Quantum process tomography of unitary maps from time-delayed measurements

Irene López Gutiérrez[1] · **Felix Dietrich**[1] · **Christian B. Mendl**[1,2]

Abstract

Quantum process tomography conventionally uses a multitude of initial quantum states and then performs state tomography on the process output. Here we propose and study an alternative approach which requires only a single (or few) known initial states together with time-delayed measurements for reconstructing the unitary map and corresponding Hamiltonian of the time dynamics. The overarching mathematical framework and feasibility guarantee of our method is provided by the Takens embedding theorem. We explain in detail how the reconstruction of a single-qubit Hamiltonian works in this setting and provide numerical methods and experiments for general few-qubit and lattice systems with local interactions. In particular, the method allows to find the Hamiltonian of a two qubit system by observing only one of the qubits.

1 Introduction

System identification refers to the estimation of the dynamics of a system from measurements of its characteristics. Its quantum analogue, quantum process tomography

✉ Irene López Gutiérrez
  irene.lopez@tum.de

  Felix Dietrich
  felix.dietrich@tum.de

  Christian B. Mendl
  christian.mendl@tum.de

[1] Department of Computer Science, School of Computation, Information and Technology, Technical University of Munich, Boltzmannstraße 3, 85748 Garching, Germany

[2] Institute for Advanced Study, Technical University of Munich, Lichtenbergstraße 2a, 85748 Garching, Germany
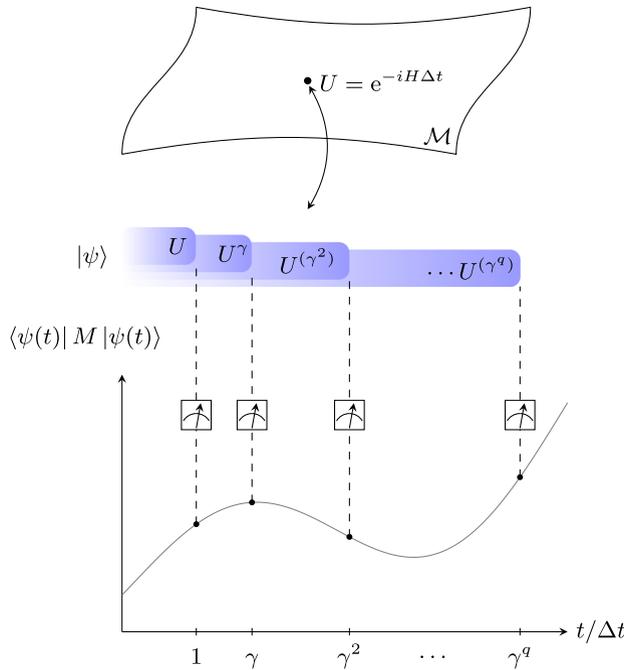
Fig. 1 Schematic illustration of the map from a unitary time step operator U to a vector of measurement averages at different time delays

(QPT), is essential for the realization and testing of quantum devices [1–3], as a benchmarking tool for quantum algorithms [4, 5], and in general for understanding the inner workings of a quantum system [6, 7]. However, textbook algorithms for QPT [8] might be difficult to realize in laboratory settings due to the required preparation of many initial states and observation of the complete target system.

In this work, we focus on the unitary time evolution of a closed quantum system. We propose and investigate an approach based on measurements with different *time delays*, which should be easily realizable in laboratory experiments. Our main contributions are algorithms and numerical procedures for identifying the corresponding time evolution operator and thus indirectly the quantum Hamiltonian. Intriguingly, we can identify the operator for the entire system even if the measurements are restricted to a subsystem (using two known initial quantum states). The Takens embedding theorem, discussed in Sect. 3, provides an overarching mathematical foundation and feasibility guarantee for our approach. Concretely, the mathematical framework starts with a manifold $\mathcal{M}$, which we take to be the special unitary group of time evolution operators. Given an initial quantum state, a time step matrix $U \in \mathcal{M}$ then determines the measurement averages at a sequence of time points, see Fig. 1. The Takens theorem states that the map from $\mathcal{M}$ to this measurement vector is actually a smooth embedding (under certain conditions), which then allows us to reconstruct U.

Related to the present work, a recent approach for quantum process tomography using tensor networks as a parametrization of the quantum channel was able to achieve

high accuracies for systems of up to 10 qubits [9]. Furthermore, in [10] the authors derive a lower bound in the number of POVMs required to fully characterize a unitary or near-unitary map. Regarding the related task of quantum state tomography, several methods, some based on machine learning techniques, have been developed [11–17]. However, except for [14], these works do not explicitly take advantage of the time trajectory of the quantum system as envisioned here. Gate set tomography [18] does not need pre-calibrated quantum states, but focuses on finding gate sets.

Our work is closely related to earlier research on the identification of quantum Hamiltonians from time series [19, 20]. The authors obtain all degrees of freedom of the Hamiltonian with known structure by solving a system of equations, mostly in the presence of noise. More recently, artificial neural networks have been used for identification and applied to experimental data [21–23]. However, the minimal number of observables needed to identify the Hamiltonian is not addressed mathematically, and we provide it here with the link to Takens theorem. In our work, the choice of initial state is not as important, as we will demonstrate.

Takens theorem has been fundamental to several system identification methods for general dynamical systems already [24], recently also involving machine learning [25–28], with a focus on PDE [29, 30] or special structure such as (classical) Hamiltonian dynamics [31, 32]. The identification of a unitary map from measurement data has been discussed by Koopman and von Neumann [33, 34], work that has been revived in a data-driven context in the last 20 years and extended to dissipative systems [35–37]. The connection to quantum systems and their special structure and challenges have not yet been addressed in the work cited above, however.

2 Physical model

We assume throughout that the quantum Hamiltonian H is time-independent and will investigate two settings: (i) a few-qubit system and H a general dense Hermitian matrix and (ii) a (tight binding) Ising-type model on a two-dimensional lattice Λ, as widely studied in condensed matter physics. For concreteness, the physical system for case (ii) consists a local spin degree of freedom at each lattice site. The Hamiltonian is then defined as

$$H = -\sum_{\langle j,\ell \rangle} J_{j,\ell}\, \sigma_j^z \sigma_\ell^z - \sum_{j \in \Lambda} \vec{h}_j \cdot \vec{\sigma}_j, \tag{1}$$

where $J_{j,\ell} \in \mathbb{R}$ and $\vec{h}_j \in \mathbb{R}^3$ are local parameters defining the interaction strength and the external field, respectively. σ_j^α for $\alpha \in \{x, y, z\}$ is the α-th Pauli matrix acting on site $j \in \Lambda$, and $\vec{\sigma}_j = (\sigma_j^x, \sigma_j^y, \sigma_j^z)$ the corresponding Pauli vector. The first sum in (1) runs over nearest neighbors on the lattice. We take Λ to contain a finite number n of sites and assume periodic boundary conditions. The site-dependent parameters allow to simulate disorder. Figure 2 illustrates the interaction and external field terms of H on a two-dimensional lattice. The precise form of H is not important for the reconstruction as long as the time dynamics it generates can be well approximated by
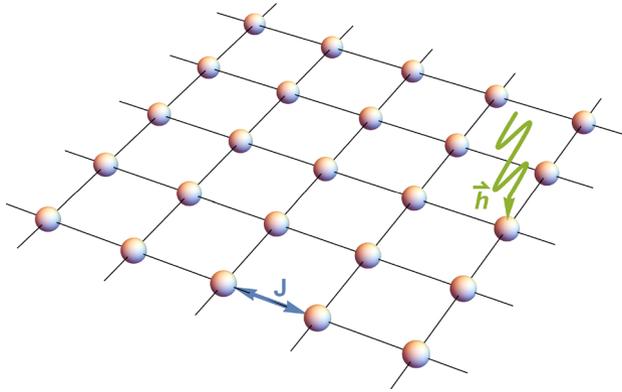
Fig. 2 Visualization of the quantum Hamiltonian in Eq. (1) on a two-dimensional lattice, with J the interaction strength and $\vec{h}$ the external field

a quantum circuit, for example via Trotterization. We assume that the overall structure of H is known, and our task is to determine the numerical values of the parameters.

