Sparse Control of Quantum Systems

Gero Friesecke*, Felix Henneke*, and Karl Kunisch**

*Faculty of Mathematics, TU Munich, Germany
gf@ma.tum.de, henneke@ma.tum.de
**Institute for Mathematics and Scientific Computing, University of Graz, Austria
karl.kunisch@uni-graz.at

Abstract. A new class of cost functionals for optimal control of quantum systems which produces controls which are sparse in frequency and smooth in time is proposed. This is achieved by penalizing a suitable time-frequency representation of the control field, rather than the control field itself, and by employing norms which are of L^1 or measure form with respect to frequency but smooth with respect to time.

A mathematical framework is developed which yields existence of optimal controls and necessary optimality conditions of the resulting nonsmooth, nonconvex optimization problem. The framework covers the important systems of physical interest, including (infinite-dimensional) Schrödinger dynamics on multiple potential energy surfaces as arising in laser control of chemical reactions.

Numerical simulations demonstrate that the optimal controls, unlike those obtained with the usual L^2 or H^1 costs, concentrate on just a few frequencies, even in the infinite-dimensional case of laser-controlled chemical reactions.

1 Introduction

This paper is motivated by application problems of current interest in quantum control, which range from steering chemical reactions [BKZB08] over creating excited or ionized states [HRG13] to faithfully storing and manipulating bits of quantum information [SKS+14].

We propose a new class of cost functionals for the optimal control of quantum systems which result in controls with a very simple time-frequency structure. This is achieved via two key ideas.

First, we do not penalize the time profile of the field amplitude but a suitable *time-frequency representation* of it. While such representations are a familiar tool to interpreting or analyzing a given field in quantum control and signal analysis, they here acquire center stage already in the design of the controls.

Second, we capitalize on recent advances in the optimal control theory of elliptic and parabolic systems built upon the basic idea [Sta09, HSW12] of *sparsity-enhancing L^1 or measure-norm costs*. More specifically, we exploit the idea of function-valued measures

to achieve directional sparsity in parabolic control [KPV14]. The novelty as compared to the latter advances is that in quantum control, unlike in parabolic control, the target for sparsity should not be the field amplitude, but its frequency structure.

These ideas result in constrained non-smooth optimization problems of the form

$$\text{Minimize } \frac{1}{2}\langle\psi(T), \mathcal{O}\psi(T)\rangle + \alpha||u||_{\mathcal{M}} \text{ over controls } u\colon \Omega \times [0,T] \to \mathbb{C} \tag{1}$$

subject to

$$i\partial_t\psi = (H_0 + (Bu)(t)H_1)\psi, \;\; \psi(0) = \psi_0. \tag{2}$$

Here H_0 is the Hamiltonian of the quantum system, B is a 'synthesis operator which assembles the control field from a time-frequency representation $u(\omega, t)$, and $||\cdot||_{\mathcal{M}}$ is an L^1 or measure norm with respect to frequency but a smoothness-promoting norm with respect to time. A prototypical choice is

$$(Bu)(t) = \int_\Omega u(\omega, t)\, e^{i\omega t}\, \mathrm{d}\omega, \tag{3}$$

where $\Omega \subset \mathbb{R}$ is a region of admissible frequencies, and

$$||u||_{\mathcal{M}} = \int_\Omega ||u(\omega, \cdot)||_{H^1(0,T)}\, \mathrm{d}\omega. \tag{4}$$

Note that the control, a priori, can use an arbitrary selection or superposition of the available continuum of frequencies, with each frequency possessing its own time profile. Equations (3)–(4) replace the standard approach in quantum optimal control initiated by [PDR88] to penalize just the L^2 or H^1 norm of the field amplitude (see [IK07, vWB08, HMMS13] for mathematical results).

Numerical simulations presented in Section 5 below show that, unlike controls obtained from standard L^2 or H^1 costs, the optimizers concentrate on just a few frequencies, even when the quantum dynamics is a full infinite-dimensional Schrödinger dynamics.

The plan of this paper is as follows. In the next section we give an overview of important examples of quantum dynamics with controls, and derive the coupling operators for atomic excitation problems and for laser-guided chemical reaction dynamics from first principles. This serves to explain basic structural features of the coupling operators such as 'forbidden transitions' and the oscillatory nature of the controls to readers who are less familiar with quantum control systems. In Section 3 we introduce our measure-norm sparsity-enhancing costs within a general functional-analytic framework, and give several examples. In particular, while the choices (3)–(4) lead to frequency-sparsity with global time profiles, appropriate modifications lead to frequency-sparsity with local time profiles. Section 4 is devoted to the mathematical analysis of the non-smooth optimal control problem (1)–(2). We establish uniqueness of mild solutions to the state equation, existence of optimal controls, and necessary optimality conditions on the optimizers, including the interesting result that optimizers have compact frequency support. Finally, in Section 5 we numerically calculate optimal controls and compare them to those obtained from the usual L^2 or H^1 penalization of the field amplitude. Specifically, we

present a 3-level example which arises in atomic excitation problems, and an example of Schrödinger dynamics on two potential energy surfaces as arising in laser-controlled chemical reaction dynamics.

2 Quantum dynamics with controls: physical examples

The evolution equations in quantum control problems typically have the structure

$$i\partial_t\psi(t) = \left(H_0 + \sum_{\ell=1}^{L} v_\ell(t)H_\ell\right)\psi(t), \tag{5}$$

where the state $\psi(t)$ belongs to some Hilbert space $\mathcal{H}$, H_0 and the H_ℓ are (bounded or unbounded) self-adjoint operators on $\mathcal{H}$, and the $v_\ell(t)$ are real-valued scalar amplitudes of components of applied electric or magnetic fields. The operator H_0 is the Hamiltonian of the system in the absence of fields, and the H_ℓ describe the system-field coupling.

In the mathematical control theory literature, problems of above form have been previously considered, but important features of the models and their behaviour fall out of thin air. This is particularly the case for

- the assumed structure of the control terms
- specific choices in model simulations
- and the highly oscillatory nature of optimal controls, unfamiliar from elliptic and parabolic problems.

On the other hand, in the physics literature on quantum control, the origin of these features is implicitly understood, but not written down. To help bridge the gap between these discourses, we give here an overview of examples from atomic physics, quantum information theory, and quantum chemistry. Moreover we derive the coupling operators for atomic excitation problems and for laser-guided chemical reaction dynamics from first principles.

We note first that equation (5) already contains two important approximations which are valid in many situations of interest. First, quantum fluctuations of the field amplitudes can be neglected, that is to say we are dealing with classical fields and do not need to move to the much more complicated framework of quantum field theory. Second, the spatial wavelength of the applied fields is much larger than the localization length of the state $\psi(t)$, so that it is sufficient to assume that the field strengths depend on time only. This is often called 'dipole approximation'.

We now give four examples of (5), all of which are of significant physical interest.

Example 2.1. (Spin of a spin 1/2 particle in a magnetic field) This is the simplest control system of physical interest. It arises as a basic example in NMR, and more recently as a model of a single qubit in quantum information theory (see [SKS$^+$14] for a recent careful experimental realization of this system and [BCG$^+$02] for rigorous

mathematical results). It already exhibits surprisingly many features of complex systems. The spin at time t is a unit vector in the Hilbert space $\mathcal{H} = \mathbb{C}^2$. The general evolution equation of a spin in a time-dependent magnetic field $B\colon \mathbb{R} \to \mathbb{R}^3$ is

$$i\partial_t \psi(t) = -\gamma B(t) \cdot S\, \psi(t) = -\gamma \sum_{\ell=1}^{3} B_\alpha(t) S_\alpha\, \psi(t) \tag{6}$$

where the component operators S_α of the spin operator S are given by $\hbar/2$ times the Pauli matrices, i.e. in atomic units ($\hbar = 1$)

$$S_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_2 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad S_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The factor γ depends on the type of particle (electron, proton, neutron, nucleus) and can be positive or negative. A typical control problem consists of taking B_3 time-independent and comparatively large, and B_1 and B_2 as time-dependent controls which are small (in NMR experiments, the control is several orders of magnitude smaller). This gives a system of form (5),

$$i\partial_t \psi = (H_0 + \sum_{\ell=1}^{2} B_\ell(t) H_\ell)\psi, \;\; \text{with } H_0 = -\gamma B_3 S_3,\; H_1 = -\gamma S_1,\; H_2 = -\gamma S_2. \tag{7}$$

Denoting the two eigenvalues of H_0 by E_1, E_2, this system can be written in the follwoing elegant form

$$i\partial_t \psi = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} \psi + \begin{pmatrix} 0 & v^*(t) \\ v(t) & 0 \end{pmatrix} \psi \tag{8}$$

with complex-valued control $v(t) = -2\gamma(B_1(t) + iB_2(t))$. The basic case of a time-harmonic control field

$$v(t) = Ae^{i\omega t} \tag{9}$$

is exactly soluble, as done in a classical paper by Rabi [Rab37]. This allows to understand mathematically the emergence of *oscillatory controls* and *Bohr frequencies*. The Bohr frequency of a transition between two quantum states is the eigenvalue difference; the time-harmonic control with this frequency, when applied over a time window of suitable length, achieves a 100 % transfer; and it is the *only* time-harmonic control (9) which achieves a 100 % transfer.

Example 2.2. (Electronic states of atoms in laser fields) A standard reference in the physics literature is [Sho90]. Consider an atom with N electrons of charge -1 and a nucleus of charge $Z = N$ clamped at the origin. The electronic state of the atom is described by a function belonging to the Hilbert space

$$\mathcal{H} = \{\psi \in L^2((\mathbb{R}^3 \times \mathbb{Z}_2)^N) \,|\, \psi \text{ antisymmetric}\}. \tag{10}$$

That is to say electronic states are functions $\psi = \psi(x_1, s_1, \ldots, x_N, s_N)$ which depend on the position coordinates $x_i \in \mathbb{R}^3$ and the spin coordinates $s_i \in \mathbb{Z}_2$ of all the electrons. The requirement that ψ be antisymmetric means that ψ must change sign under

simultaneous permutation of position and spin coordinates of any two electrons, i.e. $\psi(x_1, s_1, \ldots, x_i, s_i, \ldots, x_j, s_j, \ldots, x_N, s_N) = -\psi(x_1, s_1, \ldots, x_j, s_j, \ldots, x_i, s_i, \ldots, x_N, s_N)$ for all $i < j$. The 'exact' (non-relativistic, Born-Oppenheimer) Hamiltonian of the system is given, in atomic units $\hbar = 1$, electron mass $m = 1$, electron charge $e = -1$, by

$$H_0 = -\frac{1}{2}\Delta + V(x_1, \ldots, x_N), \tag{11}$$

where Δ is the Laplacian on $\mathbb{R}^{3N}$ and V is a many-body Coulomb potential composed of electron-nucleus attraction terms and electron-electron repulsion terms,

$$V(x_1, \ldots, x_N) = -\sum_{i=1}^{N} \frac{Z}{|x_i|} + \sum_{1 \le i < j \le N} \frac{1}{|x_i - x_j|}. \tag{12}$$

An applied electric field can be described by a function $E\colon \mathbb{R} \to \mathbb{R}^3$, with $E(t)$ denoting the electric field vector at time t. (Here quantum fluctuations of the field as well as its spatial dependence are neglected, as discussed above.) The coupling term between the electronic state and the field is given by

$$-E(t) \cdot D \tag{13}$$

with the dipole operator

$$D(x_1, \ldots, x_N) = \sum_{i=1}^{N} e\, x_i. \tag{14}$$

In atomic units, the electron charge has the value $e = -1$ but we have inserted the charge here to make apparent the simple physical origin of D: for a general system of N pairs of opposite charges in $\mathbb{R}^3$, the electric dipole moment is defined as the sum of charge times distance of each pair. Thus for N electrons of charge e located at the points $x_1, \ldots, x_N$, and one nucleus of charge $-Ne$ located at the origin (or equivalently N protons of charge $-e$ at the origin), the electric dipole moment is exactly given by the right hand side of (14). The overall evolution equation is

$$i\partial_t \psi(t) = \left(-\frac{1}{2}\Delta + V(x_1, \ldots, x_N) - E(t) \cdot D(x_1, \ldots, x_N) \right) \psi(t). \tag{15}$$

This has the form (5), as is immediate by denoting the components of $E(t)$ and x_i with respect to some orthonormal basis of $\mathbb{R}^3$ by $E_\ell(t)$ and $x_{i\ell}$ ($\ell = 1, 2, 3$) and noting that

$$-E(t) \cdot D = E(t) \cdot \sum_{i=1}^{N} x_i = \sum_{\ell=1}^{3} E_\ell(t) H_\ell, \quad \text{with } H_\ell = \sum_{i=1}^{N} x_{i\ell}.$$

