Optimal control in disordered quantum systems

Luuk Coopmans [1,2,3,*] Steve Campbell [4,5] Gabriele De Chiara [6] and Anthony Kiely [4,5]

[1]*Dublin Institute for Advanced Studies, School of Theoretical Physics, 10 Burlington Road, Dublin, Ireland*
[2]*School of Physics, Trinity College Dublin, College Green, Dublin 2, Ireland*
[3]*Quantinuum, Partnership House, Carlisle Place, London SW1P 1BX, United Kingdom*
[4]*School of Physics, University College Dublin, Belfield, Dublin 4, Ireland*
[5]*Centre for Quantum Engineering, Science, and Technology, University College Dublin, Belfield, Dublin 4, Ireland*
[6]*Centre for Quantum Materials and Technology, School of Mathematics and Physics, Queen's University Belfast,
Belfast BT7 1NN, United Kingdom*

We investigate several control strategies for the transport of an excitation along a spin chain. We demonstrate that fast high-fidelity transport can be achieved using protocols designed with differentiable programming. Building on this, we then show how this approach can be effectively adapted to control a disordered quantum system. We consider two settings: optimal control for a known unwanted disorder pattern, i.e., a specific disorder realisation, and optimal control where only the statistical properties of disorder are known, i.e., optimizing for high average fidelities. In the former, disorder effects can be effectively mitigated for an appropriately chosen control protocol. However, in the latter setting the average fidelity can only be marginally improved, suggesting the presence of a fundamental lower bound.

I. INTRODUCTION

Robust, implementable control protocols are a necessary ingredient for many quantum devices. A particularly relevant task is the routing of quantum states and information over large distances [1,2] and/or across a network of connected quantum registers [3,4]. The use of spin chains with short-range Heisenberg interactions to transmit quantum information presents a viable framework to accomplish this goal [5], applicable to a wide range of experimental platforms [6–10]. This task can be achieved via the use of an external magnetic field with a parabolic spatial profile which is adiabatically swept across the chain [11] alleviating the need for precise control over individual couplings [12]. However, a notable drawback is that such adiabatic protocols are inherently slow and, hence, susceptible to noise from the environment. This observation precipitated the development of nonadiabatic protocols, such as analytically derived shortcuts to adiabaticity (STA) [13] and quantum optimal control methods [14–27].

Although these approaches have been successfully applied in several settings, there remain drawbacks, for instance, shortcut-to-adiabaticity approaches, such as the one employed in Ref. [13], often rely on simple model descriptions and/or approximations. In this paper we augment traditional control

techniques with differentiable programming (∂P) [28,29] and critically examine their efficacy. Differential programming is an approach which is widely used in the context of deep learning [30] and has been recently applied in various quantum control setups, such as the transport of Majorana zero modes [31], the control of thermal machines [32], numerical renormalization group [33], and interacting qubits [34]. The main advantage of ∂P is that it allows to efficiently obtain the required gradients [35], making gradient-based optimization of complex quantum many-body problems computationally tractable.

We use ∂P with several different ansatzes drawn from other traditional quantum control methods, specifically the shortcut-to-adiabaticity protocol [13] and a finite Fourier basis inspired by the chopped random basis (CRAB) protocol [17]. A common difficulty which must be overcome is the inevitability of noise or imperfections in experimentally realized systems in form of spatial heterogeneities, i.e., disorder. For sufficiently large disorder, the system's state tends to exponentially localize [36] and, thus, hinders transport of the state across the system. We establish that the presence of disorder is, in and of itself, not a limiting factor in achieving high-fidelity control in the system with ∂P strategies able to find a suitable control pulse efficiently. The drawback, however, is that the ∂P gradients implicitly depend on the exact disorder pattern, which, therefore, must be known. In the absence of this information, for a disordered-averaged gradient, we find that even with a high degree of optimization, there is little advantage gained over applying the best disorder-free protocols.

The remainder of the paper is organized as follows. In the subsequent section we outline the model and the state transfer goal. Then in Sec. III we introduce the various approaches

---

*luuk.coopmans@cambridgequantum.com

used to design the dynamical control field. Following this, we demonstrate and compare the effectiveness of these methods for a clean (i.e., disorder-free) system in Sec. IV. These methods are then adapted in Sec. V to a disordered system. Finally, in Sec. VI we summarize our paper and provide some outlook on future directions.

II. MODEL AND CONTROL SETUP

We consider a one-dimensional Heisenberg model of N-interacting spin-$\frac{1}{2}$ particles. The Hamiltonian ($\hbar = 1$) is

$$\mathcal{H} = -\frac{J}{2} \sum_{i=1}^{N-1} \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_{i+1} + \sum_{i=1}^{N} B_i \sigma_i^z, \quad (1)$$

where $\boldsymbol{\sigma}_i \equiv (\sigma_i^x, \sigma_i^y, \sigma_i^z)$ are the Pauli matrices for the spin at site i. The constant isotropic nearest-neighbor coupling J describes exchange interactions between the neighboring spins and the open boundary conditions are considered.

For a single excitation transport problem we do not need access to the full many-body Hilbert space as this Hamiltonian preserves the total spin excitation number, i.e., $S_z = \sum_{i=1}^{N} \sigma_i^z$ since $[\mathcal{H}, S_z] = 0$. We can divide the Hilbert space into separate disconnected sectors, each with a fixed total number of spin excitations $S = \langle S_z \rangle$. In what follows we focus on the single spin-excitation sector where one spin is up, denoted by $|\uparrow\rangle$, whereas all the other spins are down $|\downarrow\rangle$. In total we have N such single-spin excitation states, which we label $|n\rangle = |\downarrow\rangle_1 \otimes \cdots \otimes |\uparrow\rangle_n \otimes \cdots \otimes |\downarrow\rangle_N$ where only the nth spin is up. The Hamiltonian Eq. (1) projected into this subspace becomes

$$H = \sum_{n=1}^{N} [B_n - 2J]|n\rangle\langle n| + J|1\rangle\langle 1| + J|N\rangle\langle N|$$
$$+ J \sum_{n=1}^{N-1} |n\rangle\langle n+1| + |n+1\rangle\langle n|, \quad (2)$$

which is an exponential reduction in dimension compared to the full many-body Hamiltonian.

Note that there is a simple correspondence (up to some edge effects) between this single excitation Hamiltonian Eq. (2) and a discretized version of a single particle in a potential trap, whose continuum limit has the Hamiltonian,

$$H_c = \frac{k^2}{2m} + V(x). \quad (3)$$

A. Magnons and the optimization objective

For our optimization problem we aim to transport magnons, single-spin excitations, from one position in the chain to another, see Fig. 1. Freely propagating magnons, i.e., without a confining magnetic-field B_n, disperse over time throughout the chain. Therefore, an initially localized wave packet will delocalize across the chain as a direct result of the nonlinear dispersion relation [37]. To counteract this spreading it is necessary to guide the magnon transport [11,38] by imposing the external magnetic-field $B_n(t)$. Various kinds of magnetic traps with different spatial profiles can be used for
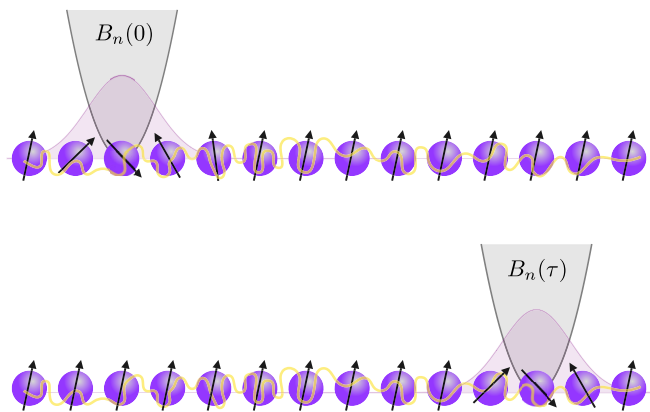

FIG. 1. Setup of the disordered magnon transport problem. The control objective is to design the parabolic magnetic-field B_n (gray shading) in such a way that a localized wave packet or magnon excitation (purple shading) is transported across the chain. The disorder is uniform across the spin chain (yellow erratic line) and is static.

this, such as the Pöschl-Teller [39] potential, a square well [37], or a harmonic trap [11].