For later reference, we state the unitary time evolution operator at time $t \in \mathbb{R}$ based on the Schrödinger equation (in units of $\hbar = 1$):

$$U(t) = \mathrm{e}^{-iHt} . \tag{2}$$

The solution to the Schrödinger equation for an initial wavefunction (statevector) $\psi \in \mathbb{C}^N$ is then $\psi(t) = U(t)\psi$.

3 Takens embedding framework

The theorems of Takens and Ruelle [38, 39], based on the embedding theorems of Whitney [40], are the theoretical foundations of time-delay embedding we use here.

The general idea we employ in this paper is to "embed" the manifold of unitary matrices (describing the dynamics of a quantum system) into the space of measurement trajectories. The primary motivation for conducting an embedding is to create a feature space for the optimization. Most QPT approaches use measurements in many different measurement bases, implicitly creating such an embedding. In our case, we instead use time-delayed measurements, considering that in some experimental setups it may be more practical to maintain the time evolution than to implement a full set of measurement bases.

We first state the mathematical theorem and then discuss the specialization for quantum time evolution. Let $k \geq d \in \mathbb{N}$, and $\mathcal{M} \subset \mathbb{R}^k$ be a d-dimensional, compact, smooth, connected, oriented manifold with Riemannian metric g induced by the embedding in k-dimensional Euclidean space. Note that this setting is sufficient for our presentation, but is more restrictive than allowed by the results cited below.

Together with the results from Packard et al. [41] and Aeyels [42], the definitions and theorems of Takens [39] describe the concept of observability of state spaces of

nonlinear dynamical systems. A dynamical system is defined through its state space (here, the manifold $\mathcal{M}$) and a diffeomorphism $\phi : \mathcal{M} \to \mathcal{M}$.

Theorem 1 *Generic delay embeddings For pairs (ϕ, y), $\phi : \mathcal{M} \to \mathcal{M}$ a smooth diffeomorphism and $y : \mathcal{M} \to \mathbb{R}$ a smooth function, it is a generic property that the map $\Phi_{(\phi,y)} : \mathcal{M} \to \mathbb{R}^{2d+1}$, defined by*

$$\Phi_{(\phi,y)}(x) = \left(y(x), y(\phi(x)), \ldots, y(\underbrace{\phi \circ \cdots \circ \phi}_{2d \ times}(x)) \right) \tag{3}$$

is an embedding of $\mathcal{M}$; here, "smooth" means at least C^2 and $x \in \mathcal{M}$.

Genericity in this context is defined as "an open and dense set of pairs (ϕ, y)" in the C^2 function space. Open and dense sets can have measure zero, so Sauer et al. [43] later refined this result significantly by introducing the concept of prevalence (a "probability one" analog in infinite dimensional spaces). See [44] for similar results with stochastic systems.

Let N denote the quantum Hilbert space dimension. In our context, $\mathcal{M}$ is the special unitary group $SU(N) \subset \mathbb{C}^{N \times N}$ when identifying $\mathbb{C} \simeq \mathbb{R}^2$. In particular, $\mathcal{M}$ (with underlying field $\mathbb{R}$) has dimension $d = N^2 - 1$. In physical terms, the elements of $SU(N)$ are the unitary time evolution matrices in Eq. (2). We assume that the Hamiltonian H is traceless, since adding multiples of the identity to H leads to a global phase factor in the time evolution, which is unobservable in subsequent measurements as envisioned here. We fix a time step Δt, which we may (without loss of generality) absorb into H, and set $U = e^{-iH}$ in the following.

Now define the diffeomorphism ϕ via a scaling of H by a factor $\gamma > 0$, $\gamma \neq 1$, such that

$$\phi(U) = e^{-i\gamma H} = U^\gamma. \tag{4}$$

Regarding y, fix a randomly chosen initial quantum state $\psi \in \mathbb{C}^N$ and an observable (Hermitian matrix) M. Now let y compute the corresponding expectation value:

$$y : \mathcal{M} \to \mathbb{R}, \quad y(U) = \langle \psi | U^\dagger M U | \psi \rangle. \tag{5}$$

With these definitions, the output of the map $\Phi_{(\phi,y)}$ in Eq. (3) becomes

$$\begin{aligned}
\Phi_{(\phi,y)}(U) = \Big(& \langle \psi | e^{iH} M e^{-iH} | \psi \rangle, \\
& \langle \psi | e^{i\gamma H} M e^{-i\gamma H} | \psi \rangle, \\
& \ldots, \langle \psi | e^{i\gamma^{2d} H} M e^{-i\gamma^{2d} H} | \psi \rangle \Big),
\end{aligned} \tag{6}$$

physically corresponding to measurements at time points $t_q = \gamma^q$ for $q = 0, \ldots, 2d$. Note that the exponential scaling of the time points requires a carefully chosen γ that adapts to the time and resolution constrains of the experiment. In practice, this may

limit the method to few qubit systems, where the required number of time points is small.

The main point here is the prospect to identify the time step matrix U, and thus indirectly the Hamiltonian, based on a single measurement time trajectory, under the assumption that the initial quantum state is known. As caveat, the relation $U = e^{-iH\Delta t}$ determines the eigenvalues of H only up to multiples of $2\pi/\Delta t$. Moreover, the map $\Phi_{(\phi,y)}$ might not be one-to-one, in the sense that two different unitary matrices give rise to the same measurement trajectory. We discuss such a case in more detail in the following. Such issues can be avoided in practice by additional assumptions on the structure of H.

4 Numerical methods

Throughout this work, we assume that it is feasible to reliably prepare a single (or when indicated several) known initial state(s) $\psi \in \mathbb{C}^N$. To demonstrate the general applicability of our methods, ψ is chosen at random in the following algorithms and numerical simulations. Specifically, the entries of ψ before normalization are independent and identically distributed (i.i.d.) random numbers sampled from the standard complex normal distribution.

4.1 Reconstruction algorithm for a single-qubit system

We now describe how to reconstruct a single-qubit Hamiltonian from a time series of measurements. The Bloch sphere picture [8] provides a geometric perspective and insight into single-qubit quantum states and operations. The Bloch vector $\vec{r} \in \mathbb{R}^3$ associated with $\psi \in \mathbb{C}^2$ is defined via the relation:

$$|\psi\rangle\langle\psi| = \frac{1}{2}(I + \vec{r} \cdot \vec{\sigma}). \tag{7}$$

For pure states as considered here, $\vec{r}$ is a unit vector. We can parametrize any single-qubit Hamiltonian (up to multiples of the identity map, which are irrelevant here) as:
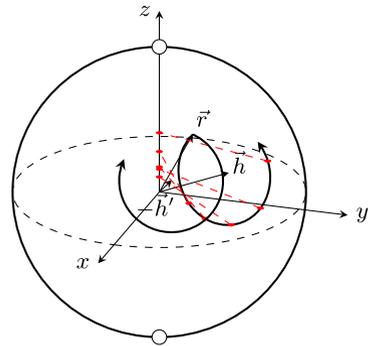
$$H = \vec{h} \cdot \vec{\sigma} \tag{8}$$

with $\vec{h} \in \mathbb{R}^3$. The corresponding time evolution operator is the rotation:

$$U(t) = e^{-iHt} = \cos(\omega t/2)I - i\sin(\omega t/2)(\vec{v} \cdot \vec{\sigma}) \in \mathrm{SU}(2), \tag{9}$$
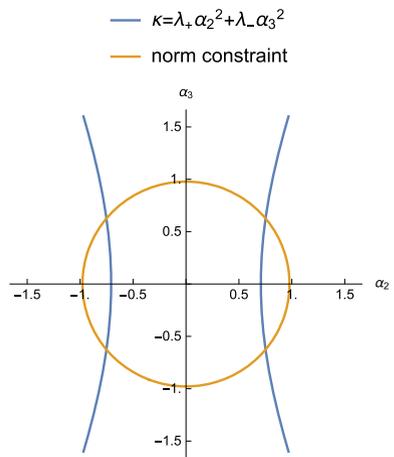
when representing $\vec{h} = \omega\vec{v}/2$ by a unit vector $\vec{v} \in \mathbb{R}^3$ and $\omega \in \mathbb{R}$. On the Bloch sphere, this operator corresponds to a classical rotation about $\vec{v}$, as illustrated in Fig. 3a and described by Rodrigues' rotation formula ($\theta = \omega t$):

$$\mathsf{U}_{\theta,\vec{v}}\,\vec{r} = \cos(\theta)\vec{r} + \sin(\theta)(\vec{v} \times \vec{r}) + (1 - \cos(\theta))(\vec{v} \cdot \vec{r})\,\vec{v}. \tag{10}$$

(a) trajectory on Bloch sphere

(b) conditions on α_2 and α_3 coefficients

$U_{\theta,\vec{v}} \in SO(3)$ is a rotation matrix (parametrized by θ and $\vec{v}$) applied to $\vec{r}$. In the above context of Takens embedding with $U = U(\Delta t)$, we may equivalently work with $U = U_{\omega \Delta t, \vec{v}}$ and matrix powers thereof.

