nature physics

SUPPLEMENTARY INFORMATION

DOI: 10.1038/NPHYS4244

In the format provided by the authors and unedited.

Efficient tomography of a quantum many-body system

B. P. Lanyon,[1,2,*] C. Maier,[1,2,*] M. Holzäpfel,[3] T. Baumgratz,[3,4,5] C. Hempel,[2,6] P. Jurcevic,[1,2] I. Dhand,[3] A. S. Buyskikh,[7] A. J. Daley,[7] M. Cramer,[3,8] M. B. Plenio,[3] R. Blatt,[1,2] and C. F. Roos[1,2]

[1]*Institut für Quantenoptik und Quanteninformation, Österreichische Akademie der Wissenschaften, Technikerstr. 21A, 6020 Innsbruck, Austria*
[2] *Institut für Experimentalphysik, Universität Innsbruck, Technikerstr. 25, 6020 Innsbruck, Austria*
[3]*Institut für Theoretische Physik and IQST, Albert-Einstein-Allee 11, Universität Ulm, 89069 Ulm, Germany*
[4]*Clarendon Laboratory, Department of Physics, University of Oxford, Oxford OX1 3PU, United Kingdom*
[5]*Department of Physics, University of Warwick, Coventry CV4 7AL, United Kingdom*
[6]*ARC Centre for Engineered Quantum Systems, School of Physics, The University of Sydney, Sydney, New South Wales 2006, Australia.*
[7]*Department of Physics and SUPA, University of Strathclyde, Glasgow G4 0NG, UK*
[8]*Institut für Theoretische Physik, Leibniz Universität Hannover, Hannover, Germany*

(Dated: July 7, 2017)

Contents

[*]These authors contributed equally to this work.

I. TRAPPED-ION QUANTUM SIMULATOR

A. Multi- and single-qubit rotations

We identify the two electronic Zeeman states $|S_{1/2}, m = +1/2\rangle$ and $|D_{5/2}, m' = +5/2\rangle$ of trapped $^{40}Ca^{+}$ ions with the $|\downarrow\rangle$ and $|\uparrow\rangle$ states of spin-1/2 particles, respectively. These atomic states are coupled by an electric quadrupole transition at the optical wavelength of 729 nm. The quantum states are coherently manipulated using a Ti:Sa CW laser with a linewidth of about 1 Hz, Two laser beam paths are employed for this. First, a *global beam* illuminates the ion string approximately equally, from a direction perpendicular to the ion string axis and at an angle approximately half way between x and y. Consider the standard Pauli spin operators on spin j: σ_x^j, σ_y^j and σ_z^j. The global beam is used to perform global σ_x and σ_y rotations simultaneously on all spins, e.g. $G_x(\theta) = \exp(-i\theta \sum_{j=1}^{N} \sigma_x^j)$ and $G_y(\theta) = \exp(-i\theta \sum_{j=1}^{N} \sigma_y^j)$. Second, a single-ion-focused beam comes in parallel to the global beam but from the opposite direction. The direction of this beam can be switched to have its focus pointing at different ions within 12 μs, using an acousto-optic deflector. This addressed beam is frequency-detuned by about 80 MHz from the spin transition and thereby performs an AC Stark rotation on the addressed ion j given by $R_z^j(\theta) = \exp(-i\theta\sigma_z^j)$. The combination of a global resonant beam and a focused detuned beam, inducing AC-Stark shifts for carrying out arbitrary single-spin rotations, has the advantage of not requiring laser beam paths whose optical path length difference is interferometrically stabilized. For an overview of the use of global and addressed beams to manipulate ionic spins (qubits) see [1].