Motivated by the correspondence with the single-particle Hamiltonian in Eq. (2) we focus on the parabolic profile,

$$B_n(t) = -\frac{\omega^2}{4J} \left[\frac{n - X_0(t)}{\Delta x} \right]^2 |n\rangle\langle n|. \quad (4)$$

$X_0(t)$ is the position of the minimum of the trap, which will serve as our control parameter. This influences the position of the magnon. This functional form allows us to exploit known analytic optimal control pulses derived with STA methods [13,40]. Note that we fix the lattice spacing as $\Delta x = 1$ hereafter.

For our magnon transport problem we initialize the parameters, such that we have a Gaussian wave packet,

$$|\psi_A\rangle = \frac{1}{\sqrt{\sum_n c_n^2(x_A)}} \sum_n c_n(x_A)|n\rangle \quad (5)$$

centered at a position x_A in the chain with $c_n(x) = \exp[-(n-x)^2/2\sigma^2]$. We vary $X_0(t)$ in such a way that we reach the target state,

$$|\psi_B\rangle = \frac{1}{\sqrt{\sum_n c_n^2(x_B)}} \sum_n c_n(x_B)|n\rangle, \quad (6)$$

after a total time τ. This target state is the same magnon but now localized at a position $x_B = x_A + d$ with transport distance d. Note that x_A, x_B are chosen sufficiently far from the boundaries to avoid any finite-size effects.

The efficacy of our magnon transport protocol, parametrized by $X_0(t)$, can be quantified by the infidelity,

$$\mathcal{I}_\tau = 1 - |\langle\psi_B|U(\tau)|\psi_A\rangle|^2. \quad (7)$$

Here, $U(\tau) = \mathcal{T} \exp[-i \int_0^\tau H(t)dt]$ is the unitary operator that solves the time-dependent Schrödinger equation with the single-spin excitation Hamiltonian $H(t)$, and $\mathcal{T}$ is the Dyson

time-ordering operator. This infidelity $\mathcal{I}_\tau$ forms the objective function which we minimize with respect to $X_0(t)$ to obtain the optimal transport control pulses. We remark that this objective function must be slightly adapted for the case of a disordered system as explained in Sec. V.

III. CONTROL PROTOCOLS

To search for the optimal magnon transport protocols that minimize the infidelity $\mathcal{I}_\tau$, we compare and contrast several different optimization approaches and ansatzes. Specifically, we examine the performance of pulse profiles obtained with two "hybrid" ∂P control approaches where other methods are augmented with ∂P with those obtained with a standard ∂P time-bin optimization. For reference we will also consider a simple linear benchmark protocol,

$$X_0(t) = x_A + \frac{x_B - x_A}{\tau} t, \quad (8)$$

as considered in Ref. [11]. Although Eq. (8) is comparatively easy to implement experimentally, achieving high final target state fidelities with this protocol typically requires long (adiabatic) transport times τ.

A. STA protocol

Here, we briefly recapitulate the main idea of Ref. [13] which forms the basis for the STA protocol ansatz we will employ. We derive optimal control protocols for a particle in a harmonic trap Eq. (3) with the inverse engineering method based on Lewis-Riesenfeld (LR) invariants. As demonstrated in Ref. [13], the approximate correspondence with the single-spin excitation Hamiltonian H allows these protocols to be highly effective for the magnon transport in a Heisenberg chain.

In order to apply the inverse engineering STA method we need to find a suitable dynamical invariant $I(t)$, which is a time-dependent Hermitian operator that satisfies $\frac{\partial I(t)}{\partial t} + i[H_c(t), I(t)] = 0$. For the single particle in a harmonic trap with Hamiltonian,

$$H_c(t) = \frac{k^2}{2m} + \frac{1}{2}m\omega_0^2[x - X_0(t)]^2, \quad (9)$$

a quadratic in momentum LR invariant [40,41] is given by

$$I(t) = \frac{1}{2m}(k - m\dot{\alpha})^2 + \frac{1}{2}m\omega_0^2(x - \alpha)^2. \quad (10)$$

The time-dependent function $\alpha \equiv \alpha(t)$ must satisfy

$$\ddot{\alpha} + \omega_0^2[\alpha - X_0(t)] = 0, \quad (11)$$

to ensure $dI/dt = 0$.

To find optimal transport protocols for the harmonic trap we require that the initial and final eigenstates of $I(t)$ match with the initial $|\psi_A\rangle$ and target states $|\psi_B\rangle$ of our control problem. This can be achieved by imposing $[I_\tau(0), H_\tau(0)] = [I(\tau), H(\tau)] = 0$ which results in the following additional boundary conditions:

$$\alpha(0) = x_A, \qquad \alpha(\tau) = x_B, \qquad \left.\frac{d^n\alpha}{dt^n}\right|_{t=0,\tau} = 0 \quad (12)$$

for $n = 1, 2$. The final shortcut protocol solutions for $X_0(t)$ are then obtained by solving the invariant equations, Eq. (11) with an arbitrary function $\alpha(t)$ that satisfies these boundary conditions. In practice we can pick any function $\alpha(t)$ which fulfill the required boundary conditions, however, as part of a hybrid ∂P approach, we will parametrize a specific subset of this family of solutions as outlined in the following subsection. Note that this formalism can be extended to include a time-dependent frequency.

B. Differentiable programming (∂P)

To optimize the transport of the wave packet in the clean system, i.e., in the absence of disorder, we minimize the infidelity with gradient-based optimization algorithms. For this we obtain the gradient of the infidelity with respect to the control parameters $\frac{d\mathcal{I}_\tau}{dX_0(t)}$ with differentiable programming [30]. This is achieved by writing the code for the infidelity $\mathcal{I}_\tau$ in the PYTHON library JAX [42] and exploiting its automatic differentiation feature. Due to the use of backpropagation (widely used in deep learning [43]), the gradient can be obtained very efficiently with similar computational complexity as forward evaluation of the infidelity [44,45]. This allows us to use many different control parameters and to have flexibility in the design of the cost functional, for example, allowing for the addition of penalty terms and the use of neural networks [31,34].

