OPTIMAL BILINEAR CONTROL OF GROSS–PITAEVSKII EQUATIONS[*]

MICHAEL HINTERMÜLLER[†], DANIEL MARAHRENS[‡], PETER A. MARKOWICH[§], AND CHRISTOF SPARBER[¶]

Abstract. A mathematical framework for optimal bilinear control of nonlinear Schrödinger equations of Gross–Pitaevskii type arising in the description of Bose–Einstein condensates is presented. The obtained results generalize earlier efforts found in the literature in several aspects. In particular, the cost induced by the physical workload over the control process is taken into account rather than the often used L^2- or H^1-norms for the cost of the control action. Well-posedness of the problem and existence of an optimal control are proved. In addition, the first order optimality system is rigorously derived. Also a numerical solution method is proposed, which is based on a Newton-type iteration, and used to solve several coherent quantum control problems.

Key words. quantum control, bilinear optimal control problem, nonlinear Schrödinger equation, Bose–Einstein condensate, Newton's method, MINRES algorithm, work induced by control

AMS subject classifications. 49J20, 81Q93, 49J50

DOI. 10.1137/120866233

1. Introduction.

1.1. Physics background. Ever since the first experimental realization of *Bose–Einstein condensates* (BECs) in 1995, the possibility of storing, manipulating, and measuring a single quantum system with extremely high precision has provided great stimulus in many fields of physical and mathematical research, among them *quantum control theory*. In the regime of dilute gases, a BEC, consisting of N particles, can be modeled by the *Gross–Pitaevskii equation* (GPE) [28], i.e., a cubically nonlinear Schrödinger equation (NLS) of the form

$$i\hbar\partial_t\psi = -\frac{\hbar^2}{2m}\Delta\psi + U(x)\psi + Ng|\psi|^2\psi + W(t,x)\psi, \quad x\in\mathbb{R}^3,\ t\in\mathbb{R},$$

with m denoting the mass of the particles, $\hbar$ denoting Planck's constant, $g = 4\pi\hbar^2 a_{\rm sc}/m$, and $a_{\rm sc}\in\mathbb{R}$ their characteristic scattering length, describing the interparticle collisions. The function $U(x)$ describes an external trapping potential which is necessary for the experimental realization of a BEC. Typically, $U(x)$ is assumed to be a *harmonic*

confinement. In situations where $U(x)$ is strongly anisotropic, one experimentally obtains a quasi–one-dimensional ("cigar shaped"), or quasi–two-dimensional ("pancake shaped") BEC; see, for instance, [20]. In the following, we shall assume $U(x)$ to be *fixed.* The condensate is consequently manipulated via a time-dependent *control potential* $W(t,x)$, which we shall assume to be of the following form:

$$W(t,x) = \alpha(t)V(x).$$

Here, $\alpha(t)$ denotes the *control parameter* (typically, a switching function acting within a certain time-interval $[0,T]$) and $V(x)$ is a given potential. In our context, the potential $V(x)$ models the spatial profile of a laser field used to manipulate the BEC and $\alpha(t)$ is its intensity [31].

The problem of quantum control, i.e., the coherent manipulation of quantum systems (in particular BECs) via external potentials $W(t,x)$, has attracted considerable interest in the physics literature; cf. [7, 12, 16, 18, 27, 29, 38, 30]. From the mathematical point of view, quantum control problems are a specific example of *bilinear control systems* [13]. It is known that linear or nonlinear Schrödinger-type equations are in general *not exactly controllable* in, say, $L^2(\mathbb{R}^3)$; cf. [32]. Similarly, approximate controllability is known to hold for only some specific systems, such as those in [23]. More recently, however, sufficient conditions for approximate controllability of linear Schrödinger equations with purely discrete spectrum have been derived in [11]. In [22] these conditions have been shown to be generically satisfied, but, to the best of our knowledge, a generalization to the case of NLS equations is still lacking.

The goal of the current paper is to consider quantum control systems within the framework of *optimal control* (cf. [35] for a general introduction) from a partial differential equation constrained point of view. The objective of the control process is thereby quantified through an objective functional $J = J(\psi, \alpha)$, which is minimized subject to the condition that the time-evolution of the quantum state is governed by the GPE. Such objective functionals $J(\psi, \alpha)$ usually consist of two parts, one being the desired physical quantity (observable) to be minimized and the other describing the cost it takes to obtain the desired outcome through the control process. In quantum mechanics, the wave function $\psi(t, \cdot)$ itself is not a physical observable. Rather, one considers self-adjoint linear operators A acting on $\psi(t, \cdot)$ and aims for a prescribed *expectation value* of A at time $t = T > 0$, the final time of the control process. Such expectation values are computed by taking the L^2-inner product $\langle \psi(T, \cdot), A\psi(T, \cdot)\rangle_{L^2(\mathbb{R}^d)}$. Note that this implies that the corresponding $\psi(t, \cdot)$ is determined only up to a constant phase. This fact makes quantum control less "rigid" when compared to classical control problems in which one usually aims to optimize for a prescribed target state.

There are many possible ways of modeling the cost it takes to reach a certain prescribed expectation value. The corresponding cost terms within $J(\psi, \alpha)$ are often given by the norm of the control $\alpha(t)$ in some function space. A typical choice is $L^2(0,T)$ or $H^1(0,T)$. However, cost terms based on the L^2-norm of α tend to yield highly oscillatory optimal controls due to the oscillatory nature of the underlying (nonlinear) Schrödinger equation. The same is true for quantum control via so-called Lyapunov tracking methods; see, e.g., [14]. In order to reduce this problem one often uses cost terms involving the H^1-norm. Any such choice, though, can be seen as ad-hoc, and it has therefore been attempted to include more general cost terms which allow for a direct physical interpretation within the underlying quantum system; see, e.g., [27]. In the present work, we shall follow these ideas and present a novel choice

for the cost terms, which is based on the corresponding physical work performed throughout the control process.

We continue this introductory section by describing the mathematical setting in more detail.

1.2. Mathematical setting. We consider a quantum mechanical system described by a wave function $\psi(t, \cdot) \in L^2(\mathbb{R}^d)$ within $d = 1, 2, 3$ spatial dimensions. The case $d = 1, 2$ models the effective dynamics within strongly anisotropic potentials (resulting in a quasi–one-or-two-dimensional BEC). The time-evolution of $\psi(t, \cdot)$ is governed by the following generalized GPE (rescaled into dimensionless form):

$$(1.1) \qquad i\partial_t \psi = -\frac{1}{2}\Delta\psi + U(x)\psi + \lambda|\psi|^{2\sigma}\psi + \alpha(t)V(x)\psi, \quad x \in \mathbb{R}^d, \ t \in \mathbb{R},$$

with $\lambda \geqslant 0$, $\sigma < 2/(d-2)$, and subject to initial data

$$\psi(0, \cdot) = \psi_0 \in L^2(\mathbb{R}^d), \quad \alpha(0) = \alpha_0 \in \mathbb{R}.$$

For physical reasons we normalize $\|\psi_0\|_{L^2(\mathbb{R}^d)} = 1$, which is henceforth preserved by the time-evolution of (1.1). In addition, the control potential is assumed to be $V \in W^{1,\infty}(\mathbb{R}^d)$, whereas for $U(x)$ we require

$$U \in C^\infty(\mathbb{R}^d) \text{ such that } \partial^k U \in L^\infty(\mathbb{R}^d) \ \forall \text{ multi-indices } k \text{ with } |k| \geqslant 2.$$

In other words, the external potential is assumed to be smooth and *subquadratic*. One of the most important examples is the harmonic oscillator $U(x) = \frac{1}{2}|x|^2$. Due to the presence of a subquadratic potential, we restrict ourselves to initial data ψ_0 in the *energy space*

$$(1.2) \qquad \Sigma := \left\{ \psi \in H^1(\mathbb{R}^d) \ : \ x\psi \in L^2(\mathbb{R}^d) \right\}.$$

In particular, this definition guarantees that the quantum mechanical energy functional

$$(1.3) \quad E(t) = \int_{\mathbb{R}^d} \frac{1}{2}|\nabla\psi(t,x)|^2 + \frac{\lambda}{\sigma+1}|\psi(t,x)|^{2\sigma+2} + (\alpha(t)V(x) + U(x))|\psi(t,x)|^2 dx$$

associated to (1.1) is well defined.

Remark 1.1. Note that $\sigma < 2/(d-2)$ allows for general power law nonlinearities in dimensions $d = 1, 2$, whereas in $d = 3$ the nonlinearity is assumed to be less than quintic. From the physics point of view a cubic nonlinearity $\sigma = 1$ is the most natural choice, but higher order nonlinearities also arise in systems with more complicated interparticle interactions, in particular in lower dimensions; cf. [20]. From the mathematical point of view, it is well known that the restriction $\sigma < 2/(d-2)$ guarantees well-posedness of the initial value problem in the energy space Σ; see [9, 10]. In addition, the condition $\lambda \geqslant 0$ (defocusing nonlinearity) guarantees the existence of global in-time solutions to (1.1); see [9]. Hence, we do not encounter the problem of finite-time blow-up in our work.

Although (1.1) conserves mass, i.e., $\|\psi(t, \cdot)\|_{L^2(\mathbb{R}^d)} = \|\psi_0\|_{L^2(\mathbb{R}^d)}$ for all $t \in \mathbb{R}$, the energy $E(t)$ is not conserved. This is in contrast to the case of time-independent potentials. In our case, rather, one finds that

$$(1.4) \qquad \frac{d}{dt}E(t) = \dot\alpha(t)\int_{\mathbb{R}^d} V(x)|\psi(t,x)|^2 dx.$$

The *physical work* performed by the system within a given time-interval $[0, T]$ is therefore equal to

$$(1.5) \qquad E(T) - E(0) = \int_0^T \dot{\alpha}(t) \int_{\mathbb{R}^d} V(x)|\psi(t,x)|^2 dx \, dt.$$

Thus a control $\alpha(t)$ acting for $t \in [0, T]$ upon a system described by (1.1) requires a certain amount of energy, which is given by (1.5). Thus, it seems natural to include such a term in the cost functional of our problem in order to quantify the control action.

Indeed, for any given final control time $T > 0$ and parameters $\gamma_1 \geqslant 0, \gamma_2 > 0$, we define the following *objective functional*:

$$(1.6) \qquad J(\psi, \alpha) := \langle \psi(T, \cdot), A\psi(T, \cdot) \rangle_{L^2(\mathbb{R}^d)}^2 + \gamma_1 \int_0^T (\dot{E}(t))^2 \, dt + \gamma_2 \int_0^T (\dot{\alpha}(t))^2 \, dt,$$

where $A : \Sigma \to L^2(\mathbb{R}^d)$ is a bounded linear operator which is assumed to be essentially self-adjoint on $L^2(\mathbb{R}^d)$. In other words, A represents a physical observable with $\text{spec}\,(A) \subseteq \mathbb{R}$. A typical choice for A would be $A = A' - a$, where $a \in \mathbb{R}$ is some prescribed expectation value for the observable A' in the state $\psi(T, x)$. For example, if $a \in \text{spec}\,(A')$ is chosen to be an eigenvalue of A', the first term in $J(\psi, \alpha)$ is zero as soon as the target state $\psi(T, \cdot)$ is, up to a phase factor, given by an associated eigenfunction of A'. However, one may consider choosing $a \in \mathbb{R}$ such that it "forces" the functional to equidistribute between, say, two eigenfunctions.

Remark 1.2. We also remark that in the case $A = P_\varphi - 1$, where P_φ denotes the orthogonal projection onto a given *target state* $\varphi \in L^2(\mathbb{R}^d)$, the first term on the right-hand side of (1.6) reads

$$(1.7) \qquad \langle \psi(T, \cdot), A\psi(T, \cdot) \rangle_{L^2(\mathbb{R}^d)} = \left| \langle \psi(T, \cdot), \varphi(\cdot) \rangle_{L^2(\mathbb{R}^d)} \right|^2 - 1,$$

using the fact that $\|\psi(T, \cdot)\|_{L^2(\mathbb{R}^d)} = 1$. Expression (1.7) is the same as that used in recent works in the physics literature; see [16].

In (1.6), the specific choice of the exponent 2 ensures nice differentiability properties of the functional, as well as its positivity, and is therefore used for the same reasons as in the case of the usual L^2 or H^1 cost terms. Using (1.4), we find that the objective functional $J(\psi, \alpha)$ explicitly reads

$$(1.8) \qquad \begin{aligned} J(\psi, \alpha) := &\langle \psi(T, \cdot), A\psi(T, \cdot) \rangle_{L^2(\mathbb{R}^d)}^2 \\ &+ \gamma_1 \int_0^T (\dot{\alpha}(t))^2 \left(\int_{\mathbb{R}^d} V(x)|\psi(t,x)|^2 \, dx \right)^2 \, dt + \gamma_2 \int_0^T (\dot{\alpha}(t))^2 \, dt. \end{aligned}$$

Here, the second line on the right-hand side displays two cost (or penalization) terms for the control: The first, involving $\gamma_1 \geqslant 0$, is given by the square of the physical work, i.e., the right-hand side of (1.4). The second is a classical cost term as used in [16]. In our case, the second term is required as a mathematical *regularization* of the optimal control problem, since for general (sign changing) potentials $V \in L^\infty(\mathbb{R}^d)$ the weight factor

$$(1.9) \qquad \omega(t) := \int_{\mathbb{R}^d} V(x)|\psi(t,x)|^2 dx$$

might vanish for some $t \in \mathbb{R}$. In such a situation, the boundedness of variations of $\alpha(t)$ is in jeopardy and the optimal control problem lacks well-posedness. Hence, we require $\gamma_2 > 0$ for our mathematical analysis, but typically take $\gamma_2 \ll \gamma_1$ in our numerics in section 5 to keep its influence small. Note, however, that in the case where the control potential satisfies the positivity condition

$$V(x) \geqslant \delta > 0 \ \ \forall x \in \mathbb{R}^d,$$

we may choose $\gamma_2 = 0$, and all of our results remain valid.

Remark 1.3. In situations where the above positivity condition on $V(x)$ does not hold, one might think of performing a time-dependent *gauge transform* of ψ, i.e.,

$$\tilde{\psi}(t,x) = \exp\left(-i\kappa \int_0^t \alpha(s) \ ds\right)\psi(t,x),$$

with a constant $\kappa > \min_{x \in \mathbb{R}^d} V(x)$, assuming that the minimum exists. This yields a GPE for the wave function $\tilde{\psi}$ with modified control potential $\tilde{V}(x) = (\kappa + V(x)) > 0$ for all $x \in \mathbb{R}^d$. Note, however, that this gauge transform leaves the expression (1.8) unchanged and hence does not improve the situation. It is only when one also changes the potential $V(x)$ within $J(\psi, \alpha)$ into $\tilde{V}(x)$ that the problem does not require any regularization term (proportional to γ_2). Note, however, that such a modification yields a control system which is no longer (mathematically) equivalent to the original problem. In fact, replacing $V(x)$ by $\tilde{V}(x)$ in the objective functional $J(\psi, \alpha)$ corresponds to increasing the parameter γ_2 by κ.

1.3. Relation to other works and organization of the paper. The mathematical research field of optimal bilinear control of systems governed by partial differential equations is by now classical; cf. [15, 21] for a general overview. Surprisingly, rigorous mathematical work on optimal (bilinear) control of quantum systems appears very limited, despite the physical significance of the involved applications (cf. the references given above). Results on simplified situations such as, e.g., for finite-dimensional quantum systems, can be found in [6] (see also the references therein). More recently, optimal control problems for linear Schrödinger equations have been studied in [2, 4, 17, 8]. In addition, numerical questions related to quantum control are studied in [3, 37, 5, 27, 33, 34]. Among these papers, the work in [17] appears closest to our effort. Indeed, in [17], the authors provide a framework for bilinear optimal control of abstract (linear) Schrödinger equations. The considered objective functional involves a cost term proportional to the L^2-norm of the control parameter $\alpha(t)$. The present work goes beyond the results obtained in [17] in several respects: First, we generalize the cost functional to account for oscillations in $\alpha(t)$ and, in particular, for the physical workload performed throughout the control process. In addition, we allow for observables A which are unbounded operators on L^2. Second, we consider NLSs of Gross–Pitaevskii type, including unbounded (subquadratic) potentials, which are highly significant in the quantum control of BECs. This type of equation makes the study of the associated control problem considerably more involved from a mathematical point of view. We finally remark that in [27] a more general class of objective functionals has been proposed in an attempt to implement the concept of energy minimization into quantum control systems. The present paper, however, seems to be the first rigorous study in this direction.

The rest of this work is organized as follows. In section 2 we clarify existence of a minimizer for our control problem. In particular, we prove that the corresponding

optimal solution $\psi_*(t,x)$ is indeed a mild (and not only a weak) solution of (1.1), depending continuously on the initial data ψ_0. Then, in section 3 the adjoint equation is derived and analyzed with respect to existence and uniqueness of a solution. It is our primary tool for the description of the derivative of the objective function reduced onto the control space through considering the solution of the GPE as a function of the control variable α. The results of section 3 are paramount for the derivation of the first order optimality system in section 4. In section 5 a gradient- and a Newton-type descent method are defined, respectively, and then used for computing numerical solutions for several illustrative quantum control problems. In particular, we consider the optimal shifting of a linear wave package, splitting of a linear wave package, and splitting of a BEC. The paper ends with conclusions on our findings in section 6.

