Quantum optimal control using the adjoint method

Abstract
Control of quantum systems is central in a variety of present and perspective applications ranging from quantum optics and quantum chemistry to semiconductor nanostructures, including the emerging fields of quantum computation and quantum communication. In this paper, a review of recent developments in the field of optimal control of quantum systems is given with a focus on adjoint methods and their numerical implementation. In addition, the issues of exact controllability and optimal control are discussed for finite- and infinite-dimensional quantum systems. Some insight is provided considering ‘two-level’ models. This work is completed with an outlook to future developments.

Keywords
Quantum systems • Schrödinger equation • Optimal control theory • Numerical optimization

PACS: 03.65.-w, 32.80.Gk, 37.90+j, 82.37.Gk, 82.53.-k
MSC: 35Q40, 49K20, 65H10, 65M06, 65N06, 90C53

© Versita sp. z o.o.

1. Introduction
The control of quantum systems is a central problem in the development of nano-devices based on quantum mechanics. Specifically, it applies to quantum optics, quantum chemistry, and semiconductor nanostructures; see, e.g., [7, 22, 24, 31, 43, 65, 66, 83]. In this development, there is the need to manipulate quantum states with the highest possible precision using, e.g., electromagnetic fields, and for this purpose optimal control theory provides a viable and effective tool. These control functions can be employed to break a bond in a molecule, to drive a certain chemical reaction [83], to control of photochemical processes [73], and to manipulate quantum dots [43]. More generally, control may be required to drive state transitions, maximize observable expectation, and obtain best performance of quantum operators; see [21] for a recent general review.

In most cases, the task of optimal control theory is to define fast control mechanisms that cannot be constructed based on perturbation theory or on a priori parametrized control fields. This fact has motivated the increasing interest in the extension of optimal control theory and related computational techniques to the quantum world, resulting in many recent successful results [13–16, 26, 60–62, 68, 79, 82]. A pioneering work in this field was done by Peirce, Dahleh, and Rabitz [64], who investigated the optimal dipole control of a diatomic molecule. The focus in the early papers (see, e.g., [22, 24]) was to validate the ability of the optimal control framework to provide suitable quantum control mechanisms.

On the other hand, the computational difficulties arising from solving quantum control problems with Schrödinger-type equations are manifold. This is partly due to the bilinear control structure of typical quantum control mechanisms and the non-convexity of the optimization problem. For this reason, most research effort in the past has focused on finite-dimensional Schrödinger equations; see, e.g., [14, 16, 61]. In this case, the computational schemes of choice have been the monotonic iterative scheme [34, 61, 65] and accelerated versions of the gradient scheme [14, 15]. These schemes perform well for finite-level quantum systems and provide acceptable results [62, 80] when applied to infinite-dimensional
systems. More recently, second-order Newton schemes and multigrid methods have been presented [13, 81, 82]. However, most of the results available in quantum optimal control refer to small-size closed systems and much remains to do to cover multi-particle and open quantum systems.

In the following section, we discuss finite-level quantum systems and illustrate the problem of exact controllability and optimal control for these systems. In Section 3, we focus on infinite-dimensional systems and address the numerical difficulties that could be encountered in solving optimal control problems governed by Schrödinger-type equations defined in space-time domains. In Section 4, we discuss the modelling and control of open quantum systems and the formulation of stochastic Schrödinger equations. A section of conclusion and outlook completes this work.

2. Finite-Level Quantum Systems

Genuine finite-dimensional quantum control problems arise in the control of spin systems and in nuclear magnetic resonance; see, e.g., [1, 48]. More generally, they arise in molecular dynamics where infinite-dimensional systems interact with external fields with given symmetries and the resulting model can be well described by finite-dimensional models. These models also are fundamental to develop theoretical and numerical tools that could be extended to infinite-dimensional models.

A finite-level quantum system is modelled by a Schrödinger equation for an n-component wave function \( \psi : [0, T] \rightarrow \mathbb{C}^n \) as follows

\[
\dot{\psi}(t) = H(u(t))\psi(t), \quad \psi(0) = \psi_0, \tag{1}
\]

for \( t \in [0, T] \), and \( T > 0 \) is a given terminal time. The Hamiltonian matrix \( H \in \mathbb{C}^{n \times n} \) depends on the external control field \( u : [0, T] \rightarrow \mathbb{C} \) and \( \psi_0 \in \mathbb{C}^n \) is a given initial condition with \( |\psi_0|^2 = 1 \).

We write \( z_0 = \Re(z) \) and \( z_\Im = \Im(z) \) for the real and imaginary part of a complex \( z \in \mathbb{C} \). Moreover, \( z^* \) stands for the complex conjugate of \( z \) and \( |z| = \sqrt{z^*z} \) for its absolute value.

In (1), the Hamiltonian \( H = H_0 + H_c(u) \) has two constitutive components: the constant free Hamiltonian \( H_0 \in \mathbb{C}^{n \times n} \) describing the uncontrolled system and the control Hamiltonian \( H_c \in \mathbb{C}^{n \times n} \) modelling the coupling of the quantum state to an external control field, typically a laser (electromagnetic) field. In particular, a dipole control mechanism can be written as \( H_c(u) = H_d u \), where \( H_d \) is a Hermitian dipole matrix.

The choice of \( T \) is a delicate issue and must be guided by physical considerations. Based on the quantum indeterminacy principle \( \Delta \lambda \Delta t \geq \hbar \) (where \( \hbar \) is the Planck constant that we set equal to one) we can state that for a small time horizon \( T \), a highly energetic \( \Delta \lambda \gg 1 \) optimal control results that involves many energy levels. On the other hand, for large \( T \) additional decoherence channels become important, which makes the control process less efficient; see, e.g., [43].

Notice that the wavefunction description given in (1) is appropriate for an isolated quantum system and in that case the governing Hamiltonian \( H \) is Hermitian. For a more realistic non isolated system with environment couplings and subject to control, a more general density-matrix description would be required. Alternatively, a non-Hermitian Hamiltonian as in [15, 78] can be constructed to account for environment losses.

In the case of \( H \) being Hermitian the evolution of \( \psi \) is unitary. This means that (1) defines an evolution on the unit sphere of \( \mathbb{C}^n \). For ease of illustration, let us assume that the spectrum of \( H_0 \) is non degenerate, with real eigenvalues \( \lambda_1, \ldots, \lambda_n \), called energy levels, and normalized eigenvectors \( \phi_1, \ldots, \phi_n \), called eigenstates, representing an orthonormal basis.

Once the eigenvalues and eigenvectors are known, the time evolution of a free state is computed as follows. Starting from an initial state \( \psi_0 = \sum_{j=1}^n c^{(0)}_j \phi_j \), we obtain the following

\[
\psi(t) = \sum_{j=1}^n c^{(t)}_j \phi_j \quad \text{with} \quad c^{(t)}_j = c^{(0)}_j \exp(-\lambda_j t),
\]

where \( |\psi(t)|^2 = \sum_{j=1}^n |c^{(t)}_j|^2 = 1 \). In this framework, \( |c^{(t)}_j|^2 \) is the probability that making a measurement of energy of the system at time \( t \), we obtain \( \lambda_j \) as its value. We denote with \( \langle \cdot, \cdot \rangle \) the complex scalar product in \( \mathbb{C}^n \) so we have \( c^{(t)}_j = \langle \psi(t), \phi_j \rangle \).

Now, at time \( t = 0 \) consider the system in an \( \phi_i \)-eigenstate, that is \( |\langle \psi(0), \phi_i \rangle|^2 = 1 \). A typical state transition problem is to design control functions \( u_1, \ldots, u_m \) such that at time \( T \) the system is in another prescribed \( \phi_j \)-eigenstate.
such that $|\langle \psi(T), \phi\rangle|^2 = 1$. Here, we have $H(u) = \sum_{k=1}^{n} u_k H_k$. Notice that because of the complex absolute function $|\cdot|$, the value of $|\langle \psi, \phi\rangle|^2 = 1$ is invariant under a phase shift $\psi \rightarrow e^{i\omega} \psi$.