The gradient can be used in specific optimization algorithms, which we will also refer to as update schemes. The simplest is traditional gradient descent (GD) where the control parameters are updated according to $X_0^{i+1} = X_0^i - \mu \nabla_{X_0} \mathcal{I}_\tau^i$ for each iteration i of the learning algorithm. The hyperparameter μ is known as the learning rate and must be set manually before optimization. It is common practice to use specific decay schedules for μ. We empirically determine an approximation to the optimal learning rate by scanning over a range of values from 10^{-4} to 10^1 and choose the value of μ corresponding to the lower value of infidelity. Another, more elaborate, update scheme we employ is known under the acronym ADAM [46] and makes, among other things, use of an adaptive momentum of the learning rate. We find that this generally leads to faster convergence to the optimal protocols.

We will first consider the setting where no additional information or knowledge about the problem is used to constrain the control protocols; it starts from a random or linear protocol, and we refer to this as ∂P free. This approach is useful for finding new unexpected protocols, for example, the jump-move-jump protocol in Ref. [31] since the learning is not influenced by potential human biases, see also Ref. [47]. A clear drawback is that it can be computationally expensive to find an optimal protocol in the large unconstrained and possibly nonconvex optimization landscape. Specifically, we discretize the control into M-individual time bins $X_0(t) \mapsto [X_0(t_1), X_0(t_2) \cdots X_0(t_M)]$ of time-width $\Delta t = \tau/M$ and we define $t_n = n\,\Delta t$ [48]. The optimization task now reduces to finding the values for each $X_0(t_n)$ that minimize $\mathcal{I}_\tau$.

Following from the demonstrations that hybrid control approaches can be highly effective [13,49,50], we consider the performance of protocols which combine ∂P and other control methods to constrain the size of the search space. In particular,

we consider a "∂P-STA" approach where we use the analytical control pulses provided by the mapping between the single excitation spin chain and the single particle as seeds for the optimization. With GD we then search over the restricted class of STA protocols which generally allows for a very fast optimization, however, as we will see, due to the restricted search space this approach may fail where others are still effective.

For this hybrid approach, in what follows, we define two control parameters $C_1 = X_0(\tau/4)$ and $C_2 = X_0(3\tau/4)$, which correspond to the positions of the trap at times $\tau/4$ and $3\tau/4$, and obtain their derivatives $\frac{\partial \mathcal{I}}{\partial C_1}$ and $\frac{\partial \mathcal{I}}{\partial C_2}$ with ∂P. We remark that this can be generalized to an arbitrary number of control parameters without any significant additional computational cost for ∂P. C_1 and C_2 can be used as control parameters by imposing them as additional constraints when solving the polynomial for $\alpha(t)$ in Eq. (11). The advantages are that imposing the STA protocol constraints need only be performed once and the resulting fixed polynomial of C_1 and C_2 can be implemented as a differentiable function for ∂P. We consider the minimal polynomial $\alpha(t) = \sum_{n=1}^{7} b_n t^n$ that satisfies the boundary conditions Eq. (12) and fixes the values C_1 and C_2. This gives a whole family of STA protocols for $X_0(t)$ in which each member corresponds to a specific configuration of the expansion coefficients b_n.

The final ansatz that we use we term "∂P Fourier", where we aim to combine differentiable programming with optimal control based on a truncated Fourier basis and inspired by the CRAB ansatz [17,51–53]. Specifically, to fulfill the boundary conditions $X_0(0) = x_A$ and $X_0(\tau) = x_B$, we define the Fourier ansatz to be the family of protocols parametrized as

$$X_0(t) = x_A + (x_B - x_A)\frac{t}{\tau} + \sum_{n=1}^{N_c} A_n \sin \omega_n t. \qquad (13)$$

The control parameters we wish to optimize are now the Fourier coefficients A_n whereas we fix the frequencies $\omega_n = \frac{n\pi}{\tau}$ to be the first $n \leqslant N_c$ harmonics. This cutoff on the maximum frequency of the protocol reflects constraints in the experimental implementation. Although often randomized frequency components are chosen to improve convergence, we remark that we do not find this necessary for our purposes.

To then find the optimal protocols in this restricted family we use ∂P to compute the derivatives $\frac{\partial \mathcal{I}_\tau}{\partial A_n}$ and employ them in standard gradient-based optimization algorithms, such as the GD algorithm defined before. The Fourier coefficients are updated from their initial values using a gradient descent update scheme with an adaptive step size. This can be performed for a large number of coefficients since the gradient is calculated via differential programming. This involves the use of reverse mode automatic differentiation (backpropagation) where the gradient is calculated using the chain rule in a reverse order. Due to backpropagation, we can scale to thousands of Fourier coefficients without significantly increasing the computational cost.

In what follows we will examine the efficacy of these ansatzes where throughout: ∂P free refers to the case where no constraints on the functional form of the control pulse are enforced, ∂P STA corresponds to optimizing over
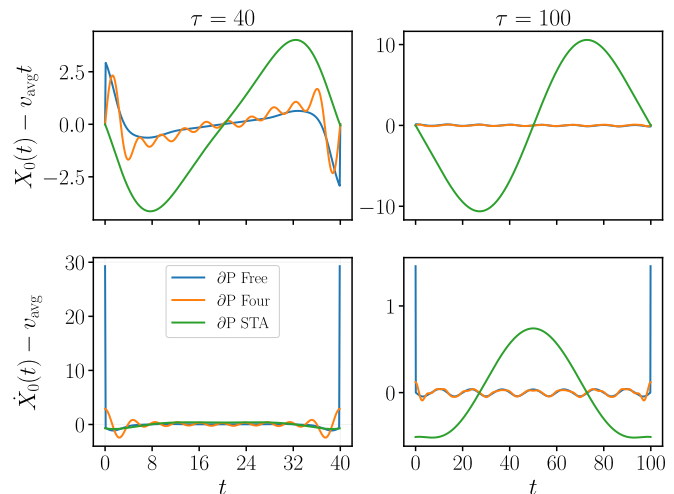

FIG. 2. Disorder-free optimization results obtained with ∂P. The three different ansatzes are free optimization, i.e., no constraints (blue), Fourier series with $N_c = \tau/2$ (orange) and the STA protocol parametrized by C_1 and C_2 (green). The results are shown relative to the linear protocol for two operation durations. Fidelities and other operations times are reported in Table. I. The system parameters are chosen as $N = 251$, $d = 50$, $\omega_0 = 0.5$, and $J = 1$.

the shortcut-to-adiabaticity protocols constrained by Eqs. (11) and (12), and finally ∂P Fourier denotes the use of Eq. (13).

IV. OPTIMAL TRANSPORT IN A DISORDER-FREE CHAIN

We begin by exploring the effectiveness of these control protocols applied to transport in the "clean" chain, which has been extensively considered in the literature [14–27,52–56]. This will allow us to first test the performance in an idealized setting, providing benchmark and reference protocols for when disorder is introduced in Sec. V. Furthermore, we establish that whereas the STA protocols have an effective speed limit [13], the ∂P method can achieve transport velocities close to the maximum magnon group velocity of $v_g = 2J$.

Our aim is to minimize the infidelity of the magnon transport $\mathcal{I}_\tau$ in Eq. (7) and find the optimal control protocols for the trapping center $X_0(t)$. For this task we apply ∂P in combination with the protocols described in the last section. To define the external constraints of the control problem we fix the total transport distance $d = 50$ and choose four different total transport times $\tau = \{40, 60, 80, 100\}J^{-1}$. To obtain the optimal protocols we minimize with the gradient descent algorithm for a maximum of 200 update steps with an empirically determined learning rate. The minimization is stopped when an infidelity lower than $\mathcal{O}(10^{-3})$ is reached.