Notation. Throughout this work we shall denote strong convergence of a sequence $(x_n)_{n\in\mathbb{N}}$ by $x_n \to x$ and weak convergence by $x_n \rightharpoonup x$. For simplicity, we shall often write $\psi(t) \equiv \psi(t,\cdot)$ and also use the shorthand notation $L_t^p L_x^q$ instead of $L^p(0,T;L^q(\mathbb{R}^d))$. Similarly, H_t^1 stands for $H^1(0,T)$, with dual $H_t^{-1} \equiv (H^1(0,T))^*$.

2. Existence of minimizers. We start by specifying the basic functional analytic framework. For any given $T > 0$, we consider $H^1(0,T)$ as the real vector space of control parameters $\alpha(t) \in \mathbb{R}$. It is known [10] that for every $\alpha \in H^1(0,T)$, there exists a unique *mild solution* $\psi \in C([0,T];\Sigma)$ of the GPE. More precisely, ψ solves

$$\psi(t,x) = S(t)\psi_0(x) - i\int_0^t S(t-s)\left(\lambda|\psi(s,\cdot)|^{2\sigma}\psi(s,\cdot) + \alpha(s)V\psi(s,\cdot)\right)(x)ds,$$

where from now on we denote by

$$(2.1) \qquad\qquad S(t) = e^{-itH}, \quad H = -\frac{1}{2}\Delta + U(x)$$

the group of unitary operators $\{S(t)\}_{t\in\mathbb{R}}$ generated by the Hamiltonian H. In other words, $S(t)$ describes the time-evolution of the linear, uncontrolled system. Next, we define

$$(2.2) \qquad\qquad \Upsilon(0,T) := L^2(0,T;\Sigma) \cap H^1(0,T;\Sigma^*),$$

where Σ^* is the dual of the energy space Σ. Then the appropriate space for our minimization problem is

$$\Lambda(0,T) := \{(\psi,\alpha) \in \Upsilon(0,T) \times H^1(0,T) : \psi \text{ is a mild solution of (1.1)}\}.$$

Since the control α is real-valued, it is natural to consider $\Lambda(0,T)$ as a *real vector space*, and we shall henceforth equip $L^2(\mathbb{R}^d)$ with the scalar product

$$(2.3) \qquad\qquad \langle \xi, \psi \rangle_{L^2(\mathbb{R}^d)} = \mathrm{Re} \int_{\mathbb{R}^d} \xi(x)\overline{\psi(x)}\,dx,$$

which is subsequently inherited by all L^2-based Sobolev spaces. (Note that this choice is also used in [10].) From what was said above, we infer that the space $\Lambda(0,T)$ is indeed nonempty.

With these definitions at hand, the optimal control problem under investigation is to find

$$(2.4) \qquad J_* = \inf_{(\psi,\alpha)\in\Lambda(0,T)} J(\psi,\alpha).$$

We are now in the position to state the first main result of this work.

THEOREM 2.1. *Let* $\lambda \geqslant 0$, $0 < \sigma < 2/(d-2)$, $V \in W^{1,\infty}(\mathbb{R}^d)$, *and* $U \in C^\infty(\mathbb{R}^d)$ *be subquadratic. Then, for any* $T > 0$, *any initial data* $\psi_0 \in \Sigma$, $\alpha_0 \in \mathbb{R}$, *and any choice of parameters* $\gamma_1 \geqslant 0$, $\gamma_2 > 0$, *the optimal control problem* (2.4) *has a minimizer* $(\psi_*, \alpha_*) \in \Lambda(0,T)$.

The proof of this theorem will be split into three steps: In subsection 2.1 we shall first prove a convergence result for minimizing or, more precisely, infimizing sequences. We consequently deduce in subsection 2.2 that the obtained limit ψ_* is indeed a mild solution of (1.1). Finally, we shall prove lower semicontinuity of $J(\psi, \alpha)$ with respect to the convergence obtained before.

2.1. Convergence of infimizing sequences. First note that there exists at least one infimizing sequence with an infimum $-\infty \leqslant J_* < +\infty$, since $\Lambda(0,T) \neq \emptyset$ and $J : \Lambda(0,T) \to \mathbb{R}$. Then we have the following result for any infimizing sequence.

PROPOSITION 2.2. *Let* $(\psi_n, \alpha_n)_{n\in\mathbb{N}}$ *be an infimizing sequence of the optimal control problem given by* (1.6). *Then under the assumptions of Theorem* 2.1 *there exist a subsequence, still denoted by* $(\psi_n, \alpha_n)_{n\in\mathbb{N}}$, *and functions* $\alpha_* \in H^1(0,T)$, $\psi_* \in L^\infty(0,T;\Sigma)$, *such that*

$$\alpha_n \rightharpoonup \alpha_* \text{ in } H^1(0,T), \qquad \alpha_n \to \alpha_* \text{ in } L^2(0,T),$$
$$\psi_n \rightharpoonup \psi_* \text{ in } L^2(0,T;\Sigma),$$
$$\psi_n \to \psi_* \text{ in } L^2(0,T;L^2(\mathbb{R}^d)) \cap L^2(0,T;L^{2\sigma+2}(\mathbb{R}^d)),$$

as $n \to +\infty$. *Furthermore it holds that*

$$(2.5) \qquad \psi_n(t) \to \psi_*(t) \text{ in } L^2(\mathbb{R}^d), \qquad \psi_n(t) \rightharpoonup \psi_*(t) \text{ in } \Sigma$$

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

Proof. By definition, $J \geqslant 0$, and thus it is bounded from below. For an infimizing sequence $(\psi_n, \alpha_n)_{n\in\mathbb{N}}$ the sequence of objective functional values $(J(\psi_n,\alpha_n))_{n\in\mathbb{N}}$ converges and is bounded on $\mathbb{R}$. Hence, it holds that $J(\psi_n,\alpha_n) \leqslant C < +\infty$ for all $n \in \mathbb{N}$. Since $\gamma_2 > 0$ it follows that

$$\int_0^T (\dot\alpha_n(t))^2 \, dt \leqslant C < +\infty.$$

For smooth $\alpha_n : [0,T] \to \mathbb{R}$ we compute

$$\alpha_n(t) = \alpha_n(0) + \int_0^t \dot\alpha_n(s) \, ds \leqslant \alpha_n(0) + \left(T \int_0^T (\dot\alpha_n(s))^2 \, ds \right)^{1/2} < +\infty,$$

and thus α_n is bounded in $L^\infty(0,T)$. By approximation (using the fact that $\alpha_n(0) = \alpha_0$ is fixed), the sequence $(\alpha_n)_{n\in\mathbb{N}}$ is uniformly bounded in $L^\infty(0,T)$, which in turn implies a uniform bound in $L^2(0,T)$ and thus in $H^1(0,T)$. Hence, there exists a subsequence, still denoted $(\alpha_n)_{n\in\mathbb{N}}$, and $\alpha_* \in H^1(0,T)$ such that

$$\alpha_n \rightharpoonup \alpha_* \in H^1(0,T).$$

Moreover, since $H^1(0,T)$ is compactly embedded into $L^2(0,T)$, we deduce that $\alpha_n \to \alpha_*$ in $L^2(0,T)$. Next, we recall that

$$\frac{d}{dt} E_n(t) = \dot{\alpha}_n(t) \int_{\mathbb{R}^d} V(x)|\psi_n(t,x)|^2 \, dx,$$

and hence

$$\|\dot{E}_n\|_{L_t^2} \leqslant \|\dot{\alpha}_n\|_{L_t^2} \|V\|_{L_x^\infty} \|\psi_0\|_{L_x^2}^2,$$

in view of mass conservation $\|\psi_n(t)\|_{L_x^2} = \|\psi_0\|_{L_x^2}$. Since $E_n(0) = E_0$ depends only on ψ_0 and α_0 (and is thus independent of $n \in \mathbb{N}$), the same argument as before yields $\|E_n\|_{L_t^\infty} \leqslant C$. Recalling the definition of the energy (1.3) and the fact that $\lambda \geqslant 0$, we obtain

$$(2.6) \qquad \frac{1}{2} \|\nabla \psi_n(t)\|_{L_x^2}^2 \leqslant \|E_n\|_{L_t^\infty} + c\|\alpha_n\|_{L_t^\infty} \|\psi_0\|_{L_x^2}^2 + C\|x\psi_n(t)\|_{L_x^2}^2,$$

again using conservation of mass $\|\psi_n(t)\|_{L_x^2} = \|\psi_0\|_{L_x^2}$. Furthermore, it holds that

$$\frac{d}{dt} \int_{\mathbb{R}^d} |x|^2 |\psi|^2 \, dx = 2 \operatorname{Re} \int_{\mathbb{R}^d} i|x|^2 \overline{\psi} \left(\frac{1}{2} \Delta \psi - \lambda |\psi|^{2\sigma} \psi - \lambda \alpha V \psi - U \psi \right) dx$$

$$= 2 \operatorname{Im} \int_{\mathbb{R}^d} x \overline{\psi} \nabla \psi \, dx \leqslant \|x\psi(t)\|_{L_x^2}^2 + \|\nabla \psi(t)\|_{L_x^2}^2,$$

which, in view of the bound (2.6) and Gronwall's inequality, yields

$$\|x\psi(t)\|_{L_x^2}^2 \leqslant C \left(\|E_n\|_{L_t^\infty} + \|\alpha_n\|_{L_t^\infty} \|\psi_0\|_{L_x^2}^2 \right)$$

for all $t \in [0,T]$. In summary we have shown

$$(2.7) \qquad \|\psi_n(t)\|_\Sigma^2 = \|\psi_n(t)\|_{H_x^1}^2 + \|x\psi_n(t)\|_{L_x^2}^2 \leqslant C,$$

where $C > 0$ is independent of $n \in \mathbb{N}$ and $t \in [0,T]$. Hence, ψ_n is uniformly bounded in $L^\infty(0,T;\Sigma)$ and in particular in $L^2(0,T;\Sigma)$. By reflexivity of $L^2(0,T;\Sigma)$, we consequently infer the existence of a subsequence (denoted by the same symbol) such that

$$\psi_n \rightharpoonup \psi_* \text{ in } L^2(0,T;\Sigma) \text{ as } n \to +\infty.$$

To obtain the strong convergence announced above, we first note that (1.1) implies $\partial_t \psi_n \in L^\infty(0,T;\Sigma^*)$. On the other hand, Σ is compactly embedded in $L^2(\mathbb{R}^d)$. Thus, we can apply the Aubin–Lions lemma to deduce

$$\psi_n \xrightarrow{n\to\infty} \psi_* \text{ in } L^2((0,T) \times \mathbb{R}^d).$$

In particular, there exists yet another subsequence (still denoted by the same symbol) such that

$$\psi_n(t) \xrightarrow{n\to\infty} \psi_*(t) \text{ in } L^2(\mathbb{R}^d) \text{ for almost all } t \in [0,T].$$

In order to obtain weak convergence in the energy space, i.e., $\psi_n(t) \rightharpoonup \psi_*(t)$ in Σ, we fix $t \in [0,T]$ such that $\psi_n(t) \to \psi_*(t)$ in $L^2(\mathbb{R}^d)$. In view of (2.7), every subsequence of $\psi_n(t)$ has yet another subsequence such that $\psi_n(t)$ converges weakly in Σ to some limit. On the other hand, this limit is necessarily given by $\psi_*(t)$, since $\psi_n(t) \to \psi_*(t)$ in $L^2(\mathbb{R}^d)$. Hence the whole sequence converges weakly in Σ to $\psi_*(t)$. By lower semicontinuity of the Σ-norm we can deduce $\|\psi_n(t)\|_\Sigma \leqslant C$ and thus $\psi_* \in L^\infty(0,T;\Sigma)$.

Finally, the announced convergence in $L^2(0,T;L^{2\sigma+2}(\mathbb{R}^d))$ is obtained by invoking the Gagliardo–Nirenberg inequality, i.e.,

$$(2.8) \qquad \|\xi\|_{L^r_x} \leqslant C\|\xi\|_{L^2_x}^{1-\delta(r)}\|\nabla\xi\|_{L^2_x}^{\delta(r)},$$

where $2 \leqslant r < \frac{2d}{d-2}$ and $\delta(r) = d(\frac{1}{2} - \frac{1}{r})$. This concludes the proof of Proposition 2.2. $\quad\square$

2.2. Minimizers as mild solutions. Next we prove that the limit ψ_* obtained in the previous subsection is indeed a mild solution of (1.1) with corresponding control α_*. From the physical point of view, this is important since it implies continuous (in time) dependence of ψ_* upon a given initial data ψ_0. To this end, one should also note that $H^1(0,T) \hookrightarrow C(0,T)$ (using Sobolev embeddings), and hence the obtained optimal control parameter $\alpha_*(t)$ is indeed a continuous function on $[0,T]$.

PROPOSITION 2.3. *Let $(\psi_*, \alpha_*) \in \Upsilon(0,T) \times H^1(0,T)$ be the limit obtained in Proposition* 2.2. *Then ψ_* is a mild solution of* (1.1) *with control α_* and*

$$\psi_* \in C([0,T];\Sigma) \cap C^1([0,T];\Sigma^*).$$

In particular, this implies that the convergence result (2.5) *holds for all $t \in [0,T]$.*

Proof. First we note that, by construction, each ψ_n satisfies

$$\psi_n(t) = S(t)\psi_0 - i\int_0^t S(t-s)\left(\lambda|\psi_n(s)|^{2\sigma}\psi_n(s) + \alpha_n(s)V\psi_n(s)\right)ds$$

for all $t \in [0,T]$. Here and in the following we shall suppress the x-dependence of ψ for notational convenience. In order to prove that ψ_* is a mild solution corresponding to the control α_*, we take the L^2-scalar product of the above equation with a test function $\chi \in C_0^\infty(\mathbb{R}^d)$. This yields

$$(2.9) \qquad \begin{aligned} \langle\psi_n(t),\chi\rangle_{L^2_x} &= \langle S(t)\psi_0,\chi\rangle_{L^2_x} - i\lambda\int_0^t \langle S(t-s)|\psi_n(s)|^{2\sigma}\psi_n(s),\chi\rangle_{L^2_x} ds \\ &\quad - i\int_0^t \langle S(t-s)\alpha_n(s)V\psi_n(s),\chi\rangle_{L^2_x} ds. \end{aligned}$$

In view of Proposition 2.2, the term on the left-hand side of this identity converges to the desired expression for almost all $t \in [0,T]$, i.e.,

$$\lim_{n\to\infty} \langle\psi_n(t),\chi\rangle_{L^2_x} = \langle\psi_*(t),\chi\rangle_{L^2_x}.$$

In order to proceed further, we note that for any $f \in \mathcal{D}'(\mathbb{R}^d)$ it holds that

$$(2.10) \qquad \langle S(t-s)f(s),\chi\rangle_{L^2_x} = \langle f(s),S(s-t)\chi\rangle_{L^2_x},$$

and we therefore define

$$(2.11) \qquad \tilde\chi : [0,t]\times\mathbb{R}^d \to \mathbb{C}, \quad \chi \mapsto \tilde\chi(\cdot,x) := S(\cdot-t)\chi(x),$$

for which we can prove the following regularity properties.