In control problems, the infinite-dimensional state equation (15) is often replaced by projecting onto finitely many eigenstates $\psi_1, \ldots, \psi_d$ of H_0 and neglecting the coupling with the rest of the system. To derive the resulting model, denote the orthogonal projector

by P, the span of the eigenstates by $\mathcal{H}'$, and the projection of ψ onto $\mathcal{H}'$ respectively its orthogonal complement $\mathcal{H}^\perp$ by ψ' and $\psi^\perp$. We have from eq. (15) that

$$i\partial_t \psi'(t) = (H_0 + P(E(t)\cdot D)P)\,\psi'(t) + P(E(t)\cdot D)\psi^\perp(t),$$

and discarding the last term yields the reduced model

$$i\partial_t \psi'(t) = (H_0 - P(E(t)\cdot D)P)\,\psi'(t). \tag{16}$$

By expanding $\psi'(t) = a_1(t)\psi_1 + \cdots + a_d(t)\psi_d$, this is equivalent to an evolution equation in $\mathbb{C}^d$. Assuming for simplicity that the field is unidirectional, i.e. $E(t) = v(t)E_0$ for some unit vector E_0 and a scalar amplitude v, and denoting the eigenvalues of H_0 corresponding to the states $\psi_1, \ldots, \psi_d$ by $E_1, \ldots, E_d$, we obtain

$$i\partial_t \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_d \end{pmatrix} = \left(\begin{pmatrix} E_1 & 0 & \cdots & 0 \\ 0 & E_2 & \cdots & 0 \\ \vdots & & \vdots & \\ 0 & \cdots & & E_d \end{pmatrix} + v(t) \begin{pmatrix} 0 & \mu_{12} & \cdots & \mu_{1d} \\ \mu_{12}^* & 0 & \cdots & \mu_{2d} \\ \vdots & \vdots & \vdots & \\ \mu_{1d}^* & \mu_{2d}^* & \cdots & 0 \end{pmatrix} \right) \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_d \end{pmatrix}, \tag{17}$$

with the coupling matrix elements

$$\mu_{mn} = \langle \psi_m, -E_0 \cdot D\,\psi_n \rangle.$$

Note that the diagonal elements of the coupling matrix vanish, by the following simple argument: the Hamiltonian H_0 commutes with the parity operator $\mathcal{P}$ given by $(\mathcal{P}\psi)(x_1, s_1, \ldots, x_N, s_N) = \psi(-x_1, s_1, \ldots, -x_N, s_N)$, so the ψ_n can be chosen to be eigenstates of $\mathcal{P}$, i.e. either even or odd functions. But in both cases $|\psi|^2$ is an even function, whereas the dipole term $-E_0 \cdot D$ is odd, and so

$$\mu_{nn} = \langle \psi_n, -E_0 \cdot D\,\psi_n \rangle = \sum_{s_1,\ldots,s_N \in \mathbb{Z}_2} \int_{\mathbb{R}^{3N}} E_0 \cdot \sum_{i=1}^{N} x_i\, |\psi(x_1, s_1, \ldots, x_N, s_N)|^2 \,\mathrm{d}x_1 \cdots \mathrm{d}x_N = 0.$$

The same reasoning shows that the off-diagonal element μ_{mn} vanishes whenever ψ_m and ψ_n have the same parity. This is a simple example of an optical selection rule.

Example 2.3. (Vibrational states of a molecule in an electric field) The central quantum mechanical model treats the nuclei as nonrelativistic quantum particles and makes the Born-Oppenheimer approximation, that is to say one assumes that the electrons, on account of their much lighter mass, can always be assumed to be in the ground state with respect to the nuclear positions. Mathematically, this corresponds to the ansatz

$$\Psi(R, x_1, s_1, \ldots, x_N, s_N, t) \approx \Phi(R, t)\psi_0^{(R)}(x_1, s_1, \ldots, x_N, s_N) \tag{18}$$

where $R = (R_1, \ldots, R_M) \in \mathbb{R}^{3M}$ is the vector of position coordinates of the nuclei and (as in Example 2.2) the $x_i \in \mathbb{R}^3$ and $s_i \in \mathbb{Z}_2$ are position respectively spin coordinates of the electrons. The governing nuclear Schrödinger equation for M atoms with nuclear

masses $m_1, \ldots, m_M$ and nuclear charges $Z_1, \ldots, Z_M$ in an electric field $E\colon \mathbb{R} \to \mathbb{R}^3$ is the following equation for the nuclear wavefunction $\Phi \in \mathcal{H} = L^2(\mathbb{R}^{3M})$:

$$i\partial_t \Phi(t) = \Big(\sum_{\alpha=1}^{M} (-\frac{1}{2m_\alpha}\Delta_{R_\alpha} + E_0(R) - E(t)\cdot\mu(R)\Big)\Phi(t). \tag{19}$$

The potential $E_0\colon \mathbb{R}^{3M} \to \mathbb{R}$ is typically called *potential energy surface* (PES) in the context of the nuclear Schrödinger equation. Not just the PES, but also the coupling potential $\mu\colon \mathbb{R}^{3M} \to \mathbb{R}^3$ can be derived from the stationary electronic Schroedinger equation with nuclei clamped at $R_1, \ldots, R_M$. While the derivation of the PES in the absence of an electric field is just the familiar Born-Oppenheimer approximation discussed in many texts, the derivation of μ appears to be unknown in the mathematical literature, so we include it here. Let N be the number of electrons, i.e. $N = \sum_{\alpha=1}^{M} Z_\alpha$ (that the overall system of electrons and atomic nuclei is charge-neutral), let $\mathcal{H}_{e\ell}$ be the Hilbert space of antisymmetric electronic wavefunctions introduced in Example 2.2, and let $H_{e\ell}$ be the associated electronic Hamiltonian

$$H_{e\ell}^{(R)} = -\frac{1}{2}\Delta + V^{(R)}(x_1, \ldots, x_N) \tag{20}$$

where, as in Example 2.2, Δ is the Laplacian on $\mathbb{R}^{3N}$, but the potential is given, instead of the atomic potential (12), by the analogous molecular potential

$$V^{(R)}(x_1, \ldots, x_N) = -\sum_{i=1}^{N}\sum_{\alpha=1}^{M} \frac{Z_\alpha}{|x_i - R_\alpha|} + \sum_{1\le i<j\le N} \frac{1}{|x_i - x_j|} + \sum_{1\le \alpha<\beta\le M} \frac{Z_\alpha Z_\beta}{|R_\alpha - R_\beta|}. \tag{21}$$

The electronic Hamiltonian (20) depends parametrically on the nuclear coordinates R, through the appearance of the latter in V. The potential energy surface $E_0(R)$ appearing in the nuclear Schrödinger equation is the lowest eigenvalue of $H_{e\ell}^{(R)}$, and the coupling potential is the *dipole moment function* constructed with the help of the associated lowest electronic eigenstate:

$$\mu(R) = \Big\langle \psi_0^{(R)}, D^{(R)}\psi_0^{(R)} \Big\rangle_{\mathcal{H}_{e\ell}}, \tag{22}$$

where $\psi_0^{(R)} \in \mathcal{H}_{e\ell}$ is the (normalized) lowest eigenstate of $H_{e\ell}$ and $D^{(R)}$ is the dipole operator associated with the joint system of electrons and atomic nuclei,

$$D^{(R)}(x_1, \ldots, x_N) = \sum_{i=1}^{N} e x_i - \sum_{\alpha=1}^{M} e Z_\alpha R_\alpha. \tag{23}$$

Here, as noted before, the electronic charge e has the value -1 in atomic units.

Example 2.4. (Laser-guided chemical reactions) Our last model is of central interest in photochemistry, but to our knowledge has not hitherto been considered at all in the mathematical literature. A careful mathematical account in the absence of control fields

and for smooth interaction potentials is given in [Teu03]. The model is a combination of the variant (17) of Example 2.2 (finitely many electronic states coupled to a laser field) and Example 2.3 (atomic nuclei moving in the potential energy surface corresponding to the lowest electronic eigenstate). The typical situation in laser control of chemical reactions is the following (see Figure 3a in the numerics section). The system starts in a stationary state of electrons and nuclei. The laser then induces a transition to a different electronic state. As a result the nuclei now see a different potential energy surface with respect to which they are no longer in equilibrium; for instance the new surface may no longer contain a barrier to a desired target position. Once the nuclei have moved barrier-free to the target position, the laser induces a transition back to the original surface so as to also put the electrons in the target state. Mathematically, this situation can be modelled by generalizing the ansatz (18) to a finite number of electronic states. Confining ourselves for simplicity to two states, one assumes

$$\Psi(R, x_1, s_1, \ldots, x_N, s_N, t) \approx \Phi_1(R,t)\psi_1^{(R)}(x_1, s_1, \ldots, x_N, s_N) + \Phi_2(R,t)\psi_2^{(R)}(x_1, s_1, \ldots, x_N, s_N), \quad (24)$$

where Φ_1, $\Phi_2 \in L^2(\mathbb{R}^{3M})$, $||\Phi_1||^2 + ||\Phi_2||^2 = 1$, and $\psi_1^{(R)}$, $\psi_2^{(R)}$ are normalized eigenstates of the electronic Hamiltonian (20)–(21). This leads to the following Schrödinger equation in the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^{3M}; \mathbb{C}^2)$

$$i\partial_t \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix}(t) = \left(\sum_{\alpha=1}^{M} -\frac{1}{2m_\alpha}\Delta + \begin{pmatrix} E_1(R) & 0 \\ 0 & E_2(R) \end{pmatrix} + \begin{pmatrix} E(t)\cdot\mu_{11}(R) & E(t)\cdot\mu_{12}(R) \\ E(t)\cdot\mu_{12}(R)^* & E(t)\cdot\mu_{22}(R) \end{pmatrix} \right) \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix}, \quad (25)$$

with the dipole moment functions

$$\mu_{ij}(R) = \left\langle \psi_i^{(R)}, -D^{(R)}\psi_j^{(R)} \right\rangle_{\mathcal{H}_{e\ell}}. \quad (26)$$

3 Cost functionals and functional analytic setting

The primary goal of optimal control theory from the point of view of applications is to identify time profiles $v_\ell \colon [0,T] \to \mathbb{R}$ of the control fields which achieve a suitable goal, such as transfer of the system from one eigenstate of H_0 to another, and are convenient to implement experimentally. A general introduction to quantum control from a mathematical perspective is given in [D'A08]. Overviews of optimal quantum control from the application point of view are given in e.g. [WG07] and [HTK+12]. The theoretical work up to now has focused on simple cost functionals like the L^2 norm of the control field (or variants thereof like the L^1 or H^1 norm). But we may ask the question: *Why* L^2 (or variants thereof)?

In fact, while L^2 at first sight looks like a relevant choice, we will argue in this paper that it is not a very good choice. This is due to the fact that the L^2 norm squared is proportional to the total field energy expended. The ensuing controls lack a simple structure, they are notoriously difficult to interpret, and have experimentally undesirable features. This is discussed in detail in Section 5. Instead, we will argue that it pays off to use more sophisticated costs which are adapted to the subtle nature of quantum dynamics. As we will see, such costs naturally deliver controls which are, at the same time, *sparse in frequency*, picking out the system's Bohr frequencies without any special ansatz, while at the same time having *slowly varying amplitude envelopes*, thereby sharing an important feature of laser pulses designed by experimentalists which are commonly and successfully used in the laboratory rather than by optimal control.

We first state the general class of costs we propose, including its functional analytic setting. Subsequently we give examples. The simplest example of such a cost which yields nice frequency-sparse controls, and which motivated the general setting, is described in Example 3.1 below.

3.1 General setting

For the quantum system

$$\begin{aligned} i\partial_t\psi(t) &= \big(H_0 + \sum_{l=1}^{L} Bu(t)_l H_l\big)\psi(t) \\ \psi(0) &= \psi_0 \end{aligned} \tag{27}$$

we consider the optimal control problem

$$\min_{\psi,u} J(\psi,u) = \frac{1}{2}\langle\psi(T), \mathcal{O}\psi(T)\rangle_{\mathcal{H}} + \alpha\|u\|_{\mathcal{M}} \quad \text{s.t. } (27) holds. \tag{28}$$

The functional J consists of the term $\frac{1}{2}\langle\psi(T), \mathcal{O}\psi(T)\rangle_{\mathcal{H}}$ which describes the expectation value that needs to be minimized, and a cost term of the form $\alpha\|u\|_{\mathcal{M}}$ which contains a measure norm explained below and forces the solution to be sparse in a suitable sense.