The time trajectory effected by the rotation applied to $\vec{r}$ results in a circle embedded within the Bloch sphere. Knowing the circle would allow to determine $\vec{v}$, and the dynamics on the circle to determine ω. By construction, we also know one point on the circle already, namely $\vec{r}$ (the initial condition), but we only have access to the expectation value of an observable M for $t > 0$. In the following, we denote the "measurement direction" by $\vec{m} \in \mathbb{R}^3$, i.e., M is parametrized as $M = \vec{m} \cdot \vec{\sigma}$. Algorithm 1 facilitates a recovery of ω and $\vec{v}$ using the projections of the time trajectory on $\vec{m}$. The main idea is to first reconstruct ω based on the time dependence and then the unit vector $\vec{v}$.

As caveat, the solution to this problem is not unique, due to the four possible signs of the coefficients α_2 and α_3 in the algorithm. Figure 3a visualizes how two rotations can generate the same projection onto the z-axis, with $\vec{h}'$ resulting from

$(\alpha_2, \alpha_3) \to -(\alpha_2, \alpha_3)$. The two equations for α_2 and α_3 are shown in Fig. 3b. (In the example, $\alpha_1 = -0.2051$, and hence, the radius of the norm constraint circle is close to 1.) In order to decide between these distinct solutions, one could perform one further measurement in a different basis, or may use a-priori knowledge about the Hamiltonian.

Algorithm 1 Reconstruction of a single-qubit Hamiltonian (measurement direction $\vec{m} \in \mathbb{R}^3$ with $\vec{m} \nparallel \vec{r}$)

Input: Measurement averages $y_q = \vec{m} \cdot (\mathsf{U}_{\omega t_q, \vec{v}} \, \vec{r})$ with $t_q = \Delta t \, \gamma^q$ for $q = 0, \ldots, 2d$

Output: Hamiltonian parameters ω and $\vec{v}$.

1: Find ω based on the dependency on ωt in Eq. (10):

$$\omega = \underset{\omega > 0}{\mathrm{argmin}} \min_{a,b,c} \sum_{q=0}^{2d} |y_q - a\cos(\omega t_q - b) - c|^2$$

2: Set $\tilde{y}_q = y_q - \cos(\omega t_q)(\vec{m} \cdot \vec{r})$ for $q = 0, \ldots, 2d$ (subtract term which is independent of $\vec{v}$)

3: Find $\alpha_1 = \vec{m} \cdot (\vec{v} \times \vec{r})$ and $\kappa = (\vec{v} \cdot \vec{r})(\vec{m} \cdot \vec{v})$ via a least squares fit:

$$\alpha_1, \kappa = \underset{\alpha_1, \kappa}{\mathrm{argmin}} \sum_{q=0}^{2d} |\tilde{y}_q - \sin(\omega t_q)\alpha_1 - (1 - \cos(\omega t_q))\kappa|^2$$

4: Represent $\vec{v}$ with respect to the orthonormal basis

$$\left\{ \vec{u}_1 = \frac{\vec{r} \times \vec{m}}{\|\vec{r} \times \vec{m}\|}, \vec{u}_2 = \frac{\vec{r} + \vec{m}}{\|\vec{r} + \vec{m}\|}, \vec{u}_3 = \frac{\vec{r} - \vec{m}}{\|\vec{r} - \vec{m}\|} \right\} :$$
$$\vec{v} = \alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \alpha_3 \vec{u}_3,$$

with to-be determined coefficients $\alpha_2, \alpha_3 \in \mathbb{R}$. Using that $\vec{u}_1, \vec{u}_2$ and $\vec{u}_3$ are eigenvectors of $K = \frac{1}{2}(\vec{m} \otimes \vec{r} + \vec{r} \otimes \vec{m})$ with respective eigenvalues $0, \lambda_+ = \frac{1}{2}(1 + \vec{m} \cdot \vec{r})$ and $\lambda_- = -\frac{1}{2}(1 - \vec{m} \cdot \vec{r})$, it follows that

$$\kappa = \vec{v} \cdot K \vec{v} = \lambda_+ \alpha_2^2 + \lambda_- \alpha_3^2.$$

Together with the normalization condition

$$\alpha_1^2 + \alpha_2^2 + \alpha_3^2 = 1,$$

this leads to $\alpha_2^2 = \kappa - (1 - \alpha_1^2)\lambda_-$ and $\alpha_3^2 = 1 - \alpha_1^2 - \alpha_2^2$. Decide between the four possible signs of α_2, α_3 using a-priori information of H or one additional measurement in a different basis.

We remark that the nonlinear optimization in the first step of the algorithm might get trapped in a local minimum, which could be resolved by restarting the optimization. On the other hand, when neglecting uncertainties associated with the measurements, a correct solution is indicated by a zero residual both in steps 1 and 3. We assume that a range of realistic frequencies ω is known beforehand, such that the Nyquist condition (given the non-uniform sampling points t_q) holds [45].

4.2 Relaxation method

As straightforward approach for determining a unitary time step matrix U which matches the measurement averages, one could start from the natural parametrization $U = \mathrm{e}^{-iH}$ (with the time step already absorbed into H) and then optimize the Hermitian matrix H directly. However, this method encounters the difficulty of a complicated optimization landscape with many local minima, in particular when using gradient descent-based approaches.

Here we discuss an alternative numerical approach tailored towards generic (dense) few-qubit Hamiltonians. As discussed in Sect. 4.1, for single qubits a transformation $\psi' = U\psi$ by a unitary matrix $U \in \mathrm{SU}(2)$ can equivalently be described by a spatial rotation in three dimensions on the Bloch sphere: $\vec{r}' = \mathsf{U}r$, with $\mathsf{U} \in \mathrm{SO}(3)$ given by Rodrigues' formula (10). An equivalent mapping from U to U is via:

$$\mathsf{U}_{\alpha,\beta} = \frac{1}{2}\,\mathrm{tr}\big[\sigma^{\alpha}U\sigma^{\beta}U^{\dagger}\big] \tag{11}$$

for $\alpha, \beta = 1, 2, 3$, where we have used the relation (7) together with the orthogonality relation of the Pauli matrices: $\frac{1}{2}\,\mathrm{tr}[\sigma^{\alpha}\sigma^{\beta}] = \delta_{\alpha\beta}$. For our purposes, the Bloch picture has the advantage that measurement averages depend linearly on the Bloch vector and hence on U, e.g., $\langle\psi'|\sigma^{z}|\psi'\rangle = \vec{e}_{z} \cdot \vec{r}' = \vec{e}_{z} \cdot (\mathsf{U}r)$. In practice, we first optimize the entries of U to reproduce the measurement data and afterwards find U related to U via Eq. (11).

Generalization to a larger number of qubits is feasible via tensor products of Pauli and identity matrices (in other words, Pauli strings). The analogue of (11) for n qubits, with $U \in \mathrm{SU}(2^{n})$, is $\mathsf{U} \in \mathrm{SO}(4^{n} - 1)$ with entries

$$\mathsf{U}_{\alpha,\beta} = \frac{1}{2^{n}}\,\mathrm{tr}\big[(\sigma^{\alpha_1} \otimes \cdots \otimes \sigma^{\alpha_n})U(\sigma^{\beta_1} \otimes \cdots \otimes \sigma^{\beta_n})U^{\dagger}\big], \tag{12}$$

where α and β are now index tuples from the set $\{0, 1, 2, 3\}^{\otimes n}\backslash\{(0, \ldots, 0)\}$, using the convention that σ^{0} is the 2×2 identity matrix. We exclude the tuple $(0, \ldots, 0)$ here since it conveys no additional information: the identity matrix is always mapped to itself by unitary conjugation.