LEMMA 2.4. *There exists a constant $C = C(T) > 0$ such that for all $t \in [0,T]$ it holds that*

$$\sup_{s \in [0,t]} \left(\|x\tilde{\chi}(s)\|_{L_x^2} + \|\nabla\tilde{\chi}(s)\|_{L_x^2} \right) \leqslant C(T) < +\infty,$$

where the function $\tilde{\chi}$ is defined in (2.11). *In particular, the function $\tilde{\chi}$ is bounded in $L^\infty(0,t; L^{2\sigma+2}(\mathbb{R}^d))$.*

Proof of Lemma 2.4. The norm $\|\tilde{\chi}(s)\|_{L_x^2} = \|S(s-t)\chi\|_{L_x^2}$ is conserved since $S(t)$ is a unitary operator on $L^2(\mathbb{R}^d)$. Furthermore, it holds that

$$i\partial_t[\nabla, S(t)] = H[\nabla, S(t)] + [\nabla, H]S(t) = H[\nabla, S(t)] + \nabla U S(t),$$

and hence

$$[\nabla, S(t)] = -i \int_0^t S(t-s)\nabla U \ S(s) \ ds.$$

We can thus estimate

$$\|\nabla\tilde{\chi}(s)\|_{L_x^2} = \|\nabla S(t-s)\chi\|_{L_x^2}$$

(2.12)
$$\leqslant \|S(t-s)\nabla\chi\|_{L_x^2} + \left\| \int_0^{t-s} S(t-s-\tau)\nabla U\tilde{\chi}(\tau) \ d\tau \right\|_{L_x^2}$$

$$\leqslant \|\nabla\chi\|_{L_x^2} + C \int_0^{t-s} \|x\tilde{\chi}(\tau)\|_{L_x^2} \ d\tau,$$

since U is subquadratic, i.e., $|\nabla U(x)| \leqslant C|x|$. Likewise, we deduce

$$[x, S(t)] = -i \int_0^t S(t-s)\nabla S(s) \ ds$$

and hence

(2.13)
$$\|x\tilde{\chi}(s)\|_{L_x^2} \leqslant \|x\chi\|_{L_x^2} + \int_0^{t-s} \|\nabla\tilde{\chi}(\tau)\|_{L_x^2} \ d\tau.$$

Combining the estimates (2.12) and (2.13) and applying Gronwall's inequality yields

$$\|x\tilde{\chi}(s)\|_{L_x^2} + \|\nabla\tilde{\chi}(s)\|_{L_x^2} \leqslant C \left(\|x\chi\|_{L_x^2} + \|\nabla\chi\|_{L_x^2} \right) < +\infty,$$

where $C > 0$. The bound in $L^\infty(0,t; L^{2\sigma+2}(\mathbb{R}^d))$ then follows from the uniform-in-time bound in $H^1(\mathbb{R}^d)$ and the Gagliardo–Nirenberg inequality (2.8). $\quad\square$

With the result of Lemma 2.4 at hand, we consider the second term on the right-hand side of (2.9). Rewriting it using (2.10), we estimate

$$\left| \int_0^t \left\langle |\psi_n(s)|^{2\sigma}\psi_n(s) - |\psi_*(s)|^{2\sigma}\psi_*(s), \tilde{\chi}(s) \right\rangle_{L_x^2} ds \right|$$

$$\leqslant \int_0^t \int_{\mathbb{R}^d} \left| |\psi_n(s,x)|^{2\sigma}\psi_n(s,x) - |\psi_*(s,x)|^{2\sigma}\psi_*(s,x) \right| |\tilde{\chi}(s,x)| \ dx \ ds$$

$$\leqslant C \int_0^t \int_{\mathbb{R}^d} \left(|\psi_n(s,x)|^{2\sigma} + |\psi_*(s,x)|^{2\sigma} \right) |\psi_n(s,x) - \psi_*(s,x)| |\tilde{\chi}(s,x)| \ dx \ ds.$$

By Hölder's inequality, it holds that

$$\int_0^t \int_{\mathbb{R}^d} \left||\psi_n(s,x)|^{2\sigma} + |\psi_*(s,x)|^{2\sigma}\right| |\psi_n(s,x) - \psi_*(s,x)||\tilde{\chi}(s)| \, dx \, ds$$

$$\leqslant \sqrt{T} \left(\|\psi_n\|^{2\sigma}_{L_t^\infty L_x^{2\sigma+2}} + \|\psi_*\|^{2\sigma}_{L_t^\infty L_x^{2\sigma+2}} \right) \|\psi_n - \psi_*\|_{L_t^2 L_x^{2\sigma+2}} \|\tilde{\chi}\|_{L_t^\infty L_x^{2\sigma+2}},$$

where, in view of Lemma 2.4, we have $\|\tilde{\chi}\|_{L_t^\infty L_x^{2\sigma+2}} < +\infty$. In addition, Proposition 2.2 implies that the factor inside the parentheses is bounded and that

$$\lim_{n\to\infty} \|\psi_n - \psi_*\|_{L_t^2 L_x^{2\sigma+2}} = 0.$$

Thus, we have shown that the second term on the right-hand side of (2.9) vanishes in the limit $n \to \infty$.

It remains to treat the last term on the right-hand side of (2.9), rewritten via (2.10). We first estimate

$$\left| \int_0^t \left\langle \alpha_n(s) V \psi_n(s) - \alpha_*(s) V \psi_*(s), \tilde{\chi}(s) \right\rangle_{L_x^2} ds \right|$$

$$\leqslant \int_0^t \int_{\mathbb{R}^d} |\alpha_n(s)| \, |V(x)| \, |\psi_n(s,x) - \psi_*(s,x)| \, |\tilde{\chi}(s,x)| \, dx \, ds$$

$$+ \int_0^t \int_{\mathbb{R}^d} |\alpha_n(s) - \alpha_*(s)| \, |V(x)| \, |\psi_*(s,x)||\tilde{\chi}(s,x)| \, dx \, ds.$$

Here, the last term on the right-hand side can be bounded by

$$\int_0^t \int_{\mathbb{R}^d} |\alpha_n(s) - \alpha_*(s)||V(x)||\psi_*(s,x)||\tilde{\chi}(s,x)| \, dx \, ds$$

$$\leqslant \|\alpha_n - \alpha_*\|_{L_t^2} \|V\|_{L_x^\infty} \|\psi\|_{L_t^2 L_x^2} \|\tilde{\chi}\|_{L_t^\infty L_x^2} \xrightarrow{n\to\infty} 0,$$

in view of the convergence of $\alpha_n \to \alpha_*$ in $L^2(0,T)$. For the remaining term we use the fact that $V \in L^\infty(\mathbb{R}^d)$ and Hölder's inequality to obtain that

$$\int_0^t \int_{\mathbb{R}^d} |\alpha_n(s)| \, |V(x)| \, |\psi_n(s,x) - \psi_*(s,x)||\tilde{\chi}(s,x)| \, dx \, ds$$

$$\leqslant \|\alpha_n\|_{L_t^2} \|V\|_{L_x^\infty} \|\psi_n - \psi_*\|_{L_t^2 L_x^2} \|\tilde{\chi}\|_{L_t^\infty L_x^2} \xrightarrow{n\to\infty} 0,$$

due to the results of Proposition 2.2 and Lemma 2.4.

In summary this proves that $\psi_* \in \Upsilon(0,T)$ satisfies, for *almost all* $t \in [0,T]$,

$$\psi_*(t) = S(t)\psi_0 - i \int_0^t S(t-s) \left(\lambda |\psi_*(s)|^{2\sigma} \psi_*(s) + \alpha_*(s) V \psi_*(s) \right) ds;$$

i.e., ψ_* is a *weak Σ-solution* in the terminology of [10, Definition 3.1.1] (where the analogous notion of weak H^1-solutions is introduced). In order to obtain that ψ_* is indeed a mild solution, we note that

$$\psi_* \in \Upsilon(0,T) \hookrightarrow C([0,T]; L^2(\mathbb{R}^d)) \cap C([0,T]; L^{2\sigma+2}(\mathbb{R}^d))$$

by interpolation and the Gagliardo–Nirenberg inequality (2.8). Classical arguments based on Strichartz estimates then yield uniqueness of the weak Σ-solution ψ_*. Arguing as in the proof of [10, Theorem 3.3.9], we infer that ψ_* is indeed a mild solution to (1.1), satisfying $\psi_* \in C([0,T]; \Sigma) \cap C^1([0,T]; \Sigma^*)$. $\qquad\square$

2.3. Lower semicontinuity of objective functional. In order to conclude that the pair $(\psi_*, \alpha_*) \in \Lambda(0, T)$ is indeed a minimizer of our optimal control problem, it remains to show lower semicontinuity of the functional $J(\psi, \alpha)$ with respect to the convergence results established in Proposition 2.2.

LEMMA 2.5. *For the sequence constructed in Proposition* 2.2*, it holds that*

$$J_* = \liminf_{n \to \infty} J(\psi_n, \alpha_n) \geqslant J(\psi_*, \alpha_*).$$

Proof. Since $A \in \mathcal{L}(\Sigma, L^2(\mathbb{R}^d))$ by assumption, the sequence $(A\psi_n(T))_{n \in \mathbb{N}}$ converges weakly to $A\psi(T)$ in $L^2(\mathbb{R}^d)$. In addition $\psi_n(T) \to \psi_*(T)$ in $L^2(\mathbb{R}^d)$ as $n \to \infty$ by Proposition 2.3, and hence the estimate

$$\left| \langle \psi_n(T), A\psi_n(T) \rangle_{L_x^2} - \langle \psi_*(T), A\psi_*(T) \rangle_{L_x^2} \right|$$
$$\leqslant \left| \langle \psi_n(T) - \psi_*(T), A\psi_n(T) \rangle_{L_x^2} \right| + \left| \langle \psi_*(T), A(\psi_n(T) - \psi_*(T)) \rangle_{L_x^2} \right|$$

yields convergence of the corresponding term in the objective functional (1.8). Next, we consider the cost term involving γ_1. In view of (1.9), we define

$$\omega_n(t) := \int_{\mathbb{R}^d} V(x)|\psi_n(t, x)|^2 dx, \qquad \omega_*(t) := \int_{\mathbb{R}^d} V(x)|\psi_*(t, x)|^2 dx$$

and estimate

$$(2.14) \quad \liminf_{n \to \infty} \int_0^T (\dot{\alpha}_n(t))^2 \omega_n^2(t) dt$$
$$\geqslant \liminf_{n \to \infty} \int_0^T (\dot{\alpha}_n(t))^2 \omega_*^2(t)\, dt + \liminf_{n \to \infty} \int_0^T (\dot{\alpha}_n(t))^2 \left(\omega_n^2(t) - \omega_*^2(t) \right) dt.$$

Note that $0 \leqslant \omega_n(t) \leqslant \|V\|_{L_x^\infty} \|\psi_0\|_{L_x^2}^2$ independently of $n \in \mathbb{N}$ and $t \in [0, T]$ and that the same holds for $\omega_*(t)$. The first term on the right-hand side of (2.14) is convex in α_n and thus satisfies

$$(2.15) \qquad \liminf_{n \to \infty} \int_0^T (\dot{\alpha}_n(t))^2 \omega_*^2(t)\, dt \geqslant \int_0^T (\dot{\alpha}_*(t))^2 \omega_*^2(t)\, dt,$$

since any convex and lower semicontinuous functional is weakly lower semicontinuous. On the other hand, the convergence (2.5) and $V \in L^\infty(\mathbb{R}^d)$ of Proposition 2.2 imply

$$(2.16) \qquad \liminf_{n \to \infty} \omega_n(t) \geqslant \omega_*(t) \geqslant 0 \quad \text{for almost all } t \in [0, T].$$

Thus, using (2.15) and (2.16) together with Fatou's lemma yields

$$\liminf_{n \to \infty} \int_0^T (\dot{\alpha}_n(t))^2 \omega_n^2(t)\, dt$$
$$\geqslant \int_0^T (\dot{\alpha}_*(t))^2 \omega_*^2(t)\, dt + \int_0^T \liminf_{n \to \infty} (\dot{\alpha}_n(t))^2 \liminf_{n \to \infty} \left(\omega_n^2(t) - \omega_*^2(t) \right) dt$$
$$\geqslant \int_0^T (\dot{\alpha}_*(t))^2 \omega_*^2(t)\, dt.$$

Finally the cost term involving γ_2 is lower semicontinuous by convexity and weak convergence of α_n in $H^1(0, T)$. $\quad \square$

In summary, we have shown that $J_* = \liminf_{n\to\infty} J(\psi_n, \alpha_n) \geqslant J(\psi_*, \alpha_*)$, and thus indeed $J_* = J(\psi_*, \alpha_*)$. In other words, $(\psi_*, \alpha_*) \in \Lambda(0, T)$ solves the optimization problem.

Remark 2.6. Note that the bound on $x\psi_n(t, \cdot)$ in $L^2(\mathbb{R}^d)$, obtained in Proposition 2.2, is indeed crucial for proving the weak lower semicontinuity of $J(\psi, \alpha)$. Without such a bound on the second moment, we would have only

$$\psi_n(t) \xrightarrow{n\to\infty} \psi(t) \quad \text{in } L^2_{\text{loc}}(\mathbb{R}^d),$$

due to the lack of compactness of $H^1(\mathbb{R}^d) \hookrightarrow L^2(\mathbb{R}^d)$. In this case, the lower semi-continuity of the term $\langle \psi(T), A\psi(T) \rangle_{L^2_x}$ is not guaranteed. A possible way to circumvent this problem would be to assume that A is *positive definite*, which, however, is not true for general observables of the form $A = A' - a$, with $a \in \mathbb{R}$. A second possibility would be to assume that A is *localizing*, i.e., for all $\psi \in H^1(\mathbb{R}^d)$: $\text{supp}_{x\in\mathbb{R}^d}(A\psi(x)) \subseteq B(R)$, for some $R < +\infty$.

3. Derivation and analysis of the adjoint equation. In order to give a characterization of a minimizer $(\psi_*, \alpha_*) \in \Lambda(0, T)$, we need to derive the first order optimality conditions for our optimal control problem (2.4). For this purpose, we shall first formally compute the derivative of the objective functional $J(\psi, \alpha)$ in the next subsection and consequently analyze the resulting adjoint problem. A rigorous justification for the derivative will be given in section 4.

3.1. Identification of the derivative of $J(\psi, \alpha)$. The mild solution of the NLS equation (1.1), corresponding to the control $\alpha \in H^1(0, T)$, induces a map

$$\psi : H^1(0, T) \to \Upsilon(0, T) : \quad \alpha \mapsto \psi(\alpha).$$

Using this map we introduce the unconstrained or *reduced functional*

$$\mathcal{J} : H^1(0, T) \to \mathbb{R}, \quad \alpha \mapsto \mathcal{J}(\alpha) := J(\psi(\alpha), \alpha).$$

For the characterization of critical points, we need to compute the derivative of $\mathcal{J}$. For this calculation let $\delta_\alpha \in H^1(0, T)$ with $\delta_\alpha(0) = 0$ be a feasible control perturbation. (Recall that $H^1(0, T) \hookrightarrow C(0, T)$, and hence it makes sense to evaluate $\delta_\alpha(t)$ at $t = 0$.) Then the chain rule yields

(3.1)
$$\begin{aligned}
\langle \mathcal{J}'(\alpha), \delta_\alpha \rangle_{H_t^{-1}, H_t^1} = {} & \langle \partial_\psi J(\psi(\alpha), \alpha), \psi'(\alpha)\delta_\alpha \rangle_{\Upsilon^*, \Upsilon} \\
& + \langle \partial_\alpha J(\psi(\alpha), \alpha), \delta_\alpha \rangle_{H_t^{-1}, H_t^1},
\end{aligned}$$

where Υ^* denotes the dual space of $\Upsilon \equiv \Upsilon(0, T)$ for any given $T > 0$. The main difficulty lies in computing $\psi'(\alpha)$ since ψ is given only implicitly through the NLS equation (1.1).

In the following, we shall write the (nonlinear) partial differential equation (1.1) in a more abstract form, i.e.,

(3.2)
$$P(\psi, \alpha) := i\partial_t \psi - H\psi - \alpha(t)V(x)\psi - \lambda|\psi|^{2\sigma}\psi = 0,$$

where $H = -\frac{1}{2}\Delta + U(x)$ denotes the linear, uncontrolled Hamiltonian operator. Setting $\psi = \psi(\alpha)$ and differentiating with respect to α formally yields

$$\frac{d}{d\alpha} P(\psi(\alpha), \alpha) = \partial_\psi P(\psi(\alpha), \alpha)\psi'(\alpha) + \partial_\alpha P(\psi(\alpha), \alpha) = 0.$$

Next, assuming that $\partial_\psi P$ is invertible, we solve for $\psi'(\alpha)$ via

$$\psi'(\alpha) = -\partial_\psi P(\psi, \alpha)^{-1} \partial_\alpha P(\psi(\alpha), \alpha).$$

Thus it holds that

$$\langle \partial_\psi J(\psi(\alpha), \alpha), \psi'(\alpha)\delta_\alpha \rangle_{\Upsilon^*, \Upsilon}$$
$$= \langle -\partial_\psi J(\psi(\alpha), \alpha), \partial_\psi P(\psi(\alpha), \alpha)^{-1} \partial_\alpha P(\psi(\alpha), \alpha)\delta_\alpha \rangle_{\Upsilon^*, \Upsilon},$$

which can be rewritten as

$$\begin{gathered}(3.3) \qquad \langle \partial_\psi J(\psi(\alpha), \alpha), \psi'(\alpha)\delta_\alpha \rangle_{\Upsilon^*, \Upsilon} \\ = \langle -\partial_\alpha P(\psi(\alpha), \alpha)^* \partial_\psi P(\psi(\alpha), \alpha)^{-*} \partial_\psi J(\psi(\alpha), \alpha), \delta_\alpha \rangle_{H_t^{-1}, H_t^1}.\end{gathered}$$

Here we abbreviate

$$\partial_\psi P(\psi(\alpha), \alpha)^{-*} := (\partial_\psi P(\psi(\alpha), \alpha)^*)^{-1} = (\partial_\psi P(\psi(\alpha), \alpha)^{-1})^*.$$

Substituting (3.3) into (3.1), we see that critical points of (2.4) satisfy

$$\begin{gathered}(3.4) \qquad 0 = \langle \mathcal{J}'(\alpha), \delta_\alpha \rangle_{H_t^{-1}, H_t^1} = \langle \partial_\alpha J(\psi(\alpha), \alpha), \delta_\alpha \rangle_{H_t^{-1}, H_t^1} \\ + \langle -\partial_\alpha P(\psi(\alpha), \alpha)^* \partial_\psi P(\psi(\alpha), \alpha)^{-*} \partial_\psi J(\psi(\alpha), \alpha), \delta_\alpha \rangle_{H_t^{-1}, H_t^1}\end{gathered}$$

for all $\delta_\alpha \in H^1(0, T)$ such that $\delta_\alpha(0) = 0$. In order to obtain (3.4) in a more explicit form, we (formally) compute the derivative

$$(3.5) \qquad \partial_\psi P(\psi, \alpha)\xi = i\partial_t \xi - H\xi - \alpha(t)V(x)\xi - \lambda(\sigma + 1)|\psi|^{2\sigma}\xi - \lambda\sigma|\psi|^{2\sigma-2}\psi^2\overline{\xi},$$

acting on $\xi \in L^2(\mathbb{R}^d) \subset \Sigma^*$. Analogously, we find

$$\partial_\alpha P(\psi, \alpha) = -V(x)\psi.$$

Next, we define

$$(3.6) \qquad \varphi := \partial_\psi P(\psi(\alpha), \alpha)^{-*} \partial_\psi J(\psi(\alpha), \alpha),$$

which, in view of (3.4), allows us to express $\mathcal{J}'(\alpha) \in (H^1(0, T))^*$ in the following form:

$$(3.7) \qquad \mathcal{J}'(\alpha) = \partial_\alpha J(\psi(\alpha), \alpha) - \partial_\alpha P(\psi(\alpha), \alpha)^* \varphi.$$

We consequently obtain $\mathcal{J}'(\alpha)$ by explicitly calculating the right-hand side of this equation (given in (4.3) below), provided we can determine φ.