We make the following very general functional-analytic assumptions on the operators and fields appearing in (27)–(28). These assumptions cover all the physical examples from Section 2, except for the infinite-dimensional version of Examples 2.3 and 2.4, which have an unbounded coupling operator.

1. Dynamics. Assume that the Hamiltonian H_0 is any self-adjoint operator on a Hilbert space $\mathcal{H}$ with domain $\mathcal{D}(H_0)$. The initial condition ψ_0 may be any element of $\mathcal{H}$. The coupling operators H_ℓ are assumed to be bounded self-adjoint operators on $\mathcal{H}$. We use the vector operator notation $v\cdot\tilde{H} = \sum_{l=1}^{L} v_l(-iH_l)$, $v\cdot\tilde{H}^* = \sum_{l=1}^{L} v_l(-iH_l)^*$, $\langle\chi_1, \tilde{H}\chi_2\rangle_{\mathcal{H}} = (\langle\chi_1, -iH_l\chi_2\rangle_{\mathcal{H}})_{l=1}^{L} \in \mathbb{R}^L$ for $v\in\mathbb{R}^L$ and $\chi_1,\chi_2\in\mathcal{H}$.

For the admissible class of controls u and control operators B_ℓ specified below, we will show that eq. (27) possesses a unique mild solution in the state space of continuous paths in the Hilbert space, $C([0,T];\mathcal{H})$.

2. Target constraint. The observable $\mathcal{O}$ specifying the target constraint can be any bounded linear operator on $\mathcal{H}$. Typically $\mathcal{O}$ is the orthogonal projection onto the subspace we want to reach. If $\mathcal{O}$ is a projection, the target constraint contribution $\frac{1}{2}\langle\psi(T), \mathcal{O}\psi(T)\rangle_{\mathcal{H}}$ to the cost lies in the interval $[0, 1/2]$. The value 0 correspond to a 100% achievement of the control objective, and the value 0.5 to a 0% achievement.

3. Control space, cost, measure norm. The control field u is assumed to belong to a measure space of form

$$\mathcal{M}(\Omega;\mathcal{U}), \tag{29}$$

where Ω is a locally compact space (typically a closed interval of admissible frequencies), and $\mathcal{U}$ is a separable Hilbert space of time-dependent functions (admitting general Hilbert spaces can be useful to obtain nice optimality conditions, see Example 3.3). The space (29) is the space of Borel measures u on Ω with values in $\mathcal{U}$ of finite mass norm $||u||_{\mathcal{M}}$. The mass norm of the measure u is the second term in the cost functional J in (28).

The space (29) is the dual of the space $C_0(\Omega;\mathcal{U})$ of continuous functions on Ω with values in $\mathcal{U}$ which can be uniformly approximated by functions with compact support. The duality pairing is given by

$$\langle u, \varphi\rangle_{\mathcal{M},\mathcal{C}_0} = \int_\Omega \langle u'(\omega), \varphi(\omega)\rangle_{\mathcal{U}} \,\mathrm{d}|u|(\omega) \tag{30}$$

(see [Mez09]) where u' is the Radon–Nikodym derivative of u with respect to the total variation measure $|u|$ (see [Lan93, VII Thm. 4.1]). Note that the inner product in the integral is the Hilbert space inner product in $\mathcal{U}$. This duality will be very useful.

4. Control operator. The control operator is assumed to be a bounded linear operator $B\colon \mathcal{M}(\Omega;\mathcal{U}) \to L^p(0,T;\mathbb{R}^L)$, for some $1 < p \leq \infty$. Moreover we assume that B has a bounded linear predual operator $B^*\colon L^q(0,T;\mathbb{R}^L) \to C_0(\Omega;\mathcal{U})$, by which we mean a bounded linear operator such that $\langle B^* f, u\rangle_{C_0,\mathcal{M}} = \langle f, Bu\rangle_{L^q,L^p}$ for all $f \in L^q(0,T;\mathbb{R}^L)$ and all $u \in \mathcal{M}(\Omega;\mathcal{U})$.

Existence of a bounded linear predual operators implies the weak-$*$–weak(-$*$) continuity of B. That is, weak-$*$ convergence of u_n to u in $\mathcal{M}(\Omega;\mathcal{U})$ implies weak convergence of Bu_n to Bu in $L^p(0,T;\mathbb{R}^L)$ if $p < \infty$, and weak-$*$ convergence if $p = \infty$.

Note that the operator B^* depends on the Hilbert space structure of $\mathcal{U}$, see examples below. Since B^* appears in the optimality system, the freedom to choose $\mathcal{U}$ can be used to generate nice optimal controls.

All spaces — possibly containing complex valued objects — are equipped with a real Banach or Hilbert space structure. That is, linear always means $\mathbb{R}$-linear and the scalar product is real-valued and $\mathbb{R}$-bilinear.

3.2 Examples

We now list some examples for choosing the frequency domain Ω, the Hilbert space $\mathcal{U}$ of time-dependent functions, and the control operator B. The first example is the prototype for generating sparse controls. It motivated the general functional analytic setting proposed above, and naturally incorporates physically relevant controls containing finitely many pulses of particular frequencies [BTS98, ABY$^+$, TLBR04, SSBK10, RSD$^+$11].

Example 3.1. (Time-smoothness with frequency-sparsity) Here u will be a time-frequency representation of the laser field amplitude and B will generate the corresponding field. The control u can a priori contain arbitrary and arbitrarily many active frequencies, each with its own smooth envelope. This can be modelled mathematically as follows. Let Ω be a closed subset of $\mathbb{R}^+$, $\mathcal{U} = H^1_0(0,T;\mathbb{C})$, and $p = \infty$. For $u \in L^1(\Omega; H^1_0(0,T;\mathbb{C}))$ we define B to be the synthesis operator

$$(Bu)(t) = \operatorname{Re} \int_\Omega u(\omega,t) e^{i\omega t} d\omega. \tag{31}$$

By approximation, the expression can be extended to measure-valued controls. This extension is important in practice, because it allows sharp concentration on a small number of frequencies, and has the following mathematically rigorous integral representation:

$$(Bu)(t) = \operatorname{Re} \int_\Omega u'(\omega,t) e^{i\omega t} \,\mathrm{d}|u|(\omega), \tag{32}$$

where u' is the Radon–Nikodym derivative of u with respect to $|u|$.

We remark that this setting naturally contains the physically motivated finite-dimensional ansatz spaces of [ABY$^+$, TLBR04, SSBK10] in which a finite number of frequencies can be switched on or off by few-parameter modulation functions: the field

$$v(t) = \sum_{j=1}^{n} v_n(t) \cos(\omega_n t + \phi_n)$$

corresponds to Bu with B as in (32) and

$$u(\omega,t) = \sum_{j=1}^{n} \delta(\omega - \omega_n) e^{i\phi_n} v_n(t).$$

We allow $\mathcal{U}$ to contain complex-valued functions. This allows phase shifts in the different frequencies without leaving the linear setting.

The predual operator B^* is the solution operator of the second-order boundary value problem

$$\frac{\partial^2}{\partial t^2} u(\omega,t) = f(t) e^{-i\omega t}, \quad u\Big|_{t=0} = u\Big|_{t=T} = 0, \tag{33}$$

i.e. $B^* f = u$. The equations (33) are not coupled for different ω. The operator B^* is continuous and has the additional regularity $B^* f \in C_0(\Omega; H^2 \cap H^1_0)$. Here it is important that Ω is closed. Otherwise B might not be weak-$*$–weak(-$*$) continuous.

Frequency constraints can easily be achieved by restricting Ω. The framework also allows for discrete Ω containing frequencies which are selected in advance by some other method, compare [LSTT09, KHK10]. An advantage of our appproach as compared to the latter works is that no a priori knowledge of suitable frequencies is needed.

Example 3.2. (Dual Gabor transform) In this example we design the control operator B and the Hilbert space $\mathcal{U}$ so that the predual operator B^* has a nice form. We take $\Omega \subset \mathbb{R}^+$, $\mathcal{U} = L^2(0,T;\mathbb{C})$, and define B by

$$(Bu)(t) = \operatorname{Re} \int_\Omega \int_0^T k(s,t) u'(\omega, s) \,\mathrm{d}s \, e^{i\omega t} \,\mathrm{d}|u|(\omega), \tag{34}$$

where $k\colon [0,T]^2 \to \mathbb{R}$ is a smooth symmetric kernel. Roughly, this operator corresponds to a 'pre-processing' of the envelopes, with only smoothed envelopes entering the equation. The predual operator B^* becomes the *Gabor transformation*

$$(B^* f)(\omega, t) = \int_0^T f(s) k(t,s) e^{-i\omega s} \,\mathrm{d}s, \tag{35}$$

which is a useful time-frequency representation of the control field.

Example 3.1 above with $\mathcal{U} = H_0^1$ led to a time-frequency representation $u = B^* f$ which is global in time. That is, it has a global window equal to the Green's function G of the one-dimensional boundary value problem (33),

$$(B^* f)(\omega, t) = \int_0^T f(s) G(t,s) e^{-i\omega s} \,\mathrm{d}s. \tag{36}$$

On the other hand, Example 3.2 leads to a time-frequency representation which is local in time, but the definition of B is somewhat complicated. Can one choose a better space $\mathcal{U}$ such that the control operator B can be defined as the simple synthesis operator of Example 3.1, but results in a local-in-time predual B^* as in Example 3.2? In fact, Example 3.2 can be brought into this form, see Example 3.3.

Example 3.3. (Time-locality with frequency-sparsity) Let Ω and B be as in Example 3.1. Furthermore let $K\colon L^2(0,T) \to L^2(0,T)$ be the compact operator given by convolution with a Gaussian kernel k,

$$(Kf)(t) = \int_0^T k(t,s) f(s) \,\mathrm{d}s. \tag{37}$$

Then K is injective and self-adjoint and has an unbounded inverse $A\colon \mathcal{D}(A) \to L^2$. Define $\mathcal{U} := \mathcal{U}_k := \mathcal{D}(A^{1/2})$ with the induced scalar product $\langle u, v\rangle_{\mathcal{U}} = \langle A^{1/2}u, A^{1/2}v\rangle_{L^2}$. Then $\mathcal{U}_k$ is a Hilbert space and the predual operator B^* of B is

$$(B^* f)(\omega, t) = \Big(K^2(f e^{-i\omega \cdot})\Big)(s).$$

This construction also works for window functions other than a Gaussian.

The equivalence to Example 3.2 follows since the dual of the map $X\colon C_0(\Omega; L^2(0,T;\mathbb{C})) \to C_0(\Omega; \mathcal{U}_k)$ defined by $(X\varphi)(\omega) = K\varphi(\omega)$ is an isometric isomorphism between the control spaces that preserves the image unter the corresponding control operators.

Our next example shows that our setting also covers interesting cases when $\mathcal{U}$ does not contain time-dependent functions.

Example 3.4. (Sparse Fourier transform) Let $\Omega \subset \mathbb{R}^+$, $\mathcal{U} = \mathbb{C}$, and let B be the Fourier synthesis operator, that is to say

$$(Bu)(t) = \operatorname{Re} \int_\Omega u(\omega) e^{i\omega t} d\omega \Big(= \operatorname{Re} \int_\Omega u'(\omega) e^{i\omega t} \,\mathrm{d}|u|(\omega) \Big). \tag{38}$$

The predual operator is, up to a constant factor, the Fourier transform of functions restricted to $[0, T]$,

$$(B^* f)(\omega) = \int_0^T f(t) e^{-i\omega t} \,\mathrm{d}t.$$

An alternative approach to achieving sparsity of a time-global frequency decomposition via an L^2 cost combined with iterative 'frequency sifting' is given in [RBKM$^+$06].

Example 3.5. (Time-frequency-sparse Gabor transform) Let $\Omega \subset \mathbb{R}^+ \times [0, T]$ be a subset of time-frequency space, $\mathcal{U} = \mathbb{C}$, and

$$(Bu)(t) = \operatorname{Re} \int_\Omega u'(\omega, s) g_{\omega,s}(t) \,\mathrm{d}|u|(\omega, t) \tag{39}$$

for the ansatz function

$$g_{\omega,s}(t) = k(t, s) e^{i\omega(t-s)} \tag{40}$$

with a Gaussian kernel k. This defines a suitable extension of the control operator from Example 3.2 to measures in the time-freqeuency plane. With this control operator, each Dirac measure $u = \delta_{\omega,t}$ corresponds to a Gaussian wave packet centered at time t with frequency ω. The predual of the control operator is given by

$$(B^* f)(\omega, t) = \int_0^T \overline{g_{\omega,t}(s)} \, f(s) \,\mathrm{d}s.$$

These examples by no means exhaust our framework, but are meant to give an idea of its flexibility.