The results for two operation times are shown in Fig. 2 and the corresponding infidelity values are reported in Table I. Note that above the STA speed limit, which is $\tau^* = 50$ for these parameters [13], we are able to obtain protocols that achieve infidelities on the of order of $\mathcal{O}(10^{-3}–10^{-4})$ with all three optimization methods, whereas the linear reference protocol has significantly larger infidelity values (we remark the reasonably good performance of the linear ramp for $\tau = 80$ is

TABLE I. Infidelity $\mathcal{I}_\tau$ results for the clean system with $N = 251$, $J = 1$, ω_0, and $d = 50$. The corresponding control protocols are shown in Fig. 2.

τ	∂P free	∂P STA	∂P Fourier	Linear
40	0.03832	0.3614	0.0176	0.9845
60	0.0015	0.00013	0.0052	0.6780
80	0.0009	0.00023	0.0003	0.0791
100	0.0024	0.0002	0.0017	0.4414

due to a special resonance effect [13]. In the continuum case, these specific operation times of high fidelity can be expressed analytically in terms of the Fourier transform of the velocity profile of the trapping potential [57]). For faster protocols, $\tau = 40$, the ∂P-free and Fourier approaches are still able to achieve low infidelities $O(10^{-2})$, whereas ∂P STA breaks down due to the fact that we have a restricted search space, although we remark it is still significantly more effective than the simple linear ramp.

Naturally, the shape of the optimal protocols $X_0(t)$ themselves depends strongly on the method employed, cf. Fig. 2. The STA protocols are drastically different taking the form of a ramp up then down protocol for the velocity where they slowly accelerate to a finite velocity and then symmetrically decelerate again to zero. In contrast, the ∂P free and Fourier protocols start and end with a small quench in position and in the middle oscillate with a constant average velocity. The size of the position quenches grows with decreasing transport time whereas the oscillations tend to become smoother with increasing time.

In order to analyze the frequencies ω_p, of these oscillations in Fig. 3(a) we show velocity protocols $\dot{X}_0(t)$ for several trapping frequencies ω_0. In the main panel we observe that the protocol frequency ω_p tends to increase with increasing ω_0.
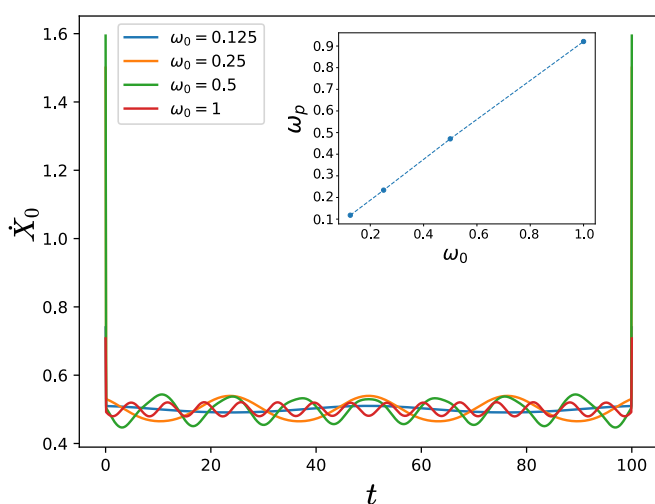

FIG. 3. Main panel: velocities of the magnon transport protocols $\dot{X}_0(t)$ obtained with the ∂P-free approach for four different trapping frequencies ω_0. The inset: the characteristic frequency ω_p of the protocols in the main panel versus ω_0. A linear relationship $\omega_p \approx \omega_0$ is observed. Other parameters: $\tau = 100$, $N = 251$, $J = 1$, and $d = 50$. The parameters for these simulations are the same as in Fig. 2.
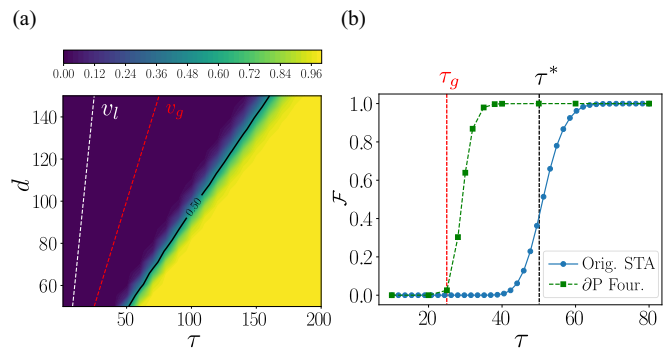

FIG. 4. (a) Target state fidelity $\mathcal{F}_\tau$ of the standard nonparametrized STA protocol defined in Eq. (14) as a function of the transport distance d and transport time τ. The black line gives the $\mathcal{F}_\tau = 0.5$ contour for which the heuristic speed limit $v_{\tau^*} = 1$ is defined, the red dashed line gives the group velocity $v_g = 2$, and the white dashed line gives the Lieb-Robinson velocity $v_l = 6$. (b) Target state fidelity as a function of the transport time τ for the original STA protocol (blue solid line) versus the optimized ∂P Fourier protocols (green line). We see that the optimized ∂P Fourier protocols keep the fidelity at one for smaller transport times τ than the original STA protocol. The group-velocity bound time is given by the red dashed line. Other parameters: $N = 251$, $\omega_0 = 0.5$, $J = 1$, and in (b) $d = 50$.

From the inset we can then see that the relationship is linear with a coefficient very close to one, i.e., $\omega_p \approx \omega_0$. Whereas such an oscillatory behavior is common in Fourier-based optimal control approaches, and often these oscillations can be excluded or suppressed without affecting the effectiveness of the protocol [58], here we find that these oscillations are crucial to the performance of the protocol, in their absence the infidelities rise significantly. Indeed, that the unconstrained ∂P free approach also converges to protocols with these characteristic oscillations indicate their importance in the protocol.

A. Approaching the group velocity speed limit with ∂P

We can recover the speed limit for the STA protocols $\tau^* \approx d/J$ that was heuristically determined based on the original (unparametrized) STA protocol for magnon transport in Ref. [13]. This protocol is given by

$$X_0(t) - x_A = d\left[6s^5 - 15s^4 + 10s^3 + \frac{s}{\omega_0^2 \tau^2}(1 - 3s + 2s^2)\right],$$
(14)

where $s = t/\tau$. This protocol is a function of the total transport time τ, the trapping frequency ω_0, and the initial and target positions x_A and x_B. In Fig. 4(a) we show the target state fidelity $\mathcal{F}_\tau = 1 - \mathcal{I}_\tau$ of this protocol for a specific range of values of d and τ. We observe a clear light-cone-like surface with a velocity of $v_{\tau^*} \approx d/J$. Below this velocity, a fidelity of, at least, $\mathcal{F}_\tau > 0.5$ can be obtained whereas for faster protocols the fidelity drops quickly to zero. We see that this speed limit is still far away from the light cones obtained by the group-velocity $v_g = 2J$ and even further from the Lieb-Robinson velocity [59] $v_l = 6J$ which bounds the timescales on which quantum correlations can

develop between different subsystems of a larger quantum system.

