# Measurement-driven navigation in many-body Hilbert space: Active-decision steering

Yaroslav Herasymenko
*Instituut-Lorentz, Universiteit Leiden, P.O. Box 9506, 2300 RA Leiden, The Netherlands*

Igor Gornyi
*Institute for Quantum Materials and Technologies, Karlsruhe Institute of Technology, 76021 Karlsruhe, Germany*
*Institut für Theorie der Kondensierten Materie, Karlsruhe Institute of Technology, 76128 Karlsruhe, Germany and*
*Ioffe Institute, 194021 St. Petersburg, Russia*

Yuval Gefen
*Department of Condensed Matter Physics, Weizmann Institute of Science, 7610001 Rehovot, Israel*



The challenge of preparing a system in a designated state spans diverse facets of quantum mechanics. To complete this task of steering quantum states, one can employ quantum control through a sequence of generalized measurements which direct the system towards the target state. In an active version of this protocol, the obtained measurement readouts are used to adjust the protocol on-the-go. This enables a sped-up performance relative to the passive version of the protocol, where no active adjustments are included. In this work, we consider such active measurement-driven steering as applied to the challenging case of many-body quantum systems. The target states of highest interest would be those with multipartite entanglement. Such state preparation in a measurement-based protocol is limited by the natural constraints for system-detector couplings. We develop a framework for finding such physically feasible couplings, based on parent Hamiltonian construction. For helpful decision-making strategies, we offer Hilbert-space-orientation techniques, comparable to those used in navigation. The first one is to tie the active-decision protocol to the greedy accumulation of the cost function, such as the target state fidelity. We show the potential of a significant speedup, employing this greedy approach to a broad family of Matrix Product State targets. For system sizes considered here, an average value of the speedup factor $f$ across this family settles about 20, for some targets even reaching a few thousands. We also identify a subclass of Matrix Product State targets, including the ground state of the Affleck-Kennedy-Lieb-Tasaki spin chain, for which the value of $f$ increases with system size. In addition to the greedy approach, the second wayfinding technique is to map out the available measurement actions onto a Quantum State Machine. A decision-making protocol can be based on such a representation, using semiclassical heuristics. This State Machine-based approach can be applied to a more restricted set of targets, where it sometimes offers advantages over the cost function-based method. We give an example of a W-state preparation which is accelerated with this method by $f \simeq 3.5$, outperforming the greedy protocol for this target.

## I. INTRODUCTION

Quantum state preparation is a prominent routine in quantum information processing toolbox [1–19]. Such procedure often implies steering a quantum system from a "simple" towards a more complex, pre-designated resourceful state, e.g. a many-body entangled state. A steering protocol is characterized by an as short as possible runtime and high resulting overlap with the target state. Constructing such protocols can be done in multiple distinct ways. One is to design the Hamiltonian of the system, such that its unitary evolution leads to a designated state. This paradigm is represented by methods like digital computation or analog simulation [1–5]. Such protocols require exact knowledge of the starting state, as well as the precise timing of the unitary evolution, to be accurate. Another strategy is to add a dissipative element to the protocol. Combined with the Hamiltonian evolution, this results in methods such as drive-and-dissipation [6, 7] and quantum channel engineering [8–11]. Finally, one can design a sequence of generalized measurements, which brings the system towards the target state via measurement back-action alone [20–22]. The system-detector coupling completely governs the relevant part of the evolution in such a protocol (see also Ref. [23]). Unlike protocols involving pre-defined unitary evolution, such measurement-driven state preparation may not require knowledge of the starting state and fine-tuning of the system Hamiltonian [22].

The above types of state-preparation strategies can be referred to as *passive*, meaning that these protocols are pre-determined and pursued regardless of how the system evolves. Given this perspective, it appears beneficial to go beyond the forms of control described above, and introduce the concept of *active* decision making. This type of steering exploits information extracted during the system's evolution to decide on the operations that follow. This is also referred to as closed-loop quantum control and is typically used to improve the Hamiltonian-based state preparation [24–29]; dissipation-based protocols are also being considered [30]. Extracting the necessary information requires introducing measurements into the protocol, which may result in an undesired back-action. Nevertheless, in many cases, closed-loop control

does yield an improvement in the speed and the fidelity of the protocol.

Another possibility, which is a subject of increasing interest, is to employ active decision in measurement-driven protocols (implying no other source of drive, Hamiltonian or dissipative) [31–33]. In such protocols, information about the running state of the system is naturally available from the employed measurements. This data can be used for the active choice of subsequent generalized measurements, such that the target state is prepared as rapidly and accurately as possible [31–33]. Some general theorems have been stated concerning such state preparation protocols [33], along with some specific protocols designed to reach single-qubit target states [31, 32]. However, it remains unclear how an active measurement-driven protocol can be effectively harnessed to engineer resourceful many-body states. In this case, the large size of the Hilbert space makes it challenging to actively steer the system evolution in the desired direction.

In this work, we establish a general framework for measurement-driven active navigation in Hilbert space. The goal of this framework is to construct active-decision protocols for measurement-only steering of many-body system to a given target state. In particular, we focus on target states manifesting genuine multipartite entanglement. The key problem here is to ensure that the protocol achieves the target state in as short time as possible. When attempting to address this problem, one is naturally constrained by a few factors. One is that only reasonably local system-detector couplings are to be used in the protocol. Moreover, we require that the number of distinct system-detectors couplings available for steering does not scale up faster than the system's size (this number should not be super extensive). This practical requirement restricts the capabilities of the protocol. From the limitations above, it naturally follows that applying one type of coupling generally leads to an update in the expected benefits from other couplings (see Sec. III for a detailed analysis). This phenomenon, which we refer to as "coupling frustration", calls for nontrivial coordination between different coupling applications. Heuristically, one can view the problem at hand as one of orienteering: it is easy to "get lost" in the many-body Hilbert space when exploring it with a limited set of tools (cf. Ref. [25]).

We note that upon the availability of indefinite computational power, one can always find an optimum sequence of measurements through dynamic programming techniques (cf. Ref. [33]). Roughly speaking, this can be done by considering all possible future quantum trajectories of the system. However, in a large Hilbert space, it is practically impossible to realize the theoretically optimal feedback policy. This is because the extensive consideration of outcome scenarios is too complex for a many-body Hilbert space of an already not very large system (it increases at least exponentially with the system's size and the runtime of the protocol). Instead, we aim at designing heuristic strategies for active decision-making. The key metric of success is the speedup coming from such active decision making. This is defined as factor $f$ between the average runtime of a passive protocol and such of a comparable protocol that employed an active decision-making strategy (see Sec. II for details). The goal of our heuristic strategies is to ensure a significant – but not necessarily optimal – value of the speedup factor $f$.

To meet these challenges, we introduce *Hilbert-space navigation techniques*. The first technique, which we term *greedy orienteering policy*, is based on the notion of a cost function. A simple example of such a cost function is the target state infidelity. Minimizing it in a greedy protocol may already yield a significant advantage compared to the passive policy. To test this approach, we study numerically the preparation of Matrix Product State (MPS) targets. We consider uniform spin-1 MPS with bond dimension 2. For such target states generated at random, we discover that the greedy policy yields speedups $f$ up to $f_{\max} \sim 10^3$ and an average value about $f_{\text{av}} \sim 10^1$ ($f_{\max} = 3400$, $f_{\text{av}} \simeq 19$ for target states sampled at system size $N = 5$). Interestingly, among these MPS targets $f$ tends to increase with the system size for states whose parent Hamiltonian has a large enough spectral gap. This in particular holds for the Affleck-Lieb-Kennedy-Tasaki state [34], which is a well-known example from the MPS family we consider.

The second navigation technique is via mapping the Hilbert space onto a colored multigraph, referred to as the Quantum State Machine. The vertices of such a graph correspond to the basis states, and the edges represent the actions of generalized measurements. Upon an appropriate choice of basis states, such Quantum State Machine representations allow for improved navigation in Hilbert space. This can be done by heuristically representing it as quantum wayfinding on the graph. To substantiate this heuristic, we introduce the notion of semiclassical coarse-graining of a Quantum State Machine graph. Optimizing the exploration of these graphs by choosing the most appropriate system-detector couplings results in advantageous active-decision protocols. Notably, Quantum State Machine approach is conceptually different to the greedy approach, and has the potential to offer a higher speedup factor $f$ for some applications. To exemplify this alternative navigation paradigm we consider the preparation of the 3-qubit W-state. A numerical study demonstrates that Quantum State Machine based approach yields $f = 3.5$, while the greedy approach results in $f = 3.1$.

Throughout the paper, we assume that we know the initial state of the system. This can be a "cheap" (say, product) and robust quantum state that does not require many resources for its preparation. However, one can directly generalize the above approaches to the case where the initial state is unknown and is therefore represented by a density matrix. In the more intricate case of a Quantum State Machine-based policy, one would then need to take a weighted combination of graph navigation protocols with different initial states.

The remainder of the paper is organized as follows. In Sec. II, we introduce the basics of measurement-induced steering. Specifically, in Sec. II A, we define the steering protocols and their elements, as well as the quantitative measure of the protocol's success. Then, in Sec. II B, we illustrate these definitions as applied to passive steering of a single qubit. The general selection criteria, including locality and extensivity, for the system-detector couplings, which are to be used for the active steering, are addressed in Sec. II C. In Sec. III, we introduce the notion of frustration of steering and discuss the possibilities of protocols' speed-up for mutually commuting (Sec. III A) and non-commuting (Sec. III B) couplings. In the latter case, we develop a parent-Hamiltonian approach. A "quantum-compass" approach to active-decision steering, based on the greedy cost-function accumulation policy, is developed in Sec. IV. where we also employ this scheme to the preparation of the MPS states. In Sec. V, we develop the Quantum State Machine framework. In Sec. V A, we introduce the generalities of this approach based on the underlying representation of the steering protocol in terms of a quantum graph. Next, we discuss the quantum parts of this graph (Sec. V B), as well as the coarse-graining procedure, with the resulting coarse-grained graph being semiclassical (Sec. V C). The advantage of this type of Hilbert-space orienteering is illustrated in Sec. V D, where an active-decision steering protocol for preparation of a three-qubit W-state is presented. Our findings are summarized and discussed in Sec. VI.

## II. MEASUREMENT-DRIVEN STATE PREPARATION

### A. Generalities

Measurement-driven state steering is a specific class of state-preparation protocols. Its basic building blocks are coupling the quantum system ($s$) to quantum detectors (ancillary systems) utilizing engineered interactions, followed by strong measurement on the detectors ($d$). The goal of designing a measurement-based steering protocol is to generate a process that prepares the desired system state by utilizing a sequence of measurement backactions.

Here, we will additionally assume that the internal evolution of the system and the detector are trivial (their Hamiltonians are kept null: $H_s = 0$, $H_d = 0$), as in Refs. [31–33], so that the dynamics in the problem is governed solely by the system-detector interaction $H_{s,d}$. For concreteness of analysis, we also constrain the detector to be a qubit initialized in a trivial state $|0\rangle$, and the system to be represented by $N$ spins. Although a general spin $S$ can be considered, we focus on the cases $S = 1/2$ and $S = 1$. We assume certain knowledge about the initial state of the system, which is described by the initial density matrix $\rho_{\rm in}$. For the sake of simplicity, we further address the target state which is a pure state $|\psi_{\rm targ}\rangle$.

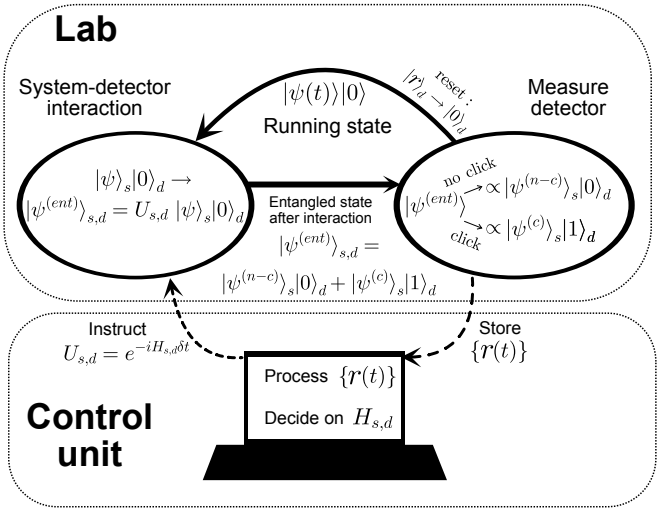


FIG. 1. Basic design of the measurement-driven state preparation. The procedure starts with a given initial state $\rho_{\rm in}$ and proceeds with a protocol, as described in Def. 1, until a good approximation of the target state $|\psi_{\rm targ}\rangle$ is achieved. The control unit decides on the system-detector interaction unitary $U_{s,d}$ based on the stored record of detector readouts. We focus on constructing an optimized policy for decision-making, such that the target state is prepared as quickly as possible.

Although the ensuing protocol can be further generalized (see Sec. VI), we now formally fix its structure as given below:

**Definition 1** *A measurement-driven state steering is a protocol that is performed to prepare a state $|\psi_{\rm targ}\rangle$, starting from the state $\rho_{\rm in}$. It runs by repeating iterative cycles of the following form (see Fig. 1):*

1. *Prepare the detector qubit in the state $|0\rangle$.*
2. *Based on the available information, select the system-detector coupling Hamiltonian $H_{s,d}$ to be used in the next step.*
3. *Perform a system-detector evolution governed by a Hamiltonian $H_{s,d}$ for a short time interval $\delta t$: $U_{s,d} = e^{-iH_{s,d}\delta t}$.*
4. *Once the system-detector evolution is over, projectively measure the detector qubit in the $Z$-basis. Store the readout $r$ for further processing.*
5. *Decide whether the protocol is to be continued or terminated. In the former case, return to step 1.*

Now, in the vast space of protocols that have such structure, we would like to emphasize the distinction between two classes of protocols: passive and active.

