), L(t, v(t), X_v(t)) - L(t, v'(t), X_{v'}(t))\rangle$$
$$+\langle Y_v(t), D_X L(t, v(t), X_v(t))(X_{v'}(t) - X_v(t))\rangle.$$

We note however that choosing at time t, $v'(t) = v(t)$ does not ensure in general that $\mathcal{D}(v, v', t)$ is zero; thus the factorization of the form $\mathcal{D}(v, v', t) = \Delta^{NL}(v, v') \cdot_E (v' - v)$ is not true any more. In particular we are not sure to be able to find a $v'(t)$ which sets this term negative. Manifestly the reason is that the adjoint is not adapted; we do not want to develop here on how to change the adjoint but we are lead to propose the following procedure: advance in time $v'(t)$ by solving for $v'(t)$ in the relation $\mathcal{D}(v, v', t) = -\theta(v'(t) - v(t))^2$ for as long as possible, say from $t_1 = 0$ to $t_2 \leq T$. Then one sets $v \leftarrow 1_{[0,t_2[}v'(t) + 1_{[t_2,T]}v(t)$, compute a new adjoint Y_v and advance again in time from t_2 to t_3, etc.

6. About the convergence of the schemes. The convergence of the sequence given by Algorithm 1 when using (4.6) has been obtained in the case of quantum control (see Section 3.1) using Łojasiewicz-Simon inequality (see [6, 15, 24, 43] and the references therein) in discrete and continuous settings in [2, 38]. 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, as is the case, e.g. in (3.4).
Note also that another proof has been presented in the framework of semi-group theory [14] using compactness arguments.

REFERENCES

[1] A. Bartana and R. Kosloff, *Laser cooling and internal degrees of freedom. II.*, J. Chem. Phys., 106 (1997), pp. 1435–1448.
[2] L. Baudouin and J. Salomon, *Constructive solution of a bilinear control problem*, Syst. Cont. Lett., 57 (2008), pp. 453–464.

[3] J.-D. Benamou and Y. Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math., 84 (2000), pp. 375–393.

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

[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] J. Bolte, A. Daniilidis, O. Ley, and L. Mazet, *Characterization of Lojasiewicz inequalities and applications*, arXiv:0802.0826, (2008).

[7] A. 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 (2009), p. 056704.

[8] R. Dautray and J.-L. 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.

[9] P. Ditz and A. Borzí, *A cascadic monotonic time-discretized algorithm for finite-level quantum control computation*, Computer Physics Communications, 178 (2008), pp. 393–399.

[10] B. Friedrich and D. Herschbach, Phys. Rev. Lett., 74 (1995), p. 4623.

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

[12] 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 (2008), p. 073002.

[13] K. Ito and K. Kunisch, *Optimal bilinear control of an abstract Schrödinger equation*, SIAM J. Control Optim., 46 (2007), pp. 274–287 (electronic).

[14] ———, *Optimal bilinear control of an abstract Schrödinger equation*, SIAM Journal on Control and Optimization, 46 (2007), pp. 274–287.

[15] M. A. Jendoubi, *A simple unified approach to some convergence theorems of L. Simon*, J. Func. Anal., 153 (1998), pp. 187–202.

[16] C. P. Koch, J. P. Palao, R. Kosloff, and F. Masnou-Seeuws, *Stabilization of ultracold molecules using optimal control theory*, Physical Review A (Atomic, Molecular, and Optical Physics), 70 (2004), p. 013402.

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

[18] 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).

[19] J.-M. Lasry and P.-L. Lions, *Jeux à champ moyen. I. Le cas stationnaire*, C. R. Math. Acad. Sci. Paris, 343 (2006), pp. 619–625.

[20] ———, *Jeux à champ moyen. II. Horizon fini et contrôle optimal*, C. R. Math. Acad. Sci. Paris, 343 (2006), pp. 679–684.

[21] ———, *Mean field games*, Jpn. J. Math., 2 (2007), pp. 229–260.

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

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

[24] S. Łojasiewicz, *Sur la géométrie semi- et sous-analytique*, Ann. Inst. Fourier, 43 (1993), pp. 1575–1595.

[25] Y. Maday, J. Salomon, and G. Turinici, *Monotonic time-discretized schemes in quantum control*, Numerische Mathematik, 103 (2006), pp. 323–338.

[26] ———, *Monotonic parareal control for quantum systems*, SIAM Journal on Numerical Analysis, 45 (2007), pp. 2468–2482.

[27] Y. Maday and G. Turinici, *New formulations of monotonically convergent quantum control algorithms*, J. Chem. Phys., 118 (2003).

[28] M. Mirrahimi, G. Turinici, and P. Rouchon, *Lyapunov control of bilinear Schrödinger equations*, Automatica, 41 (2005), pp. 1987–1994.

[29] ———, *Reference trajectory tracking for locally designed coherent quantum controls*, J. Phys. Chem. A, 109 (2005), pp. 2631–2637.

