FREQUENCY-SPARSE OPTIMAL QUANTUM CONTROL

Gero Friesecke
Faculty of Mathematics
Technical University Munich
Boltzmannstr. 3, D-85747 Garching, Germany

Felix Henneke
Institute of Mathematics
FU Berlin
Arnimallee 6, D-14195 Berlin, Germany

Karl Kunisch*
Institute of Mathematics and Scientific Computing
Karl-Franzens-Universität Graz
Heinrichstr. 36, A-8010 Graz, Austria
and
Johann Radon Institute for Computational and Applied Mathematics (RICAM)
Austrian Academy of Sciences
Altenbergerstraße 69, A-4040 Linz, Austria

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.

We prove existence of optimal controls for the resulting nonsmooth optimization problem, derive necessary optimality conditions, and rigorously establish the frequency-sparsity of the optimizers. More precisely, we show that the time-frequency representation of the control field, which a priori admits a continuum of frequencies, is supported on only finitely many frequencies. These results cover 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 confirm that the optimal controls, unlike those obtained with the usual $L^2$ 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 [11], which range from steering chemical reactions [2] over creating excited or ionized states [12] to faithfully storing and manipulating bits of quantum information [30].

2010 Mathematics Subject Classification. Primary: 49J20, 35Q40.
Key words and phrases. Optimal quantum control, frequency sparsity, multi-level Schrödinger system.

* Corresponding author: Karl Kunisch.
To identify control fields which achieve a suitable goal, such as transfer of the quantum system from one eigenstate to another, and which are convenient to implement experimentally, one typically minimizes a cost functional which promotes goal achievement and penalizes unsuitable controls [16]. Theoretical work up to now has focused on cost functionals involving the $L^2$, $L^1$ or $H^1$ norm of the control in time-space parameterization, (see e.g. [10, 7] for finite-dimensional problems and [17, 37] for infinite-dimensional ones). This approach has been very successful in achieving a wide spectrum of control goals. But it has been remarked that “[the] resulting optimized electric fields are often very complex, thus it is nearly impossible to understand the underlying processes involved in the observed control” [16]. We may therefore ask the question: Why time-space parameterization?

We demonstrate in this paper that using certain more sophisticated costs leads to “simpler” controls. The control fields are 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. The sparse time-frequency structure is achieved via two key ideas.

First, we do not penalize the time profile of the field amplitude, nor just the frequency spectrum as this would rule out that a relevant frequency can be switched on smoothly, but a suitable time-frequency representation of it. While such representations are a familiar tool to interpreting or analyzing a given field in control and signal analysis, they here acquire center stage already in the design of the controls.

Second, we build upon recent advances in the optimal control theory of elliptic and parabolic systems related to the basic idea [38, 33, 9, 5, 14] of sparsity-enhancing $L^1$ or measure-norm costs. More specifically, we build upon the idea of function-valued measures to achieve directional sparsity in parabolic control [20]. The novelty as compared to the latter advances is two-fold: first, in quantum control, unlike in parabolic control, the target for sparsity should not be the field amplitude, but its frequency structure, and second, one is dealing with a bilinear instead of a linear control problem.

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

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

subject to

$$i \partial_t \psi = (H_0 + (Bu)(t)H_1) \psi, \quad \psi(0) = \psi_0. \quad (2)$$

Here $H_0$ is the Hamiltonian of the quantum system, $H_1$ is a coupling operator, $B$ 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 the time-frequency synthesis operator

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

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

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

Note that the control, a priori, can use the whole available continuum of frequencies, with each frequency possessing its own time profile. A main result of this paper (Theorem 4.10) is that the optimizers utilize only finitely many frequencies, even
when the quantum dynamics is a full infinite-dimensional Schrödinger dynamics on multiple potential energy surfaces (Ex. 2.2 and Section 5.3). Equations (3)–(4) replace the standard approach in quantum optimal control initiated by [26] to penalize just the $L^2$ or $H^1$ norm of the field amplitude (see [17, 36, 15] for mathematical results).

Recently, there has been independent and complementary research on time-sparse control of finite dimensional quantum systems [7, 6]. There one finds control fields which are zero in subintervals of the time domain. By contrast, our approach yields sparsity in a frequency domain. This necessitates a more involved measure space setting, but results in an even more dramatic reduction of the support, from a continuum to a finite set.

Numerical simulations presented in Section 5 below illustrate that the optimizers concentrate on just a few frequencies, even when the quantum dynamics is a full infinite-dimensional Schrödinger dynamics. Thus our controls share an important feature of laser pulses designed by experimentalists.

The plan of this paper is as follows. In the following section we first introduce two important physical examples of controlled quantum dynamical systems which motivated this work and to which our sparsity results apply. Subsequently, the control theoretic setting is introduced. In Section 3 we introduce and discuss concrete cases of time-frequency control spaces and associated control operators. In particular, the choices (3)–(4) allow for frequency-sparsity with global time profiles, and appropriate modifications result in frequency-sparsity with local time profiles. Section 4 is devoted to the mathematical analysis of the non-smooth optimal control problem (1)–(2). We recall the relevant well known results on the existence of dynamics, establish existence of optimal controls, derive necessary optimality conditions, and prove that optimal controls are supported on only finitely many frequencies. Finally, in Section 5 we numerically calculate optimal controls and compare them to those obtained from the usual $L^2$ 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. Control system.

2.1. Evolution equation. We consider the evolution of general quantum systems with controls, of the form

$$i\partial_t\psi(t) = \left(H_0 + \sum_{\ell=1}^{L} v_{\ell}(t)H_{\ell}\right)\psi(t), \quad \psi(0) = \psi_0$$

(5)

where the state $\psi(t)$ belongs to some Hilbert space $\mathcal{H}$, $H_0$ is a (bounded or unbounded) self-adjoint operator on $\mathcal{H}$, the operators $H_{\ell}$ are bounded and self-adjoint on $\mathcal{H}$, and $v_{\ell}(t)$ are real-valued scalar functions. Physically, the $v_{\ell}$ are 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 operators $H_{\ell}$ describe the system-field coupling. It is well known that under very weak assumptions on the control $v$ (which are always satisfied in our optimal control setting introduced below) eq. (5) possesses a unique mild solution in the state space of continuous paths in the Hilbert space, $C([0,T];\mathcal{H})$. See Section 4 for a precise statement.
We note that eq. (5) already contains two important approximations which are valid in many situations of interest. First, quantum fluctuations of the field amplitudes have been neglected, so that one is dealing with classical fields and does not need to move to the much more complicated framework of quantum field theory. Second, spatial variations of the applied fields have been assumed to occur at much larger wavelengths than the localization length of the state $\psi$, so that it is sufficient to assume that the fields depend on time only (‘dipole approximation’, see e.g. [32]).

We now give two examples of (5) of physical interest. These examples are simulated numerically in Section 5, and motivated the general theory in Section 4.