For finite-level quantum systems with $n$ energy levels, the state space is the unit sphere $S^{2n-1} \subset \mathbb{C}^n$. These systems define right-invariant control problems on the Lie group of unitary matrices $U(n)$. If $H(u)$ has zero trace, then the resulting systems naturally evolve in $SU(n)$. In correspondence, we denote the Lie algebras $u(n)$ and $su(n)$. In particular, $su(n)$ denotes the subalgebra of zero trace matrices in $u(n)$, that is the Lie algebra of skew-Hermitian $n \times n$ matrices considered as a Lie algebra over the real field. Notice that $\dim u(n) = n^2$ and $\dim su(n) = n^2 - 1$.

The quantum controllability problem of proving that for every couple of points in $U(n)$ or $SU(n)$ one can find controls steering the system from one point to the other is well understood. For this purpose, one considers the following Lie subalgebra

$$L_0 = \text{Lie}\{iH_0, iH_1, \ldots, iH_n\}.$$ 

This represents the smallest linear subspace containing $iH_0, iH_1, \ldots, iH_n$, and their iterated commutators.

Now, we remark that the solution to (1) can be written in the following form

$$\psi(t) = X(t) \psi_0, \quad \dot{X} = i \left( H_0 + \sum_{k=1}^{n} u_k H_k \right) X,$$

where $X(0) = I$.

We have the following theorem [31, 36, 45, 67, 72, 74].

**Theorem 1.**

A necessary and sufficient condition for complete controllability of (2) on $U(n)$, is that the Lie algebra $L_0 \subset u(n)$ has dimension $n^2$. Further, the system (2) is completely controllable on $SU(n)$ if and only if $L_0 \subset su(n)$ has at least dimension $n^2 - 1$. Complete controllability of (2) implies controllability of (1) on $S^{2n-1}$.

For the purpose of illustration, consider the simplest but fundamental two-level system; see, e.g., [83]. This model consists of two orthonormal states $|0\rangle$ and $|1\rangle$ and the state vector at time $t$ is given as follows

$$|\psi(t)\rangle = c_0(t) |0\rangle + c_1(t) |1\rangle =: \begin{pmatrix} c_0(t) \\ c_1(t) \end{pmatrix}.$$ 

The time evolution of the two-level state represented on the space spanned by $|0\rangle$ and $|1\rangle$ is governed by the following Hamiltonian

$$H(u(t)) = \begin{pmatrix} \omega_0 & 0 \\ 0 & \omega_1 \end{pmatrix} + u(t) \begin{pmatrix} 0 & \mu \\ \mu & 0 \end{pmatrix}.$$ 

To determine controllability, we have to construct the Lie algebra of the skew-Hermitian operators $iH_0$ and $iH_1$ of our system. In particular, we have to show that the dimension of this algebra is $n^2 = 4$, or at least $n^2 - 1$. For this purpose, we construct a matrix $W$ which represents a basis of the Lie algebra $L_0$. This calculation becomes simpler by rewriting the Hamiltonian

$$\hat{H} = \frac{1}{2}(\omega_0 + \omega_1) I + (\omega_0 - \omega_1) \sigma_z + u(t) \mu \sigma_x,$$

where $\sigma_x, \sigma_y,$ and $\sigma_z$ are the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$ 

The columns of $W$ are calculated by transforming the $n \times n$ matrix $\hat{H}_k, k = 0, 1$, into a vector of $n^2$ components by appending the every next column of $\hat{H}$ to the end of the previous one. Thus we obtain the first two columns of $W$. The remaining columns are calculated in the following way. Calculate the commutator of $\hat{H}_0$ and $\hat{H}_1$ and add this column.
to \( W \) if it is linearly independent of the existing columns. After appending the new column, all commutators with the previous columns have to be evaluated and added if linearly independent. This process is repeated until the matrix is full or no new linearly independent columns can be found. We obtain the following

\[
W = \begin{pmatrix}
\omega_0 & 0 & 0 & 2u^2\mu^2\omega_{10} \\
0 & \omega_0 & u\mu & 0 \\
0 & u\mu & -\omega_0u\mu & 0 \\
\omega_1 & 0 & 0 & -2u^2\mu^2\omega_{10}
\end{pmatrix}
\]

where \( \omega_{10} = \omega_1 - \omega_0 \).

We see that \( \text{rank}(W) = 4 \), if \( \omega_1 \neq -\omega_0 \), (in which case \( iH_0 \notin su(n) \)) and therefore Theorem 1 states that the system is completely controllable. On the other hand, in the case \( \omega_1 = -\omega_0 \) we have \( iH_0 \in su(n) \) and \( \text{rank}(W) = 3 \) and the system is again completely controllable for the population of levels. However, if we consider (1), the control over the phase is lost, which is not a physical limitation in the case of population transfer.

We remark that the verification of the above conditions is not always straightforward, and alternative necessary and sufficient conditions for controllability have been formulated; see, e.g., [29, 65, 72]. Specifically, we refer to the controllability analysis via connectivity graphs, requiring the following no degenerate energy-transition condition

\[
|\lambda_i - \lambda_j| \neq |\lambda_k - \lambda_l|, \quad i, j, k, l = 1, \ldots, n, i \neq j, k \neq l, \{i, j\} \neq \{k, l\}.
\]

In this analysis, one considers the connectivity graph of the control Hamiltonian, with elements \( h_{ij}, i, j = 1, \ldots, n \). Two state vertices \( i \) and \( j \) are said to be connected by a path if there exists a connected set of edges \( E = \{(k, l), h_{kl} \neq 0\} \) starting in \( i \) and ending in \( j \). The graph is called connected if there exists a path between every pair of vertices. With this setting, we can state the following theorem; see [65].

**Theorem 2.**

Assume that the graph of the control Hamiltonian is connected and that the transitions of the internal Hamiltonian are non-degenerate. Then the quantum system defined by the Hamiltonians \( H_0 \) and \( H_i \) is controllable.

We remark that the issue of controllability of quantum system is very important. However, results on controllability usually do not deliver a control function and in most cases do not refer to a specific time horizon. Moreover, the theory on controllability is less developed for quantum systems modelling dissipation and other types of coupling with the environment. On the other hand, consideration of quantum systems with dissipation and subject to measurement opens new perspectives for controlling quantum models; see, in particular, [4, 78, 85].

On the other hand, in application we need to construct an optimal control that usually is required to perform its task in a short time interval and could accommodate environment interactions. For this purpose, in the following we focus on optimal control problems where it is required to reach a given target state as close as possible while minimizing some appropriate control costs.

We focus on the problem of determining an optimal control field \( u \in L^2(0, T; \mathbb{C}) \), such that (1) is fulfilled and a number of optimality criteria are met. The most frequent objectives encountered in quantum control problems require that the control sequence drives the system at a time \( T \) close to a desired target configuration \( \psi_d \in \mathbb{C}^n \). In addition, limited laser resources and smoothness of the shape of the laser field are accounted for through a minimization of appropriate norms of the control function. Furthermore, one can require suppressing population of intermediate states which suffer strong environment losses. All these requirements are implemented in a cost functional with the following structure

\[
J(\psi, u) = \frac{1}{2} ||\psi(T) - \psi_d||^2_{L^2} + \frac{\gamma}{2} ||u||^2_{L^2(0, T; \mathbb{C})} + \frac{\mu}{2} ||\dot{u}||^2_{L^2(0, T; \mathbb{C})} + \frac{1}{2} \sum_{j \in J} c_j ||\psi_j||^2_{L^2(0, T; \mathbb{C})}
\]

where the constants \( \gamma, \mu \geq 0, \gamma + \mu > 0 \), are parameters that allow us to vary the relative importance of the objectives represented by the various terms. The goal of the first term of the cost functional is to track the state \( \psi \) close to a given
terminal state at \( t = T \). The second and third terms are penalty terms that also ensure existence of at least one optimal control. In the last term of (3), which penalizes the occupation of certain states \( \psi_j \), the set \( I \subset \{1, \ldots, n\} \) denotes a subset of possible state indices and \( \sigma_j \geq 0, j \in I \), are weighting factors.