B. Preparing the Néel state

The Néel state is prepared using a combination of global and addressed pulses. The addressed operation $A_z(\theta) = \exp(-i\theta \sum_{[j=[1,3,5..N-1]} \sigma_z^j)$ is employed, corresponding to (ideally) equal rotations around the z axis of a subset of spins in the string performed sequentially. To create the Néel state from the initial state $|\downarrow\rangle^{\otimes N} = |\downarrow, \downarrow, \downarrow ...\rangle$ requires flipping every second spin to the $|\uparrow\rangle$ state. This is done by subjecting the initial state $|\downarrow\rangle^{\otimes N}$ to the following composite pulse sequence $G_x(\pi/4)G_y(\pi/4)A_z(\pi/2)G_y(\pi/4)G_x(\pi/4)\,|\downarrow\rangle^{\otimes N}$ that maps the initial state onto the Néel state $|\uparrow, \downarrow, \uparrow ...\rangle$. To first order, the state created by this sequence is insensitive to errors in the rotation angles of the x and y rotations. These errors result from the unequal coupling strength of the global beam across the string, due to its Gaussian transverse intensity profile. We prepare the Néel state for 8 and 14 ion strings. The fidelities of these states with the ideal Néel state is obtained by directly measuring in the z-basis on all spins. For 8 spins, the fidelity with the Néel state is 0.967 ± 0.006, corresponding to an average error per spin of $-\log_2(0.967)/8 = 0.006 \pm 0.001$. For 14 spins, the fidelity with the Néel state is 0.89 ± 0.01, corresponding to an average error per spin of $-\log_2(0.89)/14 = 0.012 \pm 0.001$. Clearly the error-per-particle is significantly larger for the 14 spin initial state, than for 8. Note that the aforementioned directly measured fidelities of the initial state agree well with the certified lower bounds obtained for measurements on single sites, via MPS tomography (see main text). The increase in error-per-particle for the 14-spin initial state is likely due to clipping of our single-ion-focused laser beam on optics as it tries to access ions at the ends of longer strings, leading to larger intensity fluctuations. This should be straightforward to overcome in a new optical setup.

II. MODELLING THE SIMULATOR

A. Ideal simulator model

In the main paper we give approximate spin-spin interaction power-law ranges (α values), light-like cones for information spreading (Figures 2a and 4a) and, in this supplementary material, we compare data with a theoretical model for our simulator. In this section we explain how we do this modelling.

The full Hamiltonian used for ideal modelling of our simulator dynamics is given by an Ising-type model with a transverse magnetic field:

$$H_{\mathrm{Ising}} = \hbar \sum_{i<j} J_{ij}\, \sigma_i^x \sigma_j^x + \hbar \sum_{i=1}^{N} (B + B_i)\, \sigma_i^z \quad . \tag{1}$$

The $N \times N$ spin-spin coupling matrix J_{ij}, parametrising this model, depends on the ion string vibrational mode frequencies, eigenvectors, detuning from the laser fields, ionic mass and laser-ion coupling strength (the dependency of J_{ij} on the experimental parameters is elaborated in the supplementary material of [2]). All else held constant, the interaction range (modelled by a power law, see section II B) can be changed

by a single experimental parameter: the detuning of laser fields from the motional resonances. We independently measure all the aforementioned parameters in our experimental system and thereby derive the spin-spin coupling matrix J_{ij}. The transverse field consists of an overall constant field B and site-dependent perturbations B_i. These small inhomogeneities in the transverse field act like local potential barriers to spin excitations hopping around the string. They come from e.g. electric quadrupole shifts which vary along the string, a residual magnetic field gradient and field curvature along direction of the string (in addition to our standard constant 4 Gauss field), and AC Stark shifts of the spin transitions due to the presence of laser fields with intensity gradients across the string (Gaussian beam profiles). After optimizing the experimental apparatus, the inhomogeneities are small (compared to the spin-spin coupling strength) and play little role in obtaining an accurate description of our system dynamics at the evolution times considered in the main paper.

In the limit $B \gg |J_{ij}|$, which is upheld in all our experiments, the Ising Hamiltonian (1) is equivalent to an XY model in a transverse field. Including the field inhomogenuities, the effective Hamiltonian describing our experiments is given by

$$H_{XY}=\hbar \sum_{i<j} J_{ij} \left(\sigma_i^+ \sigma_j^- + \sigma_i^- \sigma_j^+\right) + \hbar \sum_{i=1}^{N} (B + B_i)\, \sigma_i^z \quad , \tag{2}$$

with σ_i^+ (σ_i^-) the spin raising (lowering) operator for spin i.

In previous work we showed that the simulator dynamics is well described by this model in the low-excitation regime, that is, when the initial state contains at most one [2] or two [3] quasiparticle excitations. In this work we explore the dynamics of highly excited initial states: the Néel state contains $N/2$ excitations. Note that the XY Hamiltonian preserves the excitation number throughout dynamical evolutions (subspaces with different excitation number are decoupled).