**Example 2.1.** (Electronic states of atoms in laser fields) A standard reference in the physics literature is [32]. The Hilbert space is $\mathcal{H} = \mathbb{C}^d$, and the evolution equation is

$$i\partial_t \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_d \end{pmatrix} = \begin{pmatrix} E_1 & 0 & \cdots & 0 \\ 0 & E_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & E_d \end{pmatrix} + v(t) \begin{pmatrix} 0 & \mu_{12} & \cdots & \mu_{1d} \\ \mu_{12} & 0 & \cdots & \mu_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{1d} & \mu_{2d} & \cdots & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_d \end{pmatrix}. \quad (6)$$

The coordinate vectors in $\mathbb{C}^d$ represent normalized atomic eigenstates $\psi_j \in L^2((\mathbb{R}^3 \times \mathbb{Z}_2)^N)$ where $N$ is the number of electrons with associated energy levels $E_j \in \mathbb{R}$. Here finite-dimensionality means that one has approximated the atomic wavefunction by $\psi(t) \approx \sum_{j=1}^d a_j(t)\psi_j$ and has neglected the coupling with the rest of the system. The function $v(t)$ is the scalar amplitude of a unidirectional electric field $E(t) = v(t)E_0$ with direction $E_0 \in \mathbb{R}^3$, and the $\mu_{ij}$ are coupling matrix elements of the form

$$\mu_{mn} = \langle \psi_m, -E_0 \cdot D \psi_n \rangle,$$ \quad (7)

where $D$ is the dipole operator defined by $D(x_1, \ldots, x_N) = \sum_{i=1}^N e x_i$. Here $e$ is the charge of the electron and the $x_i \in \mathbb{R}^3$ are electronic position coordinates. We remark that the off-diagonal element $\mu_{mn}$ vanishes whenever $\psi_m$ and $\psi_n$ have the same parity. This is a simple example of a “forbidden transition”. A typical control goal consists in inducing such a forbidden transition, i.e. one in which the two states of interest are only coupled indirectly via other states.

**Example 2.2.** (Laser-guided chemical reactions) This model is of central interest in photochemistry, but to our knowledge has not hitherto been considered in the mathematical literature. The Hilbert space is infinite-dimensional, $\mathcal{H} = L^2(\mathbb{R}^{3M}; \mathbb{C}^2)$. The evolution is given by the two-level Schrödinger system for the nuclear wave functions $\Phi_j$

$$i\partial_t \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix}(t) = \begin{pmatrix} E_1(R) & 0 \\ 0 & E_2(R) \end{pmatrix} + \begin{pmatrix} E(t) \cdot \mu_{11}(R) \\ E(t) \cdot \mu_{12}(R) \end{pmatrix}(t) \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix} + \begin{pmatrix} 0 \\ E(t) \cdot \mu_{22}(R) \end{pmatrix}^{*}(t) \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix}. \quad (8)$$

Here $R = (R_1, \ldots, R_M) \in \mathbb{R}^{3M}$ is the position vector of the atomic nuclei, the $E_j(R)$ are electronic eigenvalues (often called potential energy surfaces) of the system which depend on the positions of the nuclei, and the $\mu_{ij}(R)$ are dipole moment
Figure 1. Schematic representation of laser-controlled chemical reaction dynamics. The nuclei of a molecule move on different potential energy surfaces depending on the electronic state, and the laser induces transitions between these states. Blue: Potential energy surfaces. Magenta: Initial wave function of the nuclei. Cyan: Target region.

functions which depend on the associated electronic eigenstates $\psi_j^{(R)} \in L^2(\mathbb{R}^3 \times \mathbb{Z}_2)^N$. More precisely, the $E_j(R)$ and $\psi_j^{(R)}$ are eigenvalues and eigenstates of the electronic Hamiltonian $H_{el}^{(R)} = -\frac{i}{2}\Delta - \sum_{i=1}^N \sum_{\alpha=1}^M e^2 Z_{\alpha}/|x_i - R_{\alpha}| + \sum_{i<j} e^2/|x_i - x_j| + \sum_{\alpha<\beta} e^2 Z_\alpha Z_\beta/|R_{\alpha} - R_{\beta}|$, and

$$\mu_{ij}(R) = \langle \psi_i^{(R)}, -D^{(R)} \psi_j^{(R)} \rangle_{L^2(\mathbb{R}^3 \times \mathbb{Z}_2)^N},$$

(9)

where $D^{(R)}$ is the dipole operator defined by $D^{(R)}(x_1, \ldots, x_N) = \sum_{i=1}^N e x_i - \sum_{\alpha=1}^M e Z_{\alpha} R_{\alpha}$. Here, as in the previous example, $e$ is the charge of the electron (−1 in atomic units), and $-eZ_\alpha$ are the nuclear charges (+$Z_\alpha$ in atomic units). Note that for a single nucleus located at the origin, $D^{(R)}$ reduces to the dipole operator in the previous example. The restriction to two potential energy surfaces means that one has approximated the joint wavefunction of electrons and nuclei by a truncated Born-Oppenheimer expansion

$$\Psi(R, z, t) \approx \Phi_1(R, t) \psi_1^{(R)}(z) + \Phi_2(R, t) \psi_2^{(R)}(z),$$

where $z = (r_1, s_1, \ldots, r_N, s_N) \in (\mathbb{R}^3 \times \mathbb{Z}_2)^N$ is the vector of electron positions and spins (for a careful mathematical account of this expansion in the absence of control fields and for smooth interaction potentials see [34]).

The typical situation in laser control of chemical reactions is the following, see Figure 1. 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.

2.2. Control space. We assume that the control field $v$ in eq. (5) is the image of a control $u$ in practice, a time-frequency representation of $v$, see Section 3 – under
a control operator $B$. The control $u$ is assumed to belong to a measure space of form
\[ M(\Omega; U), \] (10)
where $\Omega$ is a locally compact space (typically a closed interval of admissible frequencies), and $U$ is a separable Hilbert space of time-dependent functions. Admitting general separable Hilbert spaces is useful because the freedom to choose $U$ can be exploited to generate nice optimal controls. See the examples in Section 3 and in particular the discussion after Example 3.2. The space (10) is the space of Borel measures $u$ on $\Omega$ with values in $U$ of finite mass norm $\|u\|_M$. The mass norm of the measure $u$ will appear in the cost functional $J$ in (14) below.

The space (10) is the dual of the separable space $C_0(\Omega; U)$ of continuous functions on $\Omega$ with values in $U$ which can be uniformly approximated by functions with compact support. The duality pairing is given by
\[ \langle u, \varphi \rangle_{M, C_0} = \int_\Omega \langle u'(\omega), \varphi(\omega) \rangle_U d|u|\langle \omega \rangle \] (11)
(see [24]) where $u'$ is the Radon–Nikodym derivative of $u$ with respect to the total variation measure $|u|$ (see [21, VII Thm. 4.1]). Note that the inner product in the integral is the Hilbert space inner product in $U$. This duality will be very useful.

2.3. Control operator. We assume that the control field $v$ in eq. (5) is obtained from the control $u$ belonging to the above measure space via a bounded linear control operator $B$, more precisely:
\[ v = Bu, \quad B : M(\Omega; U) \to L^p(0, T; \mathbb{R}^L) \] bounded linear, for some $1 < p \leq \infty$. (12)
Moreover we assume that $B$ has a bounded linear predual operator $B^* : L^q(0, T; \mathbb{R}^L) \to C_0(\Omega; U)$, by which we mean a bounded linear operator such that $\langle B^* f, u \rangle_{C_0, M} = \langle f, Bu \rangle_{L^q, L^p}$ for all $f \in L^q(0, T; \mathbb{R}^L)$ and all $u \in M(\Omega; U)$ where $\frac{1}{p} + \frac{1}{q} = 1$.

Existence of a bounded linear predual operator implies the weak-$*$-weak(-*) continuity of $B$. That is, weak-$*$ convergence of $u_n$ to $u$ in $M(\Omega; U)$ implies weak convergence of $Bu_n$ to $Bu$ in $L^p(0, T; \mathbb{R}^L)$ if $1 < p < \infty$, and weak-$*$ convergence if $p = \infty$. The case $p = 1$ has to be excluded in our framework since we will make use of the reflexivity of $L^p(0, T; \mathbb{R}^L)$. Note that the operator $B^*$ depends on the Hilbert space structure of $U$, see the examples below.