In order to perform this calculation, we recall that the duality pairing between $\xi \in L^2(\mathbb{R}^d) \subset \Sigma^*$ and $\psi \in \Sigma$ can be expressed by the inner product defined in (2.3). Thus, (3.6) implies

$$(3.8) \qquad \langle \varphi, \partial_\psi P(\psi(\alpha), \alpha)\delta_\psi \rangle_{L_t^2 L_x^2} = \langle \partial_\psi J(\psi(\alpha), \alpha), \delta_\psi \rangle_{L_t^2 L_x^2}$$

for all test functions $\delta_\psi \in \Upsilon(0, T)$ such that $\delta_\psi(0) = 0$. This is the correct "tangent space" for ψ in view of the Cauchy data

$$\psi(0) + \delta_\psi(0) = \psi_0 \text{ and } \psi(0) = \psi_0.$$

By virtue of the symmetry of the *linearized operator* $\partial_\psi P(\psi(\alpha), \alpha)$, (3.8) corresponds to the weak formulation of the following *adjoint equation*:

$$(3.9) \quad \begin{cases} i\partial_t \varphi - H\varphi - \alpha(t)V(x)\varphi - \lambda(\sigma+1)|\psi|^{2\sigma}\varphi - \lambda\sigma|\psi|^{2\sigma-2}\psi^2\overline{\varphi} = \dfrac{\delta J(\psi, \alpha)}{\delta\psi(t)} \\[2mm] \forall t \in [0, T] \text{ and with data } \varphi(T) = i\dfrac{\delta J(\psi, \alpha)}{\delta\psi(T)}. \end{cases}$$

Here, $\frac{\delta J(\psi,\alpha)}{\delta\psi(t)}$ denotes the first variation of $J(\psi, \alpha)$ with respect to the value of $\psi(t) \in H^1(\mathbb{R}^d)$, where ψ is the solution of (1.1) with control α. Likewise, $\frac{\delta J(\psi,\alpha)}{\delta\psi(T)}$ denotes the first variation with respect to solutions of (1.1) evaluated at the final time $t = T$. Explicitly, these derivatives are given by

$$(3.10) \quad \begin{aligned} \frac{\delta J(\psi, \alpha)}{\delta\psi(t)} &= 4(\dot{\alpha}(t))^2 \left(\int_{\mathbb{R}^d} V(x)|\psi(t, x)|^2 dx \right) V(x)\psi(t, x) \\ &\equiv 4(\dot{\alpha}(t))^2 \omega(t) V(x)\psi(t, x), \end{aligned}$$

in view of the definition (1.9), and

$$(3.11) \quad \frac{\delta J(\psi, \alpha)}{\delta\psi(T)} = 4\langle \psi(T, \cdot), A\psi(T, \cdot)\rangle_{L_x^2} A\psi(T, x).$$

The system (3.9) consequently defines a Cauchy problem for φ with data given at $t = T$, the final time. Thus, one needs to solve (3.9) backwards in time, a common feature of adjoint systems for time-dependent phenomena.

Remark 3.1. In fact, φ can also be seen as a *Lagrange multiplier* within the Lagrangian formulation of the optimal control problem. In order to see this, one defines the Lagrangian

$$L(\psi, \alpha, \varphi) = J(\psi, \alpha) - \langle \varphi, P(\psi, \alpha)\rangle_{L_t^2 L_x^2},$$

where $P(\psi, \alpha)$ is the NLS equation given in (3.2). Formally, the Euler–Lagrange equations associated to $L(\psi, \alpha, p)$ yield (3.7) and (3.9). In section 5 we shall use the Lagrangian formulation to formally compute the Hessian of the reduced objective functional $\mathcal{J}(\alpha)$.

In the next subsection, we shall set up an existence theory for (3.9), which in turn will be used to rigorously justify the above derivation in section 4 below.

3.2. Existence of solutions to the adjoint equation. In order to obtain existence of solutions to (3.9), we need sufficiently high regularity of ψ, the solution of the GPE (1.1). For this purpose, for every $m \in \mathbb{N}$ we define

$$\Sigma^m := \left\{ \psi \in L^2(\mathbb{R}^d) \ : \ x^j \nabla^k \psi \in L^2(\mathbb{R}^d) \ \forall \text{ multi-indices } j \text{ and } k \text{ with } |j| + |k| \leqslant m \right\},$$

equipped with the norm (note that $\Sigma^1 \equiv \Sigma$)

$$\|\psi\|_{\Sigma^m} := \sum_{|j|+|k|\leqslant m} \left\| x^j \nabla^k \psi \right\|_{L_x^2}.$$

Remark 3.2. If the external potential $U(x)$ were in $L^\infty(\mathbb{R}^d)$, it would be enough to work in the space $H^m(\mathbb{R}^d)$ instead of Σ^m. In the presence of an external subquadratic

potential, however, we also require control of higher moments of the wave function ψ with respect to x.

LEMMA 3.3. *Let $S(t)$ be given by (2.1) with subquadratic potential $U \in C^\infty(\mathbb{R}^d; \mathbb{R})$. Then, there exists a constant $c > 0$ such that*

$$\|S(t)\psi_0\|_{\Sigma^m} \leqslant e^{ct}\|\psi_0\|_{\Sigma^m}$$

for all $t \in [0, \infty)$ and $\psi_0 \in \Sigma^m$.

The proof of this lemma can be deduced by differentiating the Σ^m-norm with respect to time and applying Gronwall's inequality. It consequently implies the following regularity result for solutions to (1.1).

LEMMA 3.4. *Let $\lambda \geqslant 0$, $\sigma \in \mathbb{N}$ with $\sigma < 2/(d-2)$, and let $U \in C^\infty(\mathbb{R}^d)$ be subquadratic. For $m > d/2$, let $\psi_0 \in \Sigma^m$ and $V \in W^{m,\infty}(\mathbb{R}^d)$. Then the mild solution of (1.1) satisfies $\psi \in L^\infty(0, T; \Sigma^m)$.*

In view of Lemma 3.3, this result can be proved by following the same arguments as in [10, Theorem 5.5.1]. Having obtained $\psi \in L^\infty(0, T; \Sigma^m)$, we infer $\psi \in L^\infty((0, T) \times \mathbb{R}^d)$ by the Sobolev embedding $H^m(\mathbb{R}^d) \hookrightarrow L^\infty(\mathbb{R}^d)$ whenever $m > d/2$. Thus, all the ψ-dependent coefficients appearing in adjoint equation (3.9) are indeed in L^∞.

Remark 3.5. Note that Lemma 3.4 requires us to impose $\sigma \in \mathbb{N}$, which together with the condition $\sigma < 2/(d-2)$ necessarily implies $d \leqslant 3$. The reason is that for general $\sigma > 0$ (not necessarily an integer) the nonlinearity $|\psi|^{2\sigma}\psi$ is not locally Lipschitz in Σ^m (cf. Lemma 4.2), and the lifespan of solution $\psi(t, \cdot) \in \Sigma^m$ is in general not known; see [10] for more details.

From now on, we shall always assume that $V \in W^{m,\infty}(\mathbb{R}^d)$ for $m > d/2$ and that $U \in C^\infty(\mathbb{R}^d)$ is subquadratic. With the above regularity result at hand, classical semigroup theory [25] allows us to construct a solution to the adjoint problem.

PROPOSITION 3.6. *Let $\lambda \geqslant 0$, $\sigma \in \mathbb{N}$ with $\sigma < 2/(d-2)$, and let $U \in C^\infty(\mathbb{R}^d)$ be subquadratic. For $m > d/2$, let $\psi_0 \in \Sigma^m$, $V \in W^{m,\infty}(\mathbb{R}^d)$. Then, (3.9) admits a unique mild solution*

$$\varphi \in C([0, T]; L^2(\mathbb{R}^d)).$$

Proof. First, we study the homogeneous equation $\partial_\psi P(\psi(\alpha), \alpha)\xi = 0$ associated to (3.9). It can be written as

$$\partial_t \xi = -iH\xi + B(t)\xi,$$

where

$$B(t)\xi := -i\left(\lambda(\sigma+1)|\psi|^{2\sigma}\xi + \lambda\sigma|\psi|^{2\sigma-2}\psi^2\overline{\xi} + \alpha(t)V(x)\xi\right).$$

The operator $-iH : \Sigma^2 \to L^2(\mathbb{R}^d)$ is simply the generator of the Schrödinger group $S(t) = e^{-iHt}$. On the other hand, for any $t \in [0, T]$, $B(t)$ is a linear operator on the real vector space $L^2(\mathbb{R}^d)$, equipped with the inner product (2.3) (the same would not be true if we were to consider $L^2(\mathbb{R}^d)$ as a complex vector space). In addition, $B(t)^* = B(t)$ is symmetric with respect to this inner product and the same is true for $iB(t)$. Since $V \in W^{m,\infty}(\mathbb{R}^d)$, $\alpha \in L^\infty(0, T)$ by assumption, and $\psi \in L^\infty((0, T) \times \mathbb{R}^d)$ in view of Lemma 3.4, we infer $B \in L^\infty(0, T; \mathcal{L}(L^2(\mathbb{R}^d)))$. The operator $B(t)$ may therefore be considered as a (time-dependent) perturbation of the generator $-iH$.

Following the construction given in Proposition 1.2, Chapter 3 of [25], we obtain the existence of a propagator $F(t, s)$, i.e., a family of bounded operators

$$\{F(t, s) : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)\}_{s,t \in [0,T]},$$

which are strongly continuous in time and satisfy $F(t, s) = F(t, r)F(r, s)$. This propagator $F(t, s)$ is implicitly given by

$$F(t, s) = e^{-iH(t-s)} + \int_s^t e^{-iH(t-\tau)} B(\tau) F(\tau, s) \, d\tau$$

and solves the homogeneous linearized equation in the sense that

$$\frac{d}{dt} F(t, s)\xi = (-iH + B(t)) F(t, s)\xi$$

weakly in $(\Sigma^2)^*$ for every $\xi \in L^2(\mathbb{R}^d)$ and almost every $t \in [0, T]$. Clearly, it provides a unique mild solution $\xi(t) = F(t, s)\varphi(s)$ of the homogeneous equation. Duhamel's formula applied to the adjoint problem (3.9) consequently yields

$$(3.12) \qquad \varphi(t) = iF(t, T)\frac{\delta J(\psi, \alpha)}{\delta \psi(T)} + i \int_t^T F(t, s)\frac{\delta J(\psi, \alpha)}{\delta \psi(s)} \, ds.$$

Under our assumptions on ψ and A we have that

$$\frac{\delta J(\psi, \alpha)}{\delta \psi(t)} \in L^1(0, T; L^2(\mathbb{R}^d)), \quad \frac{\delta J(\psi, \alpha)}{\delta \psi(T)} \in L^2(\mathbb{R}^d),$$

which in view of Duhamel's formula (3.12) implies the existence of a mild solution $\varphi \in C([0, T]; L^2(\mathbb{R}^d))$. Uniqueness follows from linearity and the uniqueness of the homogeneous equation. $\square$

4. Rigorous characterization of critical points. A classical approach for making the derivation of the adjoint system rigorous is based on the implicit function theorem. The latter is used to show that $\partial_\psi P(\psi(\alpha), \alpha)$ is indeed invertible, but it requires the identification of a linear function space X such that

$$P : \Upsilon(0, T) \times H^1(0, T) \to X \; ; (\psi, \alpha) \mapsto P(\psi, \alpha)$$

and

$$\partial_\psi P(\psi, \alpha)^{-1} : X \to \Upsilon(0, T).$$

In other words, we require the solution of (3.9) with a right-hand side in X to be in $\Upsilon(0, T)$. It seems, however, that the linearized operator $\partial_\psi P(\psi, \alpha)^{-1}$ is not sufficiently regularizing to allow for an easy identification of X. Therefore, we shall not invoke the implicit function theorem but rather calculate the Gâteaux-derivative $\mathcal{J}'(\alpha)$ directly. (We do not prove Fréchet-differentiability; see Remark 4.4 below.) To this end, we shall first show that the solution $\psi = \psi(\alpha)$ to (1.1) depends Lipschitz-continuously on the control parameter α. This will henceforth be used to estimate the error terms appearing in the derivative of $\mathcal{J}(\alpha)$.

4.1. Lipschitz-continuity with respect to the control. As a first step towards full Lipschitz-continuity, we prove local-in-time Lipschitz-continuity of $\psi = \psi(\alpha)$ with respect to the control parameter α.

PROPOSITION 4.1. *Let $\lambda \geqslant 0$, $\sigma \in \mathbb{N}$ with $\sigma < 2/(d-2)$, and let $U \in C^\infty(\mathbb{R}^d)$ be subquadratic. For $m > d/2$, let $V \in W^{m,\infty}(\mathbb{R}^d)$ and $\tilde\psi, \psi \in L^\infty(0,T;\Sigma^m)$ be two mild solutions to* (1.1), *corresponding to initial data $\tilde\psi_0, \psi_0 \in \Sigma^m$ and control parameters $\tilde\alpha, \alpha \in H^1(0,T)$, respectively. Assume that*

$$\|\tilde\alpha\|_{H^1_t}, \|\alpha\|_{H^1_t}, \|\tilde\psi(t,\cdot)\|_{\Sigma^m}, \|\psi(t,\cdot)\|_{\Sigma^m} \leqslant M$$

for some given $M \geqslant 0$. Then there exist $\tau = \tau(M) > 0$ and a constant $C = C(M) < +\infty$, such that

$$(4.1) \qquad \|\tilde\psi - \psi\|_{L^\infty(I_t;\Sigma^m)} \leqslant C\left(\|\tilde\psi(t) - \psi(t)\|_{\Sigma^m} + \|\tilde\alpha - \alpha\|_{H^1_t}\right),$$

where $I_t := [t, t+\tau] \cap [0,T]$. In particular, the mapping $\alpha \mapsto \psi(\alpha) \in \Upsilon(0,T)$ is continuous with respect to $\alpha \in H^1(0,T)$.

Proof. To simplify notation, let us assume $t + \tau \leqslant T$. By construction, there exists a $\tau > 0$ depending only on M such that $\psi|_{I_t}$ is a fixed point of the mapping

$$\psi \mapsto S(\cdot)\psi_0 - i\int_t^{\cdot} S(\cdot - s)\left(\lambda|\psi(s)|^{2\sigma}\psi(s) + \alpha(s)V\psi(s)\right)\,ds,$$

which maps the set

$$Y = \{\psi \in L^\infty(I_t;\Sigma^m) : \|\psi\|_{L^\infty(I_t;\Sigma^m)} \leqslant 2M\}$$

into itself. Of course, the same holds true for $\tilde\psi$ and $\tilde\alpha$ in place of ψ and α, respectively. In particular, the embedding $\Sigma^m(R^d) \hookrightarrow L^\infty(\mathbb{R}^d)$, $m > d/2$, yields

$$\|\psi\|_{L^\infty(I_t \times \mathbb{R}^d)} \leqslant 2CM.$$

To proceed further, we recall the following result, which can be proved along the lines of [10, Lemma 4.10.2].

LEMMA 4.2. *Let $M > 0$, $\sigma \in \mathbb{N}$, and $m > d/2$. Then there exists a constant $C(M) > 0$ such that for all $\psi, \tilde\psi \in \Sigma^m$ satisfying $\|\psi\|_{L^\infty_x}, \|\tilde\psi\|_{L^\infty_x} \leqslant M$, it holds that*

$$\left\||\psi|^{2\sigma}\psi - |\tilde\psi|^{2\sigma}\tilde\psi\right\|_{\Sigma^m} \leqslant C(M)\|\psi - \tilde\psi\|_{\Sigma^m}.$$

In other words, $\psi \mapsto |\psi|^{2\sigma}\psi$ is locally Lipschitz in Σ^m.

Subtracting the two fixed point expressions for $\tilde\psi$ and ψ gives

$$\tilde\psi(s) - \psi(s) = S(s-t)(\tilde\psi(t) - \psi(t))$$
$$- i\int_0^{s-t} S(s-r)\left(\lambda(|\tilde\psi|^{2\sigma}\tilde\psi - |\psi|^{2\sigma}\psi) + V(x)(\tilde\alpha\,\tilde\psi - \alpha\,\psi)\right)(\tau)\,d\tau$$

for all $s \in [t, t+\tau]$. Taking the $L^\infty(I_t;\Sigma^m)$-norm and recalling Lemma 3.3, together with $\|\psi(s)\|_{\Sigma^m}, \|\tilde\psi(s)\|_{\Sigma^m} \leqslant 2M$, for $s \leqslant t + \tau$, yields

$$\|\tilde\psi - \psi\|_{L^\infty(I_t;\Sigma^m)} \leqslant C\|\tilde\psi(t) - \psi(t)\|_{\Sigma^m} + 2M\|\tilde\alpha - \alpha\|_{H^1_t}\|V\|_{W^{m,\infty}_x}$$
$$+ C\tau\left(C(2M) + \|\tilde\alpha\|_{H^1_t}\|V\|_{W^{m,\infty}_x}\right)\|\tilde\psi - \psi\|_{L^\infty(t,t+\tau;\Sigma^m)},$$

where $C(2M)$ is the constant appearing in Lemma 4.2 with $2M$ replacing M. Since $\|\tilde{\alpha}\|_{H^1_t} \leqslant M$, the estimate (4.1) follows from possibly choosing τ even smaller.