Time-evolved model simulator states are calculated by brute force matrix exponentiation for up to 8 spins, e.g. $|\phi(t)\rangle = \exp(-iH_{\mathrm{Ising}}t)\,|\phi(0)\rangle$. For 14 spins, this approach takes hours and hours to run, using the computers that we have readily available. Therefore, it was time efficient for 14 spins to use the Krylov subspace projection methods (Arnoldi and Lanczos processes) which, in the case of sparse Hamiltonians, give a substantial speed up and well controlled error bounds [4].

B. Interaction range in experiments

The realised spin-spin interactions are to a good approximation described by a power law dependence falling of with distance $|i - j|$: $J_{ij} \propto |i - j|^{-\alpha}$ with decay parameter α. In the experiment there are two ways to tune the interaction range: either by varying the laser detuning from the motional resonances or by bunching up or fanning out the transverse modes in frequency space. Here the detuning is directly chosen whereas the mode-bunching depends on the effective trapping parameters (therewith also on the number of ions). Comparing the experimental coupling strengths as a function of the distance with ideal power-law decay shows that not possible to extract an unambiguous decay parameter α by a direct fit in real space [2]. However, an effective value for α can be estimated by fitting the eigenmode spectrum (or quasiparticle dispersion relation) of our system with the eigenmode spectrum for power-law interactions [2]. The power-law exponent α yielding the best fit provides an estimate for the interaction range: $\alpha = 1.58$ (8 spins) and $\alpha = 1.27$ (14 spins).

III. MEASURING AND RECONSTRUCTING LOCAL REDUCTIONS

In this section, we describe the measurements performed in the experiment and how these measurements are employed in the analysis. MPS tomography requires the ability to estimate the local reduced density matrices of all blocks of k neighbouring spins. On a linear chain of N spins, there are $N - k + 1$ such blocks. A straightforward method of reconstructing all these blocks requires a total of $(N - k + 1)4^k$ measurements, each performed on one of the $N - k + 1$ local blocks of k spins. Instead, we perform 3^k measurements, each on the entire system of N spins and use these measurements to infer the local reductions. In this section, we describe these 3^k measurements, show that they suffice and how they are used.

A. Chosen measurement setting

Here we recall the straightforward method of reconstructing the local reduced density matrices. One can obtain the density matrix of k spins from the expectation values of a linearly independent set of 4^k

observables. The set of all k-fold tensor products of the three Pauli operators $X = \sigma_x$, $Y = \sigma_y$, $Z = \sigma_z$ and the identity operator $\mathbb{1}$ provides one such set. For example, for $k = 2$ spins, the density matrix can be obtained from the following 16 expectation values

$$\begin{array}{cccc} \langle X_1X_2\rangle & \langle X_1Y_2\rangle & \langle X_1Z_2\rangle & \langle X_1\mathbb{1}_2\rangle \\ \langle Y_1X_2\rangle & \langle Y_1Y_2\rangle & \langle Y_1Z_2\rangle & \langle Y_1\mathbb{1}_2\rangle \\ \langle Z_1X_2\rangle & \langle Z_1Y_2\rangle & \langle Z_1Z_2\rangle & \langle Z_1\mathbb{1}_2\rangle \\ \langle \mathbb{1}_1X_2\rangle & \langle \mathbb{1}_1Y_2\rangle & \langle \mathbb{1}_1Z_2\rangle & \langle \mathbb{1}_1\mathbb{1}_2\rangle . \end{array} \tag{3}$$

In order to obtain the expectation value of say $\langle Z_1X_2\rangle$, the following measurement would be performed in the experiment: Spin 1 is measured in the eigenbasis of Z while spin 2 is measured in the eigenbasis of X. We refer to this measurement setting as $[Z, X]$; the eigenbasis of the i-th vector element provides the measurement basis for the i-th spin. The measurement setting $[Z, X]$ has four distinguishable outcomes (resolved on the CCD camera in our experiment). We obtain spin up $|\uparrow\rangle$ or spin down $|\downarrow\rangle$ in the Z basis for spin 1 and spin up $|\uparrow\rangle$ or spin down $|\downarrow\rangle$ in the X basis for spin 2. By repeating the measurement $[Z, X]$ many times, we can estimate the four outcome probabilities $p_{\uparrow\uparrow}$, $p_{\uparrow\downarrow}$, $p_{\downarrow\uparrow}$, $p_{\downarrow\downarrow}$.