In [14, 15] a quantum control problem with (1)–(3) is formulated as follows

\[
\min J(\psi, u), \quad (\psi, u) \in X \text{ subject to } (1).
\]

Depending on the parameter \( \mu \), we define the Hilbert space \( X \) by

\[
X = \begin{cases} 
H^1(0, T; \mathbb{C}^n) \times L^2(0, T; \mathbb{C}) & \text{if } \mu = 0, \\
H^1(0, T; \mathbb{C}^n) \times H^1_0(0, T; \mathbb{C}) & \text{if } \mu > 0
\end{cases}
\]

and supply \( X \) with the natural product topology. We remark that the choice \( H^1_0 \) provides a natural way to require that the control is zero at the beginning and the end of the control horizon.

Notice that the quadratic cost functional \( J : X \to [0, \infty) \) is twice continuously Fréchet-differentiable. Using the fact that the Schrödinger operator equation is Fréchet-differentiable and its linearization is surjective in \( X \), first-order necessary conditions for a minimum are obtained in the adjoint method by equating to zero the Fréchet derivatives of \( L \) with respect to \((\psi, u, p)\), where \( L \) is the following Lagrange function

\[
L(\psi, u, p) = J(\psi, u) + \mathbb{R} \int_0^T \left( i\dot{\psi} - H(u) \psi \right) p^* \, dt
\]

and \( p \) is the adjoint variable, also called Lagrange multiplier. Further, we denote \( L_\psi, L_u, \) and \( L_p \) the partial derivatives of \( L \) with respect to \( \psi, u, \) and \( p \), respectively. Denote \( \nabla L(\psi, u, p) = (L_\psi, L_u, L_p) \). We have the following theorem [14]. (An equivalent formulation of this theorem could be given based on Pontryagin's maximum principle; see [38].)

**Theorem 3.**
Suppose that \((\psi, u) \in X \) is a local solution to (4). Then there exists a (unique) Lagrange multiplier \( p \in H^1(0, T; \mathbb{C}^n) \) satisfying

\[
\begin{align*}
\dot{i}\psi &= H(u(\cdot))\psi \quad & \text{in } (0, T], \\
\psi(0) &= \phi_0, \\
i(\dot{p})_j &= \left( H(u(\cdot))^*p \right)_j - \sigma_j(\psi)_j \\ & \quad \text{in } (0, T], j \in I, \\
i(\dot{p})_j &= \left( H(u(\cdot))^*p \right)_j \\ & \quad \text{in } (0, T], j \notin I, \\
ip(T) &= \psi(T) - \psi_u, \\
-\mu \dot{u} + \gamma u &= \Re e \left( H_{1R} \dot{\psi} \cdot p^* \right) + \Re e \left( H_{1I} \dot{\psi} \cdot p^* \right) \\
u(T) &= u(0) = 0.
\end{align*}
\]

in case of \( \mu > 0 \). Moreover, \( u \in C^2([0, T]; \mathbb{C}) \cap C([0, T]; \mathbb{C}) \), i.e., \( u \) is a classical solution. If \( \mu = 0 \) holds, (5f) and (5g) have to be replaced by

\[
\begin{align*}
\gamma u &= \Re e \left( H_{1R} \dot{\psi} \cdot p^* \right) + \Re e \left( H_{1I} \dot{\psi} \cdot p^* \right) \quad \text{in } (0, T), \\
\end{align*}
\]

where

\[
H_i(z) = z_{\alpha_i} H_{1R} + z_{\beta_i} H_{1I} \quad \text{for } z = z_{\alpha_i} + iz_{\beta_i} \in \mathbb{C},
\]

Optimality systems as (5) can be discretized by implicit second-order Crank-Nicholson schemes and solved using nonlinear optimization methods like the cascading nonlinear CG scheme and the monotonic scheme [14, 60], and the GRAPE algorithm [49]. Numerical results show that this optimization problem has very flat minima and the resulting optimal control is sensitive to the tolerance on the norm of the reduced gradient. In the following, we discuss a specific
Table 1. Optimization results depending on optimization parameters.

| γ   | μ   | α   | |ψ(T) − ψd| | J     |
|-----|-----|-----|-----------------|-----------------|
| 10^{-7} | 10^{-7} | 0.05 | 8.6 · 10^{-4} | 2.37 · 10^{-3} |
| 10^{-7} | 10^{-9} | 0.05 | 3.7 · 10^{-4} | 5.46 · 10^{-4} |
| 10^{-7} | 0     | 0.05 | 6.9 · 10^{-5} | 1.41 · 10^{-4} |
| 10^{-7} | 0     | 0    | 1.2 · 10^{-3} | 2.33 · 10^{-6} |
| 10^{-4} | 10^{-4} | 0.05 | 3.3 · 10^{-2} | 6.52 · 10^{-2} |
| 10^{-4} | 10^{-6} | 0.05 | 4.4 · 10^{-3} | 9.03 · 10^{-3} |
| 10^{-4} | 0     | 0.05 | 2.7 · 10^{-3} | 5.68 · 10^{-3} |
| 10^{-4} | 0     | 0    | 8.3 · 10^{-3} | 3.34 · 10^{-4} |

example that extends the two-level model illustrated above to accommodate environmental coupling. See [14] for all details.

Consider a three-level quantum system whose configuration is represented by \( \psi = (\psi_1, \psi_2, \psi_3) \in H^1(0, T; C^3) \) which consists of two long-lived states \( \psi_1 \) and \( \psi_2 \), which are energetically separated by some amount \( \omega \), and a state \( \psi_3 \), which has a finite lifetime because of environment coupling. Such \( \Lambda \)-type configurations have a long-standing history in quantum optics, and more recently, similar configurations have received increasing attention also in semiconductor quantum dots; see, e.g., [43] and the references given therein.

Time evolution of this finite-level quantum system is governed by the following non-Hermitian Hamiltonian

\[
H_0 = \frac{1}{2} \begin{pmatrix}
-\omega & 0 & 0 \\
0 & \omega & 0 \\
0 & 0 & -i\Gamma_0
\end{pmatrix},
\]

(8)

where the term \(-i\Gamma_0\) accounts for environmental losses (e.g., spontaneous photon emissions). The external control field is assumed complex, \( u(t) \in C \), which corresponds to two real control fields. The coupling of this field to the system is modelled as follows

\[
H_c(u) = -\frac{1}{2} \begin{pmatrix}
0 & 0 & \mu_1 u \\
0 & 0 & \mu_2 u \\
\mu_1 u^* & \mu_2 u^* & 0
\end{pmatrix},
\]

(9)

where \( \mu_1 \) and \( \mu_2 \) describe the coupling strengths of states \( \psi_1 \) and \( \psi_2 \) to the interconnecting state \( \psi_3 \) (e.g., optical dipole matrix elements). Typical initial and final states are given by

\[
\psi_0 = \begin{pmatrix}
1 \\
0 \\
0
\end{pmatrix} \quad \text{and} \quad \psi_d = \begin{pmatrix}
0 \\
1 \\
0
\end{pmatrix},
\]

respectively.

It is useful to discuss the effect of different choices of values of the optimization parameters using the results reported in Table 1. These results are obtained with a cascadic nonlinear conjugate gradient (NCG) scheme; see [14].

In Table 1, we see that smaller values of \( |\psi(T) - \psi_d| \) are attained for smaller \( \gamma \). We remark that \( \gamma = 10^{-7} \) is quite small, and that makes the problem quite stiff and ill-conditioned. We also can see the effect of the regularization parameter \( \mu \). As \( \mu \) increases, \( |\psi(T) - \psi_d| \) increases, demonstrating that the additional smoothness of the control function (slightly) reduces the capability of tracking. Apparently, larger \( \mu \) makes the problem behaving better, resulting in less computational effort. Concerning the parameter \( \alpha_0 = \alpha \) we obtain better tracking for nonzero \( \alpha \). This is expected since we have \( |\psi_d| = 1 \), whereas \( |\psi(T)| < 1 \) whenever \( \gamma_0 > 0 \), because of dissipation. By taking \( \alpha > 0 \) dissipation is reduced and therefore better \( |\psi(T) - \psi_d| \) is possibly achieved.