4 Theory of the optimal control problem

In this section we will study the optimal control problem (28). We first show well-posedness of the problem. In contrast to [KPV14] the main difficulty does not lie in the low regularity of the control since we assume sufficient smoothing of the control operator. It rather lies in the bilinearity of the state equation together with the low regularity of the state. Subsequently we derive necessary optimality conditions. We shall show that the support of the optimal measure can be influenced. Proposition 4.7 is the natural analog of Theorem 2.12 in [KPV14]. Differences arise due to the bilinearity of the equation and the non-trivial control operator. We shall also note interesting relationships between the choices for B and $\mathcal{U}$. Throughout this section we will stay in the setting of mild solutions. This suggests to develop a derivation of the necessary optimality conditions which only requires integral manipulations and no further regularity discussion for the primal and dual state are necessary.

4.1 Existence of solutions to the state equation

In this section we study existence of mild solutions of (27) and compactness properties of the corresponding solution operator. Throughout this subsection we will consider mild solutions for the control field $v \in L^1(0,T;\mathbb{R}^L)$. Thus we study $\psi \in C([0,T];\mathcal{H})$ which for the field v, satisfies

$$\psi(t) = G(t)\psi_0 + \int_0^t G(t-s)v(s)\cdot\tilde{H}\psi(s)\,\mathrm{d}s. \tag{41}$$

Here G is the unitary group generated by the skew-adjoint operator $-iH_0$. We need the following version of [BMS82, Theorem 2.5]. It provides existence of solutions of (41) as well as their differentiability in the direction of the control.

Proposition 4.1. *Let $v \in L^1(0,T;\mathbb{R}^L)$. Then (41) possesses a unique solution $\psi \in C([0,T];\mathcal{H})$. Furthermore the mapping $F\colon L^1(0,T;\mathbb{R}^L) \to C([0,T];\mathcal{H})$ defined by $F(v) = \psi$ is continuously differentiable. The derivative is given by $F'(v)(\delta v) = \psi'$ where ψ' solves the equation*

$$\psi'(t) = \int_0^t G(t-s)v(s)\cdot\tilde{H}\psi'(s)\,\mathrm{d}s + \int_0^t G(t-s)\delta v(s)\cdot\tilde{H}\psi(s)\,\mathrm{d}s. \tag{42}$$

For fixed v one can define the evolution operator $\mathfrak{G}\colon \{\ (s,t)\mid 0\le s\le t\le T\ \}$ by $\mathfrak{G}(t,s)\psi(s) = \psi(t)$. Using the evolution operator one can reformulate (42) as

$$\psi'(t) = \int_0^t \mathfrak{G}(t,s)v(s)\cdot\tilde{H}\psi(s)\,\mathrm{d}s, \tag{43}$$

see e.g. [LY95, Chapter 2]. We will also look at solutions of the mild adjoint equation,

$$\varphi(s) = G(T-s)^*\varphi_T + \int_s^T G(t-s)^* v(s)\cdot\tilde{H}^*\varphi(t)\,\mathrm{d}t. \tag{44}$$

The solution of this backwards-in-time equation satisfy $\varphi(s) = \mathfrak{G}(t,s)^*\varphi(t)$.

The next lemma addresses the norm preservation of solutions. In contrast to the result of the preceding proposition it makes explicit use of the self-adjointness of the Hamiltonian and the coupling operators. Since we work in the setting of mild solutions the usual proof of norm preservation does not hold since the state is not differentiable in a suitable sense. One can overcome this difficulty by approximation arguments. We will give a proof based on a primal-dual argument that stays in the setting of mild solutions.

Lemma 4.2. *The solution ψ of (41) satisfy $\|\psi(t)\| = 1$ for all $t \in [0,T]$.*

Proof. We will show the time reversibility of the time evolution, that is

$$\psi_0 = \mathfrak{G}(t,0)^*\psi(t) \tag{45}$$

for all $t \in [0,T]$. Then we obtain

$$\|\psi(t)\|^2 = \langle \mathfrak{G}(t,0)\psi_0, \psi(t)\rangle = \langle \psi_0, \mathfrak{G}(t,0)^*\psi(t)\rangle = \|\psi_0\|^2 = 1.$$

To show (45) we define $\varphi(s) = \mathfrak{G}(t,s)^*\psi(t)$. Then φ solves

$$\varphi(s) = G(t-s)^*\psi(t) + \int_s^t G(r-s)^* v(r) \cdot \tilde{H}^* \varphi(r)\,\mathrm{d}r. \tag{46}$$

We show that $\psi = \varphi$. For the difference we obtain, using the group property of G and the skewadjointness of $\tilde{H}$,

$$\begin{aligned}
\psi(s) - \varphi(s) =\ & G(s)\psi_0 + \int_0^s G(s-r)v(r) \cdot \tilde{H}\psi(r)\,\mathrm{d}r \\
& - G(t-s)^*\psi(t) - \int_s^t G(r-s)^* v(r) \cdot \tilde{H}^* \varphi(r)\,\mathrm{d}r \\
=\ & G(s)\psi_0 + \int_0^s G(s-r)v(r) \cdot \tilde{H}\psi(r)\,\mathrm{d}r \\
& - G(t-s)^* \left(G(t)\psi_0 + \int_0^t G(t-r)v(s) \cdot \tilde{H}\psi(r)\,\mathrm{d}r \right) \\
& - \int_s^t G(r-s)^* v(r) \cdot \tilde{H}^* \varphi(r)\,\mathrm{d}r \\
=\ & - \int_s^t G(r-s)^* v(r) \cdot \tilde{H}\Big(\psi(r) - \varphi(r)\Big)\,\mathrm{d}r.
\end{aligned}$$

Therefore we obtain

$$\|\psi(s) - \varphi(s)\| \le \int_s^t \sum_{l=1}^{L} |v(r)_l| \|H_l\| \|\psi(r) - \varphi(r)\|\,\mathrm{d}r.$$

Applying Gronwall's inequality to $f(s) = \|\psi(t-s) - \varphi(t-s)\|$ yields $\psi(s) = \varphi(s)$ for all $s \in [0,t]$. Thus we have

$$\psi_0 = \varphi(0) = \mathfrak{G}(t,0)^*\varphi(t) = \mathfrak{G}(t,0)^*\psi(t)$$

which is (45). □

Next we establish the following compactness property for mild solutions.

Proposition 4.3. *Let $(v_n)_n$ be a sequence in $L^1(0,T;\mathbb{R}^L)$ such that $v_n \rightharpoonup v$. Then the corresponding solutions ψ_n of (41) satisfy $\psi_n \to \psi$ in $C([0,T];\mathcal{H})$, where ψ is the mild solution corresponding to v.*

Proof. The proof is split up into three steps. First we will show existence of a pointwise weak limit for all $t \in [0,T]$ for some subsequence of $(\psi_n)_n$ using the Arzelà–Ascoli theorem. Then we prove that this pointwise limit is the mild solution corresponding to the limit field. Then the uniform strong convergence of the original sequence $(\psi_n)_n$ is established.

Step 1. Let $\chi \in \mathcal{H}$. We need to verify the boundedness and equicontinuity assumptions of the Arzelà–Ascoli theorem for the family $t \mapsto \langle \chi, \psi_n(t) \rangle$. The boundedness

follows from Lemma 4.2 which gives $\langle \chi, \psi_n(t)\rangle \leq \|\chi\|$. For the equicontinuity we observe that, for $t' > t$,

$$\psi_n(t') - \psi_n(t) = (G(t'-t) - I)\psi_n(t) + \int_t^{t'} G(t'-s)v_n(s) \cdot \tilde{H}\psi_n(s)\,\mathrm{d}s. \tag{47}$$

This gives

$$\begin{aligned}|\langle \chi, \psi_n(t') - \psi_n(t)\rangle| &\leq \|(G(t'-t)^* - I)\chi\|\|\psi_n(t)\| + \int_t^{t'} \|v_n(s)\|_{\mathbb{R}^L}\|\tilde{H}\|\|\psi_n(s)\|\mathrm{d}s \\ &\leq \|(G(t'-t)^* - I)\chi\| + \|\tilde{H}\| \int_t^{t'} \|v_n(s)\|_{\mathbb{R}^L}\,\mathrm{d}s.\end{aligned}$$

Since G is strongly continuous and $(v_n)_n$ is equiintegrable as weakly converging sequence in L^1, we can for every $\epsilon > 0$ find a $\delta > 0$ such that $|t' - t| < \delta$ implies

$$|\langle \chi, \psi_n(t') - \psi_n(t)\rangle| < \epsilon.$$

for all $n \in \mathbb{N}$. The Arzelà–Ascoli theorem then gives the existence of a limit $a_\chi \in C([0,T];\mathbb{R})$ and the uniform convergence $\langle \chi, \psi_n(t)\rangle \to a_\chi$ for some subsequence still denoted by $(\psi_n)_n$. To prove pointwise weak convergence of $\psi_n(t)$ it remains to show the existence of $\psi \in L^\infty(0,T;\mathcal{H})$ such that $a_\chi = \langle \chi, \psi(t)\rangle$ for all $\chi \in \mathcal{H}$. Set $\psi(t) = \sum_\iota a_{\chi_\iota}\chi_\iota$ for some orthonormal basis (χ_ι). Then for $\chi = \sum_\iota b_\iota \chi_\iota$ we have

$$\langle \chi, \psi(t)\rangle = \sum_\iota b_\iota a_{\chi_\iota} = \lim_{n\to\infty} \langle \sum_\iota b_\iota \chi_\iota, \psi_n(t)\rangle = a_\chi(t).$$

Thus,

$$\psi_n(t) \rightharpoonup \psi(t) \tag{48}$$

for all $t \in [0,T]$.

Step 2. We will now show that ψ is the mild solution for the control field v. Let $\chi \in \mathcal{H}$. We know

$$\langle \chi, \psi_n(t)\rangle = \langle \chi, G(t)\psi_0\rangle + \sum_{l=1}^{L} \langle \chi, \int_0^t G(t-s)v_n(s)_l(-iH_l)\psi_n(s)\,\mathrm{d}s\rangle \tag{49}$$

for all $n \in \mathbb{N}$ and $t \in [0,T]$, and will pass to the limit $n \to \infty$ to obtain

$$\langle \chi, \psi(t)\rangle = \langle \chi, G(t)\psi_0\rangle + \sum_{l=1}^{L} \langle \chi, \int_0^t G(t-s)v(s)_l(-iH_l)\psi(s)\,\mathrm{d}s\rangle. \tag{50}$$

For the term on the left-hand side of (49) we use (48). For each summand of the second term on the right we define $\tilde{\chi}_l(s) := (-iH_l)^* G(t-s)^*\chi$. Then

$$\langle \chi, \int_0^t G(t-s)v_n(s)_l(-iH_l)\psi_n(s)\,\mathrm{d}s\rangle = \int_0^t v_n(s)_l\langle \tilde{\chi}_l(s), \psi_n(s)\rangle\,\mathrm{d}s.$$

We have $\langle \tilde{\chi}_l(s), \psi_n(s)\rangle \to \langle \tilde{\chi}_l(s), \psi(s)\rangle$ pointwise in s by (48), and also $\langle \tilde{\chi}_l(s), \psi_n(s)\rangle \leq \|\tilde{H}\|\|\chi\|$. Together with the assumption $(v_n)_l \rightharpoonup (v)_l$ in $L^1(0,T)$ this implies $(v_n)_l\langle \tilde{\chi}_l, \psi_n\rangle \rightharpoonup (v)_l\langle \tilde{\chi}_l, \psi\rangle$ in $L^1(0,T)$, see [FL07, Proposition 2.61]. Thus (50) holds. Since χ was arbitrary, the tested equation (50) implies

$$\psi(t) = G(t)\psi_0 + \sum_{l=1}^{L} \int_0^t G(t-s)v(s)_l(-iH_l)\psi(s)\,\mathrm{d}s$$

for all $t \in [0,T]$, i.e. $\psi \in C([0,T];\mathcal{H})$ is a mild solution of (27) for the control field $v \in L^1(0,T;\mathbb{R}^L)$.