These probabilities can be used not only for extracting the expectation value $\langle Z_1X_2\rangle$, but also for extracting the expectation values $\langle Z_1\mathbb{1}_2\rangle$ and $\langle \mathbb{1}_1X_2\rangle$. This insight generalises to $k \geq 2$ spins and to more general measurement settings, and it enables us to estimate the expectation values of the 4^k measurement observables in Eq. (3) from only 3^k measurement settings.

For each local block of k spins, the following 3^k measurement settings suffice. Each of the k spins requires measurement in the basis of the three Pauli operators, thus leading to 3^k measurement settings. For instance, consider $k = 2$. In this case, the $3^k = 9$ measurement settings are given by

$$\begin{array}{ccc} [X, X] & [X, Y] & [X, Z] \\ [Y, X] & [Y, Y] & [Y, Z] \\ [Z, X] & [Z, Y] & [Z, Z] \end{array} \tag{4}$$

Each of the 3^k measurement settings has 2^k distinguishable outcomes. In total, we estimate $3^k \times 2^k = 6^k$ outcome probabilities. Each of the 4^k expectation values required for obtaining the local reduced density matrices can be estimated from this set of 6^k outcome probabilities. Therefore, the set of 6^k outcome probabilities is sufficient to estimate a density matrix on k spins [21].

The 3^k measurement settings described above are to be repeated for each of the $N-k+1$ local blocks on the chain. Independent measurements for the local blocks would require $(N - k + 1)3^k$ measurement settings, where measurements are performed on the local blocks and remaining spins are ignored. However, a more judicious choice can provide the required information with fewer measurement settings.

We choose a set of 3^k total measurement settings such that measurements are performed on the entire spin chain rather than just the local blocks. We repeat each of the 3^k measurement settings on k spins along the chain. Specifically, for each of the 3^k measurement settings, we split the system into $\lceil N/k \rceil$ blocks and replicate the same measurement settings on each of the blocks. For instance, the case of $k = 2$ requires the measurement settings

$$\begin{array}{ccc} [X, X, X, X, \dots] & [X, Y, X, Y, \dots] & [X, Z, X, Z, \dots] \\ [Y, X, Y, X, \dots] & [Y, Y, Y, Y, \dots] & [Y, Z, Y, Z, \dots] \\ [Z, X, Z, X, \dots] & [Z, Y, Z, Y, \dots] & [Z, Z, Z, Z, \dots]. \end{array} \tag{5}$$

In our experiment, we set $k = 3$, i.e., we perform measurements in $3^3 = 27$ settings. Formally, we perform measurements in the 3^k with $k = 3$ different measurement bases

$$\left\{ [A_1, \dots, A_k, A_1, \dots, A_k, \dots]: \quad A_i \in \{X, Y, Z\}, \quad i \in \{1, \dots, k\} \right\} \tag{6}$$

on N spins. Each of the chosen 3^k measurement settings has 2^N distinguishable outcomes. $m = 1000$ outcomes (which could take any of the 2^N unique values) were recorded for each of the 27 settings.

To summarize, we choose 3^k measurement settings comprising repetitions of local 3^k Pauli measurements. This brings the total measurement setting requirement down from $(N-k+1)4^k$ local measurements to 3^k measurements on the entire chain.

B. Using measurement outcomes

Here we describe how the outcomes obtained from measurement settings in Eq. (6) are used in the subsequent analysis. The measurement data are used either (i) to reconstruct the state via certified MPS

tomography or (ii) to estimate local reduced density matrices on k spins, for instance to estimate 3-spin entanglement.