Specifying an element of $\mathrm{SO}(4^{n} - 1)$ involves more free parameters than for a unitary matrix from $\mathrm{SU}(2^{n})$ in case $n \geq 2$; thus, the optimization might find an orthogonal U which reproduces the measurement data, but does not originate from a $U \in \mathrm{SU}(2^{n})$ via (12). We circumvent this representability issue as follows, focusing on the case of two qubits here. Every $U \in \mathrm{SU}(4)$ can be decomposed as [46, 47]

$$U = \big(u^{\mathrm{a}} \otimes u^{\mathrm{b}}\big)R\big(u^{\mathrm{c}} \otimes u^{\mathrm{d}}\big) \tag{13}$$

with the "entanglement" gate

$$R = \mathrm{e}^{-\frac{i}{2}(\theta_1\,\sigma^1\otimes\sigma^1 + \theta_2\,\sigma^2\otimes\sigma^2 + \theta_3\,\sigma^3\otimes\sigma^3)}, \quad \theta_1, \theta_2, \theta_3 \in \mathbb{R} \tag{14}$$

and single-qubit unitaries $u^a, u^b, u^c, u^d \in SU(2)$. The single-qubit gates can be handled as before, i.e., represented by orthogonal rotation matrices $u^a, u^b, u^c, u^d \in SO(3)$. We find the Bloch representation of R via (12) (cf. [48, 49]); the parameters $\theta_1, \theta_2, \theta_3$ appear in the matrix entries solely as $\cos(\theta_j)$ and $\sin(\theta_j)$ for $j = 1, 2, 3$. In summary, we express the decomposition (13) in terms of to-be found orthogonal matrices in the Bloch picture.

After switching to the described Bloch representation, we "relax" the condition that an involved (real) matrix U is orthogonal, by admitting any real matrix, but adding the term

$$\mathcal{L}_{\text{orth}} = \left\| UU^T - I \right\|_F^2 \tag{15}$$

to the overall cost function, where $\|\cdot\|_F$ denotes the Frobenius norm. As advantage, we bypass a parametrization of U to enforce strict orthogonality and avoid local minima in the optimization. Although this may in principle lead to a non-unitary matrix, at the end of the optimization we retrieve an exact unitary by finding the Hamiltonian entries via the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm.

By choosing $\gamma = 2$ in Eq. (4), the time steps $t_q = \gamma^q$ are integers, and we can generate corresponding powers of U by defining $U_0 = U$, $U_q = U_{q-1}^2$ for $q = 1, \ldots, 2d$, such that

$$e^{-i\gamma^q H} = U^{\gamma^q} = U_q \quad \text{for } q = 0, \ldots, 2d. \tag{16}$$

In our case, the U_q's are separate matrices, which are set in relation via an additional cost function term

$$\mathcal{L}_{\text{steps}} = \sum_{q=1}^{2d} \left\| U_q - U_{q-1}^2 \right\|_F^2. \tag{17}$$

For the case of two-qubit Hamiltonians, we additionally substitute two-dimensional vectors (c_j, s_j) for $(\cos(\theta_j), \sin(\theta_j))$, again to avoid local minima. The condition $c_j^2 + s_j^2 = 1$ translates to another penalty term in the overall cost function:

$$\mathcal{L}_\theta = \sum_{j=1}^{3} |c_j^2 + s_j^2 - 1|^2. \tag{18}$$

In our experiments, we weighted the penalization terms with a constant factor. Thus, denoting as y_q the experimental measurements and as $\hat{y}_q$ the predictions of the model, the final cost functions are:

$$\mathcal{L} = \frac{1}{d} \sum_{q=1}^{2d} (y_q - \hat{y}_q)^2 + 0.1 \left(\sum_q \mathcal{L}_{\text{orth, q}} + \mathcal{L}_{\text{steps}} \right) \tag{19}$$
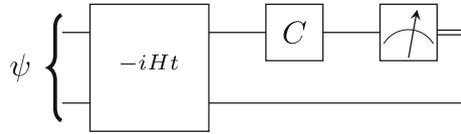
Fig. 4 Minimal example for the time-delayed partial (subsystem) measurement scenario. The goal is to reconstruct the (unknown) Hamiltonian H. We assume that one can prepare two different initial states and can select a measurement basis via the gate C

for the single-qubit case and

$$\mathcal{L} = \frac{1}{d} \sum_{q=1}^{2d} (y_q - \hat{y}_q)^2 + 0.01 \left(\sum_q \mathcal{L}_{\text{orth, q}} + \mathcal{L}_{\text{steps}} + \mathcal{L}_\theta \right) \tag{20}$$

for the two-qubit case. Here we note that, by performing the relaxation described throughout this section, we are increasing the dimensionality of the optimization landscape. However, and perhaps counter-intuitively, we find in our numerical experiments that allowing for more flexibility in the matrices considerably reduces the chances of the optimization getting stuck in local minima.

4.3 Partial (subsystem) measurements

An intriguing possibility of the time-delay measurements is the identification of the overall Hamiltonian based on measurements restricted to a subsystem. As minimal example, consider a two-qubit system, the choice between two initial states, being able to select a measurement basis via the gate C, and time-delayed measurements on one of the qubits while the other qubit is inaccessible, as illustrated in Fig. 4.

As already mentioned, the map $\Phi_{(\phi, y)}$ is not a one-to-one map, which may lead to multiple Hamiltonians producing the same set of time-delayed measurements. For a single-qubit system, it is very clear how to resolve this ambiguity (see Sect. 4.1). However, for larger systems it is not so clear. In our experiments, we observed that choosing two or three different initial states and/or two or three different measurement basis was enough for the optimizer to arrive to the ground-truth Hamiltonian, at least for the system sizes considered throughout the paper. For the numerical experiment shown in Fig. 10, we use two random initial states and X-, Y- and Z-basis measurements on the top qubit and minimize the mean-squared error between the observed measurement averages and the prediction based on a general Ansatz for the (traceless) Hamiltonian, i.e., 15 real parameters.

4.4 Lattice system with local interactions

Finally, we consider quantum systems defined on a lattice and assume an Ising-type Hamiltonian as in Eq. (1) (instead of a generic Hamiltonian) to avoid an exponential growth of the number of parameters.
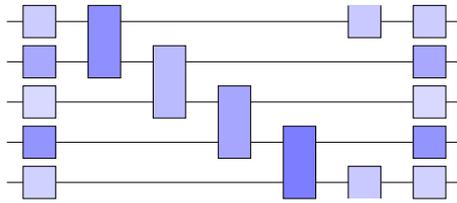
Fig. 5 One time step Δt of the quantum dynamics based on Strang splitting into interaction and local field terms, see Eq. (21), illustrated for a one-dimensional lattice with $n = 5$ sites. The single-qubit gates on the left and right are rotation gates $\exp(-i\Delta t\, \vec{h}_j \cdot \vec{\sigma}_j/2)$ for the j-th qubit, and the two-qubit gates are $\exp(-i\Delta t\, J_{j,j+1}\sigma_j^z\sigma_{j+1}^z)$. The different shades indicate different parameters at each site. The ordering of the two-qubit interaction gates among each other is not relevant since they pairwise commute

For the purpose of reconstructing the Hamiltonian parameters, the Schrödinger time evolution can be well approximated via a second-order Strang splitting method [50], i.e., the Time-Evolving Block Decimation (TEBD) algorithm [51]. Interpreting the resulting layout of single- and two-site unitary operators as forming a quantum circuit, the reconstruction task amounts to a variational circuit optimization to reproduce the reference measurement averages. Specifically, when considering the interaction and local field parts of the Hamiltonian

$$H_{\text{int}} = -\sum_{\langle j,\ell\rangle} J_{j,\ell}\,\sigma_j^z\sigma_\ell^z, \quad H_{\text{loc}} = -\sum_{j\in\Lambda} \vec{h}_j \cdot \vec{\sigma}_j \tag{21}$$

by themselves, the individual terms in each of the sums pairwise commute. Figure 5 illustrates the corresponding quantum circuit for a one-dimensional lattice; the constructing works for higher-dimensional lattices as well.