Step 3. Since ψ_n and ψ are solutions of (41) we have $\|\psi(t)\| = 1 = \|\psi(t)\|$ by Lemma 4.2. Thus for every $t \in [0,T]$, weak convergence $\psi_n(t) \rightharpoonup \psi(t)$ and convergence of the norms inplies strong convergence $\psi_n(t) \to \psi(t)$ in $\mathcal{H}$. By (47) we have for $t', t \in [0,T]$,

$$\begin{aligned}\|\psi_n(t') - \psi_n(t)\| &\leq \|(G(t'-t) - I)\psi_n(t)\| + \|\tilde{H}\| \int_t^{t'} \|v_n(s)\|_{\mathbb{R}^L}\,\mathrm{d}s \\ &\leq \|(G(t'-t) - I)\psi(t)\| + \|(G(t'-t) - I)\|\|\psi_n(t) - \psi(t)\| \\ &\qquad + \|\tilde{H}\| \int_t^{t'} \|v_n(s)\|_{\mathbb{R}^L}\,\mathrm{d}s.\end{aligned}$$

Pointwise strong convergence $\psi_n(t) \to \psi(t)$, in addition to the arguments in the first step of the proof, then restults in the following weak version of equicontinuity: For $t \in [0,T]$ and $\epsilon > 0$ there exists a $\delta > 0$ and an $N \in \mathbb{N}$ such that for all t' with $|t' - t| < \delta$ we have $\|\psi_n(t') - \psi_n(t)\| < \epsilon$. In particular, it implies that $\psi_n(t_n) \to \psi_n(t)$ for convergent sequences $t_n \to t$.

We can now conclude the proof with a contradition argument. Suppose uniform convergence $\psi_n \to \psi$ does not hold. Then there is an $\epsilon > 0$ and a subsequence still denoted by $(\psi_n)_n$ and a sequence $(t_n)_n$ in $[0,T]$ with

$$\|\psi_n(t_n) - \psi(t_n)\| > \epsilon$$

for all $n \in \mathbb{N}$. Extracting another subsequence we can assume $t_n \to t$ for some $t \in [0,T]$. Then

$$\|\psi_n(t_n) - \psi(t_n)\| \leq \|\psi_n(t_n) - \psi_n(t)\| + \|\psi_n(t) - \psi(t)\| + \|\psi(t) - \psi(t_n)\|.$$

Since the three terms on the right converge to zero for $n \to \infty$ by equicontinuity of $(\psi_n)_n$, pointwise convergence $\psi_n(t) \to \psi(t)$ and continuity of ψ, we obtain a contradiction. Therefore the uniform convergence $\psi_n \to \psi$ in $C([0,T];\mathcal{H})$ holds. □

Remark. We emphasize that we did not make use of compact embeddings to obtain this compactness result. Instead we used the Arzelà–Ascoli theorem. In the second step of the proof we made heavy use of the separability of the coupling into a time dependent control function and a spatial coupling operator.

4.2 Existence of optimal controls

We will now prove the existence of solutions to problem (28).

Theorem 4.4. *There exists a solution* $(\bar{\psi}, \bar{u}) \in C([0,T];\mathcal{H}) \times \mathcal{M}(\Omega;\mathcal{U})$ *of* (28).

Proof. Let (ψ_n, u_n) be a minimizing sequence,

$$\lim_n J(\psi_n, u_n) = \inf J(\psi, u). \tag{51}$$

Sice $\mathcal{O}$ is bounded and $\alpha > 0$, the sequence (u_n) in $\mathcal{M}(\Omega;\mathcal{U})$ is bounded. Therefore, and because $\mathcal{M}(\Omega;\mathcal{U}) = C(\Omega;\mathcal{U})^*$ is the dual of a separable Banach space, the sequence $(u_n)_n$ has a weak-$*$ convergent subsequence still denoted by $(u_n)_n$ with limit $\bar{u} \in \mathcal{M}(\Omega;\mathcal{U})$. The weak-$*$–weak(-$*$) continuity of B implies Bu_n converges to $B\bar{u}$ weakly in $L^p(0,T;\mathbb{R}^L)$ for some $p > 1$ and thus also for $p = 1$. By Proposition 4.3 the corresponding sequence of states $(\psi_n)_n$ satisfies $\psi_n(T) \to \bar{\psi}(T)$. Thus the first summand of J converges, $\langle \psi_n(T), \mathcal{O}\psi_n(T)\rangle \to \langle \bar{\psi}(T), \mathcal{O}\bar{\psi}\rangle$. The second summand of J is weak-$*$ lower semi-continuous as it is a norm in a dual space. Thus we obtain $\lim_n J(\psi_n, u_n) \geq J(\bar{\psi}, \bar{u})$. Together with (51) this implies the claim. □

4.3 Necessary optimality conditions

For theoretical and numerical purposes we will use the reduced form of (28),

$$\min_u j(u). \tag{52}$$

Here $j(u) = J(\psi(u), u)$, where $\psi(u)$ denotes the solution of (27) for the control u.

The reduced cost functional can be split into two parts. A nonlinear differentiable part

$$f(v) = \langle \psi(T), \mathcal{O}\psi(T)\rangle_{\mathcal{H}} \tag{53}$$

with $v = Bu$, and a nondifferentiable convex part

$$g(u) = \alpha \|u\|_{\mathcal{M}}.$$

In the next lemma we will see that the derivative of the differentiable part f is given by

$$f'(v) = \langle \varphi, \tilde{H}\psi\rangle_{\mathcal{H}}$$

where φ is the mild solution of the dual equation

$$\begin{aligned} i\partial_t \varphi(t) &= (H_0 + v(t) \cdot i\tilde{H})\varphi(t), \\ \varphi(T) &= \mathcal{O}\psi(T). \end{aligned} \tag{54}$$

Using the evolution operator this can be rewritten as

$$\varphi(t) = \mathfrak{G}(T,t)^* \mathcal{O}\psi(T). \tag{55}$$

Under suitable assumptions, the representation (53) of f' can be derived using a Lagrange functional approach, see [PDR88, vWBV10]. We will give a short proof in the setting of mild solutions.

Lemma 4.5. *Let $v, \delta v \in L^1(0,T;\mathbb{R}^L)$, and let ψ, ψ' and φ be the corresponding solutions of (41), (42) and (55), respectively. Then the mapping $f\colon L^1(0,T;\mathbb{R}^L) \to \mathbb{R}$ defined by*

$$f(v) = \frac{1}{2}\langle \psi(T), \psi(T)\rangle$$

is continuously differentiable with derivative

$$f'(v)(\delta v) = \int_0^T \delta v(t) \cdot \langle \varphi(t), \tilde{H}\psi(t)\rangle_{\mathcal{H}} \,\mathrm{d}t. \tag{56}$$

Proof. Proposition 4.1 and the product rule give continuous differentiability of f and

$$f'(v)(\delta v) = \langle \mathcal{O}\psi(T), \psi'(T)\rangle. \tag{57}$$

Using (43) and (55) we obtain

$$\begin{aligned}
\langle \mathcal{O}\psi(T), \psi'(T)\rangle_{\mathcal{H}} &= \langle \mathcal{O}\psi(T), \int_0^T \mathfrak{G}(T,t)\delta v(t) \cdot \tilde{H}\psi(t)\,\mathrm{d}t\rangle_{\mathcal{H}} \\
&= \int_0^T \delta v(t) \cdot \langle \mathfrak{G}(T,t)^* \mathcal{O}\psi(T), \tilde{H}\psi(t)\rangle_{\mathcal{H}} \,\mathrm{d}t \\
&= \int_0^T \delta v(t) \cdot \langle \varphi(t), \tilde{H}\psi(t)\rangle_{\mathcal{H}} \,\mathrm{d}t.
\end{aligned}$$

□

We can now derive the following optimality condition, compare [KPV14].

Proposition 4.6. *Let $\bar{u}$ be a minimizer of problem (52), and let $\bar{\psi}, \bar{\varphi} \in \mathcal{C}([0,T];\mathcal{H})$ be the corresponding solutions of (41) and (55) for the control field $B\bar{u}$. Then*

$$\alpha\|\bar{u}\|_{\mathcal{M}} = -\langle B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}}, \bar{u}\rangle_{C_0,\mathcal{M}}. \tag{58}$$

and

$$\|B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}}\|_{C_0} \le \alpha. \tag{59}$$

Proof. Since we can split our functional into a sum of a nonconvex and a nonsmooth part, the result can be deduced from the very general differential calculus of Clarke [Cla90]. Because this calculus is somewhat intricate and not common knowledge beyond the mathematical control theory community, and because of the importance of the optimality system, we prefer to give a more elementary proof.

Let $\bar{u}$ be a minimizer of problem (52) and let $\bar{\psi}$ and $\bar{\varphi}$ be the corresponding solutions of (41) and (55). We first show the variational inequality

$$g(\bar{u}) - f'(B\bar{u})(Bu - B\bar{u}) \le g(u). \tag{60}$$

Since $\bar{u}$ is optimal, we have

$$\frac{1}{h}\big(j(\bar{u} + h(u - \bar{u})) - j(\bar{u})\big) \ge 0$$

for $u \in \mathcal{M}(\Omega;\mathcal{U})$ and $h \in (0,1)$. Using the decomposition $j = f \circ B + g$ and convexity of g, this implies

$$\frac{1}{h}\big(f(B\bar{u} + h(Bu - B\bar{u})) - f(B\bar{u})\big) + g(u) - g(\bar{u}) \geq 0.$$

Since f is differentiable, taking the limit $h \to 0$ yields (60).

Testing (60) with $u = 0$ and $u = 2\bar{u}$ gives

$$g(\bar{u}) + f'(B\bar{u})(B\bar{u}) = 0. \tag{61}$$

Substituting (61) into (60) gives

$$-f'(B\bar{u})(Bu) \leq g(u) \tag{62}$$

for all $u \in \mathcal{M}(\Omega;\mathcal{U})$.

Using Lemma 4.5 on the derivative of f, equation (61) gives

$$g(\bar{u}) = -\langle\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}}, B\bar{u}\rangle_{L^q,L^p} = -\langle B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}}, \bar{u}\rangle_{C_0,\mathcal{M}}$$

which proves (58). For (62) we obtain

$$-\langle B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}}, u\rangle_{C_0,\mathcal{M}} \leq \alpha\|u\|_{\mathcal{M}}.$$

Testing this inequality with $u = -\delta_\omega(B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}})(\omega)$ for some $\omega \in \Omega$ yields

$$\|(B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}})(\omega)\|_{\mathcal{U}}^2 \leq \alpha\|(B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}})(\omega)\|_{\mathcal{U}}$$

which gives (59). □

Remark. Proposition 4.6 provides only a necessary condition for local optimality. Due to the nonlinear structure of the problem (28), we expect that there also exist non-optimal critical points of j, as well as local optima that are not global.

Proposition 4.6 implies the following interesting restrictions on the support and direction of the optimal measure.

Proposition 4.7. *Let $\bar{u}$, $\bar{\psi}$ and $\bar{\varphi}$ be as in Proposition 4.6. Then we have*

$$\operatorname{supp}|\bar{u}| \subset \{\, \omega \in \Omega \mid \|(B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}})(\omega)\|_{\mathcal{U}} = \alpha \,\}, \tag{63}$$

$$-\alpha\bar{u}'(\omega) = (B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}})(\omega), \quad |\bar{u}|\text{-almost everywhere.} \tag{64}$$

Proof. Writing equation (58) as an integral yields

$$\int_\Omega \Big(\alpha + \langle (B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}})(\omega), \bar{u}'(\omega)\rangle_{\mathcal{U}}\Big)\, \mathrm{d}|\bar{u}|(\omega) = 0. \tag{65}$$

For the integrand we obtain by the Cauchy–Bunyakovsky–Schwarz (CBS) inequality, using $\|\bar{u}'(\omega)\| = 1$ and (59),

$$\alpha + \langle (B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}})(\omega), \bar{u}'(\omega)\rangle_{\mathcal{U}} \geq \alpha - \|(B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}})(\omega)\|_{\mathcal{U}} \geq 0. \tag{66}$$

Therefore (65) yields

$$\alpha + \langle (B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}})(\omega), \bar{u}'(\omega)\rangle_{\mathcal{U}} = 0$$

for $|\bar{u}|$-almost all $\omega \in \Omega$. For those ω the CBS inequality in (66) was sharp. This implies (64) and

$$\|(B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}})(\omega)\|_{\mathcal{U}} = \alpha$$

for $|\bar{u}|$-almost all $\omega \in \Omega$. Since $\omega \mapsto \|(B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}})(\omega)\|_{\mathcal{U}}$ is continuous this implies (63). □

The relation (63) for the support of the optimal measure gives us the following corollary.