The measurement data are input to the certified MPS tomography algorithms (Section IV) after converting to one out of the following two forms. The first form is that of $(N-k+1)4^k$ local expectation values

$$\Big\{ \langle A_{s+1} \otimes \cdots \otimes A_{s+k} \rangle : \quad A_{s+i} \in \{\mathbb{1}, X, Y, Z\} \Big\}, \tag{7}$$

where $s \in \{0, \ldots, N-k\}$ indicates the position of the local block and $i \in \{1, \ldots, k\}$ labels sites within the respective local block. An alternate but equivalent form of the input to certified tomography is that of the outcome probabilities of the 6^k non-identity Pauli measurements performed on each of the $N-k+1$ local blocks. Formally, the following $(N-k+1)6^k$ local outcome probabilities are estimated:

$$\Big\{ \langle P_{s+1,a_1,b_1} \otimes \cdots \otimes P_{s+k,a_k,b_k} \rangle : \quad a_i \in \{X, Y, Z\}, \quad b_i \in \{-1, +1\}, \\ s \in \{0, \ldots, N-k\}, \quad i \in \{1, \ldots, k\} \Big\}, \tag{8}$$

where $P_{j,a_i,b_i} = |a_i b_i\rangle_j \langle a_i b_i|$ projects spin j onto the eigenvector $|a_i b_i\rangle$ of the Pauli operator a_i (= X, Y or Z) with eigenvalue b_i. The methods for estimating quantities (7) or 8 from the measurement data are detailed in Section IV.

The second use of the measurement data is to reconstruct local reduced density matrices of k neighbouring spins, or in other words, to perform full quantum state tomography of the local blocks. We use maximum likelihood estimation [5] to obtain density matrix estimates ρ^k_{qst} from the local k-spin outcome probabilities. Quantities of interest are computed from the density matrix estimates, e.g., the entanglement measures in Figures 2c,d and 4b of the main text or Entropy in FIG. S8. Error bars in quantities derived from local reconstructions are obtained from standard Monte-Carlo simulations of quantum projection noise.

IV. CERTIFIED MPS TOMOGRAPHY

In this section, we describe our method for certified MPS tomography, which is based on the results of Refs. [6] and [7]. We use the modified SVT algorithm from [6] and the scalable maximum likelihood estimation method for quantum state tomography from [7] to obtain an estimate of the unknown lab state from experimental measurement data. Because the experimentally measured observables do not contain complete information on the unknown state, an additional step is necessary to verify the correctness of the result. We use the assumption-free lower bound on the fidelity between our estimate and the unknown state in the laboratory from Ref. [6] for this purpose; we call such a lower bound on the fidelity a *certificate*. We refer to the combined procedure of MPS tomography and certification as *certified MPS tomography*.

A. Details of procedure

We discuss how to reconstruct and certify a pure quantum state on N qubits from measurements which allow for the reconstruction of all reduced density matrices of k neighbouring qubits. As there are $N-k+1$ contiguous blocks of size k on a linear chain of length N, the total measurement effort scales at most linearly with N. However, even a number of measurements settings constant in N is sufficient if suitably chosen measurements on all N qubits are performed. Here, we use the outcomes of $q = 3^k$ settings measured in the quantum simulator experiment (Sec. III). Our discussion is formulated for N qubits (spin-$\frac{1}{2}$ particles), but it equally applies to N qudits.

FIG. S1 provides a schematic overview over the following subsections. We will proceed as follows: Measurement data is split into two parts; the first part is used for MPS tomography while the second part is used for certification (Sec. IV A 1). We apply existing MPS tomography algorithms to obtain an initial estimate $|\psi_{\mathrm{est}}\rangle$ of the unknown state ρ_{lab} in the laboratory (Sec. IV A 2). From the initial estimate, a family of candidates for a so-called parent Hamiltonian is constructed and one of them is selected, denoted by H. The parent Hamiltonian H provides the certificate and the certified estimate $|\psi^k_c\rangle$ of the unknown state ρ_{lab} is given by the ground state $|\psi_{\mathrm{GS}}\rangle$ of H (Sec. IV A 3). The general approach to obtain the measurement uncertainty of the fidelity lower bound is derived (Sec. IV A 4). The fact that local probabilities have been obtained from global measurements in the experiment complicates obtaining the measurement uncertainty of the fidelity lower bound; remaining technical details related to this issue are covered at the end (Sec. IV A 5).