In order to demonstrate that we can improve on this speed limit with ∂P we focus on one specific distance slice $d = 50$ and minimize $\mathcal{I}_\tau$ for a range of different total times τ near the STA speed limit time $\tau^* = 50$. We use the Fourier ansatz with $N_c = 50$ frequency components and, to push the limits of our optimization methodology, we run 500 update steps with the ADAM update scheme. We stop the optimization when an infidelity of $O(10^{-5})$ is reached.

The resulting fidelity values as function of τ compared to the original STA protocol are shown in Fig. 4(b). Compellingly, we see that we can push the maximum velocity very close to the group-velocity time $\tau_g = d/v_g$. Taking $\mathcal{F}_\tau > 0.5$ to measure the speed limit we find $\tau_{\partial P} \approx 29$ which corresponds to $v_{\partial P} \approx 1.72J$, which is very close to the group-velocity v_g. Below τ_g the fidelity quickly drops to zero with no noticeable improvement arising from further optimization. This is in accordance with the group velocity being the maximum speed of the magnons [38] and with the speed limit found in Ref. [60] with the numerical Krotov [61] optimization method for the same system but with a different form of spin excitation.

V. MAGNON TRANSPORT IN THE PRESENCE OF DISORDER

We now consider the effects of disorder in the spin chain [11,13,37,62,63], which could arise either from inhomogeneities in applied fields or defects in fabrication processes, and, generally, leads to localization, significantly hindering the system's controllability. We consider on-site disorder corresponding to an inhomogeneous magnetic trapping field $B_n(t) \mapsto B_n(t) + \epsilon_n$. Note that whereas the trapping field is time dependent for the control of the position of the magnon, the disorder itself is static, and ϵ_n is uniformly distributed on the interval $[-\Delta, \Delta]$ with noise strength Δ.

The main effect of the disordered magnetic field is a localization of the single-spin excitation wave functions with the localization length scaling with the magnitude of the disorder as shown in Fig. 5 for the lowest-energy eigenstate of the disordered Hamiltonian H for several disorder strengths. We observe that the localization length ξ decreases approximately as $\xi \sim 1/\sqrt{\Delta}$. The wave-function localization means that the transport infidelity is expected to increase when the magnon is moved a distance of, at least, a few disorder length scales ξ. From Fig. 5(b) we see that already for $\Delta \sim 0.05$ we have $\xi \sim 40$, thus, even for this small disorder the transport distance ($d = 50$) is already greater than the localization length.

Based on the results from the previous section, we focus on ∂P in combination with the Fourier ansatz in Eq. (13) and aim to find the optimal protocol(s) in two distinct and physically relevant scenarios. In the first, we focus on one specific fixed disorder realization and determine whether effective control can still be achieved. In the second, we optimize over an ensemble of different disorder realizations to search for an optimal control strategy that, on average, works the best in all disordered cases with fixed magnitude.
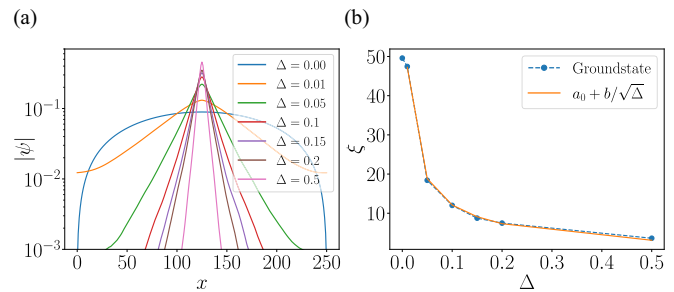

FIG. 5. (a) Disorder-averaged amplitude $|\psi|$ of the lowest-energy eigenstate of the single-spin excitation Heisenberg Hamiltonian in Eq. (2) with a disordered magnetic-field $B_n = \epsilon_n$ which is uniformly distributed over $[-\Delta, \Delta]$. For increasing Δ the amplitude $|\psi|$ becomes localized to a smaller number of lattice sites x. The results are averaged over 1000 realizations and centered about $x = 125$. Other parameters: $N = 251$, $\omega_0 = 0.5$, and $J = 1$. (b) Localization length ξ of the amplitudes in (a) obtained upon fitting a Laplace-like distribution $\frac{1}{\sqrt{\xi}} e^{-(x-x_0)/\xi}$. ξ (blue dashed line) falls off approximately as a square root of Δ as shown by the fit (orange solid line).

A. Fixed disorder patterns

We consider a particular disorder realization $\boldsymbol{\epsilon}^p = (\epsilon_1^p, \ldots, \epsilon_N^p)$ which we label by p. Note that this is simply an N-dimensional random vector from the uniform distribution $[-\Delta, \Delta]$. For this pattern p we then aim to minimize the infidelity $\mathcal{I}_\tau^p$. This infidelity measure is the same as the standard target state infidelity $\mathcal{I}_\tau$ used before but computed with the fixed disorder pattern p added to the magnetic-field $B_n(t) + \epsilon_n^p$ in the Hamiltonian. For the minimization we use $\mathcal{I}_\tau^p$ as the cost functional, and we compute its derivatives with respect to the control parameters with ∂P. These derivatives are then used in combination with ADAM to update the controls A_n of the Fourier ansatz. We repeat the complete optimization process for several different disorder patterns p and finally average the result $\langle \mathcal{I}_\tau \rangle_s = \frac{1}{p_{\max}} \sum_p \mathcal{I}_\tau^p$, where the subscript s indicates that the infidelity was optimized, individually, for each disorder pattern. In Fig. 6(a) we show the results for these optimizations as a function of the disorder strength Δ. We also show benchmark results of the optimal protocols obtained with the ∂P-STA and ∂P-Fourier methods in the clean system. We observe that the disorder ∂P optimization method (green line) is able to keep the infidelity $\langle \mathcal{I}_\tau \rangle_s$ close to zero (on the order of 10^{-3} and smaller). The infidelity $\langle \mathcal{I}_\tau \rangle_s$ of the benchmark protocols (blue and orange lines) for the same fixed disorder patterns is found to be increasing approximately quadratically with disorder strength. This shows that our ∂P optimization method is still able to obtain high-fidelity control protocols in the presence of a specific disorder pattern.