Corollary 4.8. *Let $\bar{u}$ be a local minimizer of* (52)*. Then* supp$|\bar{u}|$ *is compact.*

Proof. Since $B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle \in C_0(\Omega;\mathcal{U})$ we know that there is a compact set $K \subset \Omega$ such that $B^*\langle \bar{\varphi}, \tilde{H}\bar{\psi}\rangle \leq \alpha/2$ for all $\omega \notin K$. Using (63) this implies supp$|\bar{u}| \subset K$. Therefor supp$|\bar{u}|$ is compact as a closed subset of the compact set K. □

This corollary is of significant physical interest. It says that although the frequency domain Ω might be unbounded, optimal solutions will always have bounded support. Controls from experiments, of course, always have this property, because arbitrarily fast oscillations are not realizable. But it is nice to know that such oscillations do not occur in our theoretical controls, as well.

For specific control operators B, equation (64) implies additional regularity for $\bar{u}'(\omega)$. For example, in the case of the control operator from Example 3.1 we obtain the regularity $\bar{u}'(\omega) \in H^2(0,T)$. Additional regularity with respect to $\omega \in \Omega$ is an open problem even in the simpler case of linear optimal control problems.

5 Numerical results

In this section we apply the framework for sparse time-frequency control to different quantum systems. First we will describe our numerical approach. This includes a short discussion of the discretization and the regularization of the optimal control problem. Then we will present two examples. The first example is the control of the finite-dimensional system from Example 2.2. It serves to illustrate the basic effects of sparse control of quantum systems. The second example addresses controlling the system of one dimensional Schrödinger equations from Example 2.4. The focus in this more challenging example will be the effect of different control spaces on the resulting optimal controls.

5.1 Numerical approach

Our numerical approach relies on the following three steps. First, we discretize the measure space by a finite sum of Dirac measures with values in a finite-dimensional Hilbert space. Then, we Huber-regularize the corresponding finite-dimensional nonsmooth problem. Finally, we solve the resulting smooth optimization problem with a quasi-Newton method.

The first step depends on Ω, of course. In our examples we always have $\Omega \subset \mathbb{R}^k$ for $k \in \{1, 2\}$ and Ω is an interval or the product of two intervals. We fix a uniform (tensor) grid Ω^h and choose measures supported at those points as our ansatz space. Those measures can always be written as finite sums of Dirac measures. In this discrete setting the measure norm reduces to an ℓ^1 norm for the coefficients from $\mathcal{U}$ multiplying the Dirac measures. To obtain a discrete problem we also need to discretize the Hilbert space $\mathcal{U}$ and the quantum system. In our examples we have $\mathcal{U} = H_0^1(0, T; \mathbb{C})$ or $\mathcal{U} = L^2(0, T; \mathbb{C})$ where we use piecewise linear finite elements on a uniform grid with the appropriate discrete norms, or $\mathcal{U} = \mathbb{C}$ were we do not need to discretize. The discrete Hilbert space is denoted by $\mathcal{U}^h$. The control operator B^h maps discrete controls to piecewise linear function. The discretization of the quantum system depends heavily on the system at hand. We approximate the time evolution of the discretized quantum system by a generalized Suzuki–Trotter method, see [HL15]. Together we obtain a finite-dimensional optimization problem that is non-smooth and non-convex.

To deal with the nondifferentiability of the norm at the origin we Huber-regularize this non-smooth problem. We replace the norm of $\mathcal{U}$ in the $\ell^1(\mathcal{U}^h)$ norm by the function $h\colon \mathcal{U} \to \mathbb{R}$ given by

$$h(z) = \begin{cases} \|z\|_{\mathcal{U}} - \frac{\theta}{2}, & \text{if } \|z\|_{\mathcal{U}} > \theta, \\ \frac{1}{2\theta}\|z\|_{\mathcal{U}}^2, & \text{if } \|z\|_{\mathcal{U}} \leq \theta, \end{cases}$$

for some regularization parameter θ. The function h has the following two important properties: it is smooth, and the derivatives of h and the derivatives of the norm of $\mathcal{U}$ have the same behavior outside of a small neighborhood of zero. The first property makes the optimization problem smooth. The second property preserves the possibility of strongly peaked solutions. The smoothness, however, comes at the cost of relaxing the support condition from Proposition 4.7. Instead of (63) we obtain

$$\|(B^*\langle\bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}})(\omega)\|_{\mathcal{U}} < \alpha \Rightarrow |u|(\{\omega\}) \leq \theta$$

for all ω on the grid. In our numerical examples the regularization parameter θ is chosen at least two orders of magnitude smaller than $\max|u|(\{\omega\})$. Therefor we will in slight abuse of notation refer to the superlevel set $\{\ \omega \mid |u|(\{\omega\}) > \theta\ \}$ as the support of u. In the next section we will see the effect of θ on the support of the optimal control.

We solved the resulting smooth problem with a quasi-Newton method. Since the dimension of the control space can become quite large with our approach we chose the memory efficient L-BFGS method, see [NW06]. For the numerical realization gradients for the discretized problems were used. The optimization method was terminated as soon as the $\ell^2(\mathcal{U}^h)$ norm of the gradient relative to the largest gradient was below a tolerance of 10^{-5}.

As expected from Proposition 4.7 the support of the optimal measure will depend on the cost parameter α. For large α the support will be small while the observation term will be large. The reverse holds for small α. We are interested in a compromise where α is chosen such that we obtain a small support while still achieving about 95% of the control objective, i.e. $\frac{1}{2}\langle\psi(T), \mathcal{O}\psi(T)\rangle_{\mathcal{H}} \leq 0.025$. An automated choice of α would be

helpful to obtain useful cost parameters for a variety of problems. This might be done in future work.

The result of this nonlinear optimization problem also depends on the initial guess for the control. For small α, the initial guess u_0^h for the control uses a fixed element $\tilde{u}_0^h$ in $\mathcal{U}^h$ for all $\omega \in \Omega^h$ with a phase θ_ω randomly chosen from a uniform distribution on $[0, 2\pi 12[$, $u_{0,\omega}^h = e^{i\theta_\omega}\tilde{u}_0^h$. For larger α, where such a generic initial guess leads to convergence of the method to suboptimal critical points near the origin, we use optimal solutions for smaller α as initial guess.

An alternative approach to Huber-regularization could be to add an $L^2(\Omega;\mathcal{U})$ regularization term to the cost functional and then apply a semi-smooth Newton method to this problem [KPV14]. This approach would have the advantage of preserving a support property and therefore decreasing the degrees of freedom, but it lies beyond the scope of this paper. The strategy used here is easier to implement since standard optimization methods can be applied, and suffices to illustrate the qualitative behavior of resulting optimal controls.

5.2 Three level system

As our first example we chose a typical finite-dimensional projection of an atom in a laser field, see Example 2.2. We use the matrices

$$H_0 = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad H_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

This corresponds for instance to a 1s, 2s, and 3p state. As explained in Section 2, the transition 1s $\to$ 2s is forbidden since the states have equal parity. On the other hand the transitions 1s $\to$ 3p and 3p $\to$ 2s are allowed. The control objective is to reach the third eigenstate starting from the ground state. The initial condition and the observation operator are given by $\psi_0 = (1, 0, 0)$ and $\mathcal{O} = \mathrm{diag}(1, 1, 0)$. We use a frequency band $\Omega = [2, 5]$ discretized with 100 grid points. The expected transition frequencies $\omega_1 = 3$ and $\omega_2 = 4$ are contained in Ω. We chose a time horizon of $T = 100$ and a time grid with 4096 points. The time horizon was chosen to allow for sufficiently many oscillations with the transition frequencies. We chose a cost parameter of $\alpha = 0.1$.

In Figure 1a we show the optimal field for the control space $\mathcal{M}(\Omega; H_0^1(0, T; \mathbb{C}))$ with the control operator B given by (32). Figure 1b shows an optimal field corresponding to the $L^2(0, T; \mathbb{R})$ regularized problem as proposed in [PDR88]. At first sight both fields do not seem to have a clear structure. Looking at the achievement of the control objective we report that the observation term for the L^2 control ($2.0 * 10^{-11}$) is much smaller than the term for the measure control ($7.8 * 10^{-7}$). Both observation terms are way below the critical threshold of $2.5 * 10^{-2}$ corresponding to 95% achievement of the control objective. The great advantage of the measure space control is that we can decompose the field into simple components.

Figure 1d shows the absolute value of the coefficients of the optimal measure in the time-frequency plane. We see that only two frequencies have a visible contribution. They

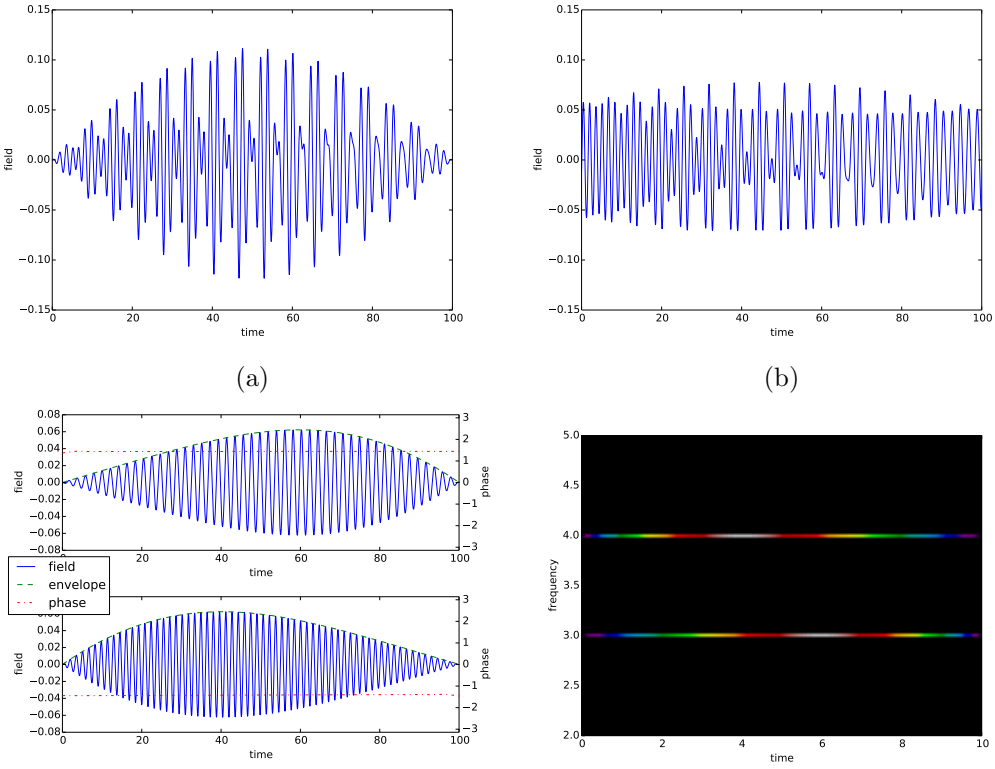
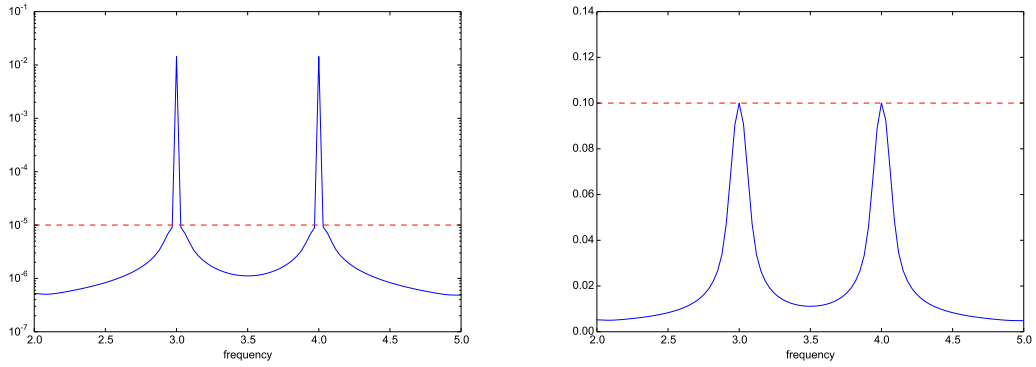


Figure 1: In (a) the optimal control field for the measure space case in $\mathcal{M}(\Omega; H_0^1(0,T;\mathbb{C}))$ and in (b) the optimal field for the Hilbert space case in $L^2(0,T;\mathbb{R})$ are plotted. In (c) we see the two main contributions of the field in (a). In (d) we give the absolute values of the coefficients of the optimal measure in the time-frequency plane.