Finally, we show the continuity of the map $H^1(0,T) \to \Upsilon(0,T)$, $\alpha \mapsto \psi(\alpha)$. Set

$$t_* := \inf \left\{ 0 \leqslant t \leqslant T : \limsup_{\tilde{\alpha} \to \alpha} \|\tilde{\psi} - \psi\|_{L^\infty(0,t;\Sigma^m)} > 0 \right\},$$

with the convention $\inf \emptyset := +\infty$. We have to show that $t_* = +\infty$. Assuming $t_* \leqslant T < +\infty$, fix $M' \geqslant M$ such that $\|\psi\|_{L^\infty(0,T;\Sigma^m)} \leqslant M'$, and let $\tau' = \tau(M'+1) > 0$ be chosen as above, with $M' + 1$ replacing M. Furthermore let $\Delta t = \tau'/2$. The definition of t_* yields

$$\limsup_{\tilde{\alpha} \to \alpha} \|\tilde{\psi} - \psi\|_{L^\infty(0,t_*-\Delta t;\Sigma^m)} = 0.$$

In particular, it holds that $\|\tilde{\psi}\|_{L^\infty(0,t_*-\Delta t;\Sigma^m)} \leqslant M' + 1$ for all $(\tilde{\alpha} - \alpha)$ small enough. But now we see that the Lipschitz-continuity (4.1) is satisfied by $\tilde{\psi}$ and ψ and such controls $\tilde{\alpha}$, α on the interval $[t_* - \Delta t, t_* - \Delta t + \tau']$. Hence

$$\limsup_{\tilde{\alpha} \to \alpha} \|\tilde{\psi} - \psi\|_{L^\infty(0,t_*-\Delta t+\tau';\Sigma^m)} = 0,$$

which is a contradiction to the definition of t_*. Hence we must have $t_* = \infty$, and continuity holds.   □

As a direct consequence of this continuity result, we obtain uniform boundedness of the solution $\psi(\alpha)$ on compact sets in $\alpha \in H^1(0,T)$. Of course, bounded sets in $H^1(0,T)$ are in general not compact, and thus we have to restrict ourselves to finite-dimensional subsets.

COROLLARY 4.3. *Under the assumptions of Proposition* 4.1, *let* $\delta_\alpha \in H^1(0,T)$ *with* $\delta_\alpha(0) = 0$ *be a direction of change for* α, *and let* $\psi(\alpha + \varepsilon\delta_\alpha)$ *be the solution to* (1.1) *with control* $\alpha + \varepsilon\delta_\alpha$ *and initial data* $\psi_0 \in \Sigma^m$, $m > d/2$. *Then there exists* $M < \infty$ *such that*

$$\|\psi(\alpha + \varepsilon\delta_\alpha)\|_{L^\infty(0,T;\Sigma^m)} \leqslant M \quad \forall\, \varepsilon \in [-1,1].$$

Remark 4.4. This bound on finite-dimensional subsets of $H^1(0,T)$ is the reason why we can prove only Gâteaux-differentiability. If we had a bound on $\psi(\alpha)$ in the Σ^m-norm, which was uniform in $t \leqslant T$ and $\|\alpha\|_{H^1_t} \leqslant M$, we could prove Fréchet-differentiability. For our further analysis, however, this will not be of any consequence.

Now we are ready to prove Lipschitz-continuity of the solution $\psi(\alpha)$ with respect to the control parameter $\alpha \in H^1(0,T)$ on the whole control interval $[0,T]$.

PROPOSITION 4.5. *Let* $\lambda \geqslant 0$, $\sigma \in \mathbb{N}$ *with* $\sigma < 2/(d-2)$, *and let* $U \in C^\infty(\mathbb{R}^d)$ *be subquadratic. For* $m > d/2$, *let* $V \in W^{m,\infty}(\mathbb{R}^d)$, $\psi_0 \in \Sigma^m$, $\alpha \in H^1(0,T)$, *and* $\psi \equiv \psi(\alpha) \in L^\infty(0,T;\Sigma^m)$ *be the solution to* (1.1). *Set* $\tilde{\psi} \equiv \psi(\tilde{\alpha})$ *where, for any* $\varepsilon \in [-1,1]$, *we let* $\tilde{\alpha} := \alpha + \varepsilon\delta_\alpha$ *with* $\delta_\alpha \in H^1(0,T)$ *such that* $\delta_\alpha(0) = 0$. *Then there exists a constant* $C > 0$ *such that*

$$(4.2) \qquad \|\tilde{\psi} - \psi\|_{L^\infty(0,T;\Sigma^m)} \leqslant C\|\tilde{\alpha} - \alpha\|_{H^1(0,T)} = C|\varepsilon|\|\delta_\alpha\|_{H^1(0,T)}.$$

In other words, the solution to (1.1) *depends Lipschitz-continuously on the control* α *for each fixed direction* δ_α.

Proof. Since Corollary 4.3 provides a uniform (in ε) bound on $\|\tilde{\psi}\|_{L^\infty(0,T;\Sigma^m)}$, the quantity τ in the local Lipschitz estimate (4.1) is now independent of ϵ and t, and the estimate indeed holds on every interval $[t, t+\tau]$, i.e.,

$$\|\tilde{\psi} - \psi\|_{L^\infty(t,t+\tau;\Sigma^m)} \leqslant C\left(\|\tilde{\psi}(t) - \psi(t)\|_{\Sigma^m} + \|\tilde{\alpha} - \alpha\|_{H^1(t,t+\tau)}\right).$$

Since both solutions $\tilde{\psi}$ and ψ coincide at $t = 0$, finite summation of this estimate over intervals $[n\tau, (n+1)\tau]$ yields (4.2). $\quad\square$

4.2. Proof of differentiability and characterization of critical points. We are now in a position to state the second main result of this work.

THEOREM 4.6. *Let $\lambda \geqslant 0$, $\sigma \in \mathbb{N}$ with $\sigma < 2/(d-2)$, and let $U \in C^\infty(\mathbb{R}^d)$ be subquadratic. In addition, let $\psi_0 \in \Sigma^m$, $V \in W^{m,\infty}(\mathbb{R}^d)$ for some $m \in \mathbb{N}$, $m \geqslant 2$, and $\alpha \in H^1(0,T)$.*

Then the solution of (1.1) satisfies $\psi \in L^\infty(0,T;\Sigma^m)$ and the functional $\mathcal{J}(\alpha)$ is Gâteaux-differentiable for all $t \in [0,T]$, with

$$(4.3) \qquad \mathcal{J}'(\alpha) = \text{Re} \int_{\mathbb{R}^d} \overline{\varphi}(t,x) V(x) \psi(t,x) \ dx - 2\frac{d}{dt}\left(\dot{\alpha}(t)\left(\gamma_2 + \gamma_1 \omega^2(t)\right)\right),$$

in the sense of distributions, where $\omega(t)$ is the weight factor defined in (1.9) and $\varphi \in C([0,T]; L^2(\mathbb{R}^d))$ is the solution of the adjoint equation

$$(4.4) \quad \begin{aligned} i\partial_t \varphi = \ & -\frac{1}{2}\Delta\varphi + U(x)\varphi + \alpha(t)V(x)\varphi + \lambda(\sigma+1)|\psi|^{2\sigma}\varphi + \lambda\sigma|\psi|^{2\sigma-2}\psi^2\overline{\varphi} \\ & + 4\gamma_1(\dot{\alpha}(t))^2\,\omega(t)V(x)\psi, \end{aligned}$$

subject to Cauchy data $\varphi(T,x) = 4i\langle\psi(T,\cdot), A\psi(T,\cdot)\rangle_{L_x^2} A\psi(T,x)$.

Remark 4.7. When compared to the assumptions of Theorem 2.1, the result of Theorem 4.6 requires additional regularity (and stronger decay) of the initial data ψ_0 and the potential V (plus, we need to restrict ourselves to $\sigma \in \mathbb{N}$). Note that the requirements $m \in \mathbb{N}$ and $m \geqslant 2$ imply $m > d/2$ for $d = 1, 2, 3$ spatial dimensions.

Proof. We need to prove that $\mathcal{J}'(\alpha)$ is of the form (4.2). For this purpose, let $\psi = \psi(\alpha)$, $\tilde{\psi} = \psi(\tilde{\alpha})$ with $\tilde{\alpha} = \alpha + \varepsilon\delta_\alpha$ satisfy the assumptions of Proposition 4.5, and consider the difference of the corresponding objective functionals $\mathcal{J}(\alpha), \mathcal{J}(\tilde{\alpha})$. This difference can be written as the sum of three terms

$$\mathcal{J}(\tilde{\alpha}) - \mathcal{J}(\alpha) = \text{I} + \text{II} + \text{III},$$

where we define

$$\text{I} := \langle\tilde{\psi}(T), A\tilde{\psi}(T)\rangle_{L_x^2}^2 - \langle\psi(T), A\psi(T)\rangle_{L_x^2}^2, \quad \text{II} := \gamma_2 \int_0^T (\dot{\tilde{\alpha}}(t))^2 - (\dot{\alpha}(t))^2 \ dt,$$

and

$$\begin{aligned} \text{III} := \ & \gamma_1 \int_0^T (\dot{\tilde{\alpha}}(t))^2 \left(\int_{\mathbb{R}^d} V(x)|\tilde{\psi}(t,x)|^2 \ dx\right)^2 dt \\ & - \gamma_1 \int_0^T (\dot{\alpha}(t))^2 \left(\int_{\mathbb{R}^d} V(x)|\psi(t,x)|^2 \ dx\right)^2 dt. \end{aligned}$$

The general strategy will be to use the Lipschitz property established in Proposition 4.5 and rewrite the terms I, II, and III in such a way that

$$\mathcal{J}(\tilde{\alpha}) - \mathcal{J}(\alpha) = \text{linear terms in } (\tilde{\alpha} - \alpha) + \mathcal{O}(\|\tilde{\alpha} - \alpha\|_{H_t^1}^2).$$

Since $\tilde{\alpha} = \alpha + \varepsilon\delta_\alpha$ and thus $\mathcal{O}(\|\tilde{\alpha} - \alpha\|_{H_t^1}^2) = \mathcal{O}(\varepsilon^2)$, the limit $\varepsilon \to 0$ then yields the desired functional derivative.

We start by considering the term I. It can be rewritten in the form

$$\begin{aligned}
\text{I} &= \langle \tilde{\psi}(T), A\tilde{\psi}(T) \rangle_{L_x^2}^2 - \langle \psi(T), A\psi(T) \rangle_{L_x^2}^2 \\
&= 2\langle \psi(T), A\psi(T) \rangle_{L_x^2} \left(\langle \tilde{\psi}(T), A\tilde{\psi}(T) \rangle_{L_x^2} - \langle \psi(T), A\psi(T) \rangle_{L_x^2} \right) \\
&\quad + \left(\langle \tilde{\psi}(T), A\tilde{\psi}(T) \rangle_{L_x^2} - \langle \psi(T), A\psi(T) \rangle_{L_x^2} \right)^2.
\end{aligned}$$

Using the essential self-adjointness of A, the terms within the parentheses yield

$$\begin{aligned}
&\langle \tilde{\psi}(T), A\tilde{\psi}(T) \rangle_{L_x^2} - \langle \psi(T), A\psi(T) \rangle_{L_x^2} \\
&= 2\langle \tilde{\psi}(T) - \psi(T), A\psi(T) \rangle_{L_x^2} + \langle \tilde{\psi}(T) - \psi(T), A(\tilde{\psi}(T) - \psi(T)) \rangle_{L_x^2}.
\end{aligned}$$

Using the Lipschitz-estimate (4.2), we obtain

$$\left| \langle \tilde{\psi}(T) - \psi(T), A(\tilde{\psi}(T) - \psi(T)) \rangle_{L_x^2} \right| \leqslant \|A\|_{\mathcal{L}(\Sigma, L_x^2)} \|\tilde{\psi}(T) - \psi(T)\|_\Sigma^2 \leqslant C\varepsilon^2 \|\delta_\alpha\|_{H_t^1}^2,$$

and hence

$$\langle \tilde{\psi}(T), A\tilde{\psi}(T) \rangle_{L_x^2} - \langle \psi(T), A\psi(T) \rangle_{L_x^2} = 2\langle \tilde{\psi}(T) - \psi(T), A\psi(T) \rangle_{L_x^2} + \mathcal{O}(\|\tilde{\alpha} - \alpha\|_{H_t^1}^2).$$

Squaring the above result and plugging it into our expression for I consequently yields

$$(4.5) \qquad \text{I} = 4\langle \psi(T), A\psi(T) \rangle_{L_x^2} \langle \tilde{\psi}(T) - \psi(T), A\psi(T) \rangle_{L_x^2} + \mathcal{O}(\|\tilde{\alpha} - \alpha\|_{H_t^1}^2).$$

Next we consider II, which can be written as

$$\begin{aligned}
\text{II} &= 2\gamma_2 \int_0^T \dot{\alpha}(t) \left(\dot{\tilde{\alpha}}(t) - \dot{\alpha}(t) \right) dt + \gamma_2 \int_0^T \left(\dot{\tilde{\alpha}}(t) - \dot{\alpha}(t) \right)^2 dt \\
&= 2\gamma_2 \int_0^T \dot{\alpha}(t) \left(\dot{\tilde{\alpha}}(t) - \dot{\alpha}(t) \right) dt + \mathcal{O}(\|\tilde{\alpha} - \alpha\|_{H_t^1}^2).
\end{aligned}$$

The first term in the second line is thereby seen to be of the form given in (4.2). Finally we consider III, which in view of definition (1.9) can be written as

$$\begin{aligned}
\text{III} &= \gamma_1 \int_0^T \left((\dot{\tilde{\alpha}}(t))^2 - (\dot{\alpha}(t))^2 \right) \omega^2(t)\, dt \\
&\quad + \gamma_1 \int_0^T (\dot{\tilde{\alpha}}(t))^2 \left(\left(\int_{\mathbb{R}^d} V(x)|\tilde{\psi}(t,x)|^2 dx \right)^2 - \omega^2(t) \right) dt.
\end{aligned}$$

As before, we can expand these terms using quadratic expansions in both $\tilde{\psi}$ and $\tilde{\alpha}$. In view of the Lipschitz-estimate (4.2), any quadratic error $\|\tilde{\psi} - \psi\|_{L_t^\infty L_x^2}^2$ is bounded

by $\mathcal{O}(\|\tilde{\alpha} - \alpha\|_{H_t^1}^2)$, and hence we obtain

$$
\text{(4.6)} \quad
\begin{aligned}
\text{III} &= 4\gamma_1 \int_0^T (\dot{\alpha}(t))^2 \, \omega(t) \left(\text{Re} \int_{\mathbb{R}^d} \left((\overline{\tilde{\psi}} - \overline{\psi}) V \psi \right)(t, x) \, dx \right) dt \\
&\quad + 2\gamma_1 \int_0^T \left(\dot{\tilde{\alpha}}(t) - \dot{\alpha}(t) \right) \dot{\alpha}(t) \, \omega^2(t) \, dt + \mathcal{O}(\|\tilde{\alpha} - \alpha\|_{H_t^1}^2).
\end{aligned}
$$

Here the second term on the right-hand side is linear in $(\tilde{\alpha} - \alpha)$ and hence of the desired form. In order to treat the first term, we note that the expression

$$
4\gamma_1 \left((\dot{\alpha}(t))^2 \int_{\mathbb{R}^d} V(x)|\psi(t,x)|^2 \, dx \right) V(x)\psi(t,x)
$$

appears as a source term in the adjoint equation (4.4). Thus we obtain

$$
\text{(4.7)} \quad
\begin{aligned}
&4\gamma_1 \int_0^T (\dot{\alpha}(t))^2 \, \omega(t) \left(\text{Re} \int_{\mathbb{R}^d} \left((\overline{\tilde{\psi}} - \overline{\psi}) V \psi \right)(t, x) \, dx \right) dt \\
&= \text{Re} \int_0^T \int_{\mathbb{R}^d} \overline{\varphi}(t, x) \left(\partial_\psi P(\psi, \alpha)(\tilde{\psi} - \psi) \right)(t, x) \, dx \\
&\quad - \text{Re} \int_{\mathbb{R}^d} i \, \overline{\varphi}(T, x) \left(\tilde{\psi}(T, x) - \psi(T, x) \right) dx,
\end{aligned}
$$

where we recall that $\partial_\psi P(\psi, \alpha)$ denotes the linearized Schrödinger operator obtained in (3.5). The last term on the right-hand side of (4.7) stems from the boundary condition at $t = T$. Note that the boundary term at $t = 0$ vanishes since $\tilde{\psi}(0) = \psi_0 = \psi(0)$ by assumption. We recall that $\varphi \in C([0, T]; L^2(\mathbb{R}^d))$ and

$$
\tilde{\psi}, \psi \in L^\infty(0, T; \Sigma^m) \cap W^{1,\infty}(0, T; \Sigma^{m-2}), \quad \text{with } m \geqslant 2,
$$

and hence the right-hand side of (4.7) is well defined. In addition, since both $\tilde{\psi}$ and ψ solve the NLS equation (1.1), we can write

$$
\text{(4.8)} \quad
\begin{aligned}
\partial_\psi P(\psi, \alpha)(\tilde{\psi} - \psi) &= i\partial_t(\tilde{\psi} - \psi) - H(\tilde{\psi} - \psi) - V(x)(\tilde{\alpha}(t)\tilde{\psi} - \alpha(t)\psi) \\
&\quad + \lambda|\psi|^{2\sigma}\psi - \lambda|\tilde{\psi}|^{2\sigma}\tilde{\psi} + (\tilde{\alpha}(t) - \alpha(t))V(x)\tilde{\psi} + \varrho(\tilde{\psi}, \psi) \\
&= (\tilde{\alpha}(t) - \alpha(t))V(x)\tilde{\psi} + \varrho(\tilde{\psi}, \psi),
\end{aligned}
$$

where the remainder $\varrho(\tilde{\psi}, \psi)$ is given by

$$
\frac{1}{\lambda}\varrho(\tilde{\psi}, \psi) = |\tilde{\psi}|^{2\sigma}\tilde{\psi} - |\psi|^{2\sigma}\psi - (\sigma+1)|\psi|^{2\sigma}(\tilde{\psi} - \psi) - \sigma|\psi|^{2\sigma-2}\psi^2 \left(\overline{\tilde{\psi}} - \overline{\psi} \right).
$$