It appears clearly that the adjoint approach provides a powerful tool to prove existence of quantum controls and to compute these controls by solving the corresponding optimality system. Assuming that the desired target state is
attainable, then it is reasonable to consider the solution of the optimal control as $\gamma, \mu \to 0$. If this procedure converges to an optimal control, then it is possible to show that this control solves the following exact controllability problem

$$\min \left( \frac{\mathcal{V}}{2} \left\| u \right\|^2_{L^2(\Omega, T, C)} + \frac{\mu}{2} \left\| \dot{u} \right\|^2_{L^2(\Omega, T, C)} \right), \quad (\psi, u) \in X \text{ subject to } (1), \quad \text{with } \psi(0) = \psi_0, \psi(T) = \psi_T. \quad (10)$$

Notice that using the Hilbert uniqueness method proposed in [59], it is possible to show that the solution of (10) is unique; see also [50].

### 3. Infinite-Dimensional Quantum Systems

In this section, we discuss infinite-dimensional quantum systems, that is, systems with an unbounded energy spectrum, as it usually occurs with space-time quantum models. For these models, much less is known on their controllability properties; see [8, 11, 16, 27, 47, 50, 74, 79] and the references given there. On the other hand, in this context ‘negative results’, stating that full control in the infinite-dimensional Hilbert sphere cannot be achieved with a finite-dimensional control Lie algebra, can be found; see [47] for more details. In [75], it is shown that exact controllability does not hold in infinite-dimensional quantum systems, and in [46] it is proven that the Hartree equation with bilinear control is not controllable in finite or infinite time. Positive results can be found in [74], where it is shown that the state of an infinite-dimensional quantum system could be controlled only within a dense subspace of the relevant Hilbert space; see [12, 54, 77, 79] for additional references and additional results.

Also concerning the adjoint methodology, optimization with infinite-dimensional quantum systems is one of the most challenging research fields in computational sciences. In fact, the complexity of multi-particle quantum systems outmatches our present computational capabilities for simulating physical processes at nanoscales; see, e.g., [70].

A representative model to start with the investigation of optimization problems of infinite-dimensional quantum systems is the control problem of one-particle state transition in a confining potential. We have $H_0 = -\Delta + V_0$, where the minus Laplacian represents the kinetic energy operator of the particle and $V_0$ is the confining potential.

In quantum mechanics, the quantum state of a particle is described by a wavefunction $\psi : \Omega \times (0, T) \to C$ that is governed by the following time-dependent Schrödinger equation (TDSE) [25]

$$i \partial_t \psi(x, t) = \{-\Delta + V(x, u(t))\} \psi(x, t), \quad (11)$$

where we choose the scaling of the Planck constant $\hbar = 1$ and the mass $m = 1/2$. We consider the potential $V(x, u(t))$ consisting of a stationary part $V_0(x)$ and a time varying control part.

Notice that quantum systems are usually defined in unbounded domains while the presence of a confinement potential results in wavefunctions $\psi$ whose support is localized in a bounded region. Therefore, with $\Omega$ we represent a spatial domain that is large enough to represent $\psi$ during evolution. In this case, periodic or homogeneous boundary conditions may be appropriate; see [5] for more sophisticated boundary conditions.

Another representative infinite-dimensional quantum system discussed in this section is the Bose–Einstein condensate (BEC) model that describes the state of matter formed by a cloud of bosons (e.g., helium-4, rubidium) cooled to temperatures very near to absolute zero. Under such conditions, the atoms with magnetic spin collapse into the lowest quantum state sharing the same wavefunction. This wavefunction describes the mean-field dynamics of a coherent BEC. It is modelled by the following Gross–Pitaevskii equation [17, 32]

$$i \partial_t \psi(x, t) = \{-\Delta + V(x, u(t)) + g |\psi(x, t)|^2\} \psi(x, t), \quad (12)$$

where $x \in \Omega$ and $t \in (0, T)$, with $g$ a coupling constant related to the scattering length of the atoms, density, and transversal confinement. In (12), the confining potential $V(x, u(t))$ is produced by magnetic microtraps whose variation is described by a control function $u : (0, T) \to \mathbb{R}$.

Among models with higher complexity, we find multi-particle Schrödinger-type equations that allow us to describe quantum systems like molecules. In particular, we mention the time-dependent Kohn–Sham equation [39, 51] and the $n$-particle Schrödinger equation; see, e.g., [86] and the references given therein. For these two classes of systems, optimization and control are topics of future research efforts.

In the following, we discuss in detail the solution of optimal control problems governed by (11) and (12). By doing this, we address some numerical optimization issues which arise in quantum control problems.
3.1. A Dipole Quantum Control Problem

In quantum mechanics, a dynamically stable system with an atom exists in the presence of a stationary confining potential with a "well" envelope [25]. The states of this system are defined based on the following eigenproblem

\[ \{-\Delta + V_0(x) - \lambda_j\} \phi_j(x) = 0, \quad j = 1, 2, \ldots \]  

(13)

The eigenfunctions \( \phi_j \) represent the stationary states and the eigenvalues \( \lambda_j \) represent the corresponding energy. Like in the finite-dimensional case, the time evolution of these states is formally given by \( \psi_j(x, t) = \phi_j(x) \exp(-i\lambda_j t) \). A representative stationary potential with various applications in semiconductor nanostructures is the following infinite barrier well potential

\[ V_0(x) = 0 \quad \text{for} \ x \in \Omega \quad \text{and} \quad V_0(\pm \ell/2) = +\infty, \quad \Omega = (-\ell/2, \ell/2). \]

The infinite barrier condition is equivalent to homogeneous Dirichlet boundary conditions for the wavefunction, and thus we have \( \lambda_j = \frac{n^2\pi^2}{\ell^2} \) and \( \phi_j(x) = \cos(j\pi x/\ell) \).

To discuss some mathematical properties of the present framework, we give some definitions. We define \( \mathcal{H} = L^2(\Omega; \mathbb{C}) \), the Hilbert space endowed with the inner product

\[ (\varphi, \psi)_{\mathcal{H}} = \int_{\Omega} \varphi(x)^* \psi(x) \, dx \quad \text{for} \ \varphi, \psi \in \mathcal{H} \]

and the induced norm \( ||\varphi||_{\mathcal{H}} \) for \( \varphi \in \mathcal{H} \). The Hilbert space \( \mathcal{V} = H^1_0(\Omega; \mathbb{C}) \) is given by

\[ \mathcal{V} = \left\{ \varphi \in \mathcal{H} \mid \||\varphi||_{\mathcal{H}} \leq \left( \int_{\Omega} |\nabla \varphi(x)|^2 \, dx \right)^{1/2} < \infty, \quad \varphi = 0 \text{ on } \partial \Omega \right\}, \]

supplied with the inner product \( (\varphi, \psi)_{\mathcal{V}} = (\nabla \varphi, \nabla \psi)_{\mathcal{H}} \) for \( \varphi, \psi \in \mathcal{V} \) and the induced norm \( ||\varphi||_{\mathcal{V}} \). We also need the Hilbert space

\[ \mathcal{W} = L^2(0, T; H^1_0(\Omega; \mathbb{C}) \cap H^2(\Omega; \mathbb{C})) \cap H^1(0, T; L^2(\Omega; \mathbb{C})). \]

For more details on the above Lebesgue and Sobolev spaces see, e.g., [37].

Now, consider the TDSE with an initial state of the quantum system given by \( \psi_0 \in \mathcal{V} \) at \( t = 0 \). One recognizes that the Schrödinger evolution operator is time-reversible (nondissipative) and therefore \( \psi \) cannot have better regularity than \( \psi_0 \) [17]. We also see that with a time-varying potential there is no energy conservation. However, we have norm conservation as stated by the following proposition; see, e.g., [25].

**Theorem 4.**

Let \( V(x, t) = V_0(x) + u(t) x \) and \( ||\psi_0(\cdot)||_{\mathcal{H}} = 1 \); then we have \( ||\psi(\cdot, t)||_{\mathcal{H}} = 1 \) for all \( t \in [0, T] \).