2.4. Cost and optimal control problem. Associated to the evolution equation (5), we consider the optimal control problem
\[ \min_{\psi, u} J(\psi, u) \quad \text{subject to (5)}, \] (13)
where
\[ J(\psi, u) = \frac{1}{2} \langle \psi(T), \mathcal{O}\psi(T) \rangle_H + \alpha \|u\|_{M(\Omega; U)} \] (14)
The functional $J$ consists of the term $\frac{1}{2} \langle \psi(T), \mathcal{O}\psi(T) \rangle_H$ which describes the control goal, and a cost term $\alpha \|u\|_{M(\Omega; U)}$ which contains a measure norm explained above and forces the solution to be sparse in a suitable sense. The observable $\mathcal{O}$ specifying the control goal can be any bounded self-adjoint operator on $H$. Typically $\mathcal{O}$ is the orthogonal projection onto the subspace we want to leave. In this case, the target constraint contribution $\frac{1}{2} \langle \psi(T), \mathcal{O}\psi(T) \rangle_H$ to the cost lies in the interval $[0, 1/2]$. The value 0 corresponds to a 100% achievement of the control objective, and the value 0.5 to a 0% achievement.
3. Examples of measure norm costs and control operators. We now list some examples for choosing the frequency domain $\Omega$, the Hilbert space $\mathcal{U}$ of time-dependent functions, and the control operator $B$. The first example is the prototype for generating frequency-sparse controls. It motivated the general functional analytic setting proposed above, and naturally incorporates physically relevant controls containing finitely many pulses of particular frequencies $[4, 1, 35, 31, 29]$.

The Hilbert space $\mathcal{U}$, and the Hilbert space $\mathcal{H}$ of quantum states, typically contain complex valued functions. However, the space $\mathcal{U}$ is mapped by the control operator to real valued functions. For this reason, all spaces below are equipped with a real Banach or Hilbert space structure. For complex spaces this amounts to always using the real part of the natural complex valued scalar product. To avoid confusion we often explicitly write $\text{Re}(\cdot, \cdot)_\mathcal{H}$ for the scalar product in $\mathcal{H}$. In particular, for mappings linear always means $\mathbb{R}$-linear and scalar products are real-valued and $\mathbb{R}$-bilinear.

Throughout this section it will be convenient to set $L = 1$.

**Example 3.1.** (Two-scale synthesis, smooth functions of time) Here $u$ will be a two-scale time-frequency representation of the laser field amplitude and $B$ will generate the corresponding field. The control $u$ can a priori contain a continuum of active frequencies, each with its own smooth envelope. This can be modeled mathematically as follows. Let $\Omega$ 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 two-scale synthesis operator

\[ (Bu)(t) = \text{Re} \int_\Omega u(\omega, t) e^{i\omega t} \, d\omega. \]  

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:

\[ (Bu)(t) = \text{Re} \int_\Omega u'(\omega, t) e^{i\omega t} \cos(\omega t) \, d|u|(\omega), \]  

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

We remark that the above space of controls naturally contains the physically motivated finite-dimensional ansatz spaces of $[1, 35, 31, 19]$ in which a finite number of distinct frequencies can be switched on or off by few-parameter modulation functions: the field $v(t) = \sum_{j=1}^n v_j(t) \cos(\omega_j t + \phi_j)$ corresponds to $Bu$ with $B$ as in (16) and $u(\omega, t) = \sum_{j=1}^n \delta(\omega - \omega_j) e^{i\phi_j} v_j(t)$. 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 $B^* : L^1(0, T) \to C_0(\Omega; H^1_0(0, T; \mathbb{C}))$ is the solution operator of the second-order boundary value problem

\[ \frac{\partial^2}{\partial t^2} u(\omega, t) = f(t) e^{-i\omega t}, \quad u \bigg|_{t=0} = u \bigg|_{t=T} = 0, \]  

i.e. $B^* f = u$. As we shall see in Section 4.3 the operator $B^*$ is essential for the characterization of the optimal controls. Note that the definition of the predual operator is based on the dual pairings $\langle \cdot, \cdot \rangle_{C_0, \mathcal{M}}$ and $\langle \cdot, \cdot \rangle_{L^1, L^\infty}$. The equations (17) are not coupled for different $\omega$. 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 $\Omega$ is closed. Otherwise $B$ might not be weak-$\ast$–weak(-$\ast$) continuous.
Example 3.2. (Gabor synthesis, $L^2$ functions of time) 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 $\Omega \subset \mathbb{R}_+$ closed, $\mathcal{U} = L^2(0,T;\mathbb{C})$, $p = \infty$, and define $B$
by
\[(Bu)(t) = \text{Re} \int_\Omega \int_0^T k(s,t)u'(\omega,s)\,ds\,e^{i\omega t}\,d|u|(\omega),\tag{18}\]
where $k: [0,T]^2 \to \mathbb{R}$ is a smooth symmetric kernel. Roughly, this operator corre-
sponds to a pre-processing of the envelopes, with only smoothed envelopes entering
the equation. The predual operator $B^*$ becomes the Gabor transformation
\[(B^*f)(\omega,t) = \int_0^T f(s,k(t,s)e^{-i\omega s}\,ds,\]
which is a time-frequency representation of the control field.

Example 3.1 above with $\mathcal{U} = H^1_0$ leads 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 (17), hence $(B^*f)(\omega,t) = \int_0^T f(s)G(t,s)e^{-i\omega s}\,ds$. 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 com-
plicated. Note that Example 3.2 can be reformulated in terms of the control
operator from Example 3.1 by using a different control space. To see this, let $K: L^2(0,T) \to \mathcal{D}(A)$ be the
compact operator given by convolution with a Gaussian kernel $k$, $(Kf)(t) = \int_0^T k(t,s)f(s)\,ds$. Then $K$ is injective and self-
adjoint and has an unbounded inverse $A: \mathcal{D}(A) \to L^2$. Define $\mathcal{U}_k := \mathcal{D}(A)$
with the induced scalar product $(u,v)_{\mathcal{U}_k} = (Au,Av)_{L^2}$. Then $\mathcal{U}_k$ is a Hilbert space
and the predual operator $B^*$ of $B$ is $(B^*f)(\omega,t) = \left(\int_0^T f(t,s)G(t,s)e^{-i\omega s}\right)^{(s)}$. This construction
also works for window functions other than a Gaussian. The equivalence to the
choice $\mathcal{U} = L^2(0,T;\mathbb{C})$, with $B$ as in (18) follows since the dual of the map $X: C^0(\Omega;L^2(0,T;\mathbb{C})) \to C^0(\Omega;\mathcal{U}_k)$ defined by $(X\varphi)(\omega) = K\varphi(\omega)$ is an isometric
isomorphism between the control spaces that preserves the image under the corre-
sponding control operators.