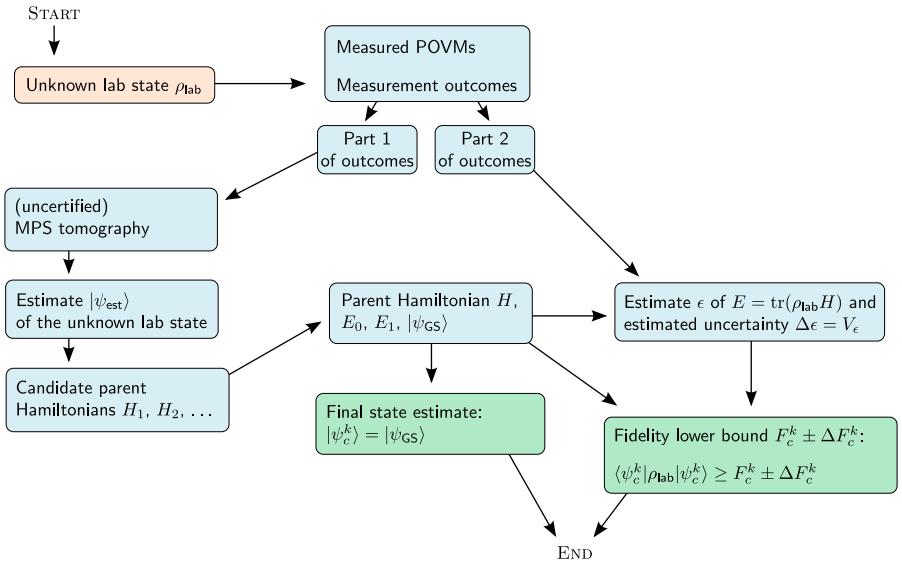


FIG. S1: **Schematic overview of certified MPS tomography.** This scheme illustrates the certified MPS tomography process discussed in Sec. IV A. in Supplementary Material. Measurement data from the unknown state ρ_{lab} is split into two parts. The first part of the data is used to obtain a certified estimate $|\psi_c^k\rangle$ and a parent Hamiltonian H, while the second part is used to obtain the value $F_c^k \pm \Delta F_c^k$ of the fidelity lower bound.

1. *Measurements*

Our method begins with measurements on the unknown lab state ρ_{lab}. We use the data from the measurements described in Sec. III. The data comprise $m = 1000$ outcomes for each of the $q = 3^k$ different measurement settings (6) on N qubits. The samples are split into two parts of 500 samples each. The first part is used to obtain an estimate of the unknown lab state, while the second part is used to obtain the certificate, i.e. the lower bound on the fidelity between the unknown lab state and our estimate of the lab state. This splitting is performed to avoid any risk of overestimating fidelity by constructing or selecting a parent Hamiltonian (see below) which is tuned to the particular set of statistical fluctuations in a single set of measurement data. Future work could study whether one can use measurement data in a more efficient way.

2. *Uncertified MPS tomography*

We obtain an estimate of the unknown lab state by combining two efficient MPS tomography algorithms [8]. We use the modified SVT algorithm from Ref. [6] to obtain a pure state. This pure state is used as start vector for the iterative likelihood maximization scheme over pure states from Ref. [7]. The computation time required for both algorithms scales polynomially with the number of qubits N.

The input for the modified SVT algorithm consists of the local expectation values from Eq. (7) (Sec. III). Alternatively, one can specify the input as estimates of the local reduced states on k neigbouring qubits (the difference is only an operator basis change in Hilbert-Schmidt space). In our implementation, we choose the latter and convert the local outcome probabilities from Eq. (8) into local reduced states with linear inversion. Linear inversion uses the Moore–Penrose pseudoinverse of the map $\mathcal{M}_s$ given in Eq. (36) on page 12.

The input for the scalable maximum likelihood estimation (MLE) scheme consists of the local outcome probabilities from Eq. (8). In principle, one could perform MLE with more information than only the local outcome probabilities. For example, one could extract all pairwise correlations available from the existing measurement data and provide them to the MLE algorithm. This is another avenue for further work. The scalable MLE algorithm returns an initial estimate of the unknown lab state, denoted by $|\psi_{\text{est}}\rangle$.