Based on perturbation theory, we obtain that in a long time horizon a time-harmonic (adiabatic) control \( u(t) \) is able to induce transition between two eigenstates \( i \rightarrow j \) if its frequency equals the difference of energy of the two states [25]. Therefore, it is relatively easy to control state transitions for long time intervals. However, the problem becomes very difficult if short time intervals are considered. Here short means that \( T \approx 2\pi/\omega \), where \( \omega = \lambda_j - \lambda_i \), and in this case the perturbation theory is inapplicable and the required control \( u(t) \) deviates greatly from an harmonic function. In this case, an optimal control approach is unavoidable.

To formulate the optimal control problem, we have to decide in which functional space the control is sought or which form of the control is required a priori; for the latter case see, e.g., [35, 52]. From the previous discussion, it appears that the control space \( \mathcal{U} = H^1_0(0, T; \mathbb{R}) \) is the most appropriate for dipole controls, as it means that the laser pulse cannot change instantaneously and it accommodates sinusoidal functions for long time controls. This choice means that the
objective of the optimization has a regularization term of the form \( ||u||_{\Omega}^2 \), where this norm is induced by the following inner product

\[
(u, v)_{\Omega} = \int_0^T (u(t) v(t) + \alpha \dot{u}(t) \dot{v}(t)) \, dt \quad \text{for } u, v \in \mathcal{U},
\]

with \( 0 < \alpha \ll 1 \). Notice that with this norm the control is a continuous function since \( H_0^1(0, T) \) is compactly embedded in \( C^0([0, T]) \) in one dimension. Use of smaller values of \( \alpha \) allows for controls with larger rates of change. Notice that the control is zero at the beginning and at the end of the time interval which is the maximum time window for the laser pulse.

The present control problem requires finding a control \( u \in \mathcal{U} \) such that a quantum system initially in the state \( \psi_0 \) evolves with (11) to a state \( \psi(\cdot, T) \) that is as close as possible to a desired target configuration \( \psi_d \). This aim is formulated by requiring the minimization of the following cost functional

\[
\min J(\psi, u) := \frac{1}{2} \left( 1 - ||P\psi(\cdot, T)||_{\mathcal{H}}^2 \right) + \frac{\gamma}{2} ||u||_{\Omega}^2,
\]

under the constraint given by the TDSE, including the initial condition, denoted as

\[
c(\psi, u) := \{ i \partial_t - H(u) \} \psi = 0,
\]

where \( H(u) = -\Delta + V_0(x) + u(t) x \) and the projector \( P\psi = (\psi_0 , \psi_d)_{i\mathbb{H}} \psi_d \). The goal of the first term of the cost functional is to track the given terminal state \( \psi_d \) up to a global phase \( e^{i\gamma^T} \), which usually cannot be specified.

A control suitable for fast quantum state transition can be obtained in the optimal control formulation given by (15) and (16). To characterize the solution to this problem, we introduce the following Lagrange function

\[
L(\psi, u, p) = J(\psi, u) + \mathbb{R} e \int_0^T \int_\Omega p^*(x, t) c(\psi, u)(x, t) \, dx \, dt.
\]

In [82] it is proved that any minima of (15) and (16) correspond to an extremal point of the Lagrangian; see also [58]. Therefore, taking the Fréchet derivatives of \( L(\psi, u, p) \) with respect to the optimization variables gives the following first-order optimality system that characterizes the optimal solution. We have

\[
\begin{align*}
\{ i \partial_t + \Delta - V_0(x) - u(t) x \} \psi(x, t) &= 0, \\
\{ i \partial_t + \Delta - V_0(x) - u(t) x \} p(x, t) &= 0, \\
-\gamma \alpha \dot{u}(t) + \gamma \alpha \ddot{u}(t) + \mathbb{R} e \int_\Omega p^*(x, t) x \psi(x, t) \, dx &= 0.
\end{align*}
\]

This system consists of the state equation, the adjoint equation, and the optimality condition, respectively, with homogeneous Dirichlet boundary conditions, and initial and terminal conditions given by

\[
\begin{align*}
\psi(x, 0) &= \psi_0(x), \\
p(x, T) &= i (\psi_0(\cdot) , \psi(\cdot, T))_i \psi_d(x), \\
u(0) &= 0, \quad u(T) = 0.
\end{align*}
\]

In [82] it is proved that there exists at least one solution to (18)–(19). It is also proved that if \( \gamma \) is sufficiently large and the norm of the projection \( ||P\psi(T)||_{\mathcal{H}} \) is sufficiently close to 1, then the second-order sufficient optimality condition holds and the optimization problem is locally strictly convex.

Notice that the control is a function of time only, and the state and adjoint variables can be seen as implicit functions of the control. Therefore the dimensionality of the optimization problem can be reduced significantly, introducing a reduced cost functional \( \bar{J}(u) = J(\psi(\cdot, T), u) \). The corresponding gradient is given by

\[
(\nabla \bar{J}(u))(t) = \gamma \dot{u}(t) - \gamma \alpha \ddot{u}(t) - \mathbb{R} e \int_\Omega p^*(x, t) x \psi(x, t) \, dx.
\]
Therefore we have that $\nabla J(u) \in H^{-1}(0, T; \mathbb{R})$, which is problematic with a gradient-based approach because the gradient is not in the same space of the solution and thus it does not provide an update to the control along the descent direction. It has been shown [80] that this problem can be solved by using the Riesz representation of the gradient in the $H_0^1(0, T)$ space as a means of Sobolev smoothing.

We remark that a main difficulty in the analysis of quantum control problems is that they may admit multiple solutions (as most bilinear control problems). This is the case of the present control problem. In fact, we have the following proposition [82].

**Theorem 5.**

Let the initial and target states be eigenfunctions and the stationary potential be symmetric. Then the reduced cost functional does not have a unique minimizer. In particular, if $u^*(t)$ is a minimizer, then so is $-u^*(t)$ and consequently $\tilde{J}(u)$ is nonconvex independently of the values of the regularization parameters.

Next, we have the problem of a suitable discretization of the optimality system. Numerical experience suggests to pursue the approach of discretize-before-optimize in order to avoid any discrepancy between the directional derivative and its approximation, and to guarantee a symmetric Hessian; see [7, 23].

Concerning discretization of the Schrödinger equation, in the case of time-varying Hamiltonians, the Crank–Nicolson (CN) scheme is not norm-preserving. For this reason, in [82] a modified Crank–Nicolson (MCN) method is proposed which is proved to be second-order accurate and unconditionally stable also in the case of time-varying potentials. Let $N_t$ be the number of time steps $t_k = k \delta t$, $k = 0, 1, \ldots, N_t$, where the size $\delta t = \frac{T}{N_t}$, and let $N_x$ be the number of grid points of the $\Omega$ discretization including the boundary points. The TDSE discretized by the MCN scheme results in the following

$$\psi_k - \psi_{k-1} = -\frac{i \delta t}{4}[H(t_k) + H(t_{k-1})] [\psi_k + \psi_{k-1}].$$  \hfill (21)

Let $y_\ell \in C^{N_x-2}$ be a vector which contains the values of $\psi$ on the interior grid points at the $k$th time step. Let $N_d = (N_x - 2)(N_t - 1)$ be the total number of degrees of freedom of the state variable, which excludes the initial and boundary points.

It is convenient to reformulate the equality constraint by introducing the following matrices [81]

$$A_1 = I + \frac{i \delta t}{2} H_0, \quad A_2 = I - \frac{i \delta t}{2} H_0, \quad B = \frac{i \delta t}{4} X,$$

where $H_0$ is the discretization of the stationary Hamiltonian $H_0 = -\Delta + V_0(x)$, using linear finite elements on a uniform grid, and $X$ is the discretized position operator. The discrete equality constraint (21) can be written as follows

$$c_\ell(y, u) = [A_1 - (u_k + u_{k-1}) B] y_{\ell} - [A_2 + (u_k + u_{k-1}) B] y_{\ell-1}.$$  \hfill (23)

We write the cost in terms of the discretized wavefunction

$$J(y, y^*, u) = \frac{1}{2} (1 - y^T P y) + \frac{1}{2} u^T W u,$$

where $P \in \mathbb{R}^{N_x \times N_d}$ is the discrete projection operator, defined by

$$P y = (0, \ldots, 0, (y_d^T y_{\ell}) y_{\ell})^T,$$

where $^T$ means transpose and the variables are understood to be column vectors. The matrix $W$ is the finite element approximation to the Helmholtz operator $I - \alpha \partial_x^2$.