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

Example 3.3. (Fourier synthesis, constant functions of time) Let $\Omega \subset \mathbb{R}_+$, $\mathcal{U} = \mathbb{C}$, and
let $B$ be the Fourier synthesis operator, that is to say
\[(Bu)(t) = \text{Re} \int_\Omega u(\omega)\,e^{i\omega t}\,d\omega = \text{Re} \int_\Omega u'(\omega)\,e^{i\omega t}\,d|u|(\omega).\]
The function $u(\omega)$ here can be viewed as a time-frequency representation $u(\omega,t)$
which is independent of time $t$. The predual operator is, up to a constant factor,
the Fourier transform of functions restricted to $[0,T]$,
\[(B^*f)(\omega) = \hat{f}(\omega) = \int_0^T f(t)e^{-i\omega t}\,dt.\]

An alternative approach to achieving sparsity of a time-global frequency decom-
position via an $L^2$ cost combined with iterative ‘frequency sifting’ is given in [28].
Example 3.4. (Time-frequency Gabor synthesis) Let $\Omega \subset \mathbb{R}^+ \times [0, T]$ be a subset of time-frequency space, $\mathcal{U} = \mathbb{C}$, and

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

for the ansatz function $g_{\omega, s}(t) = e^{-\frac{(\omega - \sigma)^2}{2\sigma^2}} e^{i\omega(t-s)}$ for some $\sigma > 0$. This defines a suitable extension of the control operator from Example 3.2 to measures in both time and frequency. With this control operator, each Dirac measure $\delta_{\omega, t}$ corresponds to a Gaussian wave packet centered at time $t$ with frequency $\omega$. The predual of the control operator is given by

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

These examples by no means exhaust our framework, but are meant to give an idea of its flexibility. We also remark that frequency constraints could easily be included by restricting $\Omega$, compare [22].

4. Theory of the optimal control problem. In this section we will study the optimal control problem (13). We first show well-posedness of the problem. In contrast to [20] 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. Theorem 4.8 is the natural analog of Theorem 2.12 in [20]. 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 and compactness properties for the state equation. Here we recall well-known properties of the state equation (5) needed to prove the existence of solutions to the optimal control problem. Throughout this subsection, for a given control field $v \in L^1(0, T; \mathbb{R}^L)$ we consider mild solutions to the state equation (5) i.e. functions $\psi$ which satisfy $\psi \in C([0, T]; \mathcal{H})$ and

$$\psi(t) = G(t)\psi_0 + \int_0^t G(t-s)v(s) \cdot \hat{H}\psi(s)\,ds.$$ (19)

Here $G$ is the unitary group generated by the skew-adjoint operator $-iH_0$, $\hat{H} = (-iH_1)_t$, and $v(s) \cdot \hat{H} = \sum v_l(s)(-iH_l)$. Also, here and below $\langle \varphi(t), \hat{H}\psi(t) \rangle_{\mathcal{H}}$ denotes the vector with components $\langle \varphi(t), \hat{H}_l\psi(t) \rangle_{\mathcal{H}}$. Existence of mild solutions as well as their differentiability in the direction of the field follow from [3, Thm. 2.5], and a convenient expression for the derivative is given in [23, Chapter 2].

Proposition 4.1. Let $v \in L^1(0, T; \mathbb{R}^L)$. Then (19) possesses a unique solution $\psi \in C([0, T]; \mathcal{H})$. Furthermore, the mapping $F: L^1(0, T; \mathbb{R}^L) \to C([0, T]; \mathcal{H})$ defined by $F(v) = \psi$ is differentiable. Its derivative is given by $F'(v)(\delta v) = \psi'$ where

$$\psi'(t) = \int_0^t G(t, s)\delta v(s) \cdot \hat{H}\psi(s)\,ds.$$ (20)
with the evolution operator $\mathcal{G}$ defined through $\mathcal{G}(t, s)\psi(s) = \psi(t)$ for $t \geq s$.

We remark that the evolution operators $\mathcal{G}(t, s)$ are unitary. This can easily be shown by working in the rotating frame, i.e. by studying the function $t \mapsto G^*(t)\psi(t)$ [27], but it will not be needed in the following.

One downside of the mild solution framework is that the typical compactness arguments based on additional spatial regularity [17, 15] cannot be used. Fortunately there is a powerful replacement that works in a very general setting. We will use the following compactness result from [3, Theorem 3.6].

**Proposition 4.2.** Let $(v_n)_n$ be a sequence in $L^1(0, T; \mathbb{R}^L)$ such that $v_n \to v$. Then the corresponding solutions $\psi_n$ of (19) satisfy $\psi_n \to \psi$ in $C([0, T]; \mathcal{H})$, where $\psi$ is the mild solution corresponding to $v$.

**Remark 4.3.** The results of Proposition 4.1 and 4.2 hold for general $C_\sigma$-semigroups. The unitarity of the evolution operator is not used for the theoretical results in this work.

### 4.2. Existence of optimal controls

We will now prove the existence of solutions to the optimal control problem (13).

**Theorem 4.4.** There exists a global solution $(\bar{\psi}, \bar{u}) \in C([0, T]; \mathcal{H}) \times \mathcal{M}(\Omega; \mathcal{U})$ of (13).

**Proof.** The result follows from standard reasoning in the calculus of variations. Let $(\psi_n, u_n)$ be a minimizing sequence,

$$\lim_n J(\psi_n, u_n) = \inf J(\psi, u). \quad (21)$$

Since $\mathcal{O}$ is bounded and $\alpha > 0$, the sequence $(u_n)$ is bounded in $\mathcal{M}(\Omega; \mathcal{U})$. Together with the fact that $\mathcal{M}(\Omega; \mathcal{U}) = C(\Omega; \mathcal{U})^*$ is the dual of a separable Banach space, this implies that 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 $\mathcal{B}$ implies that $\mathcal{B}u_n$ converges to $\mathcal{B}\bar{u}$ weakly in $L^p(0, T; \mathbb{R}^L)$ for some $p > 1$ and thus also for $p = 1$.

By Proposition 4.2 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 (21) this implies the claim.

### 4.3. Necessary optimality conditions

For theoretical and numerical purposes we will use the reduced form of (13),

$$\min_u j(u), \quad \text{where } \quad j(u) = J(\psi(u), u). \quad (22)$$

Here $\psi(u)$ denotes the solution of (5), (12) for the control $u$.

The reduced cost functional can be split into two parts, a nonlinear differentiable part

$$f(v) = \langle \psi(T), \mathcal{O}\psi(T) \rangle_\mathcal{H} \quad (23)$$

with $v = \mathcal{B}u$, 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

$$f'(v) = \text{Re}\langle \varphi, \mathcal{H}\psi \rangle_\mathcal{H}.$$
where \( \varphi \) is the mild solution of the dual equation

\[
i\partial_t \varphi(t) = (H_0 + v(t) \cdot i\tilde{H}) \varphi(t), \quad \varphi(T) = \mathcal{O}\psi(T).
\]

Using the evolution operator this can be rewritten as

\[
\varphi(t) = G(T,t)^* \mathcal{O}\psi(T).
\]

(24)

Under suitable assumptions, the representation (23) of \( f' \) can be derived using a Lagrange functional approach, see [26, 37]. Since in our setting a variational formulation is not readily available, we will give a short proof in the present setting of mild solutions.