5 Numerical experiments

Here we present numerical experiments for the methods described in Sect. 4.

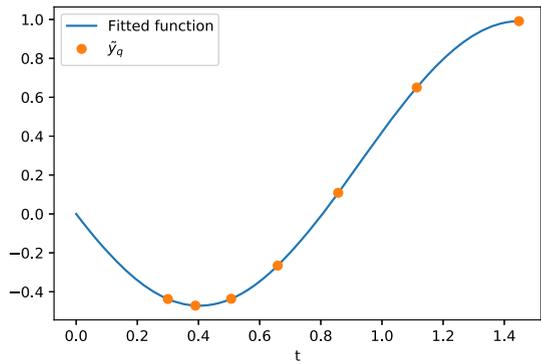## 5.1 Exact reconstruction of a single-qubit Hamiltonian

With the algorithm described in Sec. 4.1 we are able to reconstruct a single-qubit Hamiltonian up to numerical precision. Figure 6 shows the results obtained for the ω and least squares optimizations. Following the Takens framework, we use $2d + 1$ time points, where $d = 3$ is the number of Hamiltonian parameters to be reconstructed. Specifically, we have used the time points $t_q = 0.3\times(1.3)^q$ for $q = 0, \ldots, 6$ here. With this, we are able to find ω, α_1 and κ up to an error of $\sim 10^{-15}$. Thus, 7 measurements, plus a further measurement in a different basis to distinguish between the possible solutions due to symmetry, are sufficient for the reconstruction.

So far we have assumed a perfectly known initial state and noise-free measurements. In real experiments, however, this is not a reasonable assumption. One can model both

Fig. 6 **a** step 1 and **b** step 3 of Algorithm 1 using Z-basis measurements, for the time evolution of initial state $|+\rangle$ under a Hamiltonian where each entry in the vector is generated from a normal distribution
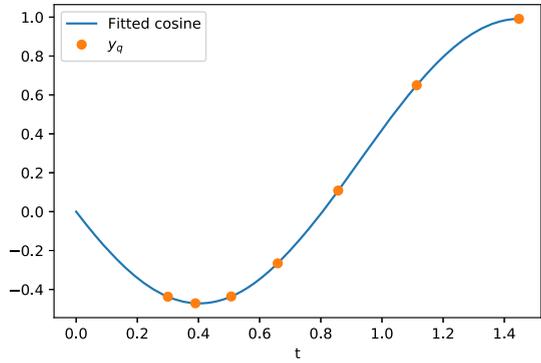
(a) cosine fit to find ω
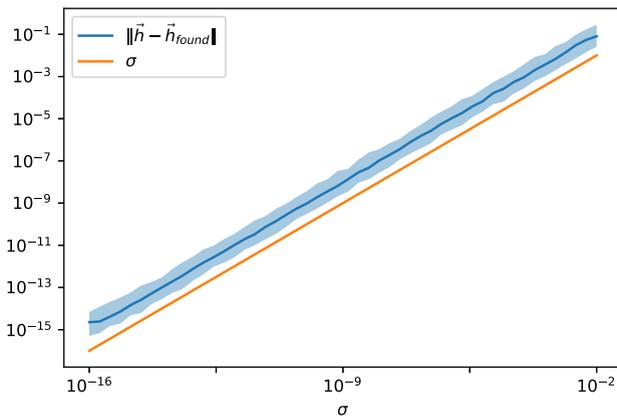

(b) function fit to find α_1 and κ

Fig. 7 Single-qubit Hamiltonian reconstruction error. The solid line shows the median, while the lighter shade shows the 10% quantiles

sources of noise by replacing the expectation values from Eq. (5) by

$$y'(U) = y(U) + \mathcal{N}(0, \sigma^2) \tag{22}$$

where the standard deviation σ quantifies the uncertainty related to the initial state and/or the measurements. Figure 7 shows that the reconstruction error is related linearly to σ. This is to be expected, as the arguments of the cost functions minimized in Algorithm 1 are related linearly to the obtained measurements.

5.2 Relaxation method

In this subsection, we consider a generic single- or two-qubit Hamiltonian H. Specifically, we draw the coefficients of H with respect to the Pauli basis from the standard normal (Gaussian) distribution.

We first focus on a single-qubit 2×2 Hamiltonian. Figure 8 a shows the result of directly fitting the vector $\vec{h} \in \mathbb{R}^3$ (see Eq. (8)), with the KL divergence between the predicted and actual measurement probabilities as cost function. This approach is compared to the "relaxation" procedure (described in Sect. 4.2) in Fig. 8b, using $\gamma = 2$ and $\Delta t = 0.1$. The manifold for the Takens embedding has dimension $d = 3$ for a single qubit, and hence, $2d + 1 = 7$ time points should be used. However, we face the difficulty that the Hamiltonian is not uniquely specified solely by Pauli-Z measurements according to the discussion in Sect. 4.1. For this reason, we include Pauli-X measurement data as well and reduce the number of time points to 4 (to compensate for this additional source of data). For both versions, we use the Adam optimizer [52]. The direct fitting method might get trapped in a local minimum and hence not be able to find the ground-truth vector $\vec{h}$, which leads to a large variation around the median. This issue is ameliorated by the relaxation method. The almost-plateau observed in Fig. 8b might be due to the constraint in Eq. (17): namely, a change of U needs to propagate to the higher powers of U. However, we remark that the loss function still decreases steadily, indicating that the optimization should continue.

Figure 8c visualizes the loss function and relative errors when applying the relaxation procedure to reconstruct the unitary time step matrix and corresponding Hamiltonian of a two-qubit system ($d = 15$ real parameters). Specifically, for parameter optimization we express Eq. (13) using the Bloch picture, i.e., in terms of orthogonal rotation matrices for the single-qubit unitaries, and likewise using the Bloch picture analogue of the entanglement gate (14), parametrized by two-dimensional vectors (c_j, s_j) for $(\cos(\theta_j), \sin(\theta_j))$. The condition $c_j^2 + s_j^2 = 1$ translates to another penalty term in the overall cost function, see Eq. (18). We set $\Delta t = 0.05$, $\gamma = 2$ and use 6 time points for this experiment. The reference measurement data at each time point are the expectation values of the nine observables $\{I \otimes \sigma^\alpha, \sigma^\alpha \otimes I, \sigma^\alpha \otimes \sigma^\alpha\}_{\alpha \in \{x,y,z\}}$. We use 54 measurement data points in total (instead of $2d + 1 = 31$) since we found that the additional information improves the reconstruction. As last step (after the main optimization), we find the Hamiltonian entries giving rise to the computed U via the BFGS algorithm.