In a passive protocol, the stored readouts $\{r(t)\}$ from step 4 may influence the decision for protocol termination or continuation at step 5, but not the choice for the

interaction Hamiltonian $H_{s,d}$ made at step 2 in the next protocol cycles. Hamiltonians $H_{s,d}$ can still be chosen differently for different iterations: e.g. for a large system, the detector qubit can be coupled to different subsystems thereof. However, $H_{s,d}$ used at each cycle in the passive protocol has to be pre-determined from the outset. If a passive protocol also has a pre-determined duration (and thus doesn't use readouts $\{r(t)\}$ at the termination step), we land in a subclass of passive protocols where the readouts don't have any influence on the protocol. We would refer to such protocols as "blind steering". For blind steering, the readouts of the detector at any given cycle can be averaged, i.e., following the measurement, the detector's density matrix is traced out. In this work, however, we will focus on the non-blind version of passive steering, where readouts are indeed employed for an informed protocol termination.

In contrast to passive protocols, in an active protocol one uses the readouts $\{r(t)\}$ to make an informed decision for the interaction Hamiltonians $H_{s,d}$ as well as for termination/continuation of the protocol. Active decision-making has to follow a certain policy, which becomes the crucial part of the protocol. For a good active policy, its adoption should result in a significant speedup of the protocol compared to its passive counterpart. Alternatively, one can also fix the protocol runtime and aim to improve the precision of the state preparation. We focus on the former target: minimizing protocol runtime for a fixed target precision. The major challenge in this work is to construct such advantageous active decision-making policies. By comparing active steering with the (non-blind) passive steering as defined above, we investigate the advantage offered specifically by the directed evolution, i.e. active decision-making for $H_{s,d}$.

Before we move on to the issue of active policy constructions, let us discuss the criteria for termination of a running protocol. In general, one cannot guarantee "perfect steering," i.e., obtaining the desired target state with the fidelity of 1 in a finite number of protocol cycles. Instead, one may consider preparing the target state with *infidelity* $R$:

$$R\left(\rho_{\text{out}}, |\psi_{\text{targ}}\rangle\right) \equiv 1 - \langle\psi_{\text{targ}}|\rho_{\text{out}}|\psi_{\text{targ}}\rangle, \tag{1}$$

where the state $\rho_{\text{out}}$ is the final state of the system once the protocol is terminated. It is worth emphasizing that the system evolution during the protocol is probabilistic and depends on the stochastic readouts $\{r(t)\}$. It follows that different runs of the same protocol may yield different $\rho_{\text{out}}$ and, thus, the infidelity R. Therefore, to characterize the protocol as a whole, we introduce the following accuracy measure:

**Definition 2** *We refer to a measurement-driven state-preparation protocol as $\epsilon$-precise, if the infidelity between the final state and the target state is bounded by $\epsilon$ for any run of the protocol:*

$$R\left(\rho_{\text{out}}, |\psi_{\text{targ}}\rangle\right) < \epsilon. \tag{2}$$

Given the knowledge of the readout sequence, we may simulate the quantum system state (the quantum trajectory) on a computer in parallel to the measurement run. Thus one can infer the running system state exactly (referred to as filtering in the literature [24]), and test inequality (2). This sets a trivial criterion for protocol termination, which we will apply by default to all passive and active protocols considered in this work. Namely, a protocol can be terminated right after the cycle when the target state infidelity becomes smaller than $\epsilon$, thus making it an $\epsilon$-precise protocol. Apart from controlling the precision, we are interested in the number of cycles $\tau$ after which the protocol has been terminated. As $\tau$ may differ greatly, depending on a specific run, we will characterize the protocol by the runtime averaged over many runs $\tau_{\text{av}} \equiv \langle\tau\rangle_{\text{run}}$. Here averaging is taken over stochastic readout sequences. In reality, steering errors as well as external noise may be further contributing factors to the stochasticity. For the given target state and target precision $\epsilon$ (cf. Def. 2), our goal is to find an $\epsilon$-precise protocol such that $\tau_{\text{av}}$ is as small as possible. We will be considering this minimization as the key goal of our constructions. In particular, when developing the active policy, we will be interested in maximizing the speedup factor

$$f \equiv \frac{\tau_{\text{av}}^{(\text{pas})}}{\tau_{\text{av}}^{(\text{act})}} \tag{3}$$

of the active protocol relative to its passive counterpart.

### B. Passive steering: Single qubit

As a simple example of a measurement-driven protocol, we consider single-qubit steering (for a more general consideration, the reader is referred to Sec. II C). For simplicity, we will assume the target state to be $|0\rangle$, and the starting state to be a perfectly mixed state: $\rho_{\text{in}} = \text{diag}(1/2, 1/2)$. A single coupling suffices to guarantee the preparation of the target state (in fact, from an arbitrary starting state) with an arbitrary precision [22]:

$$H_{s,d} = \gamma\sigma_s^-\sigma_d^+ + \text{H.c.} \tag{4}$$

Here, $\sigma_s$ and $\sigma_d$ are Pauli matrices acting in the system and detector spaces, respectively. By construction, a protocol that operates with only a single coupling Hamiltonian $H_{s,d}$, i.e., without a readout-based option of choosing different couplings, is considered passive. Nevertheless, even for passive protocols, one can introduce a policy based on the measurement outcomes, which would accelerate quantum-state steering.

Let us first address a protocol that runs for $\tau$ cycles using the coupling (4), regardless of the measurement outcomes. Under the definition given in Sec. II A, this would be an example of *blind* steering. In this case, the probability of obtaining a readout $r = 0$ decreases exponentially with the total number of cycles $\tau$. Tracing out

detector outcomes (since we are blind to measurement outcomes), this results in a density matrix:

$$\rho(\tau) = \begin{pmatrix} 1 - e^{-\tau\gamma^2\delta t^2}/2 & 0 \\ 0 & e^{-\tau\gamma^2\delta t^2}/2 \end{pmatrix}. \tag{5}$$

Given the threshold infidelity $\epsilon$, we need to run the protocol for $\tau^{(\mathrm{blind})}(\epsilon)$ cycles:

$$\tau^{(\mathrm{blind})}(\epsilon) = \frac{1}{\gamma^2\delta t^2} \log\left(\frac{1}{2\epsilon}\right) \tag{6}$$

This characterizes the efficiency of the completely blind passive protocol [22] for the single-qubit setup.

Next, we consider passive protocols where the sequence of readouts is recorded. One then needs to interpret the measurement outcomes, which for this setup is straightforward. We note that when the readout is $r = 1$ (click event), the target state is instantly prepared [cf. Eq. (10)]. Therefore, one can terminate the protocol directly after the detector clicks for the first time: in this case, all further cycles are simply redundant and do not result in any evolution of the system. This will constitute a termination-policy improvement of the passive blind protocol for this single-qubit case. If $r = 0$, i.e. no click is measured (such a null-measurement [35] event still gives the system a nudge towards the target state by measurement back-action), the protocol simply continues until a certain maximal number of cycles, $\tau_{\max}$. The target infidelity $\epsilon$ would be directly related to $\tau_{\max}$ in a way equivalent to the blind protocol runtime (6). The average runtime of the non-blind passive protocol is then given by:

$$\tau_{\mathrm{av}}^{(\mathrm{pass})} = \frac{1}{2\gamma^2\delta t^2}\left(1 - e^{-\gamma^2\delta t^2\tau_{\max}}\right) + \frac{\tau_{\max}}{2}. \tag{7}$$

This runtime is strictly smaller than the runtime for the passive blind protocol, Eq. (6), and yields a twofold speedup in the $\epsilon \to 0$ limit.

For a many-body system, the termination policy will not exhaust the possible active policies as it does for a single-qubit target. Indeed, a single click of the detector coupled to a subsystem does not guarantee that the whole system is steered to the desired state. Nevertheless, the above simple example shows that detector readouts can be used for accelerating the state preparation. In what follows, we will focus on active feedback strategies. There, instead of protocol termination, the local-measurement outcomes are employed for choosing the most efficient sequence of further measurement cycles.

### C. Selection criteria for system-detector couplings

Both for the active and passive protocols, a key feature is the choice of coupling Hamiltonians $H_{s,d}$. Given the target state $|\psi_{\mathrm{targ}}\rangle$, it is natural to constrain this choice to a certain family $\{H_{s,d}(\mathbf{p})\}$, for a set of parameters $\mathbf{p}$. A passive protocol can then take a periodic structure: couplings $\{H_{s,d}(\mathbf{p})\}$ are employed in a predefined order, and this sequence of protocol cycles is repeated once the list of couplings has been exhausted. In an active protocol, the choice of $H_{s,d}$ in each protocol cycle translates into actively selecting the value of $\mathbf{p}$. Before discussing the policies for doing so, we will consider a more basic question – how should the family $\{H_{s,d}(\mathbf{p})\}$ be constructed to yield a viable *passive* protocol?

#### 1. *Single cycle scenarios*

To understand the performance of the protocol defined by the family $\{H_{s,d}(\mathbf{p})\}$, we first analyze the change of the system state after a single protocol cycle (Def. 1). To do so, we consider a general decomposition

$$H_{s,d} = V_s\sigma_d^+ + V_s^\dagger\sigma_d^- + \tilde{V}_s\sigma_d^z. \tag{8}$$

Here $V_s$ and $\tilde{V}_s$ are arbitrary (not necessarily hermitian) system operators and matrices $\sigma_d^\pm = \frac{1}{2}(\sigma_d^x \pm i\sigma_d^y)$ act on the detector. In Eq. (8), we discard any terms of the form $\sim \mathbb{I}_d$, as those represent the internal system evolution. To simplify our further considerations, we will also impose a constraint $\tilde{V}_s = 0$.

With Eq. (8) in mind, we now consider the effect of a single protocol cycle on the system state $\rho$ when the measurement outcomes are averaged over (blind measurement). In the weak measurement limit, $\delta t \to 0$, this is represented by the map:

$$\begin{aligned}\rho &\to \Lambda_{V_s}(\rho) \\ &\equiv \left(1 - \frac{\delta t^2}{2}V_s^\dagger V_s\right)\rho\left(1 - \frac{\delta t^2}{2}V_s^\dagger V_s\right) + V_s\rho V_s^\dagger\delta t^2.\end{aligned} \tag{9}$$

The terms of order $\mathcal{O}(\delta t^2)$ in this expression represent the standard Lindbladian jump operator. Based on the map (9) for infinitesimal $\delta t$, one can derive a Lindblad equation describing the system evolution for the blind steering [22].

Let us now turn to the non-blind protocol, where the different measurement outcomes are discriminated. During step 4 of the protocol cycle (cf. Def. 1), there is a probability

$$p^{(\mathrm{cl})}(\rho, V_s) = \delta t^2\,\mathrm{tr}(V_s\rho V_s^\dagger)$$

that a qubit flip is measured in the detector (click probability). The resulting state in the limit of small $\delta t$ is then

$$\rho \to \Lambda_{V_s}^{(\mathrm{cl})}(\rho) \equiv \frac{V_s\rho V_s^\dagger}{\mathrm{tr}(V_s\rho V_s^\dagger)}. \tag{10}$$

A "no-click" scenario occurs with probability

$$p^{(\mathrm{ncl})}(\rho, V_s) = 1 - p^{(\mathrm{cl})}(\rho, V_s),$$

and results in a state

$$\rho \to \Lambda_{V_s}^{(\mathrm{ncl})}(\rho) \equiv \frac{\left(1-\frac{\delta t^2}{2}V_s^\dagger V_s\right)\rho\left(1-\frac{\delta t^2}{2}V_s^\dagger V_s\right)}{1-\delta t^2 \mathrm{tr}(V_s^\dagger V_s \rho)}. \quad (11)$$

For the weak-measurement limit considered here ($||V_s\delta t|| \ll 1$), the click probability is parametrically smaller than that for the no-click event: a qubit flip can be recorded in the detector only rarely.

#### 2. *Necessary conditions for the coupling operators*

We can now expound our considerations for the family $\{H_{s,d}(\mathbf{p})\}$ in terms of the operators $\{V_s(\mathbf{p})\}$. For a meaningful comparison between active and passive protocols, we first require that there exists a passive protocol that employs Hamiltonians $\{H_{s,d}(\mathbf{p})\}$ to reach the target state $|\psi_{\mathrm{targ}}\rangle$. To ensure this convergence of the passive protocol, it is natural to demand that none of $\{H_{s,d}(\mathbf{p})\}$ can move the system state away from the target state. Given Eqs. (10) and (11), this yields a dark-state condition $V_s(\mathbf{p})|\psi_{\mathrm{targ}}\rangle = 0$ for every $\mathbf{p}$. This is equivalent to every operator $V_s(\mathbf{p})$ taking the form

$$V_s = \sum_{\alpha=1}^{D-1} v_\alpha |\psi_{\mathrm{targ}}\rangle\langle\psi_\alpha| + \sum_{\alpha,\beta=1}^{D-1} w_{\alpha\beta}|\psi_\beta\rangle\langle\psi_\alpha|, \quad (12)$$

where $D$ is the Hilbert-space dimensionality of the system, and $\{|\psi_{\mathrm{targ}}\rangle\} \cup \{|\psi_\alpha\rangle\}$ is any basis of the system's Hilbert space that includes $|\psi_{\mathrm{targ}}\rangle$ as a basis state.

Having in mind steering of many-body states, we further require that the couplings $\{V_s(\mathbf{p})\}$ are feasible to engineer in an experimental realization of the system. In this work, we focus on the most basic aspect of this condition: locality. One may consider two types of locality: geometric and operator ($k$-locality [36]). Geometric locality of the operator $V_s$ implies that such interaction only requires coupling the system spins that are in geometrical proximity during the experiment. A $k$-local operator $V_s$ implies that only $k$ system spins are coupled at a time. It is natural to impose the locality constraint not on the full operator $V_s$, but its individual terms. For example, if $V_s$ involves all system spins, but its individual terms only couple 2 spins at a time, we will consider $V_s$ a 2-local coupling (in line with [36]). A $k$-local operator $V_s$ implies an interaction Hamiltonian $H_{s,d}$ that is $(k+1)$-local.