To take variations using the Wirtinger calculus [76], we consider the discrete Lagrangian, with Lagrange multipliers $p_1, \ldots, p_{N_t}$ and $p_{1}^*, \ldots, p_{N_t}^*$, as follows

$$L(y, y^*, u, p, \bar{p}) = J(y, y^*, u) + \sum_{k=1}^{N_t} p_k^T c_k(y, u) + p_k^{*T} c_k^*(y, u).$$  \hfill (26)
The components of the derivatives of the Lagrange function are row vectors given as follows

\[ L_y = -\frac{1}{2} y^\top P + p^\top c_y, \]  
\[ L_{y'} = -\frac{1}{2} y^\top P + p^\top c^*_{y'}, \]  
\[ L_u = y u^\top W + p^\top c_y + p^\top c^*_{u}, \]  
\[ L_p = c^\top, \]  
\[ L_{p'} = c^* c^\top. \]  

In this notation, the discrete equality constraint \( c^\top \in \mathbb{R}^{1 \times N} \) is a row vector given by \( c^\top = \left(c_{1}^\top \ c_{2}^\top \ \cdots \ c_{N}^\top \right) \).

We collect the components (27)–(31) to define the full gradient \( \nabla L = \left(L_y, \ L_{y'}, \ L_u, \ L_p, \ L_{p'}\right) \), and the Hessian is defined as the Jacobian of the Hermitian conjugate of this gradient, which we denote with \( \nabla^2 L = \nabla (\nabla L)^* \); see [82].

Notice that the complex differentiation and transposition operators do not commute, so the order of operations is critical.

Using the fact that \( c = 0 \) and \( L_y = 0 \) by construction, the full Karush-Kuhn-Tucker (KKT) Hessian and gradient equations allow to compute the differential change in the state and adjoint variables due to a differential change in the control

\[ \delta y = -c_y^\top c_y \delta u, \]  
\[ \delta p = -c_y^\top (L_y \delta y^* + L_{y'} \delta u). \]  

Using the Schur reduction on the KKT system, we can write the Newton method in terms of the reduced cost functional as follows

\[ \nabla^2 \hat{j}(u) \delta u = -\nabla^* j(u) \]  
where \( \nabla^* j(u) = L^*_u \), and the application of the reduced Hessian on \( \delta u \) is given by

\[ \nabla^2 \hat{j}(u) \delta u = L_{uu} \delta u + 2 \Re \{c_u^\top \delta p + L_{uy} \delta y\}. \]  

Based on this framework, a globalized Krylov–Newton scheme and the corresponding code are presented in [81]. In the following, we report results obtained with the Krylov–Newton scheme.

In Table 2, we report the decrease of the norm of the gradient for the method of steepest descent and for the Krylov–Newton method for successive iterations. From this table, typical Newton’s second-order convergence can be seen. The Krylov–Newton method attains a gradient which is zero to machine precision in just 8 steps. For the test of Table 2, 400 time steps and 200 spatial grid points were used with \( T = 1, \ \alpha = 10^{-2}, \ \gamma = 10^{-2}, \psi_0(x) = \cos(\pi x/2), \psi_\infty(x) = \frac{1}{\sqrt{10}} \{\cos(\pi x/2) + 3 \cos(\pi x)\}. \)

3.2. Optimal Control of Bose–Einstein Condensates

Quantum optimal control of transport of BECs in magnetic microtraps was proposed [44] to solve the problem for a trapped-atom interferometer setup where a dilute BEC should be split from a single to a double-well ground state. We recall that a coherent BEC is described by the following

\[ i \partial_t \psi(x, t) = \left\{ -\Delta + V(x, u(t)) + g |\psi(x, t)|^2 \right\} \psi(x, t), \]  

where \( x \in \Omega \) and \( t \in [0, T] \). We measure mass in units of the atom mass and length in units of micrometers. We assume that the quantum state wavefunction \( \psi(x, t) \) is normalized to one and therefore \( g \) in (35) incorporates the number of atoms \( N_h \). The fact that \( V \) is a confinement potential results in a wavefunction \( \psi \) whose support is localized in a
bounded region. Therefore, with Ω we represent a spatial domain that is large enough to represent the support of ψ during evolution and we use periodic boundary conditions.

The magnetic control potential \( V_m(x, t) = V(x, u(t)) \) is modulated by a control function \( u(t) \), and the initial and final potential configurations are given; therefore we require that \( u(t) \) takes initial and final values of zero and one, respectively. These two extremal values correspond to the case where the potential \( V_m \) is convex and to the case where it has a double-well structure. Furthermore, we require that \( V_m \) be spatially symmetric with respect to the origin of coordinates.

Suppose that initially the system is in the ground state \( ψ_0 \) for the potential \( V(x, 0) \). Upon varying \( u(t) \) in the time interval \( t \in [0, T] \) from zero to one, the system will pass through a sequence of states and will end up in the final state \( ψ(T) \). Our purpose is to determine an optimal control time evolution that allows us to channel the system from the initial state \( ψ_0 \) at time zero to a desired state \( ψ_d \) at final time \( T \) and \( ψ_d \) represents the ground state for the potential \( V(x, 1) \) at time \( T \). The ground state for a given potential \( V(x, u) \) is defined as the stationary state \( φ(x) \) with \( \int_Ω |φ(x)|^2 dx = 1 \) that minimizes the following energy [9]

\[
E_u(φ) = \frac{1}{2} \int_Ω \left( |\nabla φ(x)|^2 + V_u(x) |φ(x)|^2 + \frac{g}{2} |φ(x)|^4 \right) dx.
\]

We consider the control potential proposed in [53] to create condensates of trapped atoms coupled with a radio frequency fields. We have that

\[
V(x, u) = -\frac{u^2 d^2}{8c} x^2 + \frac{1}{c} x^4,
\]

where \( c = 40 \) and \( d \) is a parameter corresponding to twice the distance of the two minima in the double-well potential.

For our purpose, in order to define a well-defined control problem, we consider the cost functional [44]

\[
J(ψ, u) = \frac{1}{2} \left( 1 - |(ψ_d, ψ(T))|^2 \right) + \frac{γ}{2} \int_0^T (u(t))^2 dt
\]

The second term in the cost functional represents an \( L^1 \) cost and aims at penalizing fast varying confinement potentials that are more difficult to realize in real experiments.

The control problem under consideration is therefore to minimize \( J(ψ, u) \) subject to the condition that the wave function \( ψ \) fulfills the BEC equation (35) with given initial conditions. We obtain the following optimality system characterizing the solution to the BEC optimal control problem [13]

\[
i \frac{∂ψ}{∂t} = (-Δ + V_u + g |ψ|^2) ψ,
\]

\[
i \frac{∂p}{∂t} = (-Δ + V_u + 2g |ψ|^2) p + g ψ^2 p^*.
\]

\[
γ \frac{d^2 u}{dt^2} = -Re \left( ψ, \frac{∂V_u}{∂u} p \right).
\]

---

Table 2. Convergence of the steepest descent scheme and of the Krylov–Newton scheme with respect to \( ||∇J_0(u)|| \).

| Iteration | \( ||∇J_0(u)|| \) | \( ||∇J_{KN}(u)|| \) |
|-----------|------------------|------------------|
| 1         | 1.8615 × 10^{-4} | 1.8615 × 10^{-4} |
| 2         | 6.5263 × 10^{-5} | 6.5263 × 10^{-5} |
| 3         | 6.0031 × 10^{-5} | 2.4732 × 10^{-5} |
| 4         | 2.3535 × 10^{-5} | 1.5557 × 10^{-5} |
| 5         | 2.8106 × 10^{-5} | 1.2316 × 10^{-6} |
| 6         | 1.5703 × 10^{-5} | 1.0977 × 10^{-8} |
| 7         | 1.7062 × 10^{-5} | 3.5480 × 10^{-13} |
| 8         | 1.0322 × 10^{-5} | 2.0009 × 10^{-17} |
| 9         | 1.3312 × 10^{-5} | 0                |
which has to be solved together with the initial and terminal conditions

\[ \begin{align*}
\psi(0) &= \psi_0, \\
-\frac{\partial}{\partial t} \psi(T) &= -(\psi, \psi(T)) \psi, \\
u(0) &= 0, \quad u(T) = 1.
\end{align*} \] (42-44)

Because of \( H^1 \) regularization we have a natural setting to impose the required Dirichlet boundary conditions on the control function, \( u(0) = 0 \) and \( u(T) = 1 \). See [13] for a discussion concerning second-order optimality conditions.