Since $2\sigma \geqslant 2$ by assumption, $\|\tilde{\psi}\|_{L_t^\infty L_x^\infty}, \|\psi\|_{L_t^\infty L_x^\infty} \leqslant C$ in view of Corollary 4.3, and $\Sigma^m \hookrightarrow L^\infty(\mathbb{R}^d)$, the remainder can be bounded by

$$
|\varrho(\tilde{\psi}, \psi)| \leqslant C \left(|\tilde{\psi}|^{2\sigma-1} + |\psi|^{2\sigma-1} \right) |\tilde{\psi} - \psi|^2 \leqslant C|\tilde{\psi} - \psi|^2.
$$

In addition, since $\varphi \in C([0, T]; L^2(\mathbb{R}^d))$ and $\Sigma^m \subset H^m(\mathbb{R}^d) \hookrightarrow L^4(\mathbb{R}^d)$, we find that

$$
\int_{\mathbb{R}^d} |\varphi(t,x)||\tilde{\psi}(t,x) - \psi(t,x)|^2 \, dx \leqslant \|\varphi\|_{L_t^\infty L_x^2} \|\tilde{\psi} - \psi\|_{L_t^\infty L_x^4}^2 = O(\|\tilde{\alpha} - \alpha\|_{H_t^1}^2).
$$

Furthermore, the contribution of $(\tilde{\alpha}(t) - \alpha(t))V(x)\tilde{\psi}$ in (4.8) equals

$$(\tilde{\alpha}(t) - \alpha(t))V(x)\psi + (\tilde{\alpha}(t) - \alpha(t))V(x)(\tilde{\psi} - \psi),$$

where the latter term can be estimated by $O(\|\tilde{\alpha} - \alpha\|_{H_t^1}^2)$ as before. In summary, this shows that

$$(4.9) \qquad \begin{aligned} (4.7) = &\int_0^T (\tilde{\alpha}(t) - \alpha(t))\mathrm{Re} \int_{\mathbb{R}^d} \overline{\varphi}(t,x)V(x)\psi(t,x)dxdt + O(\|\tilde{\alpha} - \alpha\|_{H_t^1}^2) \\ &- 4\langle \psi(T), A\psi(T)\rangle_{L_x^2} \langle \tilde{\psi}(T) - \psi(T), A\psi(T)\rangle_{L_x^2}, \end{aligned}$$

where we have used the fact that the data of the adjoint problem at $t = T$ is given by

$$\varphi(T,x) = 4i\langle \psi(T,\cdot), A\psi(T,\cdot)\rangle_{L_x^2} \, A\psi(T,x).$$

Thus, we infer that, up to quadratic errors, the second line in (4.9) cancels with the terms obtained in (4.5). Collecting all the expressions obtained for I, II, III and taking the limit $\varepsilon \to 0$, we have shown that $\mathcal{J}(\alpha)$ is Gâteaux-differentiable with derivative $\mathcal{J}'(\alpha)$ given by (4.2). This concludes the proof of Theorem 4.6. $\square$

Equation (4.3) yields the following characterization of the critical points $\alpha_* \in H^1(0,T)$, i.e., points where $\mathcal{J}'(\alpha_*) = 0$.

COROLLARY 4.8. *Let ψ_* be the solution of* (1.1) *with control α_*. Also, let φ_* be the corresponding solution of the adjoint equation* (4.4), *and denote by ω_* the function defined in* (1.9) *with ψ replaced by ψ_*. Then $\alpha_* \in C^2(0,T)$ is a classical solution of the ordinary differential equation*

$$(4.10) \qquad \frac{d}{dt}\left(\dot{\alpha}_*(t)\left(\gamma_2 + \gamma_1\omega_*^2(t)\right)\right) = \frac{1}{2}\,\mathrm{Re}\int_{\mathbb{R}^d}\overline{\varphi_*}(t,x)V(x)\psi_*(t,x)\,dx,$$

subject to $\alpha_*(0) = \alpha_0$, $\dot{\alpha}_*(T) = 0$.

Remark 4.9. In the case $\gamma_1 = 0$ this simplifies to the expression used in the physics literature; cf. [16].

Proof. Let $\mu \in C_0^\infty(0,T)$ be a test function with compact support in $(0,T)$. Then, Theorems 2.1 and 4.6 imply that there exists $\alpha_* \in H^1(0,T)$ such that $\mathcal{J}'(\alpha_*) = 0$, satisfying (4.3) in the sense of distributions, i.e.,

$$\int_0^T \dot{\alpha}_*(t)\dot{\mu}(t)\left(\gamma_2 + \gamma_1\omega_*^2(t)\right)dt = \frac{1}{2}\mathrm{Re}\int_0^T\int_{\mathbb{R}^d}\mu(t)\overline{\varphi_*}(t,x)V(x)\psi_*(t,x)dxdt,$$

where we have used the fact that the boundary terms at $t = 0$ and $t = T$ vanish due to the compact support of $\mu(t)$. We shall show that the weak solution α_* is in fact unique. This can be seen by considering two different $\alpha_*^1(t), \alpha_*^2(t)$, satisfying $\alpha_*^1(0) = \alpha_*^2(0) = \alpha_0$. Denoting their difference by $\beta_* = \alpha_*^1 - \alpha_*^2$, we have that $\beta_*(t)$ solves

$$\int_0^T \dot{\beta}_*(t)\dot{\mu}(t)\left(\gamma_2 + \gamma_1\omega_*^2(t)\right)dt = 0 \quad \forall \mu \in C_0^\infty(0,T).$$

Since $\gamma_2 > 0$ and $\gamma_1 \geqslant 0$, this implies that $\dot{\beta}_*(t) = 0$ in the sense of distributions. However, since $\alpha_*^1, \alpha_*^2 \in H^1(0,T) \hookrightarrow C(0,T)$, we conclude that $\beta_* \in C(0,T)$ and thus $\beta_*(t) = \mathrm{const}$ for all $t \in [0,T]$. Since $\beta_*(0) = 0$ by assumption, we infer uniqueness of the weak solution $\alpha_*(t)$. On the other hand, standard arguments imply that (4.10)

admits a unique classical solution $\alpha_* \in C^2(0,T)$, provided $\omega_* \in C^1(0,T)$ and the (source term on the) right-hand side is continuous in time. The latter is obviously true in view of Propositions 2.3 and 3.6. In addition, since $V \in W^{1,\infty}(\mathbb{R}^d)$, we infer that for all $\psi(t) \in \Sigma$ it holds that $\chi(t) := (V(x)\psi(t)) \in \Sigma$. From Proposition 2.3 it follows that

$$\dot{\omega}_*(t) = 2\mathrm{Re}\int_{\mathbb{R}^d} V(x)\partial_t\psi_*(t,x)\overline{\psi_*}(t,x)\,dx = 2\langle\chi(t),\dot{\psi}_*(t)\rangle_{\Sigma,\Sigma^*} < +\infty.$$

Thus, $\omega(t) \in C^1(0,T)$, yielding the existence of a unique classical solution $\alpha_* \in C^2(0,T)$. We therefore conclude that the unique weak solution α_* obtained above is in fact a classical solution, satisfying (4.10) subject to $\alpha_*(0) = \alpha_0$, $\dot{\alpha}_*(T) = 0$. $\quad\square$

We call $\alpha_* \in H^1(0,T)$ a *critical* or *stationary* point of the problem

(4.11) $\min \mathcal{J}(\alpha)$ over $\alpha \in H^1(0,T)$

if $\mathcal{J}'(\alpha_*) = 0$, where $\mathcal{J}'$ is given as in Theorem 4.6. In order the check computationally whether α_* is critical, one needs to first solve (1.1) for $\alpha = \alpha_*$ to obtain ψ_* and then solve the adjoint equation (4.4) with $\psi = \psi_*$ and $\alpha = \alpha_*$ to compute φ_*. Inserting $(\alpha,\psi,\varphi) = (\alpha_*,\psi_*,\varphi_*)$ in (4.3) yields $\mathcal{J}'(\alpha_*)$ which has to vanish for α_* to be critical; i.e., (4.10) is satisfied. We therefore call (1.1), (4.4), and (4.10) the first order optimality conditions associated with (4.11).

5. Numerical simulation of the optimal control problem. For our numerical treatment we simplify to the case $d = \sigma = 1$. In this case, the first order optimality conditions for our optimal control problem are given by

$$\begin{cases} \dfrac{d}{dt}\left(\dot{\alpha}(t)\left(\gamma_2 + \gamma_1\omega^2(t)\right)\right) = \dfrac{1}{2}\,\mathrm{Re}\int_{-\infty}^{\infty}\overline{\varphi(t,x)}V(x)\psi(t,x)\,dx, \\[2mm] i\partial_t\psi + \dfrac{1}{2}\partial_x^2\psi = (U(x) + \alpha(t)V(x))\psi + \lambda|\psi|^2\psi, \\[2mm] i\partial_t\varphi + \dfrac{1}{2}\partial_x^2\varphi = (U(x) + \alpha(t)V(x))\varphi + 2\lambda|\psi|^2\varphi + \lambda\psi^2\overline{\varphi} + 4\gamma_1(\dot{\alpha})^2\,\omega(t)V(x)\psi, \end{cases}$$

subject to the following conditions: $\alpha(0) = \alpha_0$, $\dot{\alpha}(T) = 0$, and

$$\psi(0,x) = \psi_0(x), \quad \varphi(T,x) = 4i\langle\psi(T,\cdot),A\psi(T,\cdot)\rangle_{L_x^2}A\psi(T,x).$$

In our numerical simulations, the resulting Cauchy problems for Schrödinger-type equations are solved by a *time-splitting spectral method* of second order (Strang-splitting), as can be found in [1]. This computational approach is unconditionally stable and allows for spectral accuracy in the resolution of the wave function $\psi(t,x)$. This is needed due to the highly oscillatory nature of solutions to (nonlinear) Schrödinger–type equations. We consequently perform our simulations on a *numerical domain* $\Omega \subset \mathbb{R}$, equipped with periodic boundary conditions. The trapping potential $U(x)$ is thereby chosen such that the "effective" (i.e., the numerically relevant) support of the wave function $\psi(t,x)$ stays away from the boundary. In doing so, the boundary conditions do not significantly influence our results. We also check the accuracy of our Schrödinger-solver by monitoring the time-dependence of the L^2-norm of $\psi(t,x)$. On an analytical level this norm is preserved (i.e., mass conservation), but numerical errors will break this conservation law. In all our numerical examples (see subsection 5.3

below), we find that the L^2-norm is numerically preserved up to relative errors of the order 10^{-13}.

Remark 5.1. From the analytical point of view, the optimal control problem analyzed above (where, for $d = 1$, the spatial variable $x \in \mathbb{R}$) is not equivalent to the one posed on a finite interval with periodic boundary conditions. Strictly speaking, the periodic boundary conditions induced by our numerical method would therefore require us to redo our analysis with $x \in \mathbb{T}^d$ (the d-dimensional torus). This, however, is beyond the scope of this article, where we restrict our analysis to the physically realistic case $x \in \mathbb{R}^d$. Numerically we can mimic this case by imposing a sufficiently strong confinement potential $U(x)$.

5.1. Gradient-related descent method. Once a suitable solver for the state and the adjoint equations is at hand, our gradient-related descent scheme operates as follows. We determine a sequence of descent directions $(\delta_\alpha^k) \subset H^1(0, T)$; i.e., for every $k \in \mathbb{N}$,

$$\mathcal{J}(\alpha_k + \delta_\alpha^k) < \mathcal{J}(\alpha_k) \equiv J(\psi(\alpha_k), \alpha_k)$$

is satisfied. Note that a simple Taylor expansion of $\mathcal{J}$ around α_k shows that $\langle \mathcal{J}'(\alpha_k), \delta_\alpha^k \rangle < 0$ is sufficient for δ_α^k to be a descent direction for $\mathcal{J}$ at α_k. We are particularly interested in gradient-related descent directions which satisfy

$$M\delta_\alpha^k = -\mathcal{J}'(\alpha_k),$$

with M a suitably chosen positive definite operator.

A rather straightforward choice of M is given by $M = \partial_\alpha^2 J(\psi, \alpha_k)$. In this case δ_α^k is obtained as the solution of the ordinary differential equation (of second order)

$$\mathcal{J}'(\alpha_k) = 2 \frac{d}{dt} \left(\dot{\delta}_\alpha^k(t) \left(\gamma_2 + \gamma_1 \left(\int_{\mathbb{R}} V(x) |\psi_k(t, x)|^2 dx \right)^2 \right) \right),$$

with $\delta_\alpha^k(0) = 0$ and $\dot{\delta}_\alpha^k(T) = 0$. Here $\psi_k(t, x)$ denotes the solution of the GPE with $\alpha(t) = \alpha_k(t)$. With this choice of a descent direction, we then perform a *line search* in order to decide on the length of the step taken along δ_α^k. In fact, we seek for $\nu_k > 0$ such that

(5.1) $$\mathcal{J}(\alpha_k + \nu_k \delta_\alpha^k) \leq \mathcal{J}(\alpha_k) + \mu \nu_k \langle \mathcal{J}'(\alpha_k), \delta_\alpha^k \rangle,$$

with some fixed $\mu \in (0, 1)$. Within each line search, we determine ν_k iteratively by a backtracking strategy. Thus, the whole procedure amounts to an Armijo line search method with backtracking. Of course, more elaborate strategies based on interpolation or alternative line search criteria are possible; see, e.g., [26] for more details.

Remark 5.2. The terms $\omega(t)$ and $\lambda\psi(t, x)^2 \overline{\varphi}(t, x)$ appearing in the adjoint equation for φ are solved using simple numerical integration based on the rectangle rule and the trapezoidal rule, respectively. For more elaborate treatments of the latter term we refer to [5, 34]. Note that treatment of the term $\omega(t)$ does not require solution of an additional or nonlocal partial differential equation.

We stop the gradient descent method whenever

(5.2) $$\|\mathcal{J}'(\alpha_k)\|_{H_t^{-1}} \leq \text{TOL} \cdot \|\mathcal{J}'(\alpha_1)\|_{H_t^{-1}}$$

is satisfied for the first time. Here, TOL $\in (0, 1)$ is a given stopping tolerance and $\alpha_1 \in H^1(0, T)$ is the initial guess satisfying the boundary conditions $\alpha_1(0) = \alpha_0$ and $\dot{\alpha}_1(T) = 0$. As a safeguard, also an upper bound on the number of iterations is implemented.

In our tests, we observe the usual behavior of steepest descent–type algorithms; i.e., the method exhibits rather fast progress towards a stationary point in early iterations, but then suffers from scaling effects reducing the convergence speed. Therefore, often the maximum number of iterations is reached. Thus, we connect the first order, gradient method to a Newton-type method which relies on second derivatives or approximations thereof.

5.2. Newton method. The majority of iterations within our simulations are performed via a second order method, *Newton's method*, for which we use the full Hessian

$$M := D_\alpha^2 \mathcal{J}(\alpha_k) : H^1(0, T) \times H^1(0, T) \to \mathbb{R}$$

or a sufficiently close positive definite approximation thereof. Note that we can also consider the Hessian as a map $D_\alpha^2 \mathcal{J} : H^1(0, T) \to H^1(0, T)^*$. Recall that the gradient-related method above simply uses $M = \partial_\alpha^2 J(\psi, \alpha_k)$.