#### 3. *Possibility of spurious dark states. Room for active decision-making*

It is worth stressing that $V_s(\mathbf{p})$ following the form given by Eq. (12) for all $\mathbf{p}$ is necessary but not sufficient for $|\psi_{\mathrm{targ}}\rangle$ to be the *only* dark state of the passive protocol. For some choices of such a family $\{V_s(\mathbf{p})\}$, a spurious final state $|\psi'_{\mathrm{targ}}\rangle \neq |\psi_{\mathrm{targ}}\rangle$ might be reached. However, this would imply a dark-state condition $V_s(\mathbf{p})|\psi'_{\mathrm{targ}}\rangle = 0$ (for every $\mathbf{p}$), and this should not hold even for a single *generic* operator $V_s$, which satisfies $V_s|\psi_{\mathrm{targ}}\rangle = 0$. In other words, one does not expect $|\psi'_{\mathrm{targ}}\rangle$ to exist for generic (say, random-matrix-type) coefficients $v_\alpha$ and $\omega_{\alpha\beta}$ in (12). For that, an extra constraint is needed, such as vanishing of certain $v_\alpha$, $\omega_{\alpha\beta}$, or a specific relation between the coefficients.

One concludes that a family consisting of a single Eq. (12)-type coupling $V_s$ is sufficient to prepare $|\psi_{\mathrm{targ}}\rangle$ in a passive protocol without generating a dark space. Notably, reducing the family to a single member would leave no room for active decision-making in a protocol defined by this family (an active protocol requires at least two operators to choose from). On the other hand, such an ultimate $V_s$ would not in general satisfy the locality conditions and, thus, would be unrealistic to implement. Natural counterexample couplings $V'_s$ that have multiple dark states arise in the important case when $V'_s$ acts only on a part of the system.

To construct such a counterexample, one may start from an arbitrary operator $V_s$ that satisfies the dark-state condition $V_s|\psi_{\mathrm{targ}}\rangle = 0$ for a single state $|\psi_{\mathrm{targ}}\rangle$ in a given system. Now, consider a larger system embedding the original one and construct a different target state which is a tensor product of $|\psi_{\mathrm{targ}}\rangle$ and a certain auxiliary state: $|\Psi_{\mathrm{targ}}\rangle \equiv |\psi_{\mathrm{targ}}\rangle \otimes |\tilde{\psi}_{\mathrm{targ}}\rangle$. In this case, one may take $V_s \to V'_s$, where $V'_s = V_s \otimes \mathbb{I}_{\tilde{s}}$ still satisfies condition $V'_s|\Psi_{\mathrm{targ}}\rangle = 0$ relative to this new target state in the extended Hilbert space. Yet for a general starting state of the total system, the operator $V'_s$ is obviously not sufficient to prepare the extended target state $|\Psi_{\mathrm{targ}}\rangle$. This implies the existence of spurious dark states: in fact, all states of the form $|\psi_{\mathrm{targ}}\rangle \otimes |\tilde{\psi}_{\mathrm{targ}}\rangle$ turn out to be dark for arbitrary $|\tilde{\psi}_{\mathrm{targ}}\rangle$. This links to the previous discussion: an operator $V_s$ capable of steering to a unique dark state is expected to be highly nonlocal in general, in contrast to the limited capacity of a localized operator $V_s \otimes \mathbb{I}_{\tilde{s}}$.

From the arguments above, we conclude that a family of *multiple* operators $\{V_s(\mathbf{p})\}$ is needed to realistically prepare a target state, once that state is sufficiently complicated. Having multiple operators in $\{V_s(\mathbf{p})\}$, in turn, opens the door for gaining advantage through active decision-making.

## III. TYPES OF SYSTEM-DETECTOR COUPLINGS

The preselected family of coupling operators $\{V_s(\mathbf{p})\}$ determines both the performance of the ensuing passive protocols and the possibilities for active policy construction. In the present section, we identify the crucial role of the commutation properties of $\{V_s(\mathbf{p})\}$. We first consider $N$-qubit steering protocols which employ coupling operators $\{V_s(\mathbf{p})\}$ that are mutually commuting. As a shorthand, we call this frustration-free steering. We show that

a realistic passive protocol of this type can be designed for product states and certain graph states. Commuting couplings also allow for a simple feedback strategy, which results in a significant speedup of the respective passive protocol. Next, we move on to passive steering protocols that are frustrated. Such frustration of local couplings naturally arises for many-body target states related to local parent Hamiltonians. We propose an explicit method of constructing a family of non-commuting operators $\{V_s(\mathbf{p})\}$ that allows to prepare such a many-body target state in a passive protocol. This forms the basis for Sec. IV and V, where we move on to the active versions of frustrated steering protocols.

### A. Mutually commuting couplings

Here we focus on $N$-qubit steering protocols implemented with mutually commuting couplings $\{V_s(\mathbf{p})\}$. As will be demonstrated, a passive protocol of this type can be constructed for an arbitrary target state, yielding an asymptotically precise passive preparation. For a general target this construction yields *non-local* couplings $\{V_s(\mathbf{p})\}$, which are therefore impractical. We identify an exception to this rule – a subclass of graph states that can be obtained using *local* commuting couplings. For this, we discuss the constraints coming from both geometric locality, as well as $k$-locality. Finally, we extend the discussion from such passive protocols to their active counterparts. To achieve this, we propose a simple feedback strategy that speeds up frustration-free steering in a substantial way.

As a trivial example of frustration-free steering, consider an $N$-qubit product state as a target state, e.g., $|00..0\rangle$. The starting state will be assumed to be the perfectly mixed state. To prepare the target with a steering protocol, one can use a set of couplings parameterized by the qubit number $i = 1, ..N$:

$$V_s^{(\mathrm{prod})}(i) = \gamma \sigma_i^- . \tag{13}$$

Passively alternating between the steering cycles employing $V_s^{(\mathrm{prod})}(i)$ with different $i$ guarantees preparation of the target state with any given accuracy. This directly follows from the analysis of Sec. II B and II C. For an active version of the protocol, partial protocol termination can be applied: if a click is registered when measuring any qubit $i$, the coupling $V_s^{(\mathrm{prod})}(i)$ is dropped out from the sequence of couplings that will be applied in further cycles. In other words, the steering with this "fired" coupling is terminated at this point, whereas other couplings remain active – hence the term "partial termination". Since this implies a readout-based decision on the set of steering couplings that are used at a given cycle, we classify this as an active steering protocol. For the perfectly mixed $\rho_{\mathrm{in}}$ in the high precision limit $\epsilon \to 0$, this strategy results in the following relation between active and passive runtimes:

$$\tau_{\mathrm{av}}^{(\mathrm{act})}(\epsilon) \simeq \frac{\tau_{\mathrm{av}}^{(\mathrm{pass})}(\epsilon)}{2} + \frac{N}{2\gamma^2 \delta t^2}, \tag{14}$$

which leads to a $f \simeq 2$ speedup in the limit of small $\epsilon$ – similarly to Eq. (7).

Frustration-free steering towards any target state $|\psi_{\mathrm{targ}}\rangle$ can in principle be designed if we allow for an arbitrary coupling set. For that, consider a many-body unitary transformation to $|\psi_{\mathrm{targ}}\rangle$ from a product state $|00..0\rangle$, i.e., $|\psi_{\mathrm{targ}}\rangle = U_\psi |00..0\rangle$ (dropping subscript $_{\mathrm{targ}}$ for brevity). One may then formally construct a family of couplings:

$$V_s^{(U_\psi)}(i) = \gamma U_\psi \sigma_i^- U_\psi^\dagger. \tag{15}$$

Any protocol for $|\psi_{\mathrm{targ}}\rangle$ preparation using couplings of the form of Eq. (15) would be a unitary equivalent of the same protocol which uses couplings of Eq. (13) to prepare $|00..0\rangle$. Therefore, a passive protocol iterating over $V_s^{(U_\psi)}(i)$ for different $i$ would successfully prepare the target state $|\psi_{\mathrm{targ}}\rangle$. We also conclude that a partial-termination strategy can be applied to this coupling set with the same effect as for the product state target. Note, however, that in most cases employing $V_s^{(U_\psi)}(i)$ would not be practically feasible. Indeed, since $U_\psi$ is a general many-body operation, the couplings $V_s^{(U_\psi)}(i)$ would involve arbitrarily non-local terms. For most $N$-qubit states $|\psi_{\mathrm{targ}}\rangle$ with large $N$, one thus expects that the resulting $V_s^{(U_\psi)}(i)$ would break any requirement of geometric or $k$-locality.

This locality-violation rule can be circumvented for $U_\psi$ which is given by a shallow circuit and thus $|\psi_{\mathrm{targ}}\rangle$ which is weakly entangled. As a resourceful example of such $|\psi_{\mathrm{targ}}\rangle$, consider a graph state defined on a generic graph $G$ [37]:

$$|\psi_G\rangle = \left( \prod_{\substack{(j,k)\in \\ \mathrm{edges}(G)}} U_{(j,k)}^{(\mathrm{gr})} \right) \left( \frac{|0\rangle + |1\rangle}{\sqrt{2}} \right)^{\otimes N}, \tag{16}$$

$$U_{(a,b)}^{(\mathrm{gr})} = \exp\left( \mathrm{i}\pi |00\rangle\langle 00|_{a,b} \right), \tag{17}$$

in which case

$$U_\psi = \left( \prod_{\substack{(j,k)\in \\ \mathrm{edges}(G)}} U_{(j,k)}^{(\mathrm{gr})} \right) \left( \prod_{j\in \mathrm{qubits}} \exp\left( \mathrm{i}\frac{\pi}{4}\sigma_j^y \right) \right).$$

Since two-qubit rotations $U_{(j,k)}^{(gr)}$ all mutually commute, the coupling $V_s^{(U_\psi)}(i)$ acts only on spin $i$ and on the spins $j$ whose vertices share an edge with $i$ in the graph $G$. Therefore, this coupling is $(k+1)$-local if there are $k$ edges going out of vertex $i$. Moreover, $V_s^{(U_\psi)}(i)$ is also geometrically local, if the graph $G$ only connects the qubits which

are in geometric proximity. We conclude that for the graphs satisfying the above conditions, a realistic preparation of graph states with local frustration-free steering is possible. Such a protocol can be sped up in the same way it was possible for the product states – using active feedback via the partial-termination strategy.

For the perfectly mixed starting state, the partial-termination policy gives an *optimal* speed-up of a protocol driven by frustration-free couplings $V_s^{(U_\psi)}(i)$. Indeed, the protocols in question are then equivalent to an independent set of $N$ 1-qubit steering protocols (under the unitary transformation $U_\psi$). This picture, however, breaks down for a more general starting state. Let us first consider the trivial target $U_\psi = \mathbb{I}$, $|\psi_{\rm targ}\rangle = |00..0\rangle$, while the starting state is itself entangled (e.g. $\frac{1}{\sqrt{2}}(|00..0\rangle + |11..1\rangle)$). In this case, the click received from a single coupling $V_s(i)$ may imply that multiple couplings can be dropped from the applied sequence, and not just $V_s(i)$ itself. This would be more optimal than the partial termination strategy outlined above. The same picture extends to the more interesting case when the target state $|\psi_{\rm targ}\rangle$ is entangled itself, e.g. a graph state, while the starting state is a product state. Indeed, under the unitary mapping $U_\psi$ which takes an entangled state $|\psi_{\rm targ}\rangle$ to $|00..0\rangle$, the product starting state in turn becomes entangled. Hence, the previous reasoning applies and partial-termination would generally not be an optimal active policy in this situation (although a nontrivial speedup factor $f > 1$ is still guaranteed). To accelerate it further, one may apply one of the frustrated-coupling strategies outlined in the following sections.

### B. Frustrated system-detector couplings

Assuming locality of $\{V_s(\mathbf{p})\}$, for target states other than the product states and states prepared by a shallow circuit, we would generally need to consider noncommuting couplings. For such target states, the first question to tackle is how to design a family of local couplings $\{V_s(\mathbf{p})\}$ that are suitable for a passive protocol. In principle, this can be addressed on a case-by-case basis, tailoring a coupling set to a specific target state. (This approach will be used for the $W$-state preparation in Sec. V D.) However, this is not always a straightforward task. Therefore, it is interesting to know whether one can devise a general scheme to this end. For this, we propose an approach based on the parent-Hamiltonian construction.

The parent Hamiltonian $H_{\rm par}(\psi_{\rm targ})$ is built to have $|\psi_{\rm targ}\rangle$ as its ground state (we will only consider the nondegenerate case). Assuming that $|\psi_{\rm targ}\rangle$ hosts a limited amount of entanglement [38], $H_{\rm par}(\psi_{\rm targ})$ obeys the local projective form [39]:

$$H_{\rm par}(\psi_{\rm targ}) = \sum_j H_\psi^{(j)}. \tag{18}$$

Local projective form means that all terms $H_\psi^{(j)}$ are local in real space and that $H_\psi^{(j)}$ have $|\psi_{\rm targ}\rangle$ as their common ground state. The latter property holds even though $H_\psi^{(j)}$ in general don't commute with each other — this is possible due to the ground state degeneracy of $H_\psi^{(j)}$. In fact, the ground spaces of $H_\psi^{(j)}$ will be central to constructing our coupling family $\{V_s(\mathbf{p})\}$. For the term $H_\psi^{(j)}$ nontrivially acting on a collection of $m$ spins, denote its $m$-spin ground states as $|\phi_a^{(j)}\rangle$ and the excited states $|\theta_a^{(j)}\rangle$. Given these, we can construct the coupling operators of the following form:

$$V_s^j(\mathbf{w},\mathbf{v},\mathbf{u}) = \sum_{ab} w_{ab} \, |\phi_a^{(j)}\rangle \langle\theta_b^{(j)}| + \sum_{ab} v_{ab} \, |\theta_a^{(j)}\rangle \langle\theta_b^{(j)}| + \sum_{ab} u_{ab} \, |\phi_a^{(j)}\rangle \langle\phi_b^{(j)}| \, . \tag{19}$$

A particular example of this construction will be addressed in detail for the Matrix Product State target states in Sec. IV (see also Refs. [6, 22]).