(a) (b)

Figure 2: In (a), the norms $\|\bar{u}\|_{\mathcal{U}}$ (solid) and the regularization parameter θ (dashed) are plotted. In (b), the norms $\|(B^*\langle\varphi, \tilde{H}\psi\rangle)(\omega)\|_{\mathcal{U}}$ (solid) and the cost parameter α (dashed) are plotted.

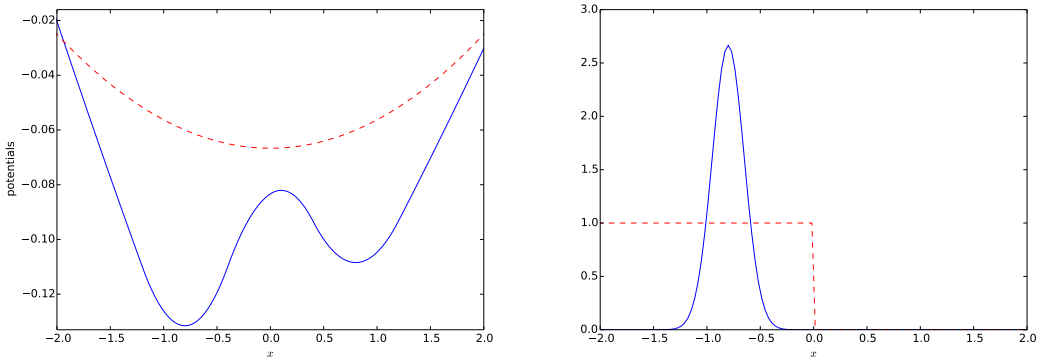
correspond to the two Bohr frequencies ω_1 and ω_2.

Figure 1c shows the decomposition of the optimal field in the two contributions of those frequencies. These two fields have a clear structure. The profile in time looks smooth and both fields are switched on during the whole time interval. This is consistent with the choice $\mathcal{U} = H_0^1(0, T; \mathbb{C})$, which promotes smoothness and non-locality in time. The field for the first transition reaches its maximum before the field for the second transition. This corresponds to the intuitive understanding that we have to induce the transition between the levels one and three before the transition between the levels three and two.

In fact, unlike all previous control-theory-based forcing fields that we are aware of for this basic system, the field obtained here bears considerable resemblance to the simple and intuitive few-parameter pulses which have been used by laser physicists for a long time. Compare, in particular, the two pulses in Figure 9 of [BTS98], whose achievement of the control objective was beautifully demonstrated experimentally.

Finally, the structure of the phase also looks very clear, except possibly for some oscillations at the boundaries. Those boundary effects are not desired in applications. However, numerical experiments show that we still achieve the control objective reasonably well after approximating the phase by a constant or an affine linear function. These approximations corresponds to a global phase shift and a global phase and frequency shift, respectively. Even neglecting the phase altogether resulted in reasonable control objectives.

In Figure 2 we can see that our optimal control is consistent with the necessary optimality conditions given in Propositions 4.6 and 4.7. We plotted $\|B^*\langle\bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}}\|_{\mathcal{U}}$ and $\|\bar{u}\|_{\mathcal{U}}$ against frequencies. We see that $\|B^*\langle\bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}}\|_{\mathcal{U}}$ is below α for all frequencies, and $\|\bar{u}\|_{\mathcal{U}}$ is below θ if $\|B^*\langle\bar{\varphi}, \tilde{H}\bar{\psi}\rangle_{\mathcal{H}}\|_{\mathcal{U}}$ is strictly smaller that α. The latter is consistent with the effect of Huber regularization with $\theta > 0$, where the support condition is not satisfied exactly.



(a) (b)

Figure 3: In (a), the two potential energy surfaces. In (b), the initial state (solid) and the coefficients for the observable (dashed).

In this example, we obtain a small support of the desired size two and a very small expectation value $7.8 * 10^{-7}$ for a cost parameter $\alpha = 10^{-1}$. For the second example we will see that a small support and a small expectation value are conflicting goals, which makes the choice of a good parameter α more important.

5.3 Schrödinger dynamics on two potential energy surfaces

In this second example we consider a Schrödinger system on two potential energy surfaces as arising in the laser control of chemical reactions, see Example 2.4. The spectral gap between the two potential energy surfaces varies depending on the position of the nuclear wave function and therefore a much larger variety of frequencies is potentially useful for succesful control. We also expect an additional time structure in the controls due to the movement of the densities in space. Therefore it is much more challenging to obtain simple controls for this example. We focus on a comparison between controls generated for different choices of Ω, $\mathcal{U}$ and B.

For simplicity we limit ourselves to one active coordinate, i.e. space dimension $d = 1$. The potential energy surfaces are plotted in Figure 3a. The control objective is to reach the potential well on the right starting from the potential well on the left. The initial state ψ_0 is a Gaussian located in the lower well. The observation operator $\mathcal{O}$ is the projection on the complement of functions with support on the lower surface on the right of the potential barrier.

The energy differences between the two potential energy surfaces measured at the local minima of the lower surface are around 0.074 and 0.048. Therefor we expect optimal controls to contain the two frequencies $\omega_1 \approx 0.074$ and $\omega_2 \approx 0.048$. We chose a time horizon of $T = 3000$ and a time grid with 2048 points. The time horizon was chosen large enough to allow for sufficiently many oscillations with the Bohr frequencies ω_1 and ω_2, and for sufficient movement of the wave function in space. For the discretization

in space we use a simple finite difference scheme on the domain $[-4, 4]$ with 256 grid points.

For this problem we compare the resulting optimal fields for different choices of the frequency domain Ω, the Hilbert space $\mathcal{U}$ of admissible time profiles, and the control operator B. We also compare them to the optimal fields for the classical Hilbert space cost functional with $L^2(0, T; \mathbb{R})$ or $H_0^1(0, T; \mathbb{R})$ norm. In particular we choose the following setups.

- **Time-smoothness with frequency-sparsity.** $\Omega = [1/30, 1/10], \mathcal{U} = H_0^1(0, T; \mathbb{C})$, B as in Example 3.1. The frequency band Ω contains the expected transition frequencies ω_1 and ω_2. It is discretized with 100 grid points. We discretize $\mathcal{U}$ with linear finite elements. The time grid corresponds to the grid of the time stepping. This results in $100 \cdot 2 \cdot 2048 = 409600$ real degrees of freedom.

- **Dual Gabor transform.** $\Omega = [1/30, 1/10]$, $\mathcal{U} = L^2(0, T; \mathbb{C})$ and B as in Example 3.2, with the Gaussian kernel k suitably adapted to generate homogeneous Dirichlet boundary conditions. We discretize Ω and $\mathcal{U}$ as before. For the evaluation of B we explicitly construct the matrix K corresponding to the kernel k.

- **Sparse Fourier transform.** $\Omega = [1/30, 1/10]$, $\mathcal{U} = \mathbb{C}$, B the Fourier synthesis operator, see Example 3.4. The frequency band Ω is discretized with 100 grid points. This results in $2 \cdot 100 = 200$ real degrees of freedom.

- **Time-frequency-sparse Gabor transform.** $\Omega = [1/30, 1/10] \times [0, T]$, $\mathcal{U} = \mathbb{C}$, B as in Example 3.5. The time-frequency cylinder Ω is discretized by a tensor grid. In frequency direction we use a regular grid with 100 grid points. In time direction we use a grid of 14 points. This results in $100 \cdot 14 \cdot 2 = 2800$ real degrees of freedom.

- **Standard L^2 cost.** This means we directly minimize over $v \in L^2(0, T; \mathbb{R})$ the funtional $J(\psi, v) = \frac{1}{2}\langle \psi(T), \mathcal{O}\psi(T)) \rangle + \alpha ||v||_{L^2}^2$. We use linear finite elements on the time grid of the time stepping method. This results in 2048 real degrees of freedom.

- **Standard H^1 cost.** As in the previous example, $B = I$, but now $v \in H^1(0, T; \mathbb{R})$, and the L^2 norm in the cost functional is replaced by the H^1 norm. Again this results in 2048 real degrees of freedom.

For the simulations we chose α such that probability to end up in the desired subspace is near the value of 95%.

In Figure 4 we plotted the fields generated by the optimal controls, their Fourier coefficients, time-frequency representation, and the absolute value of the coefficients of the optimal measure arranged in the time-frequency plane. We see that the structure of optimal controls heavily depends on the different cost functionals. In the following we will discuss the most significant differences and similarities.

From the second and third column of Figure 4 we see that all the fields, except for the $H_0^1(0, T; \mathbb{R})$ field, have two Bohr frequency regions around ω_1 and ω_2 that stand

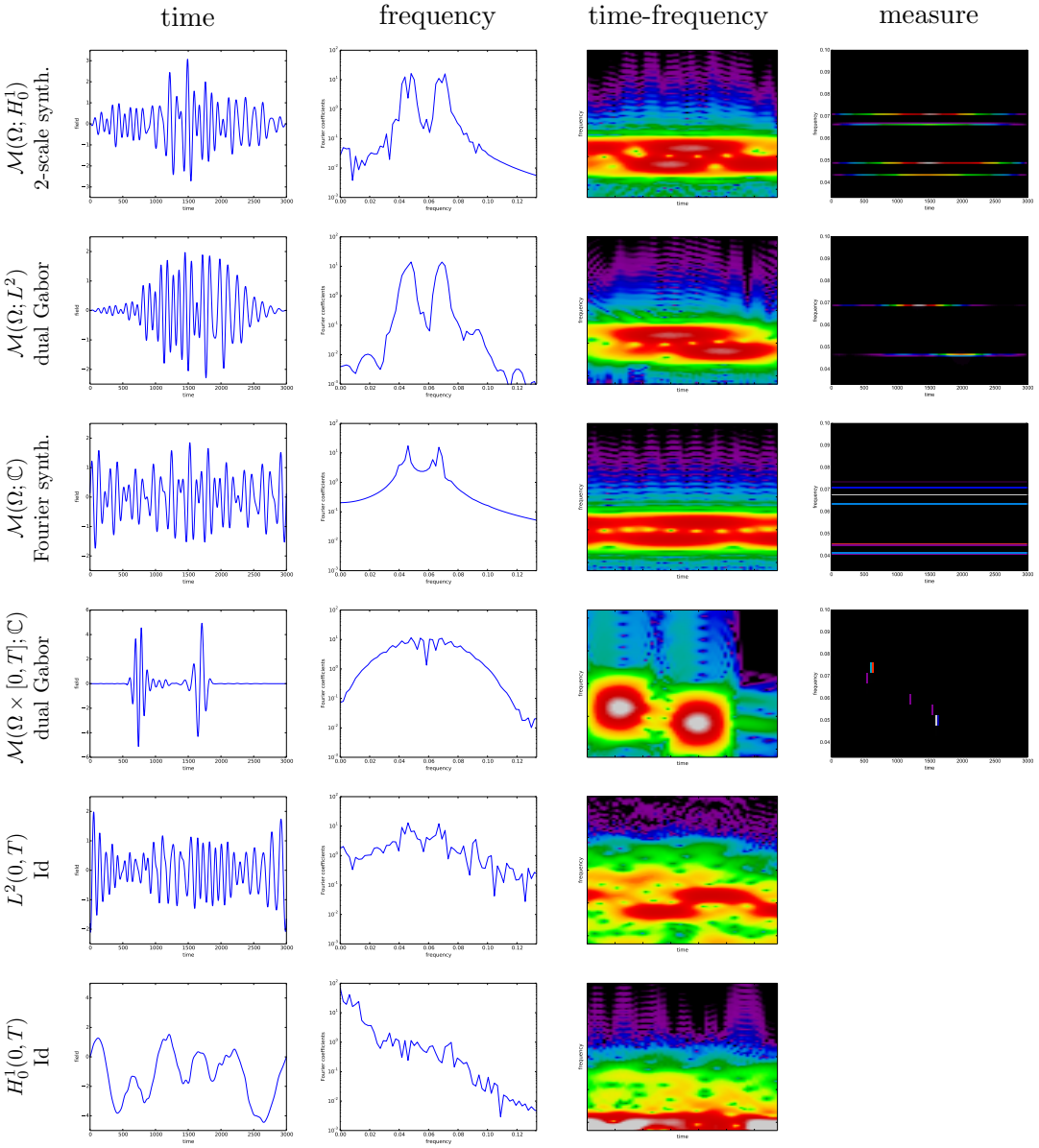


Figure 4: Different optimal controls in different representations for Schrödinger dynamics on 2 PES. Rows: cost functionals, control operators. Columns: time, frequency and time-frequency representation. In the rightmost column, the absolute values of the optimal measures are plotted in the time-frequency plane.

out. The regions correspond to the transitions up from the first well and down into the second well, respectively. This is the expected and desired behavior. The $H_0^1(0,T;\mathbb{R})$ field, on the other hand, mostly has low frequency components. In this case the wave packet stays on the first surface and is forced over the potential barrier into the upper well instead of inducing transition to the second surface. This shows that the choice of the control cost term can have an influence on the control mechanism triggered by the field. Since for this example the transitions between surfaces are of chemical interest, the use of the $H_0^1(0,T;\mathbb{R})$ norm as a cost is useless.