We derive $D_\alpha^2 \mathcal{J}$ formally from the Lagrangian formulation; see Remark 3.1. The Lagrangian is given by

$$L(\psi, \alpha, \varphi) = J(\psi, \alpha) - \langle \varphi, P(\psi, \alpha) \rangle_{L^2_{t,x}},$$

where φ is the solution to the adjoint equation (4.4) and $P(\psi, \alpha)$ is the Gross–Pitaevksii operator written in abstract form. Proceeding formally, we find

$$\langle (D_\alpha^2 \mathcal{J}) \delta_\alpha, \tilde{\delta}_\alpha \rangle_{L^2_t} = \langle (\partial_\psi^2 L) \delta_\psi, \tilde{\delta}_\psi \rangle_{L^2_{t,x}} + \langle (\partial_{\psi\alpha} L) \delta_\alpha, \tilde{\delta}_\psi \rangle_{L^2_{t,x}}$$
$$+ \langle (\partial_{\alpha\psi} L) \tilde{\delta}_\alpha, \delta_\psi \rangle_{L^2_t} + \langle (\partial_\alpha^2 L) \delta_\alpha, \tilde{\delta}_\alpha \rangle_{L^2_t},$$

where δ_ψ and $\tilde{\delta}_\psi$ solve the linearized GPE with controls $\delta_\alpha, \tilde{\delta}_\alpha$, respectively. In view of the derivation given in subsection 3.1 we have

$$\delta_\psi = \psi'(\alpha) \delta_\alpha = -\partial_\psi P(\psi(\alpha), \alpha)^{-1} \partial_\alpha P(\psi(\alpha), \alpha) \delta_\alpha,$$

and analogously for $\tilde{\delta}_\psi$. Hence we conclude that

$$(5.3) \quad \begin{aligned} \langle (D_\alpha^2 \mathcal{J}) \delta_\alpha, \tilde{\delta}_\alpha \rangle_{L^2_t} &= \langle (\partial_\psi^2 J) \psi'(\alpha) \delta_\alpha, \psi'(\alpha) \tilde{\delta}_\alpha \rangle_{L^2_{t,x}} + \langle (\partial_{\alpha\psi} J) \delta_\alpha, \psi'(\alpha) \tilde{\delta}_\alpha \rangle_{L^2_{t,x}} \\ &+ \langle (\partial_{\psi\alpha} J) \psi'(\alpha) \delta_\alpha, \tilde{\delta}_\alpha \rangle_{L^2_t} + \langle (\partial_\alpha^2 J) \delta_\alpha, \tilde{\delta}_\alpha \rangle_{L^2_t} - \langle \varphi, \left(D_\alpha^2 P(\psi, \alpha) \delta_\alpha \right) \tilde{\delta}_\alpha \rangle_{L^2_{t,x}}, \end{aligned}$$

where

$$(D_\alpha^2 P(\psi, \alpha) \delta_\alpha) \tilde{\delta}_\alpha = \left(\partial_\psi^2 P(\psi, \alpha)(\psi'(\alpha)\delta_\alpha) \right)(\psi'(\alpha)\tilde{\delta}_\alpha)$$
$$+ \left(\partial_{\alpha\psi} P(\psi, \alpha) \delta_\alpha \right)(\psi'(\alpha) \tilde{\delta}_\alpha) + \left(\partial_{\psi\alpha} P(\psi, \alpha)(\psi'(\alpha)\delta_\alpha) \right) \tilde{\delta}_\alpha,$$

since $\partial_\alpha^2 P(\psi, \alpha) = 0$. All of the terms appearing on the right-hand side of (5.3) can be evaluated by replacing $\psi'(\alpha)\delta_\alpha$ by $-\partial_\psi P(\psi, \alpha)^{-1} \partial_\alpha P(\psi, \alpha) \delta_\alpha$. Consequently, for calculating the action of the Hessian this requires solving several linearized Schrödinger-type equations with different source terms and boundary data. For example, the term involving $(\partial_{\alpha\psi} J)$ can be evaluated by using

$$\chi := \partial_\psi P(\psi, \alpha)^{-*}((\partial_{\alpha\psi} J)\delta_\alpha),$$

which solves the Cauchy problem

$$i\partial_t\chi + \frac{1}{2}\partial_x^2\chi = U(x)\chi + \alpha(t)V(x)\chi + 2\lambda|\psi|^2\chi + \lambda\psi^2\overline{\chi} + 8\gamma_1 h(t,x)\psi,$$

where $h(t,x) := \omega(t)\dot{\alpha}(t)\dot{\delta}_\alpha(t)V(x)$ and

$$\chi(T,x) = \frac{\delta^2 J(\psi,\alpha)}{\delta\psi(T,x)\,\delta\alpha(T)} = 0.$$

Bearing this in mind, we have to solve the following equation for $\delta_\alpha^k \in H^1(0,T)$:

(5.4) $$M\delta_\alpha^k \equiv D_\alpha^2 \mathcal{J}\delta_\alpha^k = -\mathcal{J}'(\alpha^k) \in H^1(0,T)^*.$$

Hence, we need to invert $D_\alpha^2\mathcal{J}$, which, in view of (5.3) is not directly possible. Rather we resort to an iterative method, the *preconditioned MINRES algorithm* (see [24]), with the preconditioner $(\partial_\alpha^2 J(\psi,\alpha))^{-1} : H^1(0,T)^* \to H^1(0,T)$.

In total, we solve one time-dependent Schrödinger-type equation per computation of ψ, one more for the gradient, and three more for the projection of the Hessian in any control direction. Let us point out that in the present paper we aim to study the qualitative behavior of solutions of our control problem rather than the optimization of the respective solution algorithm or its implementation.

5.3. Numerical examples. In all of our examples, we choose the numerical domain $\Omega = [-L, L]$ with $L = 20$ and periodic boundary conditions. The number of spatial grid points is $N = 256$. In addition, we set the final control time to be $T = 10$, and we use $M = 1024$ equidistant time steps. In order to avoid the influence of the boundary, we choose a trapping potential $U(x) = 30\left(\frac{x}{L}\right)^2$. The initial guess for the control is taken to be just $\alpha_1 \equiv 0$ in the linear case ($\lambda = 0$), whereas each algorithm in the nonlinear case ($\lambda \neq 0$) is started from the control obtained by solving the linear problem. In our tests of the first order gradient method, we choose TOL $= 10^{-8}$ in the stopping condition (5.2) for the overall algorithm, $\mu = 10^{-3}$, and a maximum number of 20,000 iterations. For the Newton method, we likewise set TOL $= 10^{-8}$ and stop the algorithm after at most 45 Newton steps.

5.3.1. Example: Shifting a linear wave packet. For validation purposes, we consider the time-evolution of a *linear* wave packet, i.e., $\lambda = 0$, whose center of mass we aim to shift towards a prescribed point $y_1 \in [-L, L]$. For this purpose consider a control potential

$$V(x) = \frac{3}{10} + \frac{3x}{200} \geqslant 0 \quad \forall x \in [-L, L]$$

and the observable

$$A(x) = 1 - e^{-(\kappa(x-y_1))^2/L^2}.$$

In this case, we find that the algorithm converges well even if we invoke only the first order gradient method. Indeed, as we decrease the regularization parameters $\gamma_1, \gamma_2 \ll 1$, we approach an optimal solution which, as it seems, cannot be improved upon. This optimal solution, or, more precisely, its spatial density $\rho = |\psi|^2$, is depicted in Figure 5.1 (right plot), where we denote by "target" the function proportional to $1 - A(x)$ with $\kappa = 0.07$ and $y_1 = -2L/8$ such that it has the same L^2-norm as ψ_0. The left plot shows the associated control.
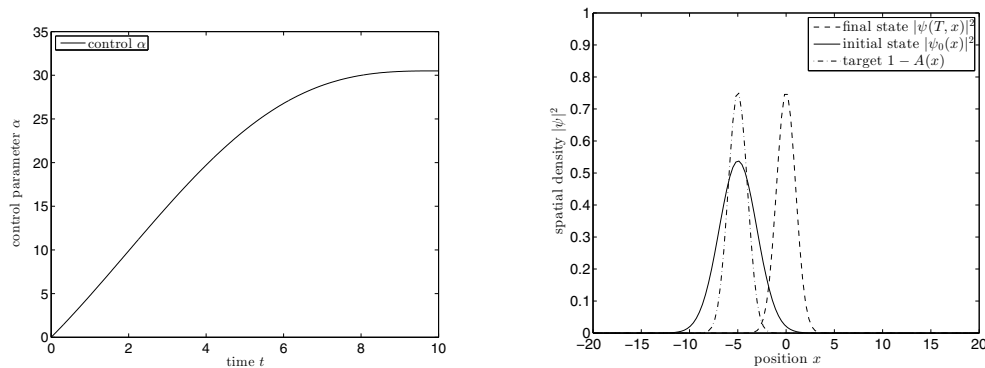
FIG. 5.1. *Shifting a linear wave packet.*

Since this solution seems optimal, the choice of γ_1, γ_2 becomes negligible below a certain threshold. Thus, it suffices to consider $\gamma_1 = 0$ and include only the cost term proportional to γ_2. Similar results hold for any other given point $y_1 \in \Omega$, provided y_1 stays sufficiently far away from the boundary.

5.3.2. Example: Splitting a linear wave packet. We still consider the linear case, i.e., $\lambda = 0$, and aim to split a given initial wave packet into two separate packets centered around y_1 and y_2, respectively. The control potential is chosen as

$$V(x) = e^{-8x^2/L^2} \geqslant 0,$$

and the observable as

$$A(x) = 1 - \left(e^{-(\kappa(x-y_1))^2/L^2} + e^{-(\kappa(x-y_2))^2/L^2} \right).$$

In the following we fix $\kappa = 0.07$, $y_1 = -2L/8$, and $y_2 = 2L/8$. In this case we find that the residual of the first order gradient method does not drop below the tolerance given in (5.2) before the maximum number of iterations is reached. With the Newton method, however, we find a (local) minimum of the objective functional $J(\psi, \alpha)$ in less than 20 Newton iterations. Of course there is no guarantee that this is a global minimum.

In order to illustrate our results, we consider the case where $\gamma_1 = 0$, $\gamma_2 = 1.5 \times 10^{-6}$. At the final control time $T = 10$ we then obtain

$$\langle A\psi(T), \psi(T) \rangle_{L_x^2}^2 \approx 2.261 \times 10^{-3}.$$

The spatial density $\rho = |\psi|^2$ of the corresponding solution is shown in the right plot of Figure 5.2. The associated control is depicted in the left plot. If, instead, we choose $\gamma_1 = 4 \times 10^{-5}$, $\gamma_2 = 1 \times 10^{-9}$, we find

$$\langle A\psi(T), \psi(T) \rangle_{L_x^2}^2 \approx 2.269 \times 10^{-3},$$

and the corresponding solution is given in Figure 5.3. Here the intermediate state is a plot of $\rho(t)$ at $t = 4 = 0.4 \times T$.

A direct comparison of the (spatial densities of the) resulting wave functions and the respective controls is given in Figure 5.4. We see that the spatial densities are
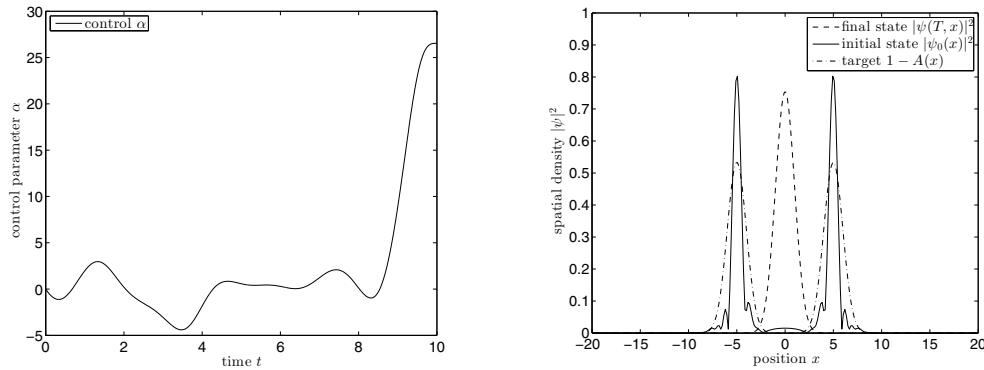
FIG. 5.2. *Splitting a linear wave packet with $\gamma_1 = 0$.*
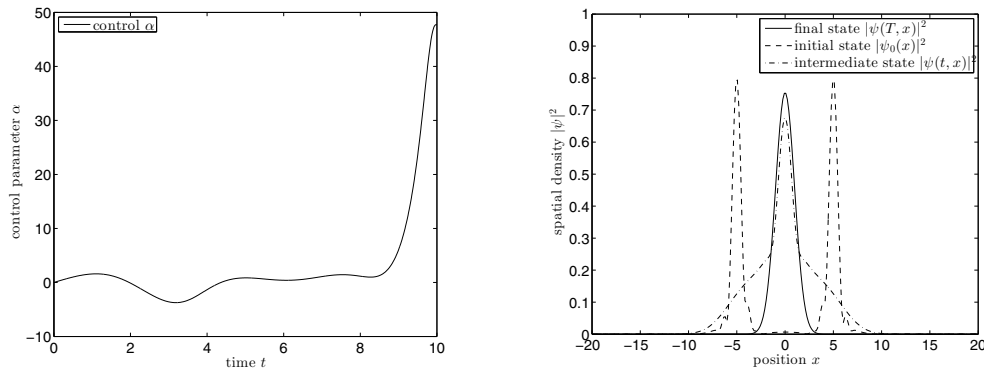
FIG. 5.3. *Splitting a linear wave packet with $\gamma_1 > 0$.*

nearly identical, but the variability of the respective control parameters is not the same. This is, of course, related to the time-evolution of the weight factor $\omega(t)$, defined in (1.9), which is shown in Figure 5.5 for the case of $\gamma_1 = 4 \times 10^{-5}$ and $\gamma_2 = 1 \times 10^{-9}$.

By construction, the time-integral of $\omega(t)$ can be interpreted as the physical work performed during the control process. We find that compared to the case $\gamma_1 > 0$, the term $\|E(\cdot)\|_{L_t^2}^2$ is approximately 30% larger (64.5 versus 49.1) and $\|\dot{E}(\cdot)\|_{L_t^2}^2$ is approximately twice as large (95.0 versus 43.4) in the case where $\gamma_1 = 0$, yielding a significant advantage of our control cost over terms considering the H^1-norm only; see [16] for the latter.

Finally, Figure 5.6 shows an example of the evolution of the objective functional $J(\psi, \alpha)$ over the number of iterations of the Newton method—here for the case where $\gamma_1 = 0$.

5.3.3. Example: Splitting a Bose–Einstein condensate. We consider the same situation as in the previous example, but with an additional (cubic) nonlinearity. More precisely, we choose $\lambda = 8 > 0$. It turns out that the conclusions are similar to those found in the linear case ($\lambda = 0$). Qualitatively, the main difference is that during the time-evolution, the wave function spreads out more because of the additionally
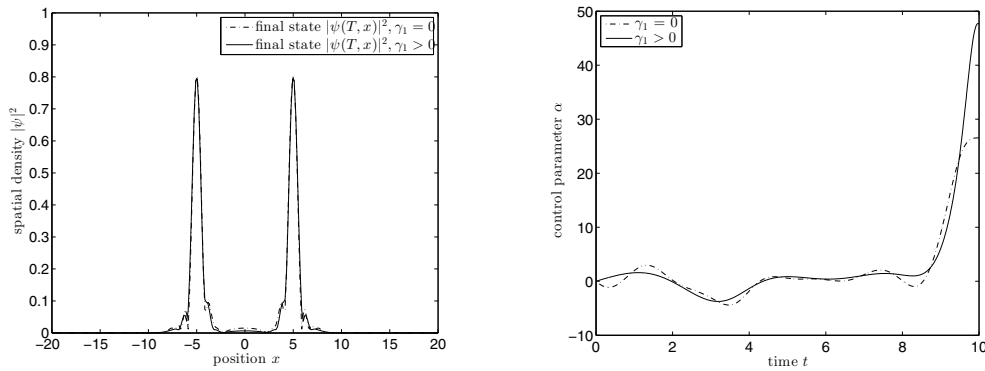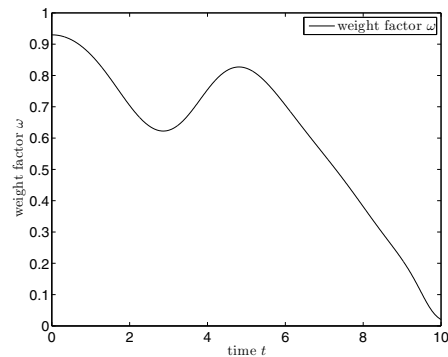
FIG. 5.4. *Direct comparison between results.*

FIG. 5.5. *The weight factor $\omega = \int V|\psi|^2 dx$ over time.*

repulsive (defocusing) nonlinearity. In the linear case, the widest extension of the wave packet is always comparable to its final value. Choosing as before $\gamma_1 = 4 \times 10^{-5}$ and $\gamma_2 = 1 \times 10^{-9}$, we obtain the solution depicted in the right plot of Figure 5.7, where we show the spatial density at the times $t = 0$, $t = T = 10$ and at the intermediate time $t = 4$. The control is shown in the left plot. In comparison to the linear case ($\lambda = 0$), the observable term in the objective functional $J(\psi, \alpha)$ is found to be slightly larger. Indeed, we obtain

$$\langle A\psi(T), \psi(T) \rangle_{L_x^2}^2 \approx 3.720 \times 10^{-3}.$$

This seems to indicate that nonlinear effects counteract the influence of the control potential.