For a generic (fixed) value of parameters $(\mathbf{w},\mathbf{v},\mathbf{u})$, a passive protocol driven by the couplings $V_s^j(\mathbf{w},\mathbf{v},\mathbf{u})$ does converge to $|\psi_{\rm targ}\rangle$. In particular, alternating different $j$ in $V_s^j(\mathbf{w},\mathbf{v},\mathbf{u})$ at consecutive protocol cycles allows to steer the system to $|\psi_{\rm targ}\rangle$ as the joint ground space of all couplings $H_\psi^{(j)}$ (see [6] for a related statement proven rigorously for AKLT model). As long as the conditions on $|\psi_{\rm targ}\rangle$ for locality of $H_{\rm par}(\psi_{\rm targ})$ are satisfied, this concludes the construction of an appropriate coupling set for $|\psi_{\rm targ}\rangle$.

Let us now consider an active-protocol construction. Unlike in the frustration-free protocol, the operators $V_s^j(\mathbf{w},\mathbf{v},\mathbf{u})$ for different values of $j$ generally don't commute. Therefore, the measurement outcome of steering by $V_s^j(\mathbf{w},\mathbf{v},\mathbf{u})$ impacts the outcomes of steering at locations close to $j$. As a result, the partial-termination strategy cannot be applied to this coupling set, as it assumes that the respective cycles of the protocol can be considered separately. Due to this difficulty, we classify steering with noncommuting couplings as frustrated [50]. The feedback strategy for frustrated steering should continuously coordinate the application of different couplings in the protocol. In a many-body context, this becomes a complicated navigation-type problem (cf. Ref. [25]). We devote the following two sections to the study of such possible Hilbert-space navigation policies.

## IV. QUANTUM COMPASS: COST-FUNCTION POLICIES

One way to enable the Hilbert-space navigation is to introduce a cost function $C(\rho)$, which is to be minimized in the protocol. The basic example would be the infidelity $C(\rho) = R(\rho, |\psi_{\rm targ}\rangle)$ of the system state $\rho$ to the target state $|\psi_{\rm targ}\rangle$, defined in Eq. (1). Achieving the

global minimum $R(\rho, |\psi_{\rm targ}\rangle) = 0$ of this cost function would be equivalent to preparing the target state. In general, to calculate $R(\rho, |\psi_{\rm targ}\rangle)$, one needs to know the state of the system $\rho$. This is, in principle, feasible, as we control the system evolution given all measurement outcomes and therefore can numerically simulate it in parallel to the experiment. However, the requirement of such a simulation being done in parallel to the experiment puts a restriction on the size of the system that one can work with. For now, we will accept this limitation; finding ways to mitigate it is among the worthwhile potential extensions of this work.

With a given cost function $C(\rho)$ at hand, we can use it to form the active decision for the coupling operator $V_s(\mathbf{p})$. The ultimate strategy is to pick $V_s(\mathbf{p})$ which brings the system to the global minimum $|\psi_{\rm targ}\rangle$ in the fastest expected time. For $C(\rho) = R(\rho, |\psi_{\rm targ}\rangle)$ this is equivalent to the ultimate strategy defined by dynamic programming [33], requiring unrealistic computation power. Instead, one can use a cheaper approach to cost function minimization – the "greedy strategy." Specifically, one can use $V_s(\mathbf{p})$ that yields the fastest expected reduction of the cost function in a single step of the evolution:

$$V_s^{(\rm greed)}(\mathbf{p}) = \text{argmin}_{V_s(\mathbf{p})} R[\Lambda_{V_s(\mathbf{p})}(\rho)], \tag{20}$$

where $\Lambda_{V_s(\mathbf{p})}(\rho)$ is defined in Eq. (9). If there are multiple minima, we will postulate that argmin returns a random representative among those. With only a small amount of computations needed to decide for the optimal next coupling $V_s^{(\rm greed)}(\mathbf{p})$, this greedy procedure allows us to avoid the complex long-term analysis of the protocol.

### A. Greedy steering: Matrix Product States

As one can see numerically from a direct implementation, the greedy minimization of the cost function can accelerate the state preparation by a large factor. To demonstrate this, we consider the Matrix Product State [40] (MPS) targets. We focus on spin-1 uniform MPS with bond dimension 2, which are defined as follows:

$$|\psi_{\rm MPS}(A)\rangle = \sum_{\alpha_k=1}^{3} \sum_{i_k=1}^{2} A^{\alpha_1}_{i_1 i_2} A^{\alpha_2}_{i_2 i_3} .. A^{\alpha_N}_{i_N i_1} |\alpha_1 \alpha_2 .. \alpha_N\rangle \,, \tag{21}$$

where $A$ in each factor is the same tensor. One example of MPS $|\psi_{\rm MPS}(A)\rangle$ that we consider is the Affleck-Lieb-Kennedy-Tasaki (AKLT) state, with:

$$A^1_{ij} = \begin{pmatrix} 0 & \sqrt{\frac{2}{3}} \\ 0 & 0 \end{pmatrix}_{ij}, \quad A^2_{ij} = \begin{pmatrix} -\sqrt{\frac{1}{3}} & 0 \\ 0 & \sqrt{\frac{1}{3}} \end{pmatrix}_{ij}, \tag{22}$$

$$A^3_{ij} = \begin{pmatrix} 0 & 0 \\ -\sqrt{\frac{2}{3}} & 0 \end{pmatrix}_{ij}. \tag{23}$$

To test the generality of our approach, we also consider random Matrix Product States as targets. In particular, the tensor $A$ is to be generated at random, from a uniform distribution $A^{\alpha_k}_{i_k i_{k+1}} \sim \mathcal{U}_{[0,1]}$ (subsequently normalizing the resulting state $|\psi_{\rm MPS}(A)\rangle$).

To design a measurement-driven preparation protocol for $|\psi_{\rm MPS}(A)\rangle$, note that a generic MPS of the form Eq. (21) admits a parent Hamiltonian with individual terms acting on pairs of sites:

$$H_{\rm par}(A) = \sum_k h^{k,k+1}, \quad N+1 \equiv 1, \tag{24}$$

$$h^{k,k+1} = \mathbb{I} - \sum_{i_k i_{k+2}} \mathbb{P}\Big(\sum_{i_{k+1}} A^{\alpha_k}_{i_k i_{k+1}} A^{\alpha_{k+1}}_{i_{k+1} i_{k+2}}\Big), \tag{25}$$

where $\mathbb{P}(v^{\alpha_k,\alpha_{k+1}})$ is the projector onto the state $\sum_{\alpha_k,\alpha_{k+1}} v^{\alpha_k,\alpha_{k+1}} |\alpha_k \alpha_{k+1}\rangle$. Defined as such, each term $h^{k,k+1}$ has $|\psi_{\rm MPS}(A)\rangle$ as its ground state. For the AKLT state, $H_{\rm par}(A)$ coincides with the AKLT Hamiltonian [34].

As discussed in Sec. III B, the parent Hamiltonian $H_{\rm par}(A)$ admits a set of couplings $V(\mathbf{p})$ defined from the local spectrum of $h^{k,k+1}$. Denoting the 4 ground states and 5 excited states of $h^{k,k+1}$ as $|\phi_a^{(j)}\rangle$ and $|\theta_b^{(k)}\rangle$ respectively, these couplings can be defined as in Eq. (19) – we will denote them as $V_s^{k,k+1}(\mathbf{w}, \mathbf{v}, \mathbf{u})$. In a passive protocol, we will cyclically alternate between different sites $k = 1, ..N$, while drawing $\mathbf{w}, \mathbf{v}$, and $\mathbf{u}$ at random for each link (each matrix element being a Gaussian random variable with zero mean and unit variance). We demonstrate numerically that this passive protocol guarantees preparation of the target state in a finite number of time-steps [51]. For an active feedback strategy to be used on top of this, we propose a greedy policy relative to $C(\rho) = R(\rho, |\psi_{\rm targ}\rangle)$ to select $\mathbf{w}, \mathbf{v}$, and $\mathbf{u}$. The key question is – does such an active policy yield any speedup relative to the passive protocol?

Our numerics demonstrates that MPS targets typically admit a significant value of speedup factor $f = \tau_{\rm av}^{(\rm pas)}/\tau_{\rm av}^{(\rm act)}$. For a particular case of AKLT target (Fig. 2), the speedup factor ranges between $f \simeq 10$ and $f \simeq 23$ for system sizes 3 to 6. For random MPS as target states, $f$ strongly fluctuates per $A$; to get an idea of the statistics of $f$ across the targets, we perform an extensive numerical study (Fig. 3). Among 820 random MPS targets sampled at $N = 5$ (Fig. 3a), the speedup ranges up to $f_{\rm max} \simeq 3400$, with the average value $f_{\rm av} \simeq 19$. In other places of the manuscript we refer to these as $f_{\rm max} \sim 10^3$ and $f_{\rm av} \sim 10^1$, to focus on the orders of magnitude and not the specific values. Such a significant speedup across many MPS targets underscores the potential of applying the cost function-based active policy to a wide variety of settings.

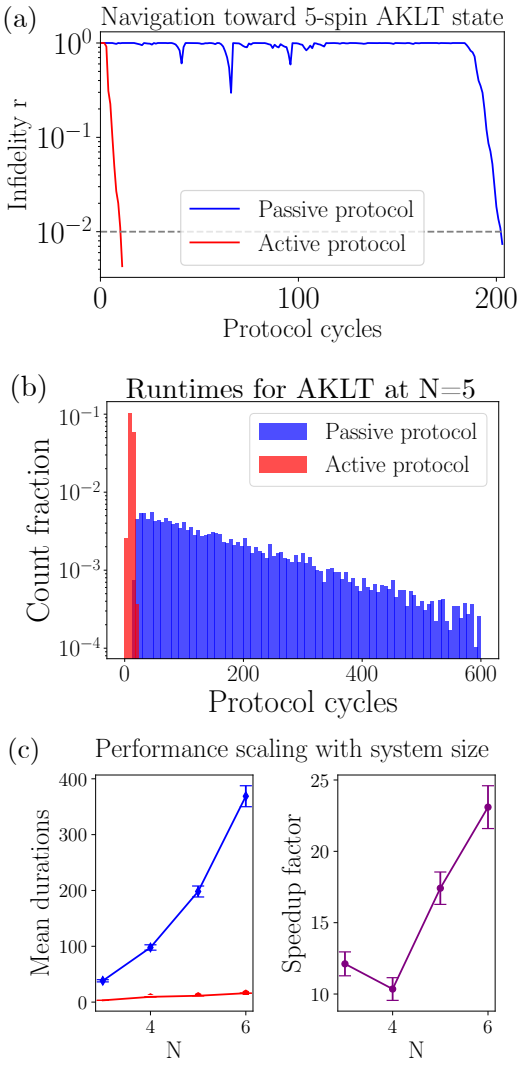


FIG. 2. Simulated measurement-driven preparation of the AKLT state up to infidelity $R < \epsilon = 0.01$. (a) Infidelity as a function of the protocol cycle for active and passive protocol runs towards a 5-spin AKLT state. These example runs are characterized by the runtimes similar to the average runtimes of respective protocols ($\tau_{\rm av}^{\rm (pas)} \simeq 200$, $\tau_{\rm av}^{\rm (act)} \simeq 11$). The passive protocol experiences setbacks in its performance at a few moments, with infidelity first decreasing and then resetting back to 1. The active protocol manages to avoid this issue. (b) Histograms of protocol runtimes $\tau$ for the five-spin AKLT state preparation. An exponential decaying profile, characteristic of a Poissonian process, can be clearly observed for the passive protocol (note the log scale). All recorded runs for an active protocol lasted far less than average passive runtime $\tau_{\rm av}^{\rm (pas)} \simeq 200$. Both histograms were compiled from $10^4$ simulated runs; the figure is truncated at 600 cycles for better presentation. (c) Scaling of the active protocol's advantage with system size $N$. The speedup factor $f$ tends to increase significantly as the system scales, with factor 23 being the estimated $f$ at 6 spins. The error bars represent 95% confidence intervals due to sampling error in numerical simulation. A sample of $10^3$ runs was collected to simulate the performance of both active and passive protocols at each system size.

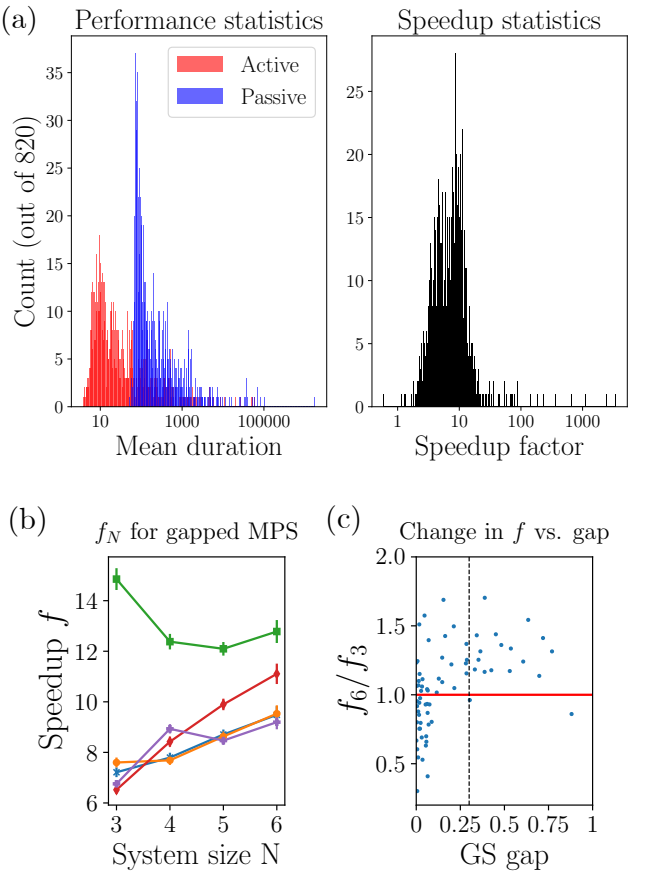


FIG. 3. Simulations of measurement-based preparation of random Matrix Product State with error $R < \epsilon = 0.01$. (a) Histograms of the average performance metrics for the preparation of the five-spin random Matrix Product State. Displayed are $\tau_{\rm av}^{\rm (pas)}$, $\tau_{\rm av}^{\rm (act)}$, and the resulting speedup $f = \tau_{\rm av}^{\rm (pas)}/\tau_{\rm av}^{\rm (act)}$. The results depend strongly per random instance of $A$ (note the log scale). Among the sampled instances, $f$ is greater than 1 for all instances but one; on the high end, $f$ ranges up to $f \simeq 3400$. Most counts are around $f \sim 10^1$, and the average value is $f \simeq 19$. Each histogram was compiled from $10^3$ simulated runs for 820 randomly generated random Matrix Product States (cf. main text). (b) Scaling of the active protocol's advantage with system size $N$ for target MPS with a large parent Hamiltonian gap. The target states are randomly generated while postselecting for the gap to be larger than a practical threshold value of 0.3 at $N = 5$. For most such targets, the speedup factor tends to increase significantly as the system scales. (cf. Fig. 2c for AKLT model, whose spectral gap at $N = 5$ is equal to 0.45.) Each data point is obtained via sampling over $10^4$ runs of the active and passive versions of the same protocol. Error bars display the respective 95% confidence interval. (c) Correlation between the speedup factor at $N = 3$ and $N = 6$ for random MPS targets with the respective parent Hamiltonian gaps. The vertical dotted line is put at the threshold gap value of 0.3. For most targets with a larger gap, the speedup $f$ increases with system size. Each gap value is calculated at $N = 5$.