We have that (39) is uniquely solvable for every \( u \in H^1(0, T; \mathbb{R}) \) such that \( V_0 \) is a symmetric double-well potential; see [71]. Thus, it is meaningful to introduce the reduced cost functional \( \tilde{J} : H^1(0, T; \mathbb{R}) \to \mathbb{R} \) given by \( \tilde{J}(u) = J(\psi(u), u) \), where \( \psi(u) \) denotes the unique solution to (39) for given \( u \). One can show that the \( (L^2) \) gradient of \( \tilde{J} \) with respect to \( u \) is given by

\[ \nabla \tilde{J}(u) = -\gamma \frac{d^2}{dt^2} - \Re e \left( \psi, \frac{\partial V_u}{\partial u} \frac{\partial}{\partial u} \right), \] (45)

where \( \psi \) and \( p \) solve the state and the adjoint equations with given \( u \). The gradient used in the optimization procedure will depend on the choice of the control space. If we choose the space \( X = L^2(0, T; \mathbb{R}) \), we have the gradient (45).

In the case \( X = H^1(0, T; \mathbb{R}) \), we can determine the gradient in \( H^1 \) by requiring

\[ \left( \nabla \tilde{J}_{L^2}(u), \varphi \right)_{L^2} = \left( \nabla \tilde{J}_{H^1}(u), \varphi \right)_{H^1}. \]

Using the definition of the \( H^1 \) inner product and integrating by parts shows that the \( H^1 \) gradient must satisfy the following one-dimensional Poisson equation with homogeneous Dirichlet conditions. That is,

\[ \frac{d^2}{dt^2} \nabla \tilde{J}(u) = -\gamma \frac{d^2}{dt^2} - \Re e \left( \psi, \frac{\partial V_u}{\partial u} \frac{\partial}{\partial u} \right), \] (46)

with \( |\nabla \tilde{J}(u)|(0) = 0 \) and \( |\nabla \tilde{J}(u)|(T) = 0 \). Two important differences between the \( L^2 \) and \( H^1 \) formulations are immediately apparent. First, the \( L^2 \) gradient does not vanish at \( t = T \), and second, the \( H^1 \) gradient possesses the same degree of smoothness as \( u \), while the \( L^2 \) gradient does not.

The discussion on which optimization space is used is very important also for the numerical implementation. As pointed out in [80], not only the gradient is function space dependent but also the construction of gradient-based optimization schemes. In [80], optimization methods in \( L^2 \) and \( H^1 \) spaces are discussed with a focus on the nonlinear conjugate gradient scheme of Hager and Zhang [42] and on the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm.

Concerning discretization of the BEC control problem, in [13, 80] an unconditionally stable explicit second-order norm-preserving time-splitting spectral scheme (TSSP) was used.

It follows that, as \( T \) becomes shorter, a more oscillatory optimal control results. However, the solutions obtained with \( H^1 \) space minimization are less oscillating.

The regularization parameter \( \gamma \) plays an important role in the \( L^2 \)-based optimization, as it forces the solution to have a finite \( H^1 \) norm, while this will always be the case by construction with the \( H^1 \)-based formulation. In the \( L^2 \) space, the limit as \( \gamma \to 0 \) small gives a control function \( u \) which has a steep gradient at \( t = T \). On the other hand, since the \( H^1 \)-based scheme produces \( u \in H^1 \) at every optimization step, the precise value of the regularization term is not essential to obtain a smooth control. Rather in this case the \( \gamma \) term serves only to penalize highly oscillatory solutions. In the limit as \( \gamma \to 0 \), the \( H^1 \)-based solution approaches a smooth curve which gives a small cost functional.

4. Open quantum systems

The progress in the development of nano systems where quantum states are used and manipulated for the encoding and processing of quantum information [18] shows, on the one hand, the enormous technological potential of envisioned quantum devices and, on the other hand, the limitation of modelling these systems as closed quantum systems. In
fact, sophisticated laboratory measurements demonstrate that any nano device suffers an undesirable coupling with the environment; see, e.g. [55]. Moreover, these devices are subject to measurements and any control mechanism also gives rise to an exterior coupling. For this reason, the modelling of the environmental impact and of measurements is essential which motivates the theory of open quantum systems [20] describing the interaction of a quantum model with another one representing the environment, while the combined system is assumed to be closed. In this theory, the dynamical evolution of the open system is then described with an effective equation of motion called the Kossakowski-Lindblad master equation [41, 57]. Notice that the Lindblad equation is based on a Markov hypothesis that does not hold if the environment has memory on a time scale comparable to the system. Furthermore, completely positive open quantum systems are modelled by Kraus operators; see, e.g., [77, 85].

In addition to these modelling strategies, once the open system is subject to continuous measurements, an extension of the master equation that includes stochastic terms is required. These terms are seen as randomness originating from the measurement process; see [10, 33].

In this section, we focus again on a two-level system in order to illustrate the open quantum framework and a newly proposed control scheme. In the case that this model is closed, it is described by a two-dimensional Hilbert space, which is a geometrical sphere with the antipodal points corresponding to mutually orthogonal state vectors. In fact, we have the spin-up and spin-down states of an electron, that we denote with $|0\rangle$ and $|1\rangle$, respectively. Other points on the sphere correspond to mixed states. A general closed two-level quantum state can be written as a complex superposition of the basis vectors as follows

$$|\psi\rangle = \cos \left(\frac{\theta}{2}\right) |0\rangle + e^{i\varphi} \sin \left(\frac{\theta}{2}\right) |1\rangle,$$

where $0 \leq \theta \leq \pi$ and $0 \leq \varphi < 2\pi$ are the spherical coordinates on the Bloch sphere.

These coordinates specify a direction, the Bloch vector given as follows

$$\vec{m} = (x, y, z) = (\sin (\theta) \cos (\varphi), \sin (\theta) \sin (\varphi), \cos (\theta)).$$

The Hamiltonian of a two-level spin system in a constant magnetic field, pointing in the $z$ direction, is given by

$$H = \frac{\hbar}{2} \omega \sigma_z - \frac{\hbar}{2} \left( \Omega^*(t) \sigma + \Omega(t) \sigma^* \right),$$

where $\omega$ is the Larmor frequency and $\sigma$ and $\sigma^*$ are the 'lowering' and 'raising' operators, respectively. These are defined as follows

$$\sigma = (\sigma_x - i \sigma_y)/2, \quad \sigma^* = (\sigma_x + i \sigma_y)/2.$$

The constant $\omega$ may represent a detuning field and the Rabi frequency $\Omega(t)$ is a complex function that is proportional to the slowly-varying complex amplitude of the control laser field (qubit) or of the RF-field (NMR).

A closed statistical ensemble of two-level systems can be described by its density matrix $\rho$ given by

$$\rho = \frac{1}{2} \left( I + \vec{m} \cdot \vec{\sigma} \right).$$

The density matrix serves to calculate the mean value of any operator of a quantum system. In particular, the expectation value of the spin orientation is given by $\vec{m} = \langle \vec{d} \rangle = Tr (\rho \vec{d})$.