This means that disorder in the setup itself is not immediately an issue as long as a suitable control protocol, designed for the specific disorder realization, is used. Note, however, that to find these protocols the learning algorithms have used information about the exact disorder patterns since the computed infidelity and the gradients implicitly depend on the disorder. As such, our ∂P "model-based" approach only works when the exact form of the pattern is known or can be
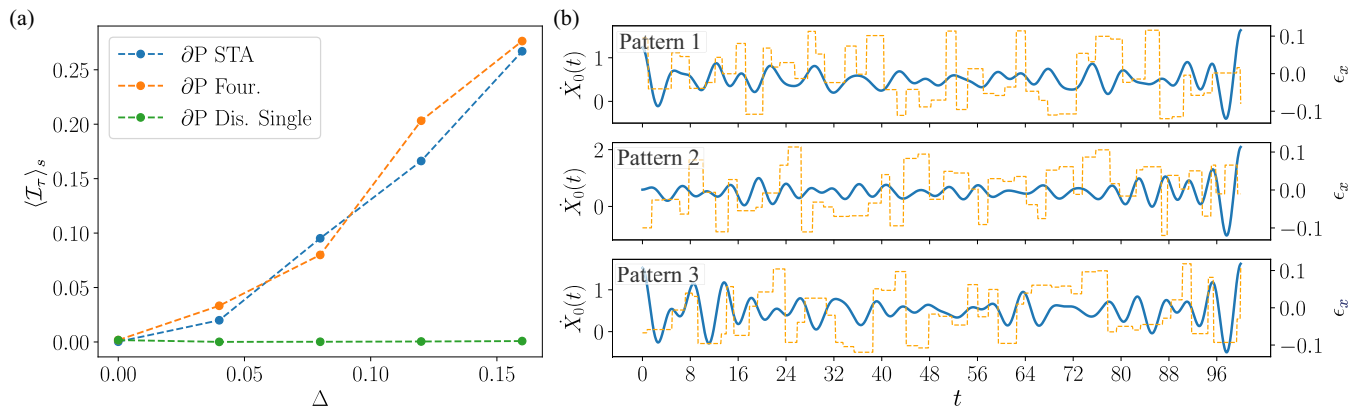
FIG. 6. (a) Disorder-averaged infidelity $\langle \mathcal{I}_\tau \rangle_s$ for five single disorder realizations versus disorder strength Δ. For the ∂P optimization method applied to a single disorder pattern (green dashed line labeled Dis. Single) we optimized the $\mathcal{I}_\tau^p$ for each pattern p individually before computing the average $\langle \mathcal{I}_\tau \rangle_s$. This optimization method is able to keep the infidelity near zero whereas the infidelity of the benchmark protocols obtained for optimization in the clean system increases quadratically. We used ADAM with 500 update steps on a Fourier ansatz of $N_c = 50$ frequency components with external constraints $(\tau, d) = (100, 50)$. All other parameters are the same as in Fig. 5. (b) The obtained velocity protocols (left y axis) with the $\langle \mathcal{I}_\tau \rangle_s$ method for three specific disorder patterns at $\Delta = 0.12$. The disorder value ϵ_x where $x = \text{int}[X_0(t)]$ is plotted in orange with the scale on the right y xis. For each different pattern we obtain a different optimal transport protocol.

derived from experimental measurements. That said, the fact that suitable control protocols exist directly motivates the use of other gradient-free optimization approaches. We refer to the conclusion for a further discussion on this.

To try to understand how the optimization method is able to achieve these low infidelity values despite the presence of disorder, we show in Fig. 6(b) the optimal velocity protocols $\dot{X}_0(t)$ for three specific disorder patterns. The obtained optimal protocols $\dot{X}_0(t)$ are different for each different disorder pattern. This implies that during the optimization process the machine indeed learns about the specific disorder realization and finds a way to correct for it. This is likely due to the fact that the information of a fixed pattern p is directly encoded in the derivatives $\frac{\partial \mathcal{I}_\tau^p}{\partial A_n}$ that the optimizer gets. The strategies being different, however, also means that there is likely no universal strategy that obtains a low infidelity value for any disorder realization, which we elucidate in the following subsection.

B. Ensemble average disorder

In the second scenario we consider ∂P optimization when we do not know the exact disorder realizations and only have access to the magnitude of the noise. For this we take the disorder averaged infidelity $\langle \mathcal{I}_\tau \rangle$ as the figure of merit to minimize with ∂P via batch gradient descent. We first fix a set of 200 different disorder realizations, divided up into 20 individual batches of 10 realizations. For each batch we compute the disorder-averaged gradients $\frac{\partial \langle \mathcal{I}_\tau \rangle}{\partial A_n}$ and use them to update the control protocol with ADAM. We do this cyclically which means we start from the first batch do an update and then move on to the second batch. When all batches have had one update we have completed one learning "episode" and repeat the cycle. During this process $\langle \mathcal{I}_\tau \rangle$ averaged over all the 200 disorder patterns goes gradually down until convergence is reached.

In Fig. 7(a) we show the resulting averaged infidelity $\langle \mathcal{I}_\tau \rangle$ values for the optimal protocols obtained with batch gradient descent for 49.5 episodes. Note that we first optimized

for 24.5 episodes on one fixed set of 200 disorder realizations and then 25 episodes on a different set of 200 disorder realizations. Here 0.5 episode means that we stopped after 100 out of the 200 disorder realizations. For these $\langle \mathcal{I}_\tau \rangle$ values we have used a new set of disorder realizations which were not used during the optimization process. This gets rid of any optimization bias to particular disorder realizations, and we can fairly evaluate the performance. We see that now the optimization method is not able to keep the infidelity close to zero for increasing disorder strength Δ. Instead it increases quadratically similarly to the benchmark protocols, with the batch optimization method providing a negligible increase in performance compared to the other protocols. We remark
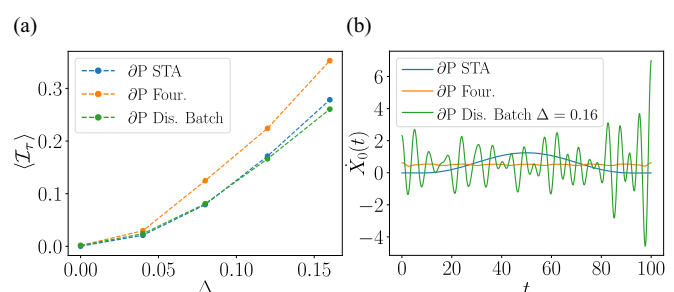

FIG. 7. (a) Disorder-averaged infidelity $\langle \mathcal{I}_\tau \rangle$ versus disorder strength Δ for protocols obtained with the ∂P batch optimization method (green dashed line) compared to the optimal clean protocols. Results are averaged over 500 disorder realizations, and all the model parameters are the same as in Fig. 5. The external constraints are set to $(\tau, d) = (100, 50)$. All the protocols increase approximately quadratically with disorder strength and the batch optimization slightly outperforms the benchmark protocols. (b) The optimal velocity protocol $\dot{X}_0(t)$ (green line) obtained with the batch optimization method at $\Delta = 0.16$. We have also shown the optimal protocols for the disorder-free system. The batch optimization protocol has much more drastic and rapid changes in velocity compared to the optimal clean protocols.

this behavior is qualitatively consistent with previous studies focusing the effect of parameter fluctuations [56,64].

As a final remark we note that although the batch optimization method does not give a significant improvement in infidelity compared to the STA protocols, the form of the pulses are different between the approaches. We show in Fig. 7(b) the optimal velocity protocol obtained at a disorder strength of $\Delta = 0.16$ compared to the optimal clean Fourier and STA protocols. We observe that the batch optimization protocol has much more rapid changes in the velocity and also obtains higher speeds, but, nevertheless, performs comparably to these protocols.

VI. CONCLUSION AND OUTLOOK

We have examined schemes to achieve high-fidelity transport of magnons in clean and disordered Heisenberg spin chains. For this optimization task we have exploited a hybrid numerical-analytical approach employing different types of analytical ansatz for the protocols, optimized with differentiable programming (∂P). In the absence of disorder, effective control can be achieved with all techniques, however, certain limiting factors relating to the maximum speed emerge with the ∂P Fourier able to approach the maximum group velocity. The relevance of external physical parameters, in particular, the frequency of the trapping potential used, was shown to manifest in the optimized protocols.