Fig. 8 **a** and **b**: Single-qubit Hamiltonian reconstruction based on four Z- and four X-basis measurements, (**a**) by direct fitting of the Hamiltonian parameters to the measurement data, and (**b**) using the "relaxation" procedure (Sect. 4.2), with the actual Hamiltonian parameters determined from U at the end (blue line at the rightmost section of the plot). Solid lines show the median, and lighter shades are 25% quantiles. (**c**) "Relaxation procedure" for two-qubit Hamiltonian reconstruction, working with the Bloch picture of the representation (13) to find the time step matrix U, and then determining a two-qubit Hamiltonian giving rise to this U
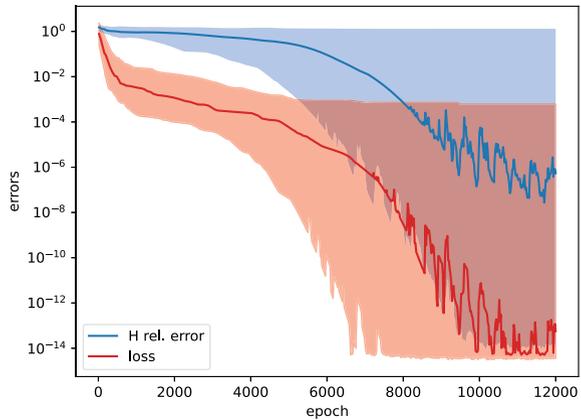
(a) direct parameter fitting (single qubit)
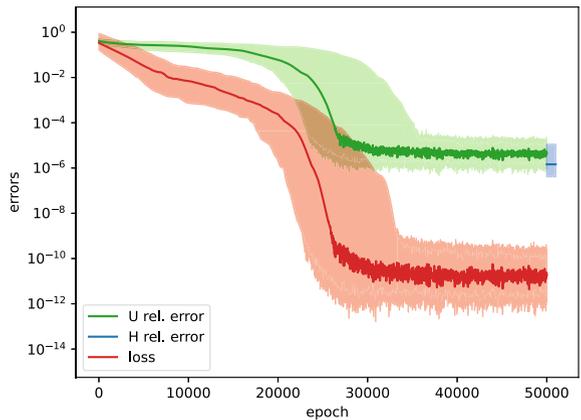
(b) "relaxation" method (single qubit)
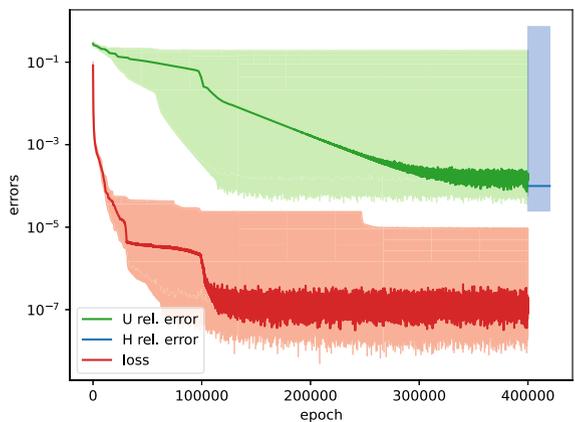
(c) "relaxation" method (two qubits)
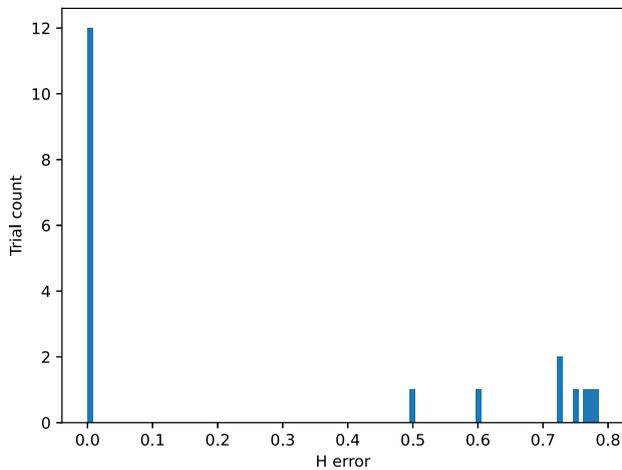
**Fig. 9** Histogram of the final error in the reconstructed Hamiltonian

According to Fig. 8c, the median relative errors of U and the Hamiltonian are below 10^{-3}, but there are still cases where the method cannot find the reference solution at all, even though the loss value is small. An explanation could be that U and H are not uniquely determined. Figure 9 shows a histogram of the final error in the reconstructed Hamiltonian. Despite most trials converging to the correct solution, there seem to be several other extrema where the optimization gets stuck. We leave a clarification of this point for future work.

5.3 Partial (subsystem) measurements

We study the scenario depicted in Fig. 4, namely performing measurements solely on one out of two qubits. The observables are the three Pauli gates here. For the reconstruction to work, we require that two different, precisely characterized initial states, can be prepared, which we choose at random for the numerical experiments.

The ground-truth Hamiltonian is similarly constructed at random, by drawing standard normal-distributed coefficients of Pauli strings, which in sum form the Hamiltonian. We use the time step $\Delta t = \frac{1}{5}$ and $\gamma = 1.15$ here and have heuristically found 12 time-delayed measurements (for each measurement basis) as viable to reliably reconstruct the Hamiltonian. The loss function is the mean-squared error between the model prediction for the measurement averages and the ground-truth values. To avoid getting trapped in local minima during the numerical optimization, we use the best out of 10 attempts (in terms of the loss function) with different random starting points.

The specific observables are the three Pauli gates applied to the first qubit. Figure 10a shows exemplary measurement time trajectories, and Fig. 10b the reconstruction error and loss function (median and 25% quantiles based on 20 random realizations of the overall setup). One observes that a reliable and precise reconstruction of the Hamiltonian (even up to numerical rounding errors) is possible in principle, when neglecting inaccuracies of the measurement process.
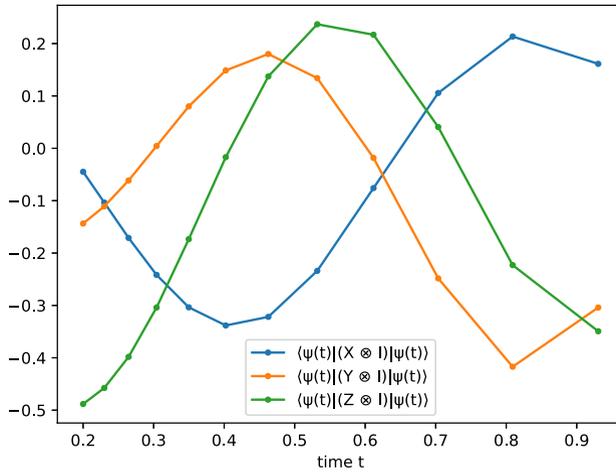
(a) example trajectory
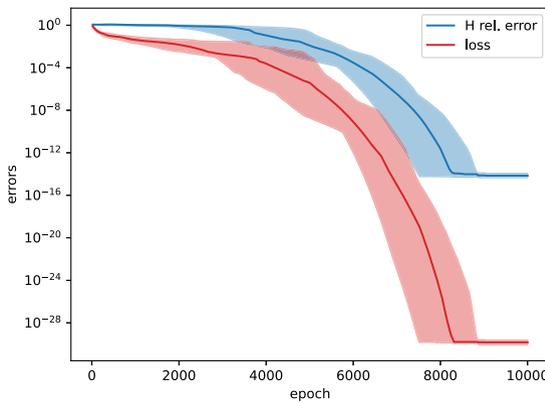
(b) reconstruction error and loss function

Fig. 10 Numerical reconstruction of a generic two-qubit Hamiltonian based on time-delayed measurements of only one of the qubits, cf. Fig. 4. **a** Exemplary measurement time trajectory of the operators $\sigma^x \otimes I$, $\sigma^y \otimes I$ and $\sigma^z \otimes I$. Two random initial states and their respective trajectories are used as input to the reconstruction algorithm. **b** Reconstruction error and loss function based on 20 random realizations

We remark that we have used only a single initial state and less time points for the "relaxation" method (see Fig. 8c), which can explain the seemingly larger errors there.

5.4 Hamiltonian on a lattice with local interactions

We first consider the case of a Hamiltonian (1) with *uniform* parameters (not depending on the lattice site), i.e.,

$$H = -\sum_{\langle j,\ell \rangle} J\, \sigma_j^z \sigma_\ell^z - \sum_{j \in \Lambda} \vec{h} \cdot \vec{\sigma}_j \tag{23}$$

with $J \in \mathbb{R}$ and $\vec{h} \in \mathbb{R}^3$. Thus, the task consists of reconstructing 4 real numbers. The ground-truth values for the following experiment are $J = 1$ and $\vec{h} = (0.5, -0.8, 1.1)$. Λ is a 3×4 lattice with periodic boundary conditions, and the initial state is a single wavefunction with complex random entries (independently normally distributed). We compute the time-evolved quantum state $\psi(t)$ via the KrylovKit Julia package [53] and use the squared entries of $\psi(t)$ as reference Born measurement averages.

For the purpose of reconstructing J and $\vec{h}$, we approximate a time step via a variational quantum circuit as shown in Fig. 5, where the single- and two-qubit gates share their to-be optimized parameters $\tilde{J}$ and $\tilde{\vec{h}}$. For simplicity, we use three uniform time points $\Delta t, 2\Delta t, 3\Delta t$ with $\Delta t = 0.2$ (instead of time points $t_q = \gamma^q \Delta t$), such that the quantum circuit for realizing a single time step can be reused. Note that such a set of time points does not correspond to the diffeomorphism described in Eq. 4. In practice, we found that using uniform time steps worked comparably well for the optimization. This suggests that the numerical method is more flexible in this respect, despite Takens theorem providing the mathematically sound argument to use time-delayed measurements as a basis for the reconstruction in the first place.