**Lemma 4.5.** Let \( v, \delta v \in L^1(0,T; \mathbb{R}^L) \), and let \( \psi, \psi' \) and \( \varphi \) be the corresponding solutions of (19), (20) and (24), respectively. Then the mapping \( f : L^1(0,T; \mathbb{R}^L) \to \mathbb{R} \) defined by

\[
f(v) = \frac{1}{2} \langle \psi(T), \mathcal{O}\psi(T) \rangle_{\mathcal{H}}
\]

is continuously differentiable with derivative

\[
f'(v)(\delta v) = \int_0^T \delta v(t) \cdot \text{Re} \langle \varphi(t), \tilde{H}\psi(t) \rangle_{\mathcal{H}} dt.
\]

**Proof.** The continuous differentiability of the state and the product rule give continuous differentiability of \( f \) and

\[
f'(v)(\delta v) = \text{Re} \langle \mathcal{O}\psi(T), \psi'(T) \rangle_{\mathcal{H}}.
\]

Using (20) and (24) we obtain

\[
\text{Re} \langle \mathcal{O}\psi(T), \psi'(T) \rangle_{\mathcal{H}} = \text{Re} \langle \mathcal{O}\psi(T), \int_0^T G(T,t) \delta v(t) \cdot \tilde{H}\psi(t) dt \rangle_{\mathcal{H}}
\]

\[
= \int_0^T \delta v(t) \cdot \text{Re} \langle G(T,t)^* \mathcal{O}\psi(T), \tilde{H}\psi(t) \rangle_{\mathcal{H}} dt
\]

\[
= \int_0^T \delta v(t) \cdot \text{Re} \langle \varphi(t), \tilde{H}\psi(t) \rangle_{\mathcal{H}} dt.
\]

\[\square\]

We can now derive the following optimality condition. It is partially inspired by the optimality condition derived in [20, Thm. 2.11] for a linear parabolic control problem of form \( \partial_t \psi - \Delta \psi = u \).

**Proposition 4.6 (Optimality conditions).** Let \( \bar{u} \) be a local minimizer of problem (22), and let \( \bar{\psi}, \bar{\varphi} \in C([0,T]; \mathcal{H}) \) be the corresponding solutions of (19) and (24) for the control field \( v = B\bar{u} \). Then

\[
\alpha \| \bar{u} \|_{\mathcal{M}} = -\langle B^* \text{Re} \langle \bar{\varphi}, \tilde{H}\bar{\psi} \rangle_{\mathcal{H}}, \bar{u} \rangle_{C_0,\mathcal{M}}.
\]

(25)

and

\[
\| B^* \text{Re} \langle \bar{\varphi}, \tilde{H}\bar{\psi} \rangle_{\mathcal{H}} \|_{C_0} \leq \alpha.
\]

(26)

**Proof.** Since we can split our functional into a sum of a nonconvex and a nonsmooth part, the result can be deduced from the general differential calculus of Clarke [8]. Because this calculus is somewhat intricate and due to the importance of the optimality system, we prefer to give a more elementary proof.
Let \( \bar{u} \) be a local minimizer of problem (22) and let \( \bar{\psi} \) and \( \bar{\varphi} \) be the corresponding solutions of (19) and (24). We first show the variational inequality
\[
g(\bar{u}) - f'(B\bar{u})(Bu - B\bar{u}) \leq g(u).
\] (27)

Since \( \bar{u} \) is locally optimal, we have
\[
\frac{1}{h}(j(\bar{u} + h(u - \bar{u})) - j(\bar{u})) \geq 0
\]
for \( u \in \mathcal{M}(\Omega; U) \) and \( h \in (0, 1) \) sufficiently small. Using the decomposition \( j = f \circ B + g \) and convexity of \( g \), this implies
\[
\frac{1}{h}(f(B\bar{u} + h(Bu - B\bar{u})) - f(B\bar{u})) + g(u) - g(\bar{u}) \geq 0.
\]
Since \( f \) is differentiable, taking the limit \( h \to 0 \) yields (27).

Testing (27) with \( u = 0 \) and \( u = 2\bar{u} \) gives
\[
g(\bar{u}) + f'(B\bar{u})(B\bar{u}) = 0.
\] (28)

Substituting (28) into (27) gives
\[
f'(B\bar{u})(Bu) \leq g(u)
\] (29)
for all \( u \in \mathcal{M}(\Omega; U) \).

Using Lemma 4.5 on the derivative of \( f \), equation (28) gives
\[
g(\bar{u}) = -\langle \text{Re}(\bar{\varphi}, \bar{H}\bar{\psi})_{\mathcal{H}}, B\bar{u} \rangle_{L^2, L^p} = -\langle B^* \text{Re}(\bar{\varphi}, \bar{H}\bar{\psi})_{\mathcal{H}}, \bar{u} \rangle_{C_0, \mathcal{M}}
\]
which proves (25). From (29) we obtain
\[
-\langle B^* \text{Re}(\bar{\varphi}, \bar{H}\bar{\psi})_{\mathcal{H}}, \bar{u} \rangle_{C_0, \mathcal{M}} \leq \alpha \|u\|_{\mathcal{M}}.
\]
Testing this inequality with \( u = -\delta_\omega(B^* \text{Re}(\bar{\varphi}, \bar{H}\bar{\psi})_{\mathcal{H}})(\omega) \) for some \( \omega \in \Omega \) yields
\[
\|(B^* \text{Re}(\bar{\varphi}, \bar{H}\bar{\psi})_{\mathcal{H}})(\omega)\|_{U}^2 \leq \alpha \|(B^* \text{Re}(\bar{\varphi}, \bar{H}\bar{\psi})_{\mathcal{H}})(\omega)\|_{U}
\]
which gives (26).

\[\Box\]

**Remark 4.7.** Proposition 4.6 provides only a necessary condition for local optimality. Due to the nonlinear structure of the problem (13), 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 4.8** (Characterization of frequency support). Let \( \bar{u}, \bar{\psi} \) and \( \bar{\varphi} \) be as in Proposition 4.6. Then we have
\[
\text{supp}[\bar{u}] \subset \{ \omega \in \Omega \mid \|(B^* \text{Re}(\bar{\varphi}, \bar{H}\bar{\psi})_{\mathcal{H}})(\omega)\|_{U} = \alpha \}.
\] (30)
\[
-\alpha \bar{u}'(\omega) = (B^* \text{Re}(\bar{\varphi}, \bar{H}\bar{\psi})_{\mathcal{H}})(\omega), \quad |\bar{u}| \text{-almost everywhere.}
\] (31)

**Proof.** Writing equation (25) as an integral yields
\[
\int_{\Omega} \left( \alpha + \langle (B^* \text{Re}(\bar{\varphi}, \bar{H}\bar{\psi})_{\mathcal{H}})(\omega), \bar{u}'(\omega) \rangle_{U} \right) d|\bar{u}|(\omega) = 0.
\] (32)

For the integrand we obtain by the Cauchy–Bunyakovsky–Schwarz (CBS) inequality, using \( \|\bar{u}'(\omega)\| = 1 \) and (26),
\[
\alpha + \langle (B^* \text{Re}(\bar{\varphi}, \bar{H}\bar{\psi})_{\mathcal{H}})(\omega), \bar{u}'(\omega) \rangle_{U} \geq \alpha - \|(B^* \text{Re}(\bar{\varphi}, \bar{H}\bar{\psi})_{\mathcal{H}})(\omega)\|_{U} \geq 0.
\]

Therefore (32) yields
\[
\alpha + \langle (B^* \text{Re}(\bar{\varphi}, \bar{H}\bar{\psi})_{\mathcal{H}})(\omega), \bar{u}'(\omega) \rangle_{U} = 0
\]
for $|\bar{u}|$-almost all $\omega \in \Omega$. For those $\omega$ the CBS inequality in (33) was sharp. This implies (31) and

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

for $|\bar{u}|$-almost all $\omega \in \Omega$. Since $\omega \mapsto \|(B^* \text{Re} \langle \bar{\varphi}, \tilde{H} \bar{\psi} \rangle_{\mathcal{H}})(\omega)\|_U$ is continuous this implies (30).

For specific control operators $B$, equation (31) implies additional regularity for $\bar{u}'(\omega)$. For example, in the case of the control operator from Example 3.1 we obtain the regularity $\bar{u}'(\omega) \in H^2(0, T)$.

The relation (30) for the support of the optimal measure gives us the following corollary.