An additional point of interest is the change of $f$ with the system size (Fig. 3b,c). Numerics demonstrates that $f$ may either tend to increase or decrease with the system size $N$, depending on the tensor $A$ of the target MPS. In particular, the direction of this tendency is highly correlated with the ground state gap of the respective parent Hamiltonian, as demonstrated in Fig. 3c for randomly generated $A$ tensors. For a large enough gap, the clear majority of targets displays an increase in speedup with system size. This is further illustrated Fig. 3b, which focuses on $f$ as a function of $N$ for targets with a large gap in respective parent Hamiltonians. The growth of $f$ with $N$ for such a broad family of targets is another compelling feature of the active policy presented here.

### B. Landscape exploration and alternative cost functions

The greedy approach defined above does not suffer from the presence of local minima in the cost-function landscape, which are typical obstacles in optimization procedures. In our approach, the target state is a dark state for each steering operator. Therefore the infidelity of the running state to the target state never increases on average. This fact and the convergence of the respective passive protocol guarantee the convergence of the greedy optimization protocol. On the other hand, there is no formal guarantee that our active protocol yields a speedup compared to the passive one. Instead, we note that the speedup factor $f > 1$ is achieved *in practice* for most examples that we tested. This includes the AKLT state targets, as well as 819 out of 820 random MPS targets that we considered (Fig. 3a). Obtaining formal guarantees may also be possible under some conditions on a target state; this could be an interesting direction for future investigation.

Despite the absence of "glassy" landscape, our greedy strategy still harbors a potential challenge. For the greedy procedure to be effective, it should always yield a nonzero bias in favor of a specific $V_s^{(\mathrm{greed})}(\mathbf{p})$ (or a small subset thereof). In other words, the landscape of the given cost function $C(\rho)$ should not be flat. It follows that applying the infidelity measure $R(\rho, |\psi_{\mathrm{targ}}\rangle)$ is in general a flawed strategy. The reason is that a $(2^N-1)$-dimensional subspace of states in the $N$-body Hilbert space is orthogonal to the target state. Consider the case when the starting state belongs to that subspace. The state would in general stay in this subspace after a single steering cycle with a local coupling $V_s(\mathbf{p})$. For our purposes, it implies that the infidelity measure $R$ is equal to 1 for a large manifold of states, and there might be no direction of increase that would allow us to choose an appropriate coupling. In this scenario, the active steering is effectively reduced to a passive one (with no preference with respect to choosing a particular coupling). Since we assume the convergence of the passive protocol, such ("local") flatness scenarios do not disable the convergence of the active protocol. However, this might still diminish the resulting speedup factor $f$.

The most direct example of this effect of flat landscapes can be observed when applying the greedy policy to frustration-free steering (see Sec. III A). For simplicity, let us again take the product state of $N$ qubits $|00..0\rangle$ as the target state, the state $|11..1\rangle$ as the starting state, and the couplings $V(i) = \sigma_i^-$ for steering. Only after such steering protocol results in $N$ successful click events, $R(\rho, |\psi_{\mathrm{targ}}\rangle)$ gains a nonzero value. Thus before $N-1$ clicks, the greedy policy for $R^{(\mathrm{inf})}$ will not be capable of providing a biased decision for the next coupling. Strongly enhanced by the system size, this phenomenon is reminiscent of Anderson's orthogonality catastrophe [41].

As a remedy to this deficiency, the full target-state fidelity can be replaced with its more localized versions. For example, a "subsystem infidelity" measure can be introduced:

$$R_{\mathcal{S}}(\rho, |\psi_{\mathrm{targ}}\rangle) = \sum_{\sigma\in\mathcal{S}} \left[1 - \mathrm{tr}\left(\sqrt{\sqrt{\rho_{\mathrm{targ},\sigma}}\rho_\sigma\sqrt{\rho_{\mathrm{targ},\sigma}}}\right)^2\right], \tag{26}$$

where $\rho_{\mathrm{targ},\sigma}$ ($\rho_\sigma$) is the reduced density matrix of the target state (current state) with respect to subsystem $\sigma$. $\mathcal{S}$ is the family of subsystems from which $\sigma$ are drawn; the choice of $\mathcal{S}$ depends per target state. In the case of the $|11..1\rangle \to |00..0\rangle$ protocol described above, the appropriate $\mathcal{S}$ would be the set of individual spins. Unlike $R$, such quantity $R_{\mathcal{S}}$ changes every time when a click occurs in this protocol. As a result, the greedy policy with respect to the local $R_{\mathcal{S}}$ would yield the partial-termination protocol of Sec. III A, significantly boosting the preparation of such a product state.

By continuity with the case of the product state target, such preference for $R_{\mathcal{S}}$ should extend to the weakly-entangled target states, and maybe to some highly-entangled targets. However, we did not see a manifestation of this in our MPS simulations, where using $R_{\mathcal{S}}$ as a cost function did not yield any improvement compared to $R$. As a likely explanation for this, the orthogonality catastrophe should become manifest only at large system sizes, where the classical simulation of the protocol is also hindered. However, we expect that some practical target states may still develop a noticeable performance difference between $R_{\mathcal{S}}$ and $R$, similarly to the case of the product state target. A further study of this question constitutes a promising direction for future work.

In addition to modifying the cost function, another way to fix the landscape flatness issue is to move away from the cost-function based policies entirely. This is one of the key motivations for an alternative (Quantum State Machine) framework we introduce in the next Section. Such an alternative framework can indeed outperform the cost function-based policy due to the landscape flatness issue the latter occasionally experiences. An example of this is shown in Sec. V D, where the measurement-based preparation of the W-state is investigated.

### C. Role of measurement imperfections

The above greedy policy is formulated for an ideal case of perfect detectors. Reducing "detection efficiency" (say, recording a click readout instead of an actual no-click one) in the active protocol could eventually reduce it to a passive one. Indeed, the choice of the further couplings will be based on a wrong position in the cost-function landscape, and, hence, might become completely random with respect to the actual landscape. Nevertheless, in our work, the set of couplings still guarantees successful passive steering. Therefore, the speedup factor for the active protocol remains generically higher than 1, even in the presence of such errors.

One can roughly describe the crossover in the speedup factor between the ideal and "imperfect" active protocols by introducing a probability of a "failed measurement" and assuming a fully passive steering after that particular measurement step. Given a typical value of the error ("measurement imperfection") time $\tau^{(\mathrm{err})}$ for switching from active to passive protocol, one can estimate the "failed" speedup factor as follows. For simplicity, we assume that the speedup factor in the active protocol is time-independent (for sufficiently long times), i.e., the distance to the target state for a given time $\tau$ in the active protocol is the same as the distance for time $f\,\tau$ in the passive one. This means that, for an arbitrary initial state and the target infidelity, the typical active-steering time, $\tau^{(\mathrm{act})}$, and passive-steering time, $\tau^{(\mathrm{pas})}$, are related by the same "ideal" speedup factor $f$. This assumption is in a good agreement with our numerical results, where the speedup factor is roughly independent of the target precision. Let us further specify that, for the passive protocol, the distance to the target state depends on time exponentially (this is the case when passive steering is described by a gapped Lindbladian, as, e.g., in the case of the AKLT model [22]). Then, the total time to reach the target state for such corrupted active steering is given by $\tilde{\tau}^{(\mathrm{act})} = \tau^{(\mathrm{pas})} - (f-1)\tau^{(\mathrm{err})}$ (with $\tau^{(\mathrm{err})} < \tau^{(\mathrm{act})}$, otherwise active steering is essentially unaffected by errors). The effective speedup factor thus becomes

$$\tilde{f} = \frac{\tau^{(\mathrm{pas})}}{\tilde{\tau}^{(\mathrm{act})}} = \frac{f\tau^{(\mathrm{act})}}{f\tau^{(\mathrm{act})} - (f-1)\tau^{(\mathrm{err})}} > 1.$$

Note that this speedup factor now depends on the target precision through $\tau^{(\mathrm{act})}$.

One can, in principle, refine this estimate, by using the full distribution function of the runtimes and error times, or by relaxing the assumption of time-inhomogeneity of the speedup factor. Further, it is interesting to study a host of non-fatal errors, when the location of the system in Hilbert space is only slightly blurred. In addition, it is noteworthy that possible imperfections in active protocols can be monitored given full information on the readouts, and the corresponding error-correction strategies can be designed. We relegate this and related questions to future work.

## V. HILBERT-SPACE ORIENTEERING MAP: QUANTUM STATE MACHINE

In this section, we present an orienteering tool that is an alternative to cost-function minimization: mapping out the steering transformations with a Quantum State Machine (QSM) construction. We then illustrate navigation in many-body Hilbert space, employing this machinery to the preparation of the highly entangled W-state of three qubits.

### A. QSM generalities

Consider the transformation of the system's state, $\Lambda^{(\mathrm{cl})}_{V_s}$ and $\Lambda^{(\mathrm{ncl})}_{V_s}$, associated to steering with a specific coupling $V_s$ in a given readout scenario (click or no-click, respectively, see Eqs. (10) and (11)). Every such steering transformation conserves the purity of the state. Therefore, it is convenient to encode transformations $\Lambda^{(\mathrm{cl,\ ncl})}_{V_s}$ in their action on Hilbert space basis states $|\phi_\alpha\rangle$:

$$\Lambda^{(\mathrm{cl,ncl})}_{V_s}(|\phi_\alpha\rangle) = \frac{1}{\sqrt{p^{(\mathrm{cl,ncl})}}} \sum_\beta L^{(\mathrm{cl,ncl})}_{\alpha\beta} |\phi_\beta\rangle \tag{27}$$

$$L^{(\mathrm{cl})}_{\alpha\beta} = \langle\phi_\beta|\, \delta t V_s \,|\phi_\alpha\rangle\,, \tag{28}$$

$$L^{(\mathrm{ncl})}_{\alpha\beta} = \langle\phi_\beta|\, 1 - \delta t^2 V_s^\dagger V_s/2 \,|\phi_\alpha\rangle\,, \tag{29}$$

where $p^{(\mathrm{cl})}$ ($p^{(\mathrm{ncl})}$) is the probability of a click (non-click) readout upon this steering action. Note that in Eq. (27), we extended the action of $\Lambda_{V_s}$ to pure states by a slight abuse of notation compared to Eq. (10).

Amplitudes $L^{(\mathrm{cl,ncl})}_{\alpha\beta}$ allow to represent the steering action $\Lambda^{(\mathrm{cl,ncl})}_{V_s}$ as a graph. The vertices in such a *steering graph* correspond to the Hilbert space basis states, and the edges describe the steering transformations. The edges are directed and weighted with complex amplitudes. Specifically, an edge $\alpha \to \beta$ is to be weighted with amplitude $L^{(\mathrm{cl,\ ncl})}_{\alpha\beta}$ (edges weighted with zero amplitudes are excluded from the graph). Implying this definition, we will use the notation $L^{(\mathrm{cl,\ ncl})}$ for the steering graphs themselves. For basic examples of steering graphs, please refer to Fig. 4.

Since the weights $L^{(\mathrm{cl})}_{\alpha\beta}$ are proportional to the matrix elements of coupling operator $V_s$ while $L^{(\mathrm{ncl})}_{\alpha\beta}$ can be expressed via $V_s$ as well, the graph $L^{(\mathrm{ncl})}$ for the no-click action can be inferred entirely from the graph $L^{(\mathrm{cl})}$ for the click action. In particular, due to the term $\propto V_s^\dagger V_s$, graph $L^{(\mathrm{ncl})}$ contains an edge $e^{(\mathrm{ncl})}_{\alpha\beta}$ from vertex $\alpha$ to $\beta$, if a graph $L^{(\mathrm{cl})}$ contains edges $e^{(\mathrm{cl})}_{\alpha\mu}$ and $e^{(\mathrm{cl})}_{\beta\mu}$ for any vertex $\mu$ (see Fig. 4). Heuristically speaking, to yield a $L^{(\mathrm{ncl})}$-edge, one has to first follow a $L^{(\mathrm{cl})}$-edge forward, and then another $L^{(\mathrm{cl})}$-edge backward. Furthermore, due to the additional identity operator term in Eq. (29), any

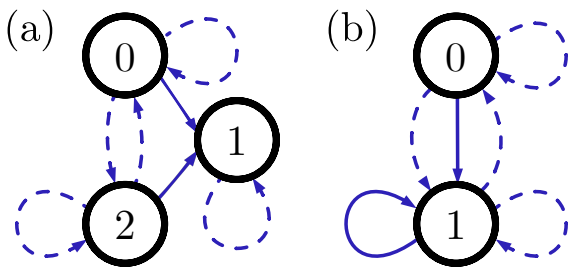


FIG. 4. Examples of steering graphs (see definition in Sec. V A): (a) Steering graphs on a 3-level system, corresponding to the coupling $V_s = \gamma(|1\rangle\langle 0| + |1\rangle\langle 2|)$. Graph $L^{(\mathrm{cl})}$ for click action is depicted with solid arrows and the graph $L^{(\mathrm{ncl})}$ for no-click action with dashed arrows. Due to the identity operator in Eq. (29), every vertex is decorated with a self-loop from the $L^{(\mathrm{ncl})}$ graph. To see how the rest of $L^{(\mathrm{ncl})}$ can be deduced from $L^{(\mathrm{cl})}$, consider the example of $e_{02}^{(ncl)}$ (dashed arrow from state 0 to 2). According to the graphical approach from Sec. V A, one is to follow edge $e_{01}^{(cl)}$ (solid arrow from 0 to 1) forward and then $e_{21}^{(cl)}$ (solid arrow from 2 to 1) backward - and thus manages to travel from state 0 to 2, in correspondence to $e_{02}^{(ncl)}$. (b) Steering graphs on a 2-level system, as defined by the coupling $V_s = \gamma(|1\rangle\langle 1| + |1\rangle\langle 0|)$. Following the same rule as above, inter-vertex edges of $L^{(\mathrm{ncl})}$ can be deduced from $L^{(\mathrm{cl})}$. For example, by following the edge $e_{11}^{(\mathrm{cl})}$ forward and then the edge $e_{01}^{(\mathrm{cl})}$ backward, one performs a transition from state 1 to state 0, thus reproducing the edge $e_{10}^{(\mathrm{ncl})}$ from $L^{(\mathrm{ncl})}$.