In both methods mentioned above, the pure state is represented as a matrix product state [9] with limited bond dimension Δ. In some cases, we observe that the fidelity lower bound obtained at the end of the scheme decreases if we allow for a larger bond dimension Δ. Presumably, this is a result of statistical noise adding spurious correlations to our state estimate, which is prevented by lowering the

bond dimension. For 8 qubits, we use $\Delta = 2$ for $t \leq 2\,\mathrm{ms}$ and $\Delta = 4$ for all remaining times. For 14 qubits, we use $\Delta = 16$ for all times.

3. *Fidelity lower bound and selection of a parent Hamiltonian*

In the last subsection, we have obtained the initial estimate $|\psi_{\mathrm{est}}\rangle$ of the unknown state ρ_{lab}. At this point, we do not know whether $|\psi_{\mathrm{est}}\rangle$ is close to the unknown state ρ_{lab}. In this section, we obtain a certified estimate $|\psi_c^k\rangle$ and a lower bound on the fidelity between $|\psi_c^k\rangle$ and ρ_{lab}. The lower bound is provided by a parent Hamiltonian H and its ground state $|\psi_{\mathrm{GS}}\rangle$ is the certified estimate, i.e. $|\psi_c^k\rangle = |\psi_{\mathrm{GS}}\rangle$.

A parent Hamiltonian of a pure state $|\psi_{\mathrm{GS}}\rangle$ is any Hermitian linear operator H such that $|\psi_{\mathrm{GS}}\rangle$ is the ground state of H. Let $E_0 \leq E_1 \leq \ldots$ be the eigenvalues of H in ascending order with degenerate eigenvalues repeated according to their multiplicities. In addition, let the ground state be non-degenerate, i.e. $E_0 < E_1$. Then, a lower bound to the fidelity between the ground state $|\psi_{\mathrm{GS}}\rangle$ and any other pure or mixed state ρ_{lab} is given by [6]

$$\langle\psi_{\mathrm{GS}}|\rho_{\mathrm{lab}}|\psi_{\mathrm{GS}}\rangle \geq 1 - \frac{E - E_0}{E_1 - E_0} \tag{9}$$

with the energy $E = \mathrm{tr}(H\rho_{\mathrm{lab}})$ of the unknown state ρ_{lab} in terms of the parent Hamiltonian H. Note that H usually is artificial and unrelated to any energy in the physical system in the laboratory. If H is a sum of local terms—that is, terms acting non-trivially only on k neighbouring spins—the measurements described above suffice to obtain the energy E and the fidelity lower bound. It remains to find such a parent Hamiltonian, given the initial estimate $|\psi_{\mathrm{est}}\rangle$.

In order to find a larger-than-zero fidelity lower bound, we have to find a parent Hamiltonian which must satisfy two conditions: (i) the ground state $|\psi_{\mathrm{GS}}\rangle$ must be close to the initial estimate $|\psi_{\mathrm{est}}\rangle$ and (ii) the gap $E_1 - E_0$ between the two smallest eigenvalues must be much bigger than the measurement uncertainty about the value of E. If condition (ii) is not satisfied, we will not learn anything about the fidelity. The following qualitative argument illustrates that condition (i) is necessary as well: If the initial estimate $|\psi_{\mathrm{est}}\rangle$ is far away from the lab state ρ_{lab}, we do not try to find a useful parent Hamiltonian because it would be unlikely to succeed. Therefore, we only consider the case where the fidelity of the initial estimate $|\psi_{\mathrm{est}}\rangle$ and the lab state ρ_{lab} is high: In this case, a high fidelity between the ground state $|\psi_{\mathrm{GS}}\rangle$ of the Hamiltonian and lab state ρ_{lab} is possible only if the fidelity of $|\psi_{\mathrm{est}}\rangle$ and $|\psi_{\mathrm{GS}}\rangle$ is high as well.