**Corollary 4.9.** Let $\bar{u}$ be a local minimizer of (22). Then $\text{supp} |\bar{u}|$ is compact.

**Proof.** Since $B^* \text{Re} \langle \bar{\varphi}, \tilde{H} \bar{\psi} \rangle \in C_0(\Omega; U)$, we know that there is a compact set $K \subset \Omega$ such that $B^* \text{Re} \langle \bar{\varphi}, \tilde{H} \bar{\psi} \rangle \leq \alpha/2$ for all $\omega \not\in K$. Using (30) this implies $\text{supp} |\bar{u}| \subset K$. Therefore $\text{supp} |\bar{u}|$ is compact as a closed subset of the compact set $K$.

It says that although the frequency domain $\Omega$ 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 remarkable that such oscillations in our theoretical controls can be ruled out rigorously.

For most of the specific control operators considered in this work, a significantly stronger result holds.

**Theorem 4.10** (Frequency sparsity). Let $\bar{u}$ be a local minimizer of (22) and let $\Omega$, $U$ and $B$ be given as in one of the Examples 3.1, 3.2, 3.3. Then $\text{supp} |\bar{u}|$ is finite.

**Proof.** The results follows from (30) by an analyticity argument. In the setting of the Examples 3.1–3.3, for every $f \in L^q(0, T)$, the map $\omega \mapsto (B^* f)(\omega)$ is real analytic. This is standard for the Fourier and Gabor transformation, and follows from analytic dependence on $\omega$ of the right hand side in (17) for the dual two-scale operator. If $\Omega \neq \mathbb{R}$ then $B^* f$ can be extended in the natural way to an analytic function on $\mathbb{R}$. In all cases, $(B^* f)(\omega) \to 0$ if $\omega \to \pm \infty$.

For $f = \text{Re} \langle \bar{\varphi}, \tilde{H} \bar{\psi} \rangle_{\mathcal{H}}$ we obtain the analyticity of $\omega \mapsto \|(B^* \text{Re} \langle \bar{\varphi}, \tilde{H} \bar{\psi} \rangle_{\mathcal{H}})(\omega)\|_U^2$, which implies that the value $\alpha^2$ is attained either everywhere in $\mathbb{R}$ or only at discrete points. Since the function vanishes at infinity and the support is compact by Corollary 4.9, the support condition (30) implies the finiteness of $\text{supp} |\bar{u}|$.

This result is of significant physical interest. It implies that the optimal control is in fact a finite sum of Dirac measures in frequency with values in $U$. We thus recover controls of the form used by experimentalists without explicitly requiring any simple structure.

Comparing Theorem 4.10 to [20, Corollary 2.13], we do not assume that the support is discrete but derive it from identifying and using smoothing properties of the adjoint control operator. Example 3.4 cannot be treated with the same technique. Although the map $\omega \mapsto (B^* f)(\omega)$ is complex analytic for $f \in L^q(0, T)$ if $\Omega$ is identified with a subset of $\mathbb{C}$, the analyticity does not carry over to the absolute values.
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 finite-dimensional system from Example 2.1. It serves to illustrate basic effects of sparse control of quantum systems. The second example is the two-level Schrödinger system from Example 2.2. 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. The regularization turns regions with zero frequency amplitude into regions with amplitude below an explicit threshold, see Theorem 5.1 below. Finally, we solve the resulting smooth optimization problem with a quasi-Newton method.

The first step depends on $\Omega$, of course. In our examples we always have $\Omega \subset \mathbb{R}^k$ for $k \in \{1, 2\}$ and $\Omega$ is an interval or the product of two intervals. We fix a uniform (tensor) grid $\Omega^h$ 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 $\ell^1$ norm for the coefficients from $U$ multiplying the Dirac measures. To obtain a discrete problem we also need to discretize the Hilbert space $U$ and the quantum system. In our examples we have $U = H^1(0, T; \mathbb{C})$ or $U = L^2(0, T; \mathbb{C})$ where we use piecewise linear finite elements on a uniform grid with the appropriate discrete norms, or $U = \mathbb{C}$ were we do not need to discretize. The discrete Hilbert space is denoted by $U^h$. The control operator $B^h$ maps discrete controls to piecewise linear functions. The discretization of the quantum system depends on the system at hand. We approximate the time evolution of the discretized quantum system by a generalized Suzuki–Trotter method, see [13]. 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 $U$ in the $\ell^1(U^h)$ norm by the function $h^\theta : U \to \mathbb{R}$ given by

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

for some regularization parameter $\theta$. The function $h$ has the following two important properties: it is differentiable, and the derivatives of $h^\theta$ and the derivatives of the norm of $U$ have the same behavior outside of a small neighborhood of zero. The first property makes the cost functional differentiable. The second property preserves the sparsity of solutions, in the sense that frequency regions with zero control amplitude turn into regions with control amplitude below the explicit threshold $\theta$. More precisely we obtain the following theorem which holds e.g. in the case $\Omega = [\omega_-, \omega_+] \cap h\mathbb{Z}$ for some mesh size $h > 0$.

**Theorem 5.1** (Characterization of frequency support after frequency discretization and Huber regularization). Let $\Omega$ be a discrete set of admissible frequencies and choose any Huber regularization parameter $\theta > 0$. Consider the optimal control
problem (13) with the norm $\|u\|_{M^1(\Omega; H)}$ replaced by the Huber regularized $\ell^1(U)$ norm
$\sum_{\omega \in \Omega} h^\theta(u(\omega))$. Then optimal controls $\bar{u}$ satisfy
\begin{equation}
\{ \omega \in \Omega \mid \|\bar{u}(\omega)\|_U \geq \theta \} = \{ \omega \in \Omega \mid \|(B^* \text{Re}(\bar{\varphi}, \bar{H}\bar{\psi}))(\omega)\|_U = \alpha \}.
\end{equation}

**Proof.** In this case the cost is differentiable and the control is a function rather
than a measure with respect to frequency, and the first order optimality condition
$0 = j'(\bar{u})(\delta u)$ acquires the simple form
\begin{equation}
0 = (B^* \text{Re}(\bar{\varphi}, \bar{H}\bar{\psi}))(\omega) + \alpha \nabla h^\theta(\bar{u}(\omega))
\end{equation}
for all $\omega \in \Omega$. Since $\nabla h^\theta(z)$ equals $z/\|z\|_U$ for $\|z\|_U > \theta$ and $z/\theta$ for $\|z\|_U \leq \theta$, the assertion
follows. \hfill \Box

In the numerical examples below, only a few frequencies were contained in the
superlevel set on the left which numerically replaces the frequency support of $u$.
See e.g. Figure 3.

We solved the resulting $C^1$-smooth problem with a quasi-Newton method,
accepting the fact that the Dennis–Moré condition cannot be expected to hold, since
the functional is not guaranteed to be $C^2$ at the solution. This would be required for
local superlinear convergence, see e.g. [18]. Since the dimension of the control space
can become quite large with our approach we chose the memory efficient L-BFGS
method, see [25]. For the numerical realization gradients for the discretized prob-
lems were used. The optimization method was terminated as soon as the
$\ell^2(U^h)$ norm of the gradient relative to the largest gradient was below a tolerance of at
least $10^{-6}$.

The cost parameter is always chosen such that one achieves at least 95% of the
control objective, i.e. $\frac{1}{2} \langle \psi(T), \mathcal{O}\psi(T) \rangle_H \leq 0.025$. An automated choice of $\alpha$ would
be helpful to obtain useful cost parameters for a variety of problems.