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

[31] Y. Ohtsuki, G. Turinici, and H. Rabitz, *Generalized monotonically convergent algorithms for solving quantum optimal control problems*, J. Chem. Phys., 120 (2004), pp. 5509–5517. also published in the Virtual Journal of Ultrafast Science.

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

[33] J. P. Palao and R. Kosloff, *Optimal control theory for unitary transformations*, Phys. Rev. A, 68 (2003), p. 062308.

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

[35] H. Rabitz, G. Turinici, and E. Brown, *Control of quantum dynamics: Concepts, procedures and future prospects*, in Computational Chemistry, Special Volume (C. Le Bris Editor) of Handbook of Numerical Analysis, vol X, P. G. Ciarlet, ed., Elsevier Science B.V., 2003, pp. 833–887.

[36] M. Rabitz, R. Hsieh and C. Rosenthal, *Quantum optimally controlled transition landscapes*, Science, 303 (2004), p. 1998.

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

[38] J. Salomon, *Convergence of the time-discretized monotonic schemes*, M2AN, 41 (2007), pp. 77–93.

[39] J. Salomon and G. 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.

[40] J. Salomon, C. Dion, and G. Turinici, *Optimal molecular alignment and orientation through rotational ladder climbing*, J. Chem. Phys., 123 (2005), p. 144310.

[41] J. Salomon and G. Turinici, *On the relationship between the local tracking procedures and monotonic schemes in quantum optimal control*, J. Chem. Phys., 124 (2006), p. 074102.

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

[43] L. Simon, *Asymptotics for a class of non-linear evolution equations, with applications to geometric problems*, Ann. of Math., 118 (1983), pp. 525–571.

[44] S. E. Sklarz and D. 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 (2002), p. 053619.

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

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

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

[48] R. Tehini and D. D. Sugny, Phys. Rev. A, 77 (2008), p. 023407.

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

[50] T. Weinacht, J. Ahn, and P. Bucksbaum, *Controlling the shape of a quantum wavefunction*, Nature, 397 (1999), pp. 233–235.

[51] O. Yukiyoshi, T. Yoshiaki, G. Turinici, and H. Rabitz, *Monotonically convergent algorithms for solving quantum optimal control problems described by an integro-differential equation of motion*, PRA, 75 (2007), p. 033407.

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

[53] 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 (1998), pp. 385–391.

[54] ———, *Uniform rapidly convergent algorithm for quantum optimal control of objectives with a positive semi-definite Hessian matrix*, Phys. Rev. A., 58 (1998), pp. 4741–4748.

Appendix: Time discretized case. This section is devoted to the time-discretization.

Setting. In order to reproduce at the discrete level the computation involved in the monotonic algorithms, one has to define a time discretized version of J and a scheme devoted to numerical resolution of (2.1–2.2).

Note that our optimization method does not impose any restrictions thus any scheme with standard numerical properties (consistency, stability, convergence) is compatible with our procedure.

Since we only deal with optimizations problems, we consider arbitrary time-discretizations of the functional (2.4):

$$J_{\Delta t}(v) = \Delta t \sum_{n=0}^{N-1} F(v_n, x_n) + G(x_N),$$

together with the general numerical scheme

$$x_{n+1} = A_{\Delta t}(v_n) x_n + B_{\Delta t}(v_n), \tag{6.1}$$

where N is a positive integer, $\Delta t = T/N$ and $v = (v_n)_{n=0...N-1}$. We assume that the functions F and G have the same properties as in Section 2.

Discrete adjoint and factorization. As in the continuous case, the adjoint operator definition directly follows from the state equation evolution. Given a numerical solver (6.1), the discrete adjoint operator is defined by :

$$\begin{aligned} y_n &= A^*_{\Delta t}(v_n) y_{n+1} + \Delta t \nabla_x F(v_n, x_n) \\ y_N &= \nabla_x G\big(x_N\big). \end{aligned}$$

With this definition, a factorization similar to the one of Lemma 4.1 can be obtained.

Lemma 6.1. *For any $v' = (v'_n)_{n=0...N-1}$, $v = (v_n)_{n=0...N-1}$,*

$$\begin{aligned} J_{\Delta t}(v') - J_{\Delta t}(v) \leq \sum_{n=0}^{N-1} & \langle y_n, (A_{\Delta t}(v'_{n-1}) - A_{\Delta t}(v_{n-1})) x'_{n-1} \rangle \\ & + \langle y_{n+1}, B_{\Delta t}(v'_n) - B_{\Delta t}(v_n) \rangle \\ & + \Delta t \left(F(v'_n, x'_n) - F(v_n, x'_n) \right). \end{aligned}$$

Thanks to this lemma, we obtain a discrete version of monotonicity condition (4.5). Depending on the way the functions A, B and F depend on v, the computation of a v'_n satisfying the discrete monotonic condition may requires an inner iterative solver. In many cases this computation can anyway be parallelized. During an optimization step, at a given time step n, the terms of the previous sum can be factorized with respect to each component of the vector $v'_n - v_n$ and made negative independently.
The fact that the computation of v'_n requires x'_n makes anyway the time resolution sequential. To solve this problem, some time parallelizations have been designed in the case of quantum control [26].