In the case of an open two-level statistical ensemble, the Liouville - von Neumann Master equation, describing the evolution of $\rho$, is augmented with an additional ‘dissipator’ term $D(\rho)$ as follows

$$i\hbar \frac{\partial \rho}{\partial t} = [H, \rho] + i\hbar D(\rho),$$

where dissipation due to spin-electromagnetic field interaction is given by

$$D(\rho) = g \sigma \rho \sigma^* - \frac{g}{2} \{ \sigma^* \sigma, \rho \}. $$
where $\{A, B\} = AB + BA$, and $g$ is a phenomenological damping constant.

This is the Kossakowski–Lindblad master equation, which becomes the following equation for the expectation values of the spin components, $m = (x, y, z)$. We have

$$
\begin{align*}
\dot{x} &= -d x - \omega y + \Omega z \\
\dot{y} &= \omega x - d y + \Omega z \\
\dot{z} &= -\Omega x - \Omega y - g (z + 1)
\end{align*}
$$

where $\Omega = \Omega_i + i \Omega_r$ is the control field, and $d = g/2$.

The next step in the modelling of open quantum systems is to develop a theory of measurements, that is, to model the action of a measurement operation on the system's dynamics. For this purpose, two types of stochastic Schrödinger equations are considered that correspond to measurements in continuous time (diffusion process) and to measurements at different instants of time (jumping process), respectively. In both cases, the result of the experiment and its effect on the dynamics of the system is considered as a random variable within the set of possible values; see, e.g., [10].

Next, we discuss the model of an open two-level system subject to heterodyne measurements; see [84]. This setting corresponds to the following stochastic Schrödinger equation (SSE) [40]. We have

$$
d\psi = -\left( i H + \frac{g}{2} \sigma^T \sigma - g(\sigma^T)_{\sigma} \right) \psi dt + \sqrt{g} dW_{\sigma} \psi
$$

where $dW$ represents a complex Wiener process.

As shown in [10], in correspondence to a SSE it is possible to formulate a stochastic master equation. This formulation generalizes the procedure illustrated above to obtain the Lindblad equation. Specifically, in the two-level case, we derive the stochastic master equation using polar coordinates on the sphere. We have $\varphi = \arctan(y/x)$ and $\theta = \arccos(z)$. The Bloch equations, with controls $\Omega_i = -au$ and $\Omega_r = av$, and including stochastic noise, are as follows

$$
\begin{align*}
\frac{d\varphi(t)}{dt} &= B_\varphi(\varphi, \theta) dt + \alpha_{11}(\varphi, \theta) dW_1 + \alpha_{12}(\varphi, \theta) dW_2 \\
\frac{d\theta(t)}{dt} &= B_\theta(\varphi, \theta) dt + \alpha_{21}(\varphi, \theta) dW_1 + \alpha_{22}(\varphi, \theta) dW_2,
\end{align*}
$$

where

$$
\begin{align*}
B_\varphi(\varphi, \theta) &= A_\varphi(\varphi, \theta) \\
B_\theta(\varphi, \theta) &= A_\theta(\varphi, \theta) + g \frac{1 + \cos(\theta)}{\sin(\theta)} \left(1 - \frac{1 + \cos(\theta) \cos(\theta)}{4}\right) \\
\alpha_{11}(\varphi, \theta) &= -\sqrt{\frac{1}{2} \frac{1 + \cos(\theta)}{\sin(\theta)}} \sin(\varphi) \\
\alpha_{12}(\varphi, \theta) &= \sqrt{\frac{1}{2} \frac{1 + \cos(\theta)}{\sin(\theta)}} \cos(\varphi) \\
\alpha_{21}(\varphi, \theta) &= \sqrt{\frac{1}{2} \frac{1 + \cos(\theta)}{\sin(\theta)}} \cos(\varphi) \\
\alpha_{22}(\varphi, \theta) &= \sqrt{\frac{1}{2} \frac{1 + \cos(\theta)}{\sin(\theta)}} \sin(\varphi)
\end{align*}
$$

and

$$
\begin{align*}
A_\varphi(\varphi, \theta, u, v) &= \omega + a \cot(\theta) (u \sin(\varphi) + v \cos(\varphi)) \\
A_\theta(\varphi, \theta, u, v) &= -a (u \cos(\varphi) - v \sin(\varphi))
\end{align*}
$$

These equations model the evolution of the orientation of the mean magnetic moment of a two-level spin system with dissipation due to environmental losses and subject to continuous measurements [84].

A consequence of modelling an open quantum system as a stochastic differential model is that we can view this system from a statistical point of view and notice that the state of a stochastic process can be completely characterized by the shape of its statistical distribution that is represented by the probability density function. The next important consequence is that the evolution of the PDF associated to a stochastic process is modelled by a Fokker–Planck (FP) equation [69].
Corresponding to our Bloch system (48) on the sphere, we obtain the following (Bloch) FP equation, where we include control terms. We have

\[
\frac{\partial f}{\partial t} = -\frac{\partial}{\partial \phi}(B_\phi(\phi, \theta, u, v) f) - \frac{\partial}{\partial \theta}(B_\theta(\phi, \theta, u, v) f) + \frac{g}{4} \frac{\partial^2}{\partial \phi^2} \left( \frac{1 + \cos(\theta)}{1 - \cos(\theta)} f \right) + \frac{g}{4} \frac{\partial^2}{\partial \theta^2} \left( (1 + \cos(\theta))^2 f \right) \tag{50}
\]

where \( \phi \in [0, 2\pi] \), \( \theta \in (0, \pi) \), and the PDF solution \( f = f(\phi, \theta, t) \) is required to be positive and its integral on the computational domain be conserved and normalized, i.e.

\[
\int_0^{2\pi} \int_0^\pi f(\phi, \theta, t) d\theta d\phi = 1 \tag{51}
\]

An initial condition \( f(\phi, \theta, 0) = \rho(\phi, \theta) \) completes the setting of the FP problem.

The formulation of the FP problem suggests an effective control strategy that, on the one hand, exploits the relationship between the stochastic Schrödinger equation and the stochastic Lindblad master equation, and on the other hand introduces the FP equation that models the evolution of the PDF of the stochastic processes. The use of the PDF allows us to consider deterministic objectives and the use of the FP equation makes possible to compute robust controls using the adjoint method that take into account the statistics of the stochastic evolution. The FP control strategy was proposed in [2, 3] to control the PDF of financial and biological models. In [4], this strategy is extended to the control of two-level open quantum systems.

5. Conclusions

In this contribution, a review of some recent developments in the field of optimal control of quantum systems was presented with a focus on adjoint methods and their numerical implementation. The issues of exact controllability and optimal control were discussed for finite- and infinite-dimensional quantum systems and some insight was provided considering 'two-level' models.

However, many other topics in the field of quantum control remain to be discussed. In application, feedback-optimized laser pulses to control chemical reactions are becoming viable technologies [6]: further, emission of atom beams into a waveguide potential is now possible through control of BEC [19], this result demonstrates the possibility to use ultracold atomic gases as sources for quantum matter–wave optics. Among other important emerging topics, we mention quantum system Hamiltonian identification [56], and the analysis of low dimensional nanostructures as a multiscale multiphysics problem [63]. Indeed, these are only few examples of the tremendously growing field of nanoscience research.

Acknowledgments

I would like to gratefully acknowledge the continuous collaboration with Mario Annunziato, Ugo Boscain, Gunther Dirr, Ulrich Hohenester, Julien Salomon, Georg Stadler, Stefan Volkwein, Greg von Winckel. I also would like to thank Roderick Melnik for encouraging this review work. This work was supported in part by DFG Project "Controllability and Optimal Control of Interacting Quantum Dynamical Systems".

References


[34] I. Degani and A. Zanna, Optimal Quantum Control by an Adapted Coordinate Ascent Algorithm, *SIAM J. Sci.
Alfio Borzì

Comput. 34, 1488 (2012).

[63] R. Melnik, Multiple scales and coupled effects in modelling low dimensional semiconductor nanostructures: Between