Introducing disorder, we showed that it is still possible to achieve near perfect optimal magnon transport for a fixed on-site disorder pattern, thus, demonstrating that the presence of disorder is, in and of itself, not a limiting factor in using quantum systems for communication. However, we also established that even with optimization one cannot achieve a "one size fits all" protocol if only the disorder strength and not the precise pattern is known. We focused explicty on the common assumption of Gaussian initial states. Although this is useful for the approximate STA protocol, it is not critical for the implementation of the other protocols. We expect that other initial states will not qualitatively change our results. These results naturally lead to the question if the same protocols (or protocols with same performance) can be found with non-gradient-based optimization approaches, such as reinforcement learning [19,65–68] or natural evolution strategies [31,69]. This will be, particularly, intriguing for the fixed disorder pattern results as this potentially eliminates the need to know the exact disorder pattern.

From an optimization perspective, several additional interesting avenues could be investigated. For example, how the control landscape properties (smoothness, convexity, etc.) [70] for the different protocol ansatzes change with the presence of disorder. It is also relevant to explore potential relationships between the complexity of the pulse, e.g., in terms of its spectral bandwidth, and the disorder. However, to explore these kind of questions about the effects of disorder often many computationally expensive numerical simulations must be performed. To resolve this and bring the computational complexity down, a final interesting question would be to see if one can train a neural network to predict the value of the cost function as was explored for a different problem in Ref. [71].

[1] P. Zoller *et al.*, Quantum information processing and communication, Eur. Phys. J. D **36**, 203 (2005).

[2] H. J. Kimble, The quantum internet, Nature (London) **453**, 1023 (2008).

[3] J. Majer *et al.*, Coupling superconducting qubits via a cavity bus, Nature **449**, 443 (2007).

[4] B. Çakmak, S. Campbell, B. Vacchini, Ö. E. Müstecaplıoğlu, and M. Paternostro, Robust multipartite entanglement generation via a collision model, Phys. Rev. A **99**, 012319 (2019).

[5] S. Bose, Quantum Communication through an Unmodulated Spin Chain, Phys. Rev. Lett. **91**, 207901 (2003).

[6] L.-M. Duan, E. Demler, and M. D. Lukin, Controlling Spin Exchange Interactions of Ultracold Atoms in Optical Lattices, Phys. Rev. Lett. **91**, 090402 (2003).

[7] A. Romito, R. Fazio, and C. Bruder, Solid-state quantum communication with josephson arrays, Phys. Rev. B **71**, 100501(R) (2005).

[8] P. Cappellaro, C. Ramanathan, and D. G. Cory, Dynamics and control of a quasi-one-dimensional spin system, Phys. Rev. A **76**, 032317 (2007).

[9] S. Hild, T. Fukuhara, P. Schauß, J. Zeiher, M. Knap, E. Demler, I. Bloch, and C. Gross, Far-from-Equilibrium Spin Transport in Heisenberg Quantum Magnets, Phys. Rev. Lett. **113**, 147205 (2014).

[10] H. Qiao, Y. P. Kandel, K. Deng, S. Fallahi, G. C. Gardner, M. J. Manfra, E. Barnes, and J. M. Nichol, Coherent Multispin Exchange Coupling in a Quantum-Dot Spin Chain, Phys. Rev. X **10**, 031006 (2020).

[11] V. Balachandran and J. Gong, Adiabatic quantum transport in a spin chain with a moving potential, Phys. Rev. A **77**, 012303 (2008).

[12] K. Eckert, O. Romero-Isart, and A. Sanpera, Efficient quantum state transfer in spin chains via adiabatic passage, New J. Phys. **9**, 155 (2007).

[13] A. Kiely and S. Campbell, Fast and robust magnon transport in a spin chain, New J. Phys. **23**, 033033 (2021).

[14] J. Werschnik and E. K. U. Gross, Quantum optimal control theory, J. Phys. B: At., Mol. Opt. Phys. **40**, R175 (2007).

[15] S. J. Glaser *et al.*, Training schrödinger's cat: quantum optimal control, Eur. Phys. J. D **69**, 279 (2015).

[16] C. P. Koch *et al.*, Quantum optimal control in quantum technologies. strategic report on current status, visions and goals for research in europe, EPJ Quantum Technol. **9**, 19 (2022).

[17] M. M. Müller, R. S. Said, F. Jelezko, T. Calarco, and S. Montangero, One decade of quantum optimal control in the chopped random basis, Rep. Prog. Phys. **85**, 076001 (2022).

[18] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and Jeffrey Mark Siskind, Automatic differentiation in machine learning a survey, J. Machine Learn. Res. **18**, 1 (2018).

[19] M. Bukov, A. G. R. Day, D. Sels, P. Weinberg, A. Polkovnikov, and P. Mehta, Reinforcement Learning in Different Phases of Quantum Control, Phys. Rev. X **8**, 031086 (2018).

[20] T. Caneva, M. Murphy, T. Calarco, R. Fazio, S. Montangero, V. Giovannetti, and G. E. Santoro, Optimal Control at the Quantum Speed Limit, Phys. Rev. Lett. **103**, 240501 (2009).

[21] X. Wang, A. Bayat, S. G. Schirmer, and S. Bose, Robust entanglement in antiferromagnetic heisenberg chains by single-spin optimal control, Phys. Rev. A **81**, 032312 (2010).

[22] T. Caneva, T. Calarco, and S. Montangero, Chopped random-basis quantum optimization, Phys. Rev. A **84**, 022326 (2011).

[23] P. M. Poggi and D. A. Wisniacki, Optimal control of many-body quantum dynamics: Chaos and complexity, Phys. Rev. A **94**, 033406 (2016).

[24] V. I. Gurman, I. S. Guseva, and O. V. Fesko, Optimization of excitation transfer in a spin chain, in *International Conference of Numerical Analysis and Applied Mathematics 2015 (ICNAAM 2015)*, *AIP Proc. Conf. No. 1738* (AIP, Melville, NY, 2016), p. 400005.

[25] X.-P. Zhang, B. Shao, S. Hu, J. Zou, and L.-A. Wu, Optimal control of fast and high-fidelity quantum state transfer in spin-1/2 chains, Ann. Phys. (NY) **375**, 435 (2016).

[26] D. S. A. Coden, S. S. Gómez, A. Ferrón, and O. Osenda, Controlled quantum state transfer in XX spin chains at the quantum speed limit, Phys. Lett. A **387**, 127009 (2021).

[27] M. Iversen, R. E. Barfknecht, A. Foerster, and N. T. Zinner, State transfer in an inhomogeneous spin chain, J. Phys. B: At., Mol. Opt. Phys. **53**, 155301 (2020).

[28] R. E. Wengert, A simple automatic derivative evaluation program, Commun. ACM **7**, 463 (1964).

[29] H.-J. Liao, J.-G. Liu, L. Wang, and T. Xiang, Differentiable Programming Tensor Networks, Phys. Rev. X **9**, 031041 (2019).

[30] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, Automatic differentiation in machine learning: A survey, J. Mach. Learn. Res. **18**, 5595 (2017).

[31] L. Coopmans, D. Luo, G. Kells, B. K. Clark, and J. Carrasquilla, Protocol discovery for the quantum control of majoranas by differentiable programming and natural evolution strategies, PRX Quantum **2**, 020332 (2021).

[32] I. Khait, J. Carrasquilla, and D. Segal, Optimal control of quantum thermal machines using machine learning, Phys. Rev. Res. **4**, L012029 (2022).