The result of this nonlinear optimization problem also depends on the initial
guess for the control. For small $\alpha$, the initial guess for the control uses a fixed
element in $U^h$ together with a random complex phase for the different frequencies.
For larger $\alpha$, 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 $\alpha$
as initial guess.

5.2. **Three level system.** As our first example we chose a typical finite-dimen-
sional model of an atom in a laser field, see Example 2.1. We use the matrices
\[
H_0 = \begin{pmatrix}
-2 & 0 & 0 \\
0 & -1 & 0 \\
0 & 0 & 2
\end{pmatrix}, \quad H_1 = \begin{pmatrix}
0 & 0 & 1 \\
0 & 1 & 0 \\
1 & 1 & 0
\end{pmatrix}
\]
This corresponds for instance to a 1s, 2s, and 3p state. Note that the coupling
matrix elements (7) between the 1s and 2s state vanishes, i.e. the transition 1s
$\rightarrow$ 2s is “forbidden” since these states have equal parity. On the other hand the
transitions 1s $\rightarrow$ 3p and 3p $\rightarrow$ 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 then given by $\psi_0 = (1, 0, 0)$ and $\mathcal{O} = \text{diag}(1, 1, 0)$. 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 $\Omega$. We chose a time horizon of
$T = 100$ and a time grid with 4096 points. The time horizon was chosen large
enough to allow for sufficiently many periods with the transition frequencies. We
chose the cost from Example 3.1, i.e. $\|u\|_{\mathcal{M}(\Omega; H^1_0(0,T; C))}$ with the control operators
The control objective was achieved to more than 99.999%.

The great advantage of the measure space control is that we can decompose the field into simple components. Figure 2 shows the optimal control. We see that only two frequencies have a visible contribution. They correspond to the two Bohr frequencies $\omega_1$ and $\omega_2$. The time profiles are smoothly switched on and off and remain active during the whole time. This is consistent with the choice $U = H^1_{10}(0,T; C)$, 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 reflects the intuitive idea that we have to induce the transition between levels one and three before that between levels three and two. In fact, this decomposition into two pulses 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 [4], whose achievement of the control objective was beautifully demonstrated experimentally.

Figure 3 gives a more detailed numerical illustration of the frequency sparsity of the optimal control. Due to the effect of Huber regularization with $\theta > 0$, the numerical analog of the support of $u$ is the superlevel set $\{ \omega \mid \| u(\omega) \|_U > \theta \}$ (see equation (34)), which is seen to consist of just two frequencies.

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 2.2. 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 successful 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 compared to the previous one. We focus on a comparison between controls generated for different choices of the space $U$ of time profiles and the control operator $B$.

For simplicity we limit ourselves to one reaction coordinate, i.e. space dimension $d = 1$. The control objective is to reach the potential well on the right starting from the potential well on the left, see Figure 1. The initial state $\psi_0$ is a Gaussian located...
in the lower well. The observation operator $O$ is the projection on the complement of functions with support on the lower surface on the right of the potential barrier. Instead of the physical coupling given by (9) we use a coupling operator $H_1$ given by the multiplication with the reaction coordinate on the diagonal and the identity on the off-diagonal.

The energy differences between the two potential energy surfaces measured at the local minima of the lower surface are approximately 0.074 and 0.048. Therefore 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 periods with the Bohr frequencies $\omega_1$ and $\omega_2$, 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. We validated that this discretization is sufficient for the range of parameters relevant to this application.

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

- **Two-scale synthesis, smooth functions of time.** $\Omega = [1/30, 1/10]$, $\mathcal{U} = H^1_0(0, T; \mathbb{C})$, $B$ as in Example 3.1. The frequency band $\Omega$ contains the expected transition frequencies $\omega_1$ and $\omega_2$. 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.
For smooth controls $u$ similarities of frequencies. In the following we will discuss the most significant differences and similarities. Gabor time-frequency representation leads to an extremely small number of active controls, and after discretization, the contribution becomes

$$\|u\|_{L^2}^2 = 4.49 \cdot 10^{-3}$$

as $\|u\|_{L^2}$.

Thus $\|u\|_{L^2}$ is near the value of $95\%$, see Table 1.

### Table 1. Cost parameters $\alpha$ and expectation values $f(\bar{u})$ for different control spaces. A value of $f(\bar{u}) = 2.5 \cdot 10^{-2}$ corresponds to a $95\%$ achievement of the control goal.

| control space                  | $\alpha$ | $f(\bar{u})$  |
|-------------------------------|----------|---------------|
| $\mathcal{M}(\Omega; H^1_0)$  | 0.03     | 4.49 $\cdot$ 10^{-3} |
| $\mathcal{M}(\Omega; L^2)$    | 0.03     | 2.69 $\cdot$ 10^{-2} |
| $\mathcal{M}(\Omega; \mathbb{C})$ | 0.06   | 2.24 $\cdot$ 10^{-2} |
| $\mathcal{M}(\Omega \times [0, T]; \mathbb{C})$ | 0.02  | 1.89 $\cdot$ 10^{-2} |
| $L^2(0, T)$                    | 0.0001  | 5.39 $\cdot$ 10^{-6} |

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

• **Fourier synthesis, constant functions of time.** $\Omega = [1/30, 1/10]$, $\mathcal{U} = \mathbb{C}$, $B$ the Fourier synthesis operator, see Example 3.3. The frequency band $\Omega$ is discretized with 100 grid points. This results in $2 \cdot 100 = 200$ real degrees of freedom.

• **Time-frequency Gabor synthesis.** $\Omega = [1/30, 1/10] \times [0, T]$, $\mathcal{U} = \mathbb{C}$, $B$ as in Example 3.4. The time-frequency cylinder $\Omega$ 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 functional $J(\bar{v}, v) = \frac{1}{2} \langle \psi(T), O\bar{v}(T) \rangle + \alpha \|v\|_{L^2}^2$. We use linear finite elements on the time grid of the time stepping method. This results in 2048 real degrees of freedom.

For the simulations we chose $\alpha$ such that probability to end up in the desired subspace is near the value of $95\%$, see Table 1.

To compare with the standard $L^2$ control $\bar{v}$, we also computed a sparse representation $u$ of the latter, by a posteriori minimization of $\|Bu - \bar{v}\|^2_{L^2(0,T)} + \alpha \|u\|_{M(\Omega; H^1(0,T;\mathbb{C}))}^2$ with the two-scale synthesis operator $B$ and a value $\alpha = 10^{-4}$ chosen to achieve a good approximation of the optimal field.

Figure 4 shows the optimal controls. We see that their structure depends on the different cost functionals, and that the measure cost (14) for the two-scale and Gabor time-frequency representation leads to an extremely small number of active frequencies. In the following we will discuss the most significant differences and similarities.

In the second column of Figure 4 we show the total magnitude of the control as function of frequency. For smooth controls $u = u(\omega, t)$, and before discretization, this contribution is defined as $\|u(\omega, \cdot)|\|_{\mathcal{U}}$ for the first three choices of the cost, as $\int_0^T |u(\omega, \cdot)| dt$ for the time-frequency Gabor approach, and as $\|u(\omega, \cdot)||_{H^1(0,T;\mathbb{C})}$ for the $L^2$ control in order to compare it to the first cost. For the frequency-sparse controls, and after discretization, the contribution becomes $\|u(\omega, \cdot)||_{\ell^1}$, see Section 5.1, respectively — in the time-frequency Gabor case — the discrete $\ell^1$ norm of $u(\omega, \cdot)$ on the time grid. We focus our attention to the regions around
**Figure 4.** Optimal controls for Schrödinger dynamics on two potential energy surfaces (Example 2.2). Rows: Different choices of cost functionals and control operators. Columns: Time, frequency, and time-frequency representation of the optimal controls (i.e. $(Bu)(t), |u|(|\omega|)$ and $u(\omega, t)$). The dashed red line in the middle column, rows 1 to 3, indicates the Huber threshold, and the nonzero contributions below it are an artifact of the Huber regularization (see Theorem 5.1). In the rightmost column, the absolute values of the optimal measures are plotted in the time-frequency plane. Note that in column 1, 2, and 4 far fewer frequencies are active compared to the standard $L^2$ control (column 5).
the two Bohr frequencies $\omega_1$ and $\omega_2$. The regions correspond to the transitions up from the first well and down into the second well, respectively. The expected and desired behavior is that only those regions contain active frequencies. This is clearly obtained in the four choices of the sparsity promoting cost functionals. The time-frequency representation of the $L^2$ control contains many more active frequencies compared to the other controls.