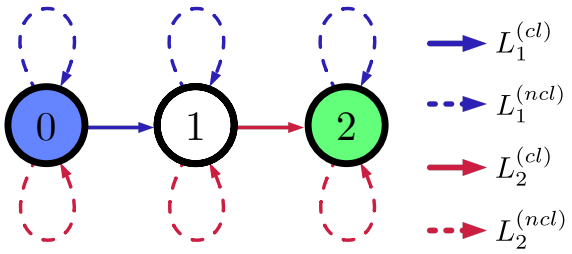


FIG. 5. A basic example of the QSM multigraph, describing the available coupling family for a three-state system. The steering options are represented by the coupling operators $V_1 = \gamma_1 |1\rangle\langle 0|$ and $V_2 = \gamma_2 |2\rangle\langle 1|$. The starting state is 0, marked in blue, and the target state is 2, marked in green. The optimal coordination policy of the two steering operations is straightforward: one needs to first repeatedly apply the $V_1$-steering until a click is obtained, and then the $V_2$-steering until a click is obtained. Compared to the passive steering which iterates between $V_1$ and $V_2$ regardless of measurement outcomes, this directly yields a 2-fold speedup in the average performance.

graph for the no-click steering action will also include self-loops on each vertex.

The steering graphs introduced above can now be used to create a Quantum State Machine. For $n_V$ elements $V_s(\mathbf{p})$ in the available coupling family, there exist $2n_V$ graphs corresponding to steering maps $\Lambda_{V_s(\mathbf{p})}^{(\mathrm{cl,\ ncl})}$, because of the two possible measurement outcomes for each of the couplings. The QSM for the steering protocol is then obtained as a collection of these graphs. It can be represented as a colored multigraph, where each steering graph is represented as a single-color subgraph (Fig. 5). Consequently, in a QSM multigraph there may be multiple edges going from any vertex $\alpha$ into any other vertex $\beta$ (making it a multigraph rather than a simple graph), but at most one such edge for each color.

Let us now consider our original task of finding the accelerated navigation protocol. To make use of the QSM construction in this context, we will restrict our consideration to bases $\{|\phi_\beta\rangle\}$ where one of the basis states is the target state $|\psi_{\mathrm{targ}}\rangle$ itself. In such a case, state $|\psi_{\mathrm{targ}}\rangle$ corresponds to a marked vertex in the graph, and the goal of the steering protocol becomes to drive the system state to that vertex. The goal of optimizing this protocol may then look similar to a known problem of finding the shortest path to the marked vertex on a weighted graph. This problem is standard in graph theory and can be solved as such. Can such a solution be used to design the navigation protocol?

As we will see in Sec. V B, this analogy is not complete, since the quantum evolution on the graph goes beyond the classical path-on-the-graph picture. This aspect creates an obstacle to directly applying the graph exploration algorithms to facilitate our protocol speed-up. Fortunately, this difficulty can be properly accounted for in some cases, as we will see in Sec. V C. In those cases, the "semi-classical heuristics" of graph exploration may indeed be applied. Finally, in Sec. V D, we will apply this approach to actively prepare the $W$-state, with a factor $f = 3.5$ improvement compared to the passive protocol.

### B. Quantum subgraphs in a QSM

Let us now compare our QSM navigation task to the standard problem of graph exploration. Our goal is to identify the differences between the two, which prevent us from applying the graph exploration techniques directly to QSM navigation. First of all, the state of the system in graph exploration is at all times represented by a single vertex. The system in a QSM, on the other hand, is generally represented by a superposition over multiple vertices. Furthermore, in graph exploration, the state is modified by following one of the edges. A steering action in a QSM, in contrast, corresponds to a whole collection of edges – i.e., a single-colored steering graph in the QSM multigraph.

Some steering graphs may induce quantum effects, such as superposition and interference. For instance, the steering action whose graph contains two outgoing edges from a given vertex (e.g., vertex 0 for graph $L_1^{(\mathrm{cl})}$ in Fig. 6a), can create a nontrivial quantum superposition. If a state is given by a superposition of multiple

vertex states, it may further undergo quantum interference. In particular, this can be facilitated by a steering action whose graph contains a vertex with two incoming edges (e.g., vertex 4 for graph $L_2^{(\mathrm{cl})}$ in Fig. 6a). In general, a notion of "superposition subgraphs" and "interference subgraphs" of a steering graph can be defined:

1. Superposition subgraph is a subgraph of a steering graph span by multiple (more than one) edges outgoing from a single vertex.

2. Interference subgraph is a subgraph of a steering graph span by multiple edges incoming to a single vertex.

Collectively, we will refer to such interference and superposition subgraphs of a single steering graph as its quantum subgraphs. If the quantum subgraphs are absent in the QSM, we will refer to it as a classical QSM. In other words, in a classical QSM, each vertex has at most one outgoing and at most one incoming edge of any given color.

If a QSM is classical, optimization of the navigation protocol can essentially be reduced to classical graph exploration. For a simple example of a classical QSM and the way to optimize the respective state preparation, consider the 3-level steering actions described in Fig. 5. Note that optimization of the classical QSM also applies to the case when the starting state is a superposition of multiple vertex states. If the steering operations contain no quantum subgraphs, the quantum superposition is equivalent to a probabilistic mixture for the sake of the protocol optimization, and the optimal navigation pattern can be extracted accordingly.

As the form of the steering graph depends on the choice of basis, it is conceivable that the number of quantum subgraphs in such a graph in some cases can be reduced by changing the basis (compare Fig. 6a and b). However, using a change of basis to remove all the quantum subgraphs in an arbitrary QSM is generally impossible (see Fig. 6).

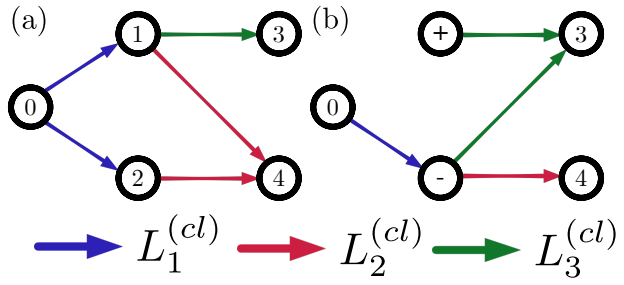


FIG. 6. Possible configurations of quantum subgraphs in a QSM, exemplified by the 5-vertex subgraph of a hypothetical QSM. (a) The click-action graphs for the three operators $V_{1,2,3}$ that form the family of couplings. The operators have the form $V_1 = \gamma_1(|1\rangle - |2\rangle)\langle 0|$, $V_2 = \gamma_2 |4\rangle(\langle 1| - \langle 2|)$, $V_3 = \gamma_3 |3\rangle\langle 1|$. The graphs for the no-click actions are not shown, as their form can be deduced from the graphs for click actions. In the present basis, the $V_1$-click is manifest as a superposition, the $V_2$-click – as an interference, and the $V_3$-click corresponds to a semiclassical evolution. (b) Quantum State Machine for the coupling family from the previous panel, depicted in a different basis. The basis transformation is $|\pm\rangle = (|1\rangle \pm |2\rangle)/\sqrt{2}$. In this case, the basis transformation removes the quantum elements in the $L_{1,2}^{(cl)}$ graphs, however, it turns $L_3^{(cl)}$ into an interference element. Note that there is no basis transformation that would turn such a QSM into a classical one. This statement follows from the uniqueness of the Jordan canonical form for operators $V_2$ and $V_3$.

### C. Coarse-grained QSM. Semiclassical heuristic for navigation

We now focus on the steering protocols whose QSM cannot be made classical via a basis transformation. In such a case, it may still be possible to optimize it via a classical graph exploration heuristic. For that, we propose to coarse-grain the QSM by grouping subsets of its vertices into single block-vertices. The coarse-grained QSM would consist of graphs drawn between such block-vertices. The block-vertex containing the target vertex can be considered as the target block-vertex.

An inter-block edge between two block-vertices is drawn, if the original QSM has at least one edge connecting the vertices inside the respective block-vertices. For the coarse-graining to be useful for our purposes, it should be done in such a way that all of the resulting QSM graphs have a classical structure. Namely, the coarse-grained graph should not have quantum subgraphs, e.g. realizing superposition or interference between the block-vertices (in analogy to Sec. V B). To satisfy this requirement, the following rule for vertex grouping can be employed (cf. Fig. 7): *if two edges of the same color are simultaneously coming in or out of a given vertex, the two vertices at the other ends of these edges should be grouped within one effective block-vertex.* This rule manifestly yields basis-dependent groupings, since the very presence of quantum subgraphs in a QSM is basis-dependent. Thus, a smart choice of the basis may allow for an efficient and simpler coarse-grained graph. Designing a general explicit algorithm for finding the optimum basis for an arbitrary QSM is a highly nontrivial task. Heuristically speaking, a convenient choice of the basis should be the one that results in the minimum number of quantum subgraphs in a QSM before coarse-graining.

For the coarse-grained graph to be effectively classical, we desire to ignore details of the system evolution inside the subspace of a given block-vertex. Specifically, we aim to view every block-vertex as an effective single state of the system and assume that every edge allows transporting the system between such block-vertex states with no obstacles. If this was directly possible, and since the coarse-grained QSM by definition contains no quantum subgraphs, optimization of its exploration

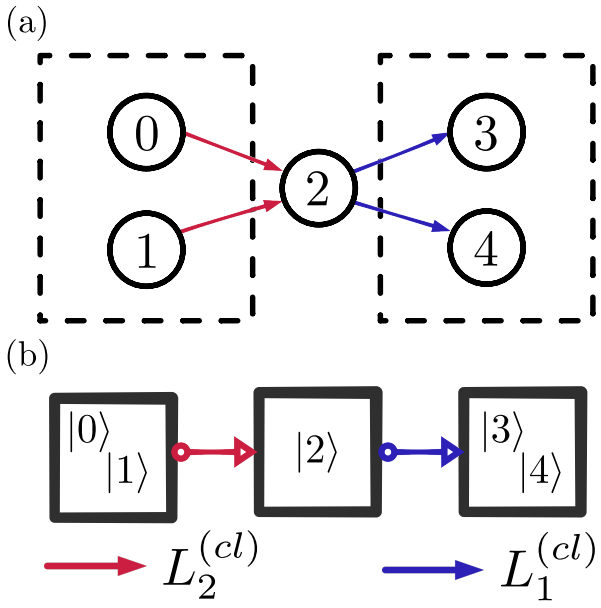


FIG. 7. Semiclassical coarse-graining applied to a QSM. (a) A 5-state part of a QSM with two quantum subgraphs: interference subgraph realized by $L_2^{(cl)}$ and a superposition subgraph realized by $L_1^{(cl)}$. Since pairs of states $\{|0\rangle, |1\rangle\}$ and $\{|3\rangle, |4\rangle\}$ fall under conditions described in Sec. V C, these are to be grouped together in a coarse-grained QSM. (b) Simplified depiction of a coarse-grained QSM, obtained from (a).

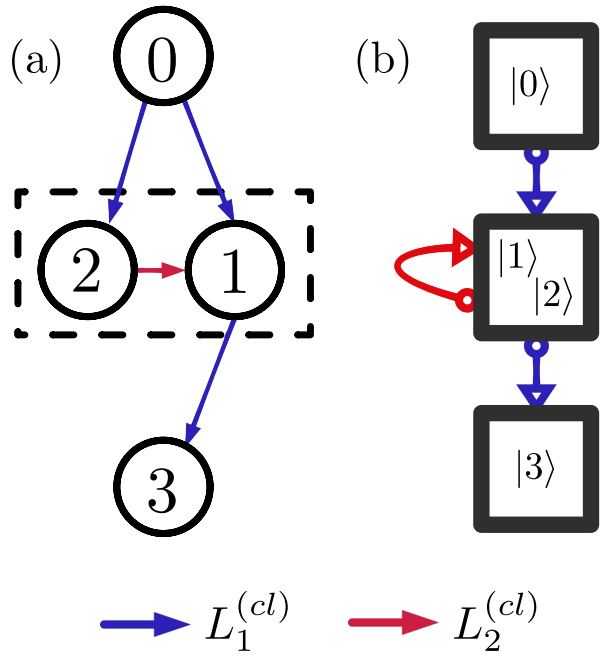


FIG. 8. Illustration of ancillary couplings in the context of QSM coarse-graining (a) A 4-state part of a QSM that is subject to coarse-graining, featuring non-trivial actions by couplings denoted as $V_1$ and $V_2$. States $|1\rangle$ and $|2\rangle$ are to be grouped together since they are both targets in a superposition subgraph. (b) The coarse-grained version of the same QSM. The block $\{|1\rangle, |2\rangle\}$ is connected to state $|3\rangle$ through an outgoing edge of $L_1^{(\mathrm{cl})}$. However, from microscopic point of view exemplified in the first panel, no population can be transferred from state $|2\rangle$ to $|3\rangle$ unless the click action $\Lambda_2^{(\mathrm{cl})}$ is realized first. Therefore, including and applying $V_2$ as an ancillary coupling is required for a valid semiclassical coarse-graining of this QSM.

would have become a classical task. However, such an approximation scheme needs more careful justification. Every block fundamentally corresponds to a Hilbert subspace, and an inter-block edge is given by a $D_1 \times D_2$ matrix of coefficients (where $D_1$ and $D_2$ are the internal dimensionalities of the linked blocks). Characterizing these effectively with single amplitudes may lead to erroneous navigation policies. In particular, one state internal to a block-vertex might be untouched by an inter-block edge, i.e., it only yields zero matrix elements in a matrix characterizing the edge. If the edge is outgoing, a system initialized in the said state would not be able to escape the block-vertex using that edge alone (see Fig. 8). This is in direct conflict with characterizing blocks and inter-block edges with single amplitudes. For an incoming edge, a similar problem may arise: some states inside a block-vertex might not get populated when that edge is activated. This may become detrimental for the navigation protocol based on a coarse-grained QSM, especially if the unavailable state in question is the final target of the protocol.