We quantify the deviation between the exact Born measurement probabilities, $p_j(t) = |\psi_j(t)|^2$ (which in practice would be obtained from the experimental data), and the probabilities $\tilde{p}_j(t)$ resulting from the trotterized circuit Ansatz, via the Kullback–Leibler divergence:

$$D_{\text{KL}}\left(p(t) \parallel \tilde{p}(t)\right) = \sum_j p_j(t) \log \frac{p_j(t)}{\tilde{p}_j(t)}. \tag{24}$$

Figure 11a visualizes the optimization progress, i.e., the relative errors of J and $\vec{h}$, and the loss function (24). The darker curves show medians over 100 trials of random initial states, and the lighter shades 25% quantiles. The dashed horizontal line is the loss function evaluated at the ground-truth values of J and $\vec{h}$; it is nonzero due to the Strang splitting approximation of a time step. Interestingly, the optimization arrives at an even smaller loss value starting around training epoch 70. We interpret this as an artefact of the Strang splitting approximation—note that around this epoch, the relative error of $\vec{h}$ slightly increases again. We have used the RMSProp optimizer [54] with a learning rate of 0.005 here. In summary, the reachable relative error is around 0.02, and higher accuracy would likely require a smaller time step to reduce the Strang splitting error.
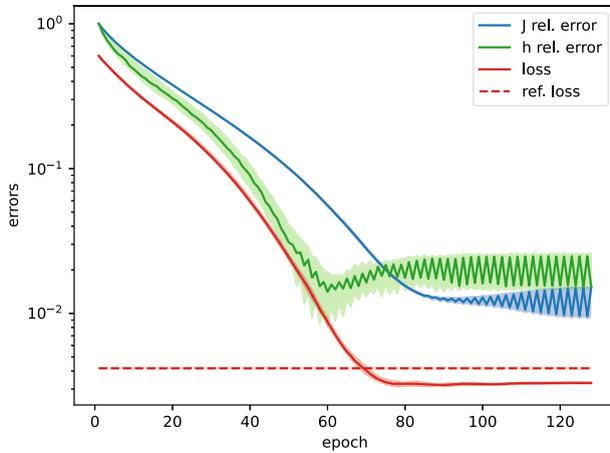
(a) uniform coefficients
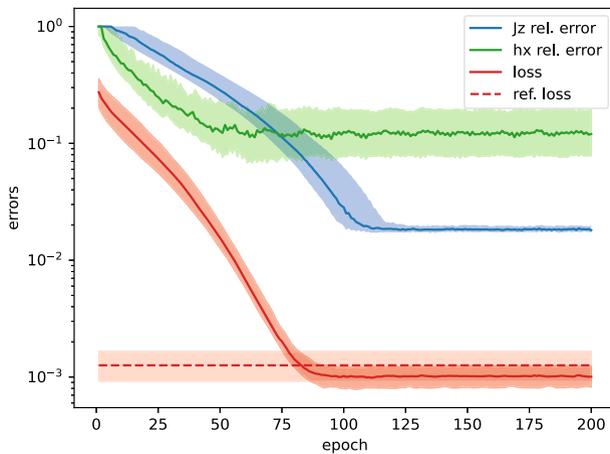
(b) random coefficients (disorder)

Fig. 11 Relative reconstruction error of the Hamiltonian parameters for the 3×4 lattice model in (a) Eq. (23) and (b) Eq. (25), and loss function during training, when fitting a quantum circuit representation of a time step (Fig. 5) to the exact Born measurement averages

Next, we consider a Hamiltonian (1) with disorder (and external field solely in x-direction):

$$H = -\sum_{\langle j, \ell \rangle} J_{j,\ell}\, \sigma_j^z \sigma_\ell^z - \sum_{j \in \Lambda} h_j \cdot \sigma_j^x \qquad (25)$$

with i.i.d. random coefficients $J_{j,\ell}$ and h_j. For the numerical experiment, $J_{j,\ell}$ is uniformly distributed in the interval $[0.8, 1.2]$, and $h_j \sim \mathcal{N}(0, 1/4)$ (normal distribution).

The results are shown in Fig. 11b. As before, we evaluate 100 realizations of the experiment, with a set of coefficients and random initial state drawn independently for each trial. We take the maximum over the relative errors of $J_{j,\ell}$ for all j, ℓ to arrive at the relative error reported in Fig. 11b, and likewise for h_j. The "reference" loss function is the KL divergence evaluated at the ground-truth coefficients. It is nonzero due to Strang splitting errors and fluctuates due to the different random coefficients at each trial.

6 Conclusions and outlook

Our work demonstrates the feasibility of using measurements at different time points to reconstruct the time evolution operator. The approach presented in this work has two main benefits: First, it requires only a single (or few) initial state(s) and reduces the number of different kinds of measurements for the reconstruction of the Hamiltonian. From an experimental point of view, this would be helpful when the experimental setup makes it easier to maintain the time evolution of the system than to implement a variety of measurements. Second, as demonstrated in Sect. 5.3, the method allows for the reconstruction of a Hamiltonian even when only part of the system can be observed.

A natural extension of our work is QPT in the situation of dissipation and noise processes, i.e., a system governed by a Lindblad equation. This would pose the additional challenge of a limited time window for extracting information and a larger number of free parameters. An interesting alternative approach could be a mapping to a unitary evolution with an unobserved environment [55], which would fit into the present framework.

The relaxation method in Sect. 4.2 can be regarded as tool to smoothen the optimization landscape, but an open question is how to guarantee convergence to the correct solution. This becomes particularly relevant for larger systems and the likewise increasing number of free parameters.

A related task for future work is a sensitivity analysis with respect to inaccuracies and noise in the measurement data. A modified version of Takens theorem still applies in this situation [44], but it is not clear how accurate our algorithm can reconstruct H. A first step could be a gradient calculation of the measurement averages with respect to the Hamiltonian parameters.

Data availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Childs, A.M., Chuang, I.L., Leung, D.W.: Realization of quantum process tomography in NMR. Phys. Rev. A **64**, 012314 (2001)
2. O'Brien, J.L., et al.: Quantum process tomography of a controlled-NOT gate. Phys. Rev. Lett. **93**, 080502 (2004). https://doi.org/10.1103/PhysRevLett.93.080502
3. Bialczak, R.C., et al.: Quantum process tomography of a universal entangling gate implemented with Josephson phase qubits. Nat. Phys. **6**(409), 1745–2481 (2010). https://doi.org/10.1038/nphys1639
4. Weinstein, Y.S., et al.: Quantum process tomography of the quantum Fourier transform. J. Chem. Phys. **121**(6117), 0021–9606 (2004). https://doi.org/10.1063/1.1785151
5. Kampermann, H., Veeman, W.S.: Characterization of quantum algorithms by quantum process tomography using quadrupolar spins in solid-state nuclear magnetic resonance. J. Chem. Phys. **122**, 214108 (2005). https://doi.org/10.1063/1.1904595
6. Jullien, T., Roulleau, P., Roche, B., Cavanna, A., Jin, Y., Glattli, D.C.: Quantum tomography of an electron. Nature **514**(603), 1476–4687 (2014)
7. Bisognin, R., et al.: Quantum tomography of electrical currents. Nat. Commun. **10**, 3379 (2019). https://doi.org/10.1038/s41467-019-11369-5
8. Nielsen, M.A., Chuang, I.L.: Quantum Computation and Quantum Information. Cambridge University Press, Cambridge (2010)
9. Torlai, G., Wood, C.J., Acharya, A., Carleo, G., Carrasquilla, J., Aolita, L.: Quantum process tomography with unsupervised learning and tensor networks, arXiv:2006.02424 (2020)
10. Baldwin, C.H., Kalev, A., Deutsch, I.H.: Quantum process tomography of unitary and near-unitary maps. Phys. Rev. A **90**, 012110 (2014)
11. Huang, H., Kueng, R., Preskill, J.: Predicting many properties of a quantum system from very few measurements. Nat. Phys. **16**, 1050 (2020). https://doi.org/10.1038/s41567-020-0932-7
12. Cramer, M., et al.: Efficient quantum state tomography. Nat. Commun. **1**, 149 (2010). https://doi.org/10.1038/ncomms1147
13. Torlai, G., et al.: Neural-network quantum state tomography. Nat. Phys. **14**, 447 (2018). https://doi.org/10.1038/s41567-018-0048-5
14. Flurin, E., Martin, L.S., Hacohen-Gourgy, S., Siddiqi, I.: Using a recurrent neural network to reconstruct quantum dynamics of a superconducting qubit from physical observations. Phys. Rev. X **10**, 011006 (2020)
15. Knee, G.C., Bolduc, E., Leach, J., Gauger, E.M.: Quantum process tomography via completely positive and trace-preserving projection. Phys. Rev. A **98**, 062336 (2018). https://doi.org/10.1103/PhysRevA.98.062336
16. Surawy-Stepney, T., Kahn, J., Kueng, R., Guta, M.: Projected least-squares quantum process tomography. Quantum **6**, 844 (2022). https://doi.org/10.22331/q-2022-10-20-844
17. Rodionov, A.V., Veitia, A., Barends, R., Kelly, J., Sank, D., Wenner, J., Martinis, J.M., Kosut, R.L., Korotkov, A.N.: Compressed sensing quantum process tomography for superconducting quantum gates. Phys. Rev. B **90**, 144504 (2014). https://doi.org/10.1103/PhysRevB.90.144504
18. Nielsen, E., Gamble, J.K., Rudinger, K., Scholten, T., Young, K., Blume-Kohout, R.: Gate set tomography. Quantum **5**, 557 (2021)
19. Zhang, J., Sarovar, M.: Quantum Hamiltonian identification from measurement time traces. Phys. Rev. Lett. **113**, 080401 (2014)
20. Zhang, J., Sarovar, M.: Identification of open quantum systems from observable time traces. Phys. Rev. A **91**, 052121 (2015)
21. Xin, T., Lu, S., Cao, N., Anikeeva, G., Lu, D., Li, J., Long, G., Zeng, B.: Local-measurement-based quantum state tomography via neural networks. NPJ Quant. Inf. **5**, 109 (2019)
22. Che, L., Wei, C., Huang, Y., Zhao, D., Xue, S., Nie, X., Li, J., Lu, D., Xin, T.: Learning quantum Hamiltonians from single-qubit measurements. Phys. Rev. Res. **3**, 023246 (2021)