In the third column of Figure 4 we see that the other controls all show a particular time structure in their frequency content. It corresponds to the expected pump-dump picture, which means that we first induce the transition up from the lower well with frequency ω_1 and then down into the upper well with frequency ω_2. The time lag between the peaks for the frequencies ω_1 and ω_2 corresponds to the time it takes the wave package to travel on the upper surface between the positions of the two wells on the lower surface. The Gabor ansatz corresponds closest to pump-dump picture of two pulses separated in time and frequency. The other controls fill the whole time interval, the basic time-frequency control, the Fourier control and the L^2 control basically performing two pump-dump sequences. The Gabor ansatz is the only approach that seems to be least dependent on the chosen time horizon.

The controls generated using the measure space setting all show a well-defined time-frequency structure. In comparison the time-frequency representation of the $L^2(0,T;\mathbb{R})$ control looks very complicated and difficult to analyze. It also contains unwanted low frequency components. It is interesting that the space $\mathcal{M}(\Omega;H_0^1(0,T;\mathbb{C}))$ leads to more active frequencies compared to the approach with space $\mathcal{M}(\Omega;L^2(0,T;\mathbb{C}))$ with the Gabor synthesis operator.

We will now take a closer look at the support of the different optimal controls.

In Figure 5 we plotted the achievement of the control objective against the size of the support for different control spaces for varying cost parameter α. The vertical line marks the value of $2.5*10^{-2}$ corresponding to a probability of 95%. The figure illustrates the conflicting goals of minimizing the cost versus achieving the control objective. In contrast to the first example we do not obtain almost perfect achievement of the control objective for the desired support size two. Using the control spaces $\mathcal{M}(\Omega;H_0^1(0,T;\mathbb{C}))$ or $\mathcal{M}(\Omega;\mathcal{U}_k)$ leads to good achievement of the control objective for relatively small support. This is because those approaches separate frequency and time degrees of freedom. In the former sparsity is favored while we still have enough flexibility in time. For Example 3.2, we achieve a remarkably low number of three contributing frequencies. Two of those frequencies lie on neighboring grid points, which could well be a consequence of an unadapt discretization. If we increase the cost parameter to a point where the size of the support actually is forced to become two, we see that that control goal is not achieved in a satisfactory way anymore.

The Fourier and Gabor approaches lead to larger but still reasonably small supports. This is remarkable because the size of the support directly corresponds to the number of active degrees of freedom. Both approaches need way less degrees of freedom compared to the other approaches. We are, however, more interested in a simple frequency structure

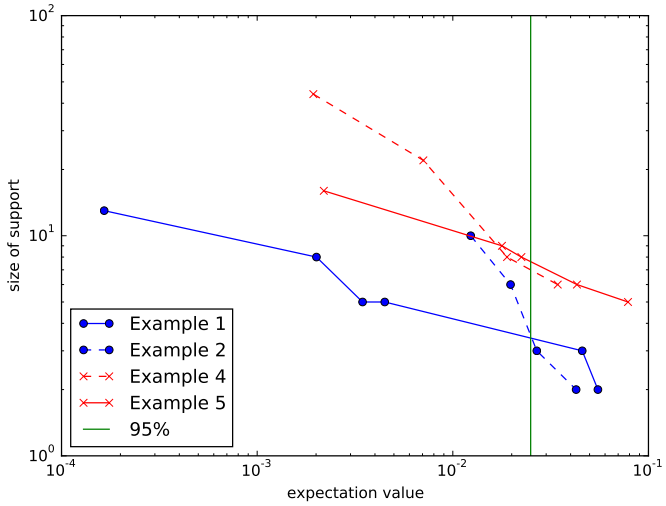


Figure 5: For different control spaces and cost terms, the size of the support of optimal measures is plotted against the corresponding expectation value. The vertical line marks the expectation value corresponding to a probability of 95% to achieve the control goal.

than in a low number of degrees of freedom. The size of the support corresponds to those degrees of freedom which contribute to the frequency structure. The Gabor ansatz does not exploit a separation of time and frequency structure. Since the shape of the pulses is fixed beforehand, it cannot adapt to the problem at hand, leading to additional active coefficients.

For the Fourier ansatz, additional frequencies become active to obtain some time structure of the resulting field.

6 Conclusion

In summary, measure valued costs imposed on time-frequency representations of the electric field as introduced in this paper produce far simpler, sparse, and physically more intuitive controls than the standard L^2 and H^1 costs. We hope that the measure-space, time-frequency approach will reduce the current gap between numerical controls on the one side and experimental implementation and physical intuition on the other. The flexibility of the approach can, in principle, be exploited to construct controls suited to concrete experimental setups.

Acknowledgements. The work of F. Henneke was funded by DFG through the International Research Training Group IGDK 1754.

References

[ABY+] A. Auger, A. Ben, Haj Yedder, E. Cances, C. Le Bris, C. M. Dion, A. Keller, and O. Atabek, *Optimal laser control of molecular systems: Methodology and results*, Math. Models Methods Appl. Sci, 1281–1315.

[BCG+02] U. Boscain, G. Charlot, J.-P. Gauthier, Stéphane Guérin, and Hans-Rudolf Jauslin, *Optimal control in laser-induced population transfer for two- and three-level quantum systems*, J. Math. Phys. **43** (2002), no. 5, 2107–2132.

[BKZB08] G. G. Balint-Kurti, S. Zou, and A. Brown, *Optimal control theory for manipulating molecular processes*, pp. 43–94, John Wiley & Sons, Inc., 2008.

[BMS82] J. M. Ball, J. E. Marsden, and M. Slemrod, *Controllability for distributed bilinear systems*, SIAM J. Control Optim. **20** (1982), no. 4, 575–597.

[BTS98] H. Bergmann, H. Theuer, and B. W. Shore, *Coherent population transfer among quantum states of atoms and molecules*, Rev. Mod. Phys. **70** (1998), no. 3, 1003–1025.

[Cla90] F. H. Clarke, *Optimization and nonsmooth analysis*, second ed., Classics in Applied Mathematics, vol. 5, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990.

[D'A08] D. D'Alessandro, *Introduction to quantum control and dynamics*, Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series, Chapman & Hall/CRC, Boca Raton, FL, 2008.

[FL07] I. Fonseca and G. Leoni, *Modern methods in the calculus of variations: L^p spaces*, Springer Monographs in Mathematics, Springer, New York, 2007.

[HL15] F. Henneke and M. Liebmann, *A generalized Suzuki–Trotter type method in optimal control of coupled Schrödinger equations*, pre-print, 2015.

[HMMS13] M. Hintermüller, D. Marahrens, P. A. Markowich, and C. Sparber, *Optimal bilinear control of Gross-Pitaevskii equations*, SIAM J. Control Optim. **51** (2013), no. 3, 2509–2543.

[HRG13] M. Hellgren, E. Räsänen, and E. K. U. Gross, *Optimal control of strong-field ionization with time-dependent density-functional theory*, Phys. Rev. A **88** (2013), 013414.

[HSW12] R. Herzog, G. Stadler, and G. Wachsmuth, *Directional sparsity in optimal control of partial differential equations*, SIAM J. Control Optim. **50** (2012), no. 2, 943–963.

[HTK+12] P. von den Hoff, S. Thallmair, M. Kowalewski, R. Siemering, and R. de Vivie-Riedle, *Optimal control theory - closing the gap between theory and experiment*, Phys. Chem. Chem. Phys. **14** (2012), 14460–14485.

[IK07] K. Ito and K. Kunisch, *Optimal bilinear control of an abstract Schrödinger equation*, SIAM Journal on Control and Optimization **46** (2007), no. 1, 274–287.

[KHK10] K. Kormann, S. Holmgren, and H.O. Karlsson, *A Fourier-coefficient based solution of an optimal control problem in quantum chemistry*, Journal of Optimization Theory and Applications **147** (2010), 491–506.

[KPV14] K. Kunisch, K. Pieper, and B. Vexler, *Measure valued directional sparsity for parabolic optimal control problems*, SIAM J. Control Optim. **52** (2014), no. 5, 3078–3108.

[Lan93] S. Lang, *Real and functional analysis*, third ed., Graduate Texts in Mathematics, vol. 142, Springer-Verlag, New York, 1993.

[LSTT09] M. Lapert, D. Sugny, R. Tehini, and G. Turinici, *Monotonically convergent optimal control theory of quantum systems with spectral constraints on the control field*, Physical Review A: Atomic, Molecular and Optical Physics **79** (2009), 063411.

[LY95] X. J. Li and J. M. Yong, *Optimal control theory for infinite-dimensional systems*, Systems & Control: Foundations & Applications, Birkhäuser Boston Inc., Boston, MA, 1995.

[Mez09] L. Meziani, *On the dual space $C_0^*(S, X)$*, Acta Math. Univ. Comenian. (N.S.) **78** (2009), no. 1, 153–160.

[NW06] J. Nocedal and S. J. Wright, *Numerical optimization*, second ed., Springer Series in Operations Research and Financial Engineering, Springer, New York, 2006.

[PDR88] A. P. Peirce, M. A. Dahleh, and H. Rabitz, *Optimal control of quantum-mechanical systems: existence, numerical approximation, and applications*, Phys. Rev. A (3) **37** (1988), no. 12, 4950–4964.

[Rab37] I. I. Rabi, *Space quantization in a gyrating magnetic field*, Phys. Rev. **51** (1937), 652–654.

[RBKM+06] Q. Ren, G. G. Balint-Kurti, F. R. Manby, M. Artamonov, T.-S. Ho, and H. Rabitz, *Quantum control of molecular vibrational and rotational excitations in a homonuclear diatomic molecule: A full three-dimensional treatment with polarization forces*, The Journal of Chemical Physics **124** (2006), no. 1, 014111.

[RSD+11] S. Ruetzel, C. Stolzenberger, F. Dimler, D. J. Tannor, and T. Brixner, *Adaptive coherent control using the von Neumann basis*, Phys. Chem. Chem. Phys. **13** (2011), 8627–8636.

[Sho90] B. W. Shore, *The theory of coherent atomic excitation*, Wiley, New York, NY, 1990.

[SKS+14] J. Scheuer, X. Kong, R. S. Said, J. Chen, A. Kurz, L. Marseglia, J. Du, P. R. Hemmer, S. Montangero, T. Calarco, B. Naydenov, and F. Jelezko, *Precise qubit control beyond the rotating wave approximation*, New Journal of Physics **16** (2014), no. 9, 093022.

[SSBK10] S. Sharma, H. Singh, and G. G. Balint-Kurti, *Genetic algorithm optimization of laser pulses for molecular quantum state excitation*, The Journal of Chemical Physics **132** (2010), no. 6, 064108.

[Sta09] G. Stadler, *Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices*, Comput. Optim. Appl. **44** (2009), no. 2, 159–181.

[Teu03] S. Teufel, *Adiabatic perturbation theory in quantum dynamics*, Lecture Notes in Mathematics, vol. 1821, Springer-Verlag, Berlin, 2003.

[TLBR04] G. Turinici, C. Le Bris, and H. Rabitz, *Efficient algorithms for the laboratory discovery of optimal quantum controls*, Phys. Rev. E **70** (2004), 016704.

[vWB08] G. von Winckel and A. Borzì, *Computational techniques for a quantum control problem with H^1-cost*, Inverse Problems **24** (2008), no. 3, 034007.

[vWBV10] G. von Winckel, A. Borzì, and S. Volkwein, *A globalized Newton method for the accurate solution of a dipole quantum control problem*, SIAM J. Sci. Comput. **31** (2009/10), no. 6, 4176–4203.

[WG07] J. Werschnik and E. K. U. Gross, *Quantum optimal control theory*, Journal of Physics B: Atomic, Molecular and Optical Physics **40** (2007), no. 18, R175–R211.