[33] J. B. Rigo and A. K. Mitchell, Automatic differentiable numerical renormalization group, Phys. Rev. Res. **4**, 013227 (2022).

[34] F. Schäfer, M. Kloc, C. Bruder, and N. Lörch, A differentiable programming method for quantum control, Mach. Learn.: Sci. Technol. **1**, 035009 (2020).

[35] N. Leung, M. Abdelhafez, J. Koch, and D. Schuster, Speedup for quantum optimal control from automatic differentiation based on graphics processing units, Phys. Rev. A **95**, 042318 (2017).

[36] P. W. Anderson, Absence of diffusion in certain random lattices, Phys. Rev. **109**, 1492 (1958).

[37] M. Ahmed, Confined magnon transport in low dimensional ferromagnetic structures, Ph.D Thesis, RMIT University, Melbourne, Australia (2017), https://researchrepository.rmit.edu.au/esploro/outputs/doctoral/Confined-magnon-transport-in-low-dimensional/9921863638601341#file-0..

[38] M. H. Ahmed and A. D. Greentree, Guided magnon transport in spin chains: Transport speed and correcting for disorder, Phys. Rev. A **91**, 022306 (2015).

[39] M. I. Makin, J. H. Cole, C. D. Hill, and A. D. Greentree, Spin Guides and Spin Splitters: Waveguide Analogies in One-Dimensional Spin Chains, Phys. Rev. Lett. **108**, 017207 (2012).

[40] E. Torrontegui, S. Ibáñez, X. Chen, A. Ruschhaupt, D. Guéry-Odelin, and J. G. Muga, Fast atomic transport without vibrational heating, Phys. Rev. A **83**, 013415 (2011).

[41] H. R. Lewis and W. B. Riesenfeld, An exact quantum theory of the time-dependent harmonic oscillator and of a charged particle in a time-dependent electromagnetic field, J. Math. Phys. **10**, 1458 (1969).

[42] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, and S. Wanderman-Milne, JAX: composable transformations of PYTHON+MUMPY programs (2018).

[43] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Learning representations by back-propagating errors, Nature (London) **323**, 533 (1986).

[44] W. Baur and V. Strassen, The complexity of partial derivatives, Theor. Comput. Sci. **22**, 317 (1983).

[45] A. Griewank, On automatic differentiation, Mathematical Programming: Recent Developments and Applications, **6**, 83 (1989), https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=034212770af42af1db1018807398dbde1308a991.

[46] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings (2015).

[47] R. Porotti, D. Tamascelli, M. Restelli, and E. Prati, Coherent transport of quantum states by deep reinforcement learning, Commun. Phys. **2**, 61 (2019).

[48] In our simulations we take $\Delta t = 0.1$ which is sufficient to ensure good convergence.

[49] S. Campbell, G. De Chiara, M. Paternostro, G. M. Palma, and R. Fazio, Shortcut to Adiabaticity in the Lipkin-Meshkov-Glick Model, Phys. Rev. Lett. **114**, 177206 (2015).

[50] H. Saberi, T. Opatrný, K. Mølmer, and A. del Campo, Adiabatic tracking of quantum many-body dynamics, Phys. Rev. A **90**, 060301(R) (2014).

[51] B. Bartels and F. Mintert, Smooth optimal control with floquet theory, Phys. Rev. A **88**, 052315 (2013).

[52] S. Meister, J. T. Stockburger, R. Schmidt, and J. Ankerhold, Optimal control theory with arbitrary superpositions of waveforms, J. Phys. A: Math. Theor. **47**, 495002 (2014).

[53] T. E. Skinner and N. I. Gershenzon, Optimal control design of pulse shapes as analytic functions, J. Magn. Reson. **204**, 248 (2010).

[54] X. Wang, D. Burgarth, and S. Schirmer, Subspace controllability of spin-$\frac{1}{2}$ chains with symmetries, Phys. Rev. A **94**, 052319 (2016).

[55] J. Lee, C. Arenz, H. Rabitz, and B. Russell, Dependence of the quantum speed limit on system size and control complexity, New J. Phys. **20**, 063002 (2018).

[56] M. H. Goerz, E. J. Halperin, J. M. Aytac, C. P. Koch, and K. B. Whaley, Robustness of high-fidelity rydberg gates with single-site addressability, Phys. Rev. A **90**, 032329 (2014).

[57] A. Couvert, T. Kawalec, G. Reinaudi, and D. Guéry-Odelin, Optimal transport of ultracold atoms in the non-adiabatic regime, Europhys. Lett. **83**, 13001 (2008).

[58] A. Larrouy, S. Patsch, R. Richaud, J.-M. Raimond, M. Brune, C. P. Koch, and S. Gleyzes, Fast Navigation in a Large Hilbert Space Using Quantum Optimal Control, Phys. Rev. X **10**, 021058 (2020).

[59] J. M. Epstein and K. B. Whaley, Quantum speed limits for quantum-information-processing tasks, Phys. Rev. A **95**, 042314 (2017).

[60] M. Murphy, S. Montangero, V. Giovannetti, and T. Calarco, Communication at the quantum speed limit along a spin chain, Phys. Rev. A **82**, 022318 (2010).

[61] V. F. Krotov, Global Methods in Optimal Control Theory, in *Advances in Nonlinear Dynamics and Control: A Report from Russia*, Progress in Systems and Control Theory, edited by A. B. Kurzhanski (Birkhäuser, Boston, MA, 1993), pp. 74–121.

[62] D. Burgarth and S. Bose, Perfect quantum state transfer with randomly coupled quantum chains, New J. Phys. **7**, 135 (2005).

[63] G. De Chiara, D. Rossini, S. Montangero, and R. Fazio, From perfect to fractal transmission in spin chains, Phys. Rev. A **72**, 012323 (2005).

[64] S. D. Mishra, R. Trivedi, A. H. Safavi-Naeini, and J. Vučković, Control Design for Inhomogeneous-Broadening Compensation in Single-Photon Transducers, Phys. Rev. Appl. **16**, 044025 (2021).

[65] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction* (Bradford Books, Cambridge, MA, USA, 2018).

[66] X.-M. Zhang, Z. Wei, R. Asad, X.-C. Yang, and X. Wang, When does reinforcement learning stand out in quantum control? A comparative study on state preparation, npj Quantum Inf. **5**, 1 (2019).

[67] P. A. Erdman and F. Noé, Identifying optimal cycles in quantum thermal machines with reinforcement-learning, npj Quantum Inf. **8**, 1 (2022).

[68] P. A. Erdman and F. Noé, Driving black-box quantum thermal machines with optimal power/efficiency trade-offs using reinforcement learning, arXiv:2204.04785 (2022).

[69] D. Wierstra, T. Schaul, T. Glasmachers, Y. Sun, J. Peters, and J. Schmidhuber, Natural evolution strategies, J. Mach. Learn. Res. **15**, 949 (2014).

[70] A. G. R. Day, M. Bukov, P. Weinberg, P. Mehta, and D. Sels, Glassy Phase of Optimal Quantum Control, Phys. Rev. Lett. **122**, 020601 (2019).

[71] M. Dalgaard, F. Motzoi, and J. Sherson, Predicting quantum dynamical cost landscapes with deep learning, Phys. Rev. A **105**, 012402 (2022).