23. Zhao, D., Wei, C., Xue, S., Huang, Y., Nie, X., Li, J., Ruan, D., Lu, D., Xin, T., Long, G.: Characterizing quantum simulations with quantum tomography on a spin quantum simulator. Phys. Rev. A **103**, 052403 (2021)
24. Sauer, T.: Time Series Prediction: Forecasting The Future And Understanding The Past. In: Chapter: Time series prediction by using delay coordinate embedding, pp. 175–193. Addison-Wesley, Harlow, UK (1994)
25. Berry, T., Cressman, J.R., Greguríc-Ferenĉek, Z., Sauer, T.: Time-scale separation from diffusion-mapped delay coordinates. SIAM J. Appl. Dyn. Syst. **12**, 618 (2013)
26. Dietrich, F., Köster, G., Bungartz, H.-J.: Numerical model construction with closed observables. SIAM J. Appl. Dyn. Syst. **15**, 2078 (2016)
27. Dietrich, F., Kooshkbaghi, M., Bollt, E.M., Kevrekidis, I.G.: Manifold learning for organizing unstructured sets of process observations. Chaos Interdisc. J. Nonlinear Sci. **30**, 043108 (2020)
28. Giannakis, D.: Delay-coordinate maps, coherence, and approximate spectra of evolution operators. Res. Math. Sci. **8**, 8 (2021)
29. Kemeth, F.P., Haugland, S.W., Dietrich, F., Bertalan, T., Hohlein, K., Li, Q., Bollt, E.M., Talmon, R., Krischer, K., Kevrekidis, I.G.: An emergent space for distributed data with hidden internal order through manifold learning. IEEE Access **6**, 77402–77413 (2018)
30. F. P. Kemeth, T. Bertalan, T. Thiem, F. Dietrich, S. J. Moon, C. R. Laing, and I. G. Kevrekidis, Learning emergent PDEs in a learned emergent space, arXiv:2012.12738 (2020)
31. Bertalan, T., Dietrich, F., Mezíc, I., Kevrekidis, I.G.: On learning Hamiltonian systems from data. Chaos Interdisc. J. Nonlinear Sci. **29**, 121107 (2019)
32. Greydanus, S., Dzamba, M., Yosinski, J.: Hamiltonian Neural Networks. In: Wallach, H., Larochelle, H., Beygelzimer, A., Alché-Buc, F.D., Fox, E., Garnett, R. (eds.) Advances in Neural Information Processing Systems. Curran Associates Inc (2019)
33. Koopman, B.O.: Hamiltonian systems and transformation in Hilbert space. Proc. Natl. Acad. Sci. United States Am. **17**, 315 (1931)
34. Koopman, B.O., Neumann, J.V.: Dynamical systems of continuous spectra. Proc. Natl. Acad. Sci. **18**, 255 (1932)
35. Mezíc, I.: Spectral properties of dynamical systems, model reduction and decompositions. Nonlinear Dyn. **41**, 309 (2005)
36. Mezíc, I.: Spectrum of the Koopman operator, spectral expansions in functional spaces, and state-space geometry. J. Nonlinear Sci. **30**, 2091 (2019)
37. Dietrich, F., Thiem, T.N., Kevrekidis, I.G.: On the Koopman operator of algorithms. SIAM J. Appl. Dyn. Syst. **19**, 860 (2020)
38. Ruelle, D., Takens, F.: On the nature of turbulence. Commun. Math. Phys. **20**, 167 (1971)
39. Takens F.: Detecting strange attractors in turbulence, Lecture Notes in Mathematics pp. 366–381 (1981)
40. Whitney, H.: Differentiable manifolds. Ann. Math. **37**, 645 (1936)
41. Packard, N.H., Crutchfield, J.P., Farmer, J.D., Shaw, R.S.: Geometry from a time series. Phys. Rev. Lett. **45**, 712 (1980)
42. Aeyels, D.: Generic observability of differentiable systems. SIAM J. Control Optim. **19**, 595 (1981)
43. Sauer, T., Yorke, J.A., Casdagli, M.: Embedology. J. Statistical Phys. **65**, 579 (1991)
44. Stark, J., Broomhead, D., Davies, M., Huke, J.: Takens embedding theorems for forced and stochastic systems. Nonlinear Anal. Theory Methods Appl. **30**, 5303 (1997). https://doi.org/10.1016/S0362-546X(96)00149-6
45. Marvasti, F. (ed.): Nonuniform Sampling: Theory and Practice, Nonuniform Sampling: Theory and Practice. Springer, Boston, MA (2001)
46. Kraus, B., Cirac, J.I.: Optimal creation of entanglement using a two-qubit gate. Phys. Rev. A **63**, 062309 (2001)
47. Zhang, J., Vala, J., Sastry, S., Whaley, K.B.: Geometric theory of nonlocal two-qubit operations. Phys. Rev. A **67**, 042313 (2003)
48. Gamel, O.: Entangled Bloch spheres: Bloch matrix and two-qubit state space. Phys. Rev. A **93**, 062320 (2016)
49. Huang, Q., Mendl, C.B.: Efficient quantum circuit simulation using a multi-qubit Bloch vector representation of density matrices, arXiv:2103.13962 (2021),
50. McLachlan, R.I., Quispel, G.R.W.: Splitting methods. Acta Numer. **11**, 341 (2002)
51. Vidal, G.: Efficient simulation of one-dimensional quantum many-body systems. Phys. Rev. Lett. **93**, 040502 (2004)

52. Kingma, D.P. Ba, J.: Adam: A method for stochastic optimization, in *3rd International Conference on Learning Representations, ICLR 2015* (2015), arXiv:1412.6980
53. https://github.com/Jutho/KrylovKit.jl (2021)
54. Hinton, G.: Neural networks for machine learning—Lecture 6a: Overview of mini-batch gradient descent, (2014) unpublished, https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf
55. Pokorny, F., Zhang, C., Higgins, G., Cabello, A.: Tracking the dynamics of an ideal quantum measurement. Phys. Rev. Lett. **124**, 080401 (2020)