It is remarkable how simple the structure of the control becomes in the two cases where the Gabor transform is used. The two-scale synthesis and the Fourier cases lead to slightly larger number of active frequencies.

The choice of a specific cost functional may be guided by the concrete application under consideration.

6. Conclusion. Measure valued costs imposed on time-frequency representations of the electric field as introduced in this paper systematically produce frequency sparse controls, in contrast to standard $L^2$ costs. The resulting controls resemble physically intuitive controls designed by experimentalists. 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 be further exploited to construct controls suited for concrete experimental setups.

Acknowledgments. The work of F. Henneke was funded by DFG through the International Research Training Group IGDK 1754. Karl Kunisch gratefully acknowledges support by the ERC advanced grant 668998 (OCLOC) under the EU’s H2020 research program.

REFERENCES

[1] A. Auger, A. Ben, H. Yedder, E. Cances, C. L. Bris, C. M. Dion, A. Keller and O. Atabek, Optimal laser control of molecular systems: Methodology and results, Math. Models Methods Appl. Sci., 12 (2012), 1281–1315.

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

[3] J. M. Ball, J. E. Marsden and M. Slemrod, Controllability for distributed bilinear systems, SIAM J. Control Optim., 20 (1982), 575–597.

[4] H. Bergmann, H. Theuer and B. W. Shore, Coherent population transfer among quantum states of atoms and molecules, Rev. Mod. Phys., 70 (1998), 1003–1025.

[5] E. Casas, R. Herzog and G. Wachsmuth, Optimality conditions and error analysis of semilinear elliptic control problems with $L^1$ cost functional, SIAM J. Optim., 22 (2012), 795–820.

[6] G. Ciaramella and A. Borzì, A \{LONE\} code for the sparse control of quantum systems, Computer Physics Communications, 200 (2016), 312–323.

[7] G. Ciaramella and A. Borzì, Quantum optimal control problems with a sparsity cost functional, Numer. Funct. Anal. Optim., 37 (2016), 938–965.

[8] F. H. Clarke, Optimization and Nonsmooth Analysis, vol. 5 of Classics in Applied Mathematics, 2nd edition, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990.

[9] C. Clason and K. Kunisch, A duality-based approach to elliptic control problems in non-reflexive Banach spaces, ESAIM Control Optim. Calc. Var., 17 (2011), 243–266.

[10] 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.

[11] S. J. Glaser, U. Boscain, T. Calarco, C. P. Koch, W. Köckenberger, R. Kosloff, I. Kuprov, B. Luy, S. Schirmer, T. Schulte-Herbrüggen, D. Sugny and F. K. Wilhelm, Training Schrödinger’s cat: Quantum optimal control, The European Physical Journal D, 69 (2015), 1–24.
12. 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.

13. F. Henneke and M. Liebmann, A generalized Suzuki–Trotter type method in optimal control of coupled Schrödinger equations, Computing and Visualization in Science, 17 (2015), 277–293.

14. R. Herzog, G. Stadler and G. Wachsmuth, Directional sparsity in optimal control of partial differential equations, SIAM J. Control Optim., 50 (2012), 943–963.

15. M. Hintermüller, D. Marahrens, P. A. Markowich and C. Sparber, Optimal bilinear control of Gross-Pitaevskii equations, SIAM J. Control Optim., 51 (2013), 2509–2543.

16. P. v. d. Hoff, S. Thallmair, M. Kowalewski, R. Siemerling and R. D. Vivie-Riedle, Optimal control theory - closing the gap between theory and experiment, Phys. Chem. Chem. Phys., 14 (2012), 14460–14485.

17. K. Ito and K. Kunisch, Optimal bilinear control of an abstract Schrödinger equation, SIAM Journal on Control and Optimization, 46 (2007), 274–287.

18. A. F. Izmailov, A. L. Pogosyan and M. V. Solodov Semismooth SQP method for equality-constrained optimization problems with an application to the lifted reformulation of mathematical programs with complementarity constraints Optimization Methods and Software, 26 (2011), 847–872.

19. K. Kormann, S. Holmgren and H. Karlsson, A Fourier-coefficient based solution of an optimal control problem in quantum chemistry, Journal of Optimization Theory and Applications, 174 (2010), 491–506.

20. K. Kunisch, K. Pieper and B. Vexler, Measure valued directional sparsity for parabolic optimal control problems, SIAM J. Control Optim., 52 (2014), 3078–3108.

21. S. Lang, Real and Functional Analysis, vol. 142 of Graduate Texts in Mathematics, 3rd edition, Springer-Verlag, New York, 1993.

22. 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.

23. X. J. Li and J. M. Yong, Optimal Control Theory for Infinite-Dimensional Systems, Systems & Control: Foundations & Applications, Birkhäuser Boston Inc., Boston, MA, 1995.

24. L. Meziani, On the dual space $C^*$ of $(S, X)$, Acta Math. Univ. Comenian. (N.S.), 78 (2009), 153–160.

25. J. Nocedal and S. J. Wright, Numerical Optimization, 2nd edition, Springer Series in Operations Research and Financial Engineering, Springer, New York, 2006.

26. 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), 4950–4964.

27. M. Reed and B. Simon, Methods of Modern Mathematical Physics. II. Fourier analysis, Self-adjointness, Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975.

28. 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), 014111.

29. S. Ruetzel, C. Stokenberger, F. Dimler, D. J. Tannor and T. Brixner, Adaptive coherent control using the von Neumann basis, Phys. Chem. Chem. Phys., 13 (2011), 8627–8636.

30. 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), 093022.

31. 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), 064108.

32. B. W. Shore, The Theory of Coherent Atomic Excitation, Wiley, New York, NY, 1990.

33. G. Stadler, Elliptic optimal control problems with $L^1$-control cost and applications for the placement of control devices, Comput. Optim. Appl., 44 (2009), 159–181.

34. S. Teufel, Adiabatic Perturbation Theory in Quantum Dynamics, vol. 1821 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2003.

35. G. Turinici, C. Le Bris and H. Rabitz, Efficient algorithms for the laboratory discovery of optimal quantum controls, Phys. Rev. E, 70 (2004), 016704.
[36] G. von Winckel and A. Borzì, Computational techniques for a quantum control problem with $H^1$-cost, *Inverse Problems*, 24 (2008), 034007, 23pp.

[37] 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), 4176–4203.

[38] G. Vossen and H. Maurer, On $L^1$-minimization in optimal control and applications to robotics, *Optimal Control Appl. Methods*, 27 (2006), 301–321.

Received April 2017; revised October 2017.

E-mail address: gf@ma.tum.de
E-mail address: felix.henneke@fu-berlin.de
E-mail address: karl.kunisch@uni-graz.at