We again compare the present case with that where $\gamma_1 = 0$ (i.e., no cost term proportional to the physical work) and $\gamma_2 = 1.5 \times 10^{-6}$. First, we find that

$$\langle A\psi(T), \psi(T) \rangle_{L_x^2}^2 \approx 3.382 \times 10^{-3}.$$

Moreover, $\|\dot{E}\|_{L_t^2}^2$ is approximately 150% larger (172.1 versus 68.5) than in the case where $\gamma_1 \neq 0$. Similarly, the total energy $\|E\|_{L_t^2}^2$ is approximately 15% larger (91.8 versus 79.5).
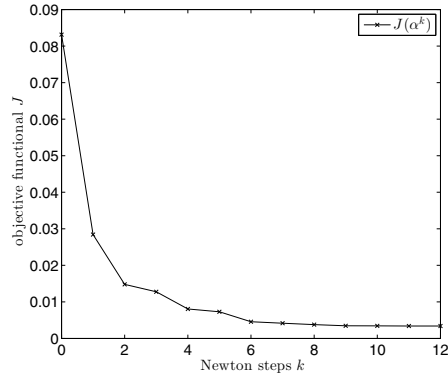
Fig. 5.6. *Value of $J(\psi, \alpha)$ over number of iterations.*
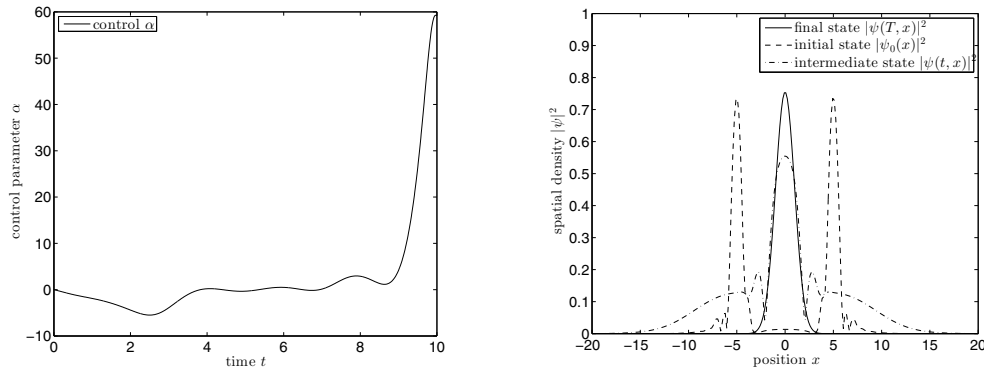
Fig. 5.7. *Splitting a condensate with $\gamma_1 > 0$.*

Finally, we show the rate of convergence of the algorithm measured in terms of the $H^{-1}(0, T)$-norm of the gradient $\mathcal{J}'(\alpha^k)$ in Table 5.1. The parameters are as in subsections 5.3.2 and 5.3.3 with $\gamma_1 = 4 \times 10^{-5}$ and $\gamma_2 = 10^{-9}$. The iteration 0 shows the gradient at the initial guess for the control α, the next row shows the gradient after the initial 50 line-search steps (LS), and the iteration k shows the gradient after k Newton steps. While for $\lambda = 0$, the convergence of Newton's method towards the end of the iteration becomes fast, the local convergence rate is clearly more delicate in the nonlinear setting with $\lambda = 8$. This can be attributed to the effect of the nonlinearity in the state equation on the Hessian of the objective, and consequently the second order sufficient optimality condition, as well as the associated primal and adjoint equations. To this end, we recall that for the Hessian-times-vector product, several (linearized) Schrödinger-type equations with different source terms and data (also depending on ψ, α, and adjoint states) need to be solved. Within such an environment, the interplay of the various iterative solvers and their associated stopping tolerances becomes an issue and requires sophisticated fine-tuning to yield locally superlinear or even quadratic convergence, as otherwise only linear convergence can be expected; see, e.g., [19, Thm. 2.3.4]. For such a tuning for dipole quantum control problems involving the linear Schrödinger equation and a quadratic objective, see [34]. The latter clearly goes beyond the scope of the present work. Here, we note only that the convergence of

TABLE 5.1
Comparison of the rate of convergence.

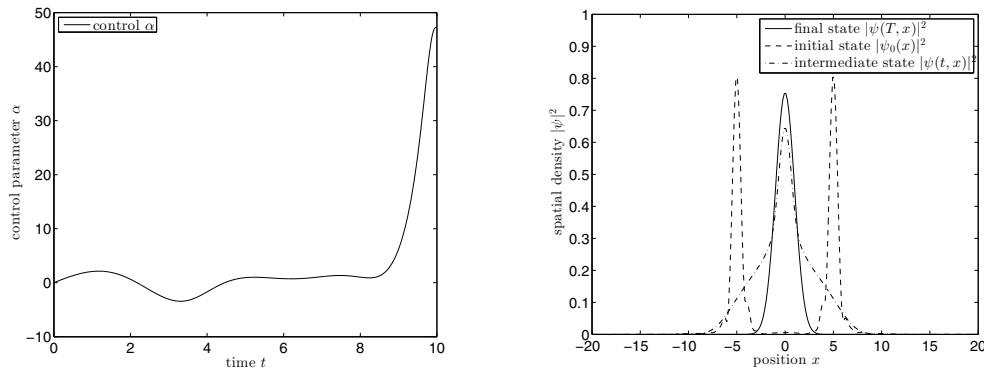| Iteration k	$\|\mathcal{J}'(\alpha^k)\|_{H^{-1}}$ with $\lambda = 0$	$\|\mathcal{J}'(\alpha^k)\|_{H^{-1}}$ with $\lambda = 8$
0	3.392×10^{-2}	2.557×10^{-1}
After 50 LS	1.696×10^{-2}	1.345×10^{-3}
1	5.411×10^{-3}	3.165×10^{-4}
2	1.148×10^{-4}	1.306×10^{-3}
3	4.247×10^{-4}	1.053×10^{-4}
4	3.896×10^{-4}	7.925×10^{-3}
5	9.837×10^{-6}	1.040×10^{-4}
6	5.186×10^{-6}	7.258×10^{-6}
7	3.104×10^{-8}	3.382×10^{-6}
8	1.198×10^{-10}	4.813×10^{-7}
9		2.183×10^{-7}
10		1.617×10^{-7}
11		4.061×10^{-6}
12		1.067×10^{-8}
13		1.328×10^{-7}
14		2.247×10^{-8}
15		1.439×10^{-8}
16		2.038×10^{-9}

FIG. 5.8. *Splitting a focusing condensate with $\gamma_1 > 0$.*

our Newton-type scheme is still substantially faster than that of the gradient-related descent method outlined in subsection 5.1.

5.3.4. Example: Splitting an attractive Bose–Einstein condensate. Our numerical method allows us to go beyond the rigorous mathematical theory developed earlier. In particular we may try to control the behavior of attractive condensates, which are modeled by (1.1) with $\lambda < 0$, i.e., a focusing nonlinearity. Here we choose $\lambda = -1$, whereas the parameters $\gamma_1 = 4 \times 10^{-5}$, $\gamma_2 = 1 \times 10^{-9}$ are the same as before. The results are shown in Figure 5.8 (control in the left plot and the state at times $t = 0, 10, 4$ in the right plot). The observable part of the objective functional satisfies

$$\langle A\psi(T), \psi(T)\rangle^2_{L^2_x} \approx 2.143 \times 10^{-3}.$$

In comparison to the case of a repulsive (defocusing) nonlinearity, the final value for the observable term $\langle A\psi(T), \psi(T)\rangle^2_{L^2_x}$ is much smaller, confirming the basic

intuition that an attractive condensate does not tend to spread out as much as in the repulsive case.

6. Concluding remarks. In this work, we have introduced a rigorous mathematical framework for optimal quantum control of linear and nonlinear Schrödinger equations of Gross–Pitaevskii type. We remark that in the physics literature, $L^2(\mathbb{R}^d)$ is usually considered as a complex Hilbert space, equipped with the inner product $\langle \varphi, \xi \rangle = \int_{\mathbb{R}^d} \varphi(x)\overline{\xi(x)}dx$, whereas we consider $L^2(\mathbb{R}^d)$ as a real Hilbert space (of complex functions), equipped with (2.3). Note, however, that the expectation value of any physical observable A and thus also $J(\psi, \alpha)$ is the same for both choices. We also remark that it would have been possible to use a complex-valued Hilbert space for our analysis, but the real-valued setting is more convenient for our purposes.

Let us briefly discuss possible generalizations for which our results remain valid. First, we point out that in our analysis above, we did not take advantage of the fact that $\gamma_1 > 0$, and hence all of our results remain true in the case $\gamma_1 = 0$. However, the example in subsection 5.3.2 shows a significant quantitative difference in the behavior of the cost functionals with and without the term proportional to γ_1.

Second, it is straightforward to extend our analysis to the case of *several* control parameters, i.e.,

$$V(t,x) = \sum_{k=1}^{K} \alpha_k(t)V_k(x), \quad K \geqslant 2.$$

Clearly, for $V_k \in W^{m,\infty}(\mathbb{R}^d)$, $m \geqslant 2 > d/2$, all of our results remain valid. In addition, it is not difficult to extend our framework to cases of *more general control potentials* $V(\alpha(t), x)$, not necessarily given in the form of a product. Such potentials are of physical significance; cf. [16]. From the mathematical point of view, all of our results still apply provided that

$$\|V(\alpha, \cdot)\|_{W_x^{m,\infty}} \leqslant C_1, \quad \|\partial_\alpha^s V(\alpha, \cdot)\|_{L_x^\infty} \leqslant C_2 \ \forall |s| \leqslant 2.$$

Note that in this case, the cost term in $J(\psi, \alpha)$, which is proportional to the physical work performed throughout the control process, reads

$$\int_0^T (\dot{E}(t))^2 dt = \int_0^T (\dot{\alpha}(t))^2 \left(\int_{\mathbb{R}^d} \partial_\alpha V(\alpha(t), x)|\psi(t, x)|^2 \ dx \right)^2 dt.$$

It is more problematic to provide a rigorous mathematical framework for control potentials $V(\alpha, x)$ which are *unbounded* with respect to $x \in \mathbb{R}^d$. Only in the cases where $V(\alpha, x)$ is subquadratic with respect to x uniformly in α can existence of a minimizer be proved along the lines of the proof of Theorem 2.1. From the physics point of view, however, a general subquadratic control potential has the possibility of undoing the confinement and henceforth destroying the condensate. In order to avoid such a scenario, additional constraints on $\alpha(t)$ would be required.

The inclusion of even more general, i.e., superquadratic, control potentials definitely requires novel mathematical techniques. Note that in this case, even the existence and uniqueness of solutions to the adjoint equation is not obvious, since (local-in-time) Strichartz estimates are more difficult to obtain (and are the subject of current mathematical research). A weaker version of Strichartz estimates for potentials of the form $O(|x|^m)$ at infinity has been derived in [36] and might be a starting point for showing existence of solutions to the NLS equation in a future work. An

analogous problem arises if we would want to replace the subquadratic trap potential U by some superquadratically growing confinement.

Finally, we want to mention that it is possible to extend our results (with some technical effort) to the case of *focusing nonlinearities*, $\lambda < 0$, provided $\sigma < 2/d$. The latter prohibits the appearance of finite-time blow-up in the dynamics of the GPE. Clearly, the optimal control problem ceases to make sense if the solution to the underlying partial differential equation no longer exists.

REFERENCES

[1] W. BAO, D. JAKSCH, AND P. MARKOWICH, *Numerical solution of the Gross–Pitaevskii equation for Bose–Einstein condensation*, J. Comput. Phys., 187 (2003), pp. 318–342.

[2] L. BAUDOUIN, O. KAVIAN, AND J.-P. PUEL, *Regularity for a Schrödinger equation with singular potentials and application to bilinear optimal control*, J. Differential Equations, 216 (2005), pp. 188–222.

[3] L. BAUDOUIN AND J. SALOMON, *Constructive solution of a bilinear optimal control problem for a Schrödinger equation*, Systems Control Lett., 57 (2008), pp. 454–464.

[4] B. BONNARD, N. SHCHERBAKOVA, AND D. SUGNY, *The smooth continuation method in optimal control with an application to quantum systems*. ESAIM Control Optim. Calc. Var., 17 (2011), pp. 267–292.

[5] A. BORZÌ AND U. HOHENESTER, *Multigrid optimization schemes for solving Bose–Einstein condensate control problems*, SIAM J. Sci. Comput., 30 (2008), pp. 441–462.

[6] U. BOSCAIN, G. CHARLOT, AND J.-P. GAUTHIER, *Optimal control of the Schrödinger equation with two or three levels*, in Nonlinear and Adaptive Control Lecture Notes in Control and Inform. Sci. 281, Springer, Berlin, 2003, pp. 33–43.

[7] V. BULATOV, B. E. VUGMEISTER, AND H. RABITZ, *Nonadiabatic control of Bose–Einstein condensation in optical traps*, Phys. Rev. A, 60 (1999), pp. 4875–4881.

[8] E. CANCÈS, C. LE BRIS, AND M. PILOT, *Contrôle optimal bilinéaire d'une équation de Schrödinger*, C. R. Acad. Sci. Paris Sér. I Math., 330 (2000), pp. 567–571.

[9] R. CARLES, *Nonlinear Schrödinger equation with time dependent potential*, Comm. Math. Sci., 9 (2011), pp. 937–964.

[10] T. CAZENAVE, *Semlinear Schrödinger Equations*, Courant Lect. Notes in Math. 10, New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2003.

[11] T. CHAMBRION, P. MASON, M. SIGALOTTI, AND U. BOSCAIN, *Controllability of the discrete-spectrum Schrödinger equation driven by an external field*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 26 (2009), pp. 329–349.

[12] S. CHOI AND N. P. BIGELOW, *Initial steps towards quantum control of atomic Bose–Einstein condensates*, J. Optics B, 7 (2005), pp. 413–420.

[13] J.-M. CORON, *Control and Nonlinearity*, Math. Surveys Monogr. 136, American Mathematical Society, Providence, RI, 2007.

[14] J.-M. CORON, A. GRIGORIU, C. LEFTER, AND G. TURINICI, *Quantum control design by Lyapunov trajectory tracking for dipole and polarizability coupling*, New J. Phys., 11 (2009), pp. 105034–105055.

[15] H. FATTORINI, *Infinite Dimensional Optimization and Control Theory*, Cambridge University Press, Cambridge, UK, 1999.

[16] U. HOHENESTER, P. K. REKDAL, A. BORZI, AND J. SCHMIEDMAYER, *Optimal quantum control of Bose Einstein condensates in magnetic microtraps*, Phys. Rev. A, 75 (2007), pp. 023602–023613.

[17] K. ITO AND K. KUNISCH, *Optimal bilinear control of an abstract Schrödinger equation*, SIAM J. Control Optim., 46 (2007), pp. 274–287.

[18] M. HOLTHAUS, *Toward coherent control of Bose–Einstein condensate in a double well*, Phys. Rev. A, 64 (2001), pp. 011601–011608.

[19] C. T. KELLEY, *Iterative Methods for Optimization*, Frontiers in Appl. Math. 18, SIAM, Philadelphia, 1999.

[20] E. B. KOLOMEISKY, T. J. NEWMAN, J. P. STRALEY, AND X. QI, *Low-dimensional Bose liquids: Beyond the Gross–Pitaevskii approximation*, Phys. Rev. Lett., 85 (2000), pp. 1141–1149.

[21] J.-L. LIONS, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, New York, 1971.

[22] P. MASON AND M. SIGALOTTI, *Generic controllability properties for the bilinear Schrödinger equation*, Comm. Partial Differential Equations, 35 (2010), pp. 685–706.

[23] M. MIRRAHIMI AND P. ROUCHON, *Controllability of quantum harmonic oscillators*, IEEE Trans. Automat. Control, 49 (2004), pp. 745–747.

[24] C. C. PAIGE AND M. A. SAUNDERS, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.

[25] A. PAZY, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.

[26] J. NOCEDAL AND S. J. WRIGHT, *Numerical Optimization*, 2nd ed., Springer, New York, 2006.

[27] A. P. PEIRCE, M. A. DAHLEH, AND H. RABITZ, *Optimal control of quantum-mechanical systems: Existence, numerical approximation, and applications*, Phys. Rev. A, 37 (1988), pp. 4950–4967.

[28] L. PITAEVSKII AND S. STRINGARI, *Bose–Einstein Condensation*, Oxford Science Publications, Oxford, UK, 2003.

[29] R. RADHA, V. R. KUMAR, AND K. PORSEZIAN, *Remote controlling the dynamics of Bose–Einstein condensates through time-dependent atomic feeding and trap*, J. Phys. A, 41 (2008), 315209.

[30] V. RAMAKRISHNA, M. SALAPAKA, M. DAHLEH, AND H. RABITZ, *Controllability of molecular systems*, Phys. Rev. A, 51 (1995), pp. 960–966.

[31] G. THEOCHARIS, P. G. KEVREKIDIS, D. J. FRANTZESKAKIS, AND P. SCHMELCHER, *Symmetry breaking in symmetric and asymmetric double-well potentials*, Phys. Rev. E, 74 (2006), 056608.

[32] G. TURINICI, *On the controllability of bilinear quantum systems*, in Mathematical Models and Methods for Ab Initio Quantum Chemistry, M. Defranceschi and C. Le Bris, eds., Lecture Notes in Chem. 74, Springer-Verlag, New York, 2000.

[33] G. VON WINCKEL AND A. BORZÌ, *Computational techniques for a quantum control problem with H^1-cost*, Inverse Problems, 24 (2008), 034007.

[34] 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), pp. 4176–4203.

[35] J. WERSCHNIK AND E. GROSS, *Quantum optimal control theory*, J. Phys. B, 40 (2007), pp. 175–211.

[36] K. YAJIMA, *Smoothness and non-smoothness of the fundamental solution of time dependent Schrödinger equations*, Comm. Math. Phys., 181 (1996), pp. 605–629.

[37] B. YILDIZ, O. KILICOGLU, AND G. YAGUBOV, *Optimal control problem for nonstationary Schrödinger equation*, Numer. Methods Partial Differential Equations, 25 (2009), pp. 1195–1203.

[38] W. ZHU AND H. RABITZ, *A rapid monotonically convergent iteration algorithm for quantum optimal control over the expectation value of a positive definite operator*, J. Chem. Phys., 109 (1998), pp. 385–391.