Such difficulties may be overcome, if some of the couplings given in a QSM allow for an internal mixing of the subspace (represented by a self-loop on the block-vertex in the respective $L^{(\mathrm{cl})}$-graph). Applying such a coupling in the protocol would allow one to make the block-vertex accessible to all the edges that are connected to it (see Fig. 8), via a sufficient number of clicks. In the scenarios described above, where additional couplings are needed to turn a block-vertex into an effective single vertex, we will refer to such couplings as ancillary couplings. Note that given a coupling family, there is no guarantee that the ancillary couplings needed for the exploration of every block-vertex, are available. In this work, we restrict our further consideration to the coarse-grained QSMs, where the ancillary couplings happen to be present wherever needed. Every block-vertex can then be made accessible to the outgoing edges, and the target state is ensured to be reachable once the target block is reached. In this case, we consider the coarse-grained QSM as effectively semiclassical.

To design an active steering policy within the coarse-grained approach, we note that the navigation protocol has the following structure. The system state can be transported between block-vertices, and eventually steered to the target block-vertex. After that, either the target state is reached already (one can obtain this information from the simulated copy of the system), or it can be reached after applying ancillary couplings on the

target block-vertex. The cost of the protocol can now be broken into two parts. The first is the cost of exploring the coarse-grained graph using the inter-vertex edges. The second is the dwell time inside the block-vertices, which is spent applying the ancillary couplings. If we could find the route through the graph that minimizes the combination of these two components, it would solve our optimization problem exactly. There is a problem, however: both the inter-vertex travel time and the block-vertex dwell time depend on the microscopic details of the evolution internal to the block-vertices. The coarse-grained geometrical information would therefore not suffice to *exactly* derive the optimal policy. On the other hand, fully accounting for quantum-mechanical microscopics is prohibitively hard. Instead, we will use the semiclassical QSM to obtain an *approximately* optimal active policy.

Let us assign every inter-block edge a characteristic traversal time, and every block-vertex a characteristic dwell time. For this, we use the matrices for click transitions between blocks $i$ and $j$ (the case of ancillary couplings given by $i = j$). Denote these as $L^{(\mathrm{cl})}_{i,\alpha;j,\beta}$, implying that only matrix elements with states from blocks $i$ and $j$ are included. In that case, the effective transition amplitude between blocks $i$ and $j$ can be defined as operator norm $L^{(\mathrm{cl})}_{i,j} = \|L^{(\mathrm{cl})}_{i,\alpha;j,\beta}\|$, and characteristic traversal (dwell if $i = j$) time $\Delta\tau_{i,j} = (L^{(\mathrm{cl})}_{i,j})^{-2}$. This reduces to the average traversal time for the case of a genuinely classical graph, with an amplitude $\gamma\delta t$ connecting two states implying duration of $\Delta\tau = (\gamma\delta t)^{-2}$ for traversal (cf. Sec. II B).

With characteristic times $\Delta\tau_{i,j}$ assigned, the time-cost of following a specific path through this graph can be estimated as a combined characteristic time of all the edges and vertices crossed along the way. The desired path will be the one that optimizes this expected time. As previously discussed, this semiclassical calculation may not yield an exactly optimal navigation policy. However, in many practical cases such an active protocol should still be quicker compared to its completely passive version. One example of such a practical improvement is given below.

### D. W-state preparation

To illustrate the principles of the QSM framework, we consider the coarse-graining approach to the navigation of a 3-qubit system from $|111\rangle$ to a so-called W-state [42]. This state has the following form:

$$W = \frac{1}{\sqrt{3}}(|100\rangle + |010\rangle + |001\rangle). \tag{30}$$

To define the measurement-based protocols, we choose the following 2-local family of couplings (assuming labels A, B, and C for the qubits):

$$V_1 = \sigma^-_A - \sigma^-_B, \tag{31}$$
$$V_2 = \sigma^+_A\sigma^+_B - \sigma^+_B\sigma^+_C, \tag{32}$$
$$V_3 = \sigma^-_A\sigma^+_B - P^0_A P^1_B, \tag{33}$$
$$V_4 = \sigma^+_B\sigma^-_C - P^1_B P^0_C. \tag{34}$$

Here, $\sigma^{\pm} = \frac{1}{2}(\sigma^x \pm i\sigma^y)$ and $P^a = |a\rangle\langle a|$, $a = 0, 1$. A passive version of the protocol would amount to blindly alternating between the steering actions with different $V_i$. This does yield the target state if the steering is applied a sufficient number of times (Fig. 9.b). A cost-function based *active* policy (Sec. IV) can also be introduced, greedily choosing between $V_i$ based on the expected gain in the target state fidelity.

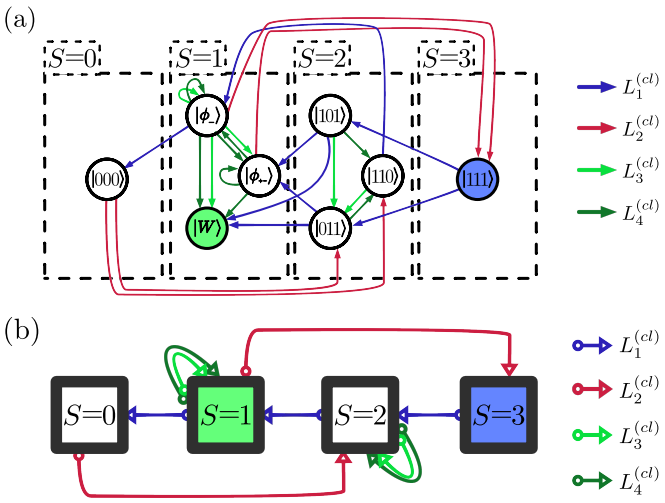


FIG. 9. Measurement-driven navigation towards the 3-qubit $W$-state: QSM representation. (a) Steering with couplings Eqs. (31)-(34). The vertices in the single-excitation subspace are given by states $|W\rangle$, $|\phi_-\rangle \equiv \frac{1}{\sqrt{2}}(|100\rangle - |001\rangle)$, and $|\phi_{+-}\rangle \equiv \frac{1}{\sqrt{6}}(|100\rangle - 2\,|010\rangle + |001\rangle)$. (b) The coarse-grained version of the above QSM. The vertices are labeled by the excitation number. From perspective of Sec. V C, couplings 2 and 3 play the ancillary role. Indeed, those couplings mix the internal structure of the block-vertices, allowing one to eventually steer the state to the target $|W\rangle$.

To design a QSM-based active policy, consider a multi-graph representation of the coupling family. It is shown in Fig. 9a. Note that this QSM has multiple quantum subgraphs. Therefore, to employ a feedback policy, it should be subjected to the coarse-graining technique as outlined in Sec. V C. It proves useful to coarse-grain the Hilbert space by the total excitation number $S \in \{0, 1, 2, 3\}$ – this results in a semiclassical QSM, as desired (Fig. 9b). Given the all-up starting state of the evolution, one designs an active policy that leads to the target state in a classically optimal way:

1. Repeat $V_1$-steering until a click is obtained;
2. Repeat $V_1$-steering until another click is obtained;

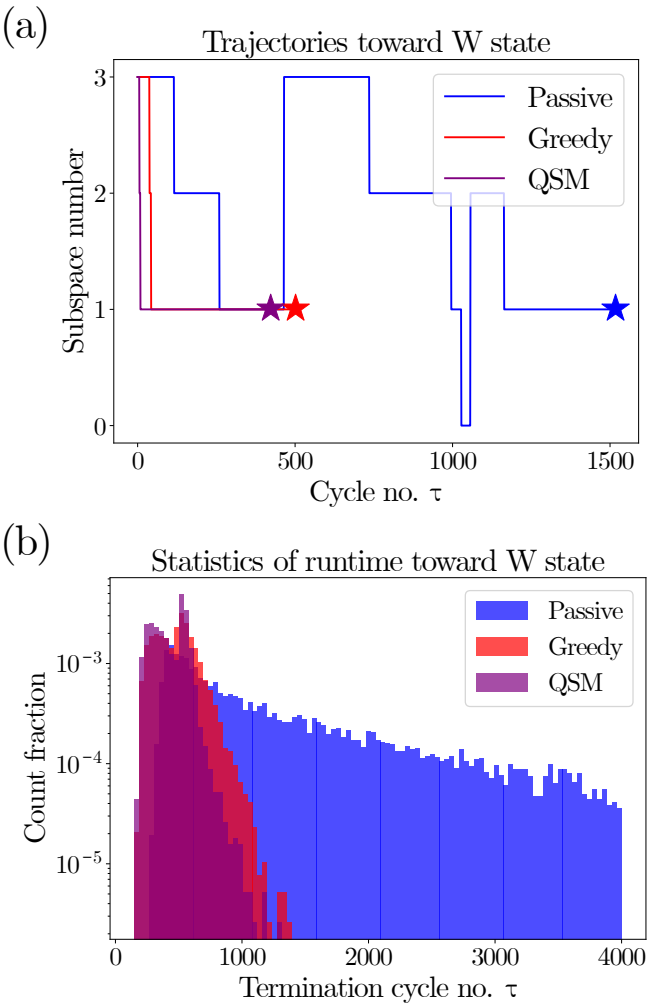


FIG. 10. Performance of the QSM-based protocol compared to its passive and greedy counterparts. The protocols are simulated with $\delta t = 0.1$ and the target precision $\epsilon = 0.01$. (a) Typical trajectories of the passive and active protocols, across the excitation number sectors $S$ occupied by the running state. Displayed are trajectories that yield the runtimes approximately equal to average runtimes $\tau_{\rm av}^{(\rm QSM)} \simeq 430$, $\tau_{\rm av}^{(\rm pas)} \simeq 1500$, and $\tau_{\rm av}^{(\rm greedy)} \simeq 490$. Passive protocol switches between different $S$ multiple times before eventually reaching the target. The greedy protocol behaves similarly to QSM-based one, but stays longer in the $S = 3$ state, accounting for its relative slowdown. (b) The logarithmic histogram over the protocol runtimes. It can be seen that the QSM-based protocol consistently outperforms the passive protocol and often outperforms the greedy one. Each histogram was obtained from $10^4$ numerical simulations, and truncated at $\tau = 4000$ for better presentation.

3. Alternate $V_3$- and $V_4$-steering until the target state is reached (with fidelity error below $\epsilon$).

This protocol moves the state of the system from the triple excitation state to the double-excitation subspace (stage 1), then to single-excitation subspace (stage 2), and then takes the system to the W-state in that subspace (part 3).

The performance of the above QSM-based active protocol can be numerically compared to its passive and greedy counterparts (Fig. 10). The average runtimes of these are, respectively, $\tau_{\rm av}^{(\rm QSM)} \simeq 430$, $\tau_{\rm av}^{(\rm pas)} \simeq 1500$, and $\tau_{\rm av}^{(\rm greedy)} \simeq 490$. To understand the reason for slower performance of the passive protocol, note that it can move the system in and out of one excitation subspace before the target state is reached (Fig. 10a). The greedy cost-function based policy avoids this issue and thus offers a speedup $f_{\rm greedy} \simeq 3.1$, however, it underperforms compared to QSM-based policy ($f_{\rm QSM} = 3.5$). The reason is that it suffers from the landscape flatness (Sec. IV B) issue. Before the system moves from $S = 3$ to $S = 2$, no coupling action is capable of directly achieving nonzero target state fidelity. The resulting delay can be seen in the typical protocol trajectories, see Fig. 10a. This disparity between $f_{\rm greedy}$ and $f_{\rm QSM}$ highlights the complementary nature of the two navigation approaches presented in this work. To apply measurement-driven navigation in the best way possible, one has to identify the better approach based on the target state and the coupling operators available.

## VI. DISCUSSION AND CONCLUSIONS

In this work, we have put forward the concept of measurement-driven active-decision steering of quantum states. We have developed steering protocols in which the measurement readouts are used to adjust the measurement protocol on-the-go, yielding significant acceleration of state preparation relative to passive steering. The possibility of exploiting the readouts explored here is the great advantage of measurement-based steering over drive-and-dissipation state preparation (which is largely equivalent to "blind" steering). While our approach has sweeping applicability, here we have chosen to focus on active measurement-driven steering as applied to the most challenging case of many-body quantum systems with entangled target states.

To satisfy physical (locality) constraints on system-detector couplings, we have proposed a scheme, based on parent Hamiltonian construction, for identifying feasible couplings. Employing such couplings, we have developed and analyzed Hilbert-space-orientation techniques for measurement-driven steering. A central ingredient here has been to develop feedback policies based on detector readouts.

The first Hilbert-space path-finding technique is based on a cost function, evaluating the running fidelity to the target state. We have shown a substantial (up to 23-fold) speedup of steering, employing this approach for preparation of the ground state of the AKLT model. For randomly generated MPS targets, the speedup from this method ranges at least up to $f_{\rm max} \sim 10^3$, with an average value of $f_{\rm av} \sim 10^1$ ($N = 5$ spin-1 system). Intriguingly, our numerics strongly suggests the growth of speedup

with system size for MPS targets whose parent Hamiltonians have a significant spectral gap.

A second method comprises mapping out the available measurement actions onto a Quantum State Machine (QSM), using a coarse-grained version of the corresponding graphs in Hilbert space. This approach is of conceptual significance complementary to the greedy method, being distinct in its principle and potentially surpassing the performance of greedy protocol for certain targets. We have shown an example of $W$-state preparation, where a QSM-based method provides a speedup $f_{\rm QSM} = 3.5$ that is higher compared to the greedy approach ($f_{\rm greedy} = 3.1$).

While we have limited ourselves here to specific examples, our schemes are of general applicability. They open the door to the design of efficient and high-quality state engineering, adiabatic state manipulation, and, possibly, quantum information processing. Moreover, steering protocols are subject to errors, both "static" (choice of steering parameters) and "dynamics" (noise) [43], in addition to a reduced "detection efficiency" discussed in Sec. IV C. Active decision-making steering may be designed to reduce the effect of such errors, by including self-corrections based on the recorded sequence of readouts. Importantly, compared to the greedy protocols, the effect of "measurement imperfections" is expected to be reduced for QSM strategies, as these operate with coarse-grained objects at the semiclassical level.

One may envision a host of further directions to generalize and develop the ideas of active steering. For example, the greedy minimization of our cost function may be further improved by finding other metrics of local "steepest decent." Further, one may systematically investigate less local (less greedy) optimization of the cost function, e.g., looking multiple cycles ahead. Another potential advantage of our protocols relies on the following observation: in the context of passive steering, one imposes constraints concerning locality (e.g., how many spins can be coupled to a local detector), and certain types of coupling terms. Given such constraints, not all target states are reachable. The introduction of active steering may overcome this handicap of target-state accessibility.

One may also combine the dynamics incorporated here with the inherent unitary evolution of the system at hand (due to a system-only Hamiltonian). Consider the context of passive (blind) measurement-induced steering, which, in the continuum time limit, leads to Lindbladian dynamics. Then, the addition of Hamiltonian dynamics enriches the variability of steering, allowing, for example, to obtain mixed states by design [44]. It is intriguing to investigate how the addition of Hamiltonian dynamics extends or improves active steering, thus marrying the frameworks of closed-loop quantum control for Hamiltonian-based state preparation and active-decision measurement-based steering. We expect, in particular, that active-decision strategies would allow one to steer the system to a pure target state even in those cases where the passive protocols yield mixed states. Systematic study of the combined action of active-decision measurement protocols and system-only Hamiltonian is an extremely interesting challenge for the future.

Further extensions of our approach include applications of QSM protocols to larger and more complex systems, going beyond a three-qubit setup. Optimizing such protocols may involve automatization of the creation and analysis of QSMs, e.g., for finding an optimal basis automatically, in similarity with quantum annealing, but now at the level of measurement operators. One may foresee a protocol, where combining local rotation of the basis states with a renormalization-group procedure, a structure of "quantum vertices" that are interconnected semiclassically emerges. This would then, in particular, admit a QSM engineering of MPS targets.

Finally, one may envision using machine learning to find more optimized navigation protocols (see [45–47] for related work). Given the delayed-reward setting at hand, a reinforcement learning strategy such as Q-learning [48] or SARSA [49] might be the most appropriate choice.

## ACKNOWLEDGMENTS

We thank D. Bondarenko, X. Bonet-Monroig, J. Chalker, R. Egger, C. Koch, P. Kumar, E. Medina Guerra, S. Morales, G. Morigi, S. Polla, R. Raussendorf, S. Roy, K. Snizhko, and A. Zazunov for discussions. The work was supported by the Deutsche Forschungsgemeinschaft (DFG): Project No. 277101999 – TRR 183 (Project C01) and Grants No. EG 96/13-1 and No. GO 1405/6-1, as well as by the Israel Science Foundation and the Helmholtz International Fellow Award, the National Science Foundation through award DMR-2037654, the US-Israel Binational Science Foundation (BSF), and the Netherlands Organization for Scientific Research (NWO/OCW).

[1] I. Bloch, *Quantum coherence and entanglement with ultracold atoms in optical lattices*, Nat. **453**, 1016 (2008).

[2] J.-W. Pan, C. Simon, C. Brukner, A. Zeilinger, *Entanglement purification for quantum communication* , Nature **410**, 1067 (2001).

[3] X.-Y. Luo, Y.-Q. Zou, L.-N. Wu, Q. Liu, M.-F. Han, M. K. Tey, and L. You, *Deterministic entanglement generation from driving through quantum phase transitions*, Science **355** 620 (2017).

[4] J.G. Bohnet, B.C. Sawyer, J. W. Britton, M.L. Wall, A.M. Rey, M. Foss-Feig, and J. J. Bollinger, *Quantum spin dynamics and entanglement generation with hundreds of trapped ions*, Science **352**, 1297 (2016).

[5] R. Stockill, M.J. Stanley, L. Huthmacher, E. Clarke, M. Hugues, A.J. Miller, C. Matthiesen, C. Le Gall, and M. Atatüre, *Phase-Tuned Entangled State Generation between Distant Spin Qubits*, Phys. Rev. Lett. **119**, 010503 (2017).

[6] B. Kraus, H. P. Büchler, S. Diehl, A. Kantian, A. Micheli, and P. Zoller, *Preparation of entangled states by quantum Markov processes*, Phys. Rev. A **78**, 042307 (2008).

[7] A. Pechen and H. Rabitz, *Teaching the environment to control quantum systems*, Phys. Rev. A **73**, 062102 (2006).

[8] C. Shen, K. Noh, V. V. Albert, S. Krastanov, M. H. Devoret, R. J. Schoelkopf, S. M. Girvin, and L. Jiang, *Quantum Channel Construction with Circuit Quantum Electrodynamics*, Phys. Rev. B **95**, 134501 (2017).

[9] R. Iten, R. Colbeck, M. Christandl, *Quantum Circuits for Quantum Channels*, Phys. Rev. A **95**, 052316 (2017).

[10] L. Piroli, G. Styliaris, and J. I. Cirac, *Quantum Circuits Assisted by Local Operations and Classical Communication: Transformations and Phases of Matter*, Phys. Rev. Lett. **127**, 220503 (2021).

[11] T.-C. Lu, L. A. Lessa, I. H. Kim, T. H. Hsieh, *Measurement as a shortcut to long-range entangled quantum matter*, arXiv:2206.13527.

[12] J. Wang, F. Sciarrino, A. Laing, and M. G. Thompson, *Integrated photonic quantum technologies*, Nat. Photonics **14**, 273 (2020).

[13] R. van Handel, J. K. Stockton, and H. Mabuchi, *Modelling and feedback control design for quantum state preparation*, J. Opt. B **7**, 10 (2005).

[14] L. C. Kwek, Z. Wei, and B. Zeng, *Measurement-Based Quantum Computing with Valence-Bond-Solids*, Int. J. Mod. Phys. B (2012).

[15] J.K. Stockton, R. van Handel, and H. Mabuchi, *Deterministic Dicke-state preparation with continuous measurement and control*, Phys. Rev. A **70**, 022106 (2004).

[16] C. Marr, A. Beige, and G. Rempe, *Entangled state preparation via dissipation-assisted adiabatic passages*, Phys. Rev. A **68**, 033817 (2003).

[17] X.-Q. Shao, T.-Y. Zheng, and S. Zhang, *Engineering steady three-atom singlet states via quantum-jump-based feedback*, Phys. Rev. A **85**, 042308 (2012).

[18] Y. Lin, J.P. Gaebler, F. Reiter, T.R. Tan, R. Bowler, Y. Wan, A. Keith, E. Knill, S. Glancy, K. Coakley, A.S. Sørensen, D. Leibfried, and D.J. Wineland, *Preparation of Entangled States through Hilbert Space Engineering*, Phys. Rev. Lett. **117**, 140502 (2016).

[19] Z. Liu, L. Kuang, K. Hu, L. Xu, S. Wei, L. Guo, and X.-Q. Li, *Deterministic creation and stabilization of entanglement in circuit QED by homodyne-mediated feedback control*, Phys. Rev. A **82**, 032335 (2010).

[20] A. Pechen, N. Il'in, F. Shuang, and H. Rabitz, *Quantum control by von Neumann measurements*, Phys. Rev. A **74**, 052102 (2006).

[21] L. Roa, A. Delgado, M. L. Ladrón de Guevara, and A. B. Klimov, *Measurement-driven quantum evolution*, Phys. Rev. A **73**, 012322 (2006).

[22] S. Roy, J.T. Chalker, I.V. Gornyi, and Y. Gefen, *Measurement-induced steering of quantum systems*, Phys. Rev. Research **2**, 033347 (2020).

[23] M. Ippoliti, M. J. Gullans, S. Gopalakrishnan, D. A. Huse, and V. Khemani, *Entanglement phase transitions in measurement-only dynamics*, Phys. Rev. X **11**, 011030, (2021).

[24] J. Zhang, Y.-X. Liu, R.-B. Wu, K. Jacobs, F. Nori, *Quantum feedback: theory, experiments, and applications*, Phys. Rep. **679**, 1 (2017).

[25] 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).

[26] F. Ticozzi, L. Viola, *Analysis and synthesis of attractive quantum Markovian dynamics*, Automatica **45**, 2002 (2009).

[27] H.M. Wiseman, *Quantum theory of continuous feedback*, Phys. Rev. A **49**, 2133 (1994).

[28] V. P. Belavkin, *Towards the theory of control in observable quantum systems*, Autom. Remote Control **44**, 178 (1983).

[29] H. M. Wiseman and G. J. Milburn, *Quantum theory of optical feedback via homodyne detection*, Phys. Rev. Lett. **70**, 548 (1993).

[30] T. Grigoletto and F. Ticozzi, *Stabilization via feedback switching for quantum stochastic dynamics*, arXiv:2012.08712.

[31] K. Jacobs, *Feedback control using only quantum backaction*, New J. Phys. **12**, 043005 (2010).

[32] S. Ashhab and F. Nori, *Control-free control: Manipulating a quantum system using only a limited set of measurements*, Phys. Rev. A **82**, 062103 (2010).

[33] S. Fu, G. Shi, A. Proutiere, and M. R. James, *Feedback Policies for Measurement-based Quantum State Manipulation*, Phys. Rev. A **90**, 062328 (2014).

[34] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki *Rigorous results on valence-bond ground states in antiferromagnets*, Phys. Rev. Lett. **59**, 799 (1987).

[35] O. Zilberberg, A. Romito, D. J. Starling, G. A. Howland, C. J. Broadbent, J. C. Howell, and Y. Gefen, *Null Values and Quantum State Discrìmination*, Phys. Rev. Lett. **110**, 170405 (2013).

[36] A.Y. Kitaev, A. Shen, and M. N.Vyalyi, *Classical and Quantum Computation*, Amer. Math. Soc., 2002.

[37] M. Hein, J. Eisert, and H. J. Briegel, *Multi-party entanglement in graph states*, Phys. Rev. A **69**, 062311 (2004).

[38] M. Fannes, B. Nachtergaele, and R. F. Werner, *Finitely correlated states on quantum spin chains*, Commun. Math. Phys. **144**, 443 (1992).

[39] M. B. Hastings, *Solving gapped Hamiltonians locally*, Phys. Rev. B **73**, 085115 (2006).

[40] D. Perez-Garcia, F. Verstraete, M. M. Wolf, J. I. Cirac, *Matrix Product State Representations*, Quantum Inf. Comput. **7**, 401 (2007).

[41] P. W. Anderson, *Infrared Catastrophe in Fermi Gases with Local Scattering Potentials*, Phys. Rev. Lett. **18**, 1049 (1967).

[42] W. Dür, G. Vidal, and J. I. Cirac, *Three qubits can be entangled in two inequivalent ways*, Phys. Rev. A **62**, 062314 (2000).

[43] Edward Medina Guerra, *Robustness of a quantum steering protocol to errors*, MSc Thesis, Weizmann Institute of Science, Rehovot, Israel (2022); Edward Medina Guerra, Parveen Kumar, Igor Gornyi, and Yuval Gefen, unpublished.

[44] P. Kumar, K. Snizhko, and Y. Gefen, *Engineering two-qubit mixed states with weak measurements*, Phys. Rev. Research **2**, 042014(R) (2020).

[45] S. Borah, B. Sarma, M. Kewming, G. J. Milburn, and J. Twamley, *Measurement-Based Feedback Quantum Con-*

*trol with Deep Reinforcement Learning for a Double-Well Nonlinear Potential*, Phys. Rev. Lett. **127**, 190403 (2021).

[46] D. Bondarenko, P. Feldmann, *Quantum autoencoders to denoise quantum data*, Phys. Rev. Lett. **124**, 130502 (2020).

[47] H. Dehghani, A. Lavasani, M. Hafezi, and M. J. Gullans, *Neural-Network Decoders for Measurement Induced Phase Transitions*, arXiv:2204.10904.

[48] C.J.C.H. Watkins, *Learning from delayed rewards*, Ph.D. dissertation, Psychology Dept. Univ. of Cambridge, UK (1989).

[49] G.A. Rummery, M. Niranjan, *On-line Q-learning using connectionist systems*, Engineering Dept. Univ. of Cambridge, UK (1994).

[50] This contrasts with the common definition of frustration in Hamiltonians, by which the locally projective $H_{\mathrm{par}}(\psi_{\mathrm{targ}})$ is frustration-free despite the non-commutation of its terms.

[51] In this section, we consider the all-down product state as our starting state. In both passive and active protocol, we assume each coupling to be applied multiple times until one either receives a click, or no-click for an asymptotically long time. Such a repeated application of a single coupling is then counted as a single protocol cycle. We take this approach for a practical purpose because simulating such protocols is more accessible numerically.