First, we attempt to construct a parent Hamiltonian whose ground state $|\psi_{\mathrm{GS}}\rangle$ is equal to $|\psi_{\mathrm{est}}\rangle$. If the matrix product state $|\psi_{\mathrm{est}}\rangle$ belongs to the class of *injective MPS* for a certain number k of neighbouring spins [10, 11], then the operator

$$H = \sum_{s=1}^{N-k+1} \mathbb{1}_{1,\ldots,s-1} \otimes P_{\ker(\rho_s)} \otimes \mathbb{1}_{s+k,\ldots,N}, \qquad \rho_s = \mathrm{tr}_{1,\ldots,s-1,s+k,\ldots,n}(|\psi_{\mathrm{est}}\rangle\langle\psi_{\mathrm{est}}|) \tag{10}$$

has $|\psi_{\mathrm{est}}\rangle$ as its unique ground state [10, 11]; $P_{\ker(\rho_s)}$ is the orthogonal projection onto the kernel of the reduced density matrix ρ_s on the k neighbouring spins from s to $s+k-1$. The injectivity property implies certain restrictions on possible combinations between the bond dimension D of $|\psi_{\mathrm{est}}\rangle$ and the number of neighbours k. However, the initial estimate $|\psi_{\mathrm{est}}\rangle$ generally is not an injective MPS for our given k and Eq. (10) provides a parent Hamiltonian with degenerate ground state. This violates condition (ii) and leads to a zero fidelity lower bound.

To mitigate the problem of ground-state degeneracy of H from Eq. (10) we construct a parent Hamiltonian in a different way. Specifically, we introduce a threshold $\tau \geq 0$ and obtain candidates for parent Hamiltonians from

$$H = \sum_{s=1}^{N-k+1} \mathbb{1}_{1,\ldots,s-1} \otimes P_{\ker(T_\tau(\rho_s))} \otimes \mathbb{1}_{s+k,\ldots,N} \tag{11}$$

where the thresholding function T_τ replaces eigenvalues of ρ_s smaller than or equal to some threshold τ by zero. Depending on the value of the threshold, the initial estimate $|\psi_{\mathrm{est}}\rangle$ may or may not be a ground state of H. We construct a set of candidates $H_1, H_2, \ldots$ for parent Hamiltonians by considering all possible values of $\tau \geq 0$. We then try to find a compromise between conditions (i) and (ii) from above, $|\psi_{\mathrm{est}}\rangle$ and $|\psi_{\mathrm{GS}}\rangle$ being similar and a large gap, by choosing the operator H which minimizes

$$c\mathcal{D}(|\psi_{\mathrm{est}}\rangle, |\psi_{\mathrm{GS}}\rangle) - (E_1 - E_0) \tag{12}$$

where $c > 0$ is some constant and

$$\begin{aligned}\mathcal{D}(|\psi\rangle, |\tilde{\psi}\rangle) &\stackrel{\mathrm{def}}{=} \| |\psi\rangle\langle\psi| - |\tilde{\psi}\rangle\langle\tilde{\psi}| \|_1/2 \\ &= \sqrt{1 - |\langle\psi|\tilde{\psi}\rangle|^2}\end{aligned} \tag{13}$$

is the trace distance [12]. We obtain a valid fidelity lower bound for any value of the constant c. However, we may obtain a very small lower bound or a lower bound associated with a large measurement uncertainty for some choices of this constant. We have used the value $c = 5$ and we do not observe significantly higher fidelity lower bounds for other values of this constant. The results presented in the main text, in particular Fig. 3a, show that this approach can provide useful fidelity lower bounds. Modifying Eq. (12) or choosing a more optimal value for c in connection with the discussion in Corollary 5 on page 31 has the potential to provide improved fidelity lower bounds.

We use the following numerical tools to compute eigenvalues and eigenvectors of parent Hamiltonians. For 8 qubits, the full spectrum of the parent Hamiltonian candidates is computed with the library function `numpy.linalg.eigh()` from SciPy [13]. For 14 qubits, a DMRG-like iterative MPS ground state search with local optimization on two neighbouring qubits [9, Section 6.3] is used to obtain two eigenvectors of the smallest eigenvalue(s). We use the functions `mineig()` and `mineig_sum()` available from the Python library mpnum [14]. The second eigenvector is obtained as ground state of $H' = H + 5\,|\psi_{\mathrm{GS}}\rangle\langle\psi_{\mathrm{GS}}|$. For both eigenvectors, the quantity $\langle\psi|H^2|\psi\rangle - (\langle\psi|H|\psi\rangle)^2$ is monitored to ensure sufficient convergence of the iterative search. It should be noted that, strictly speaking, DMRG-like algorithms only provide upper bounds on smallest eigenvalues, but in practice they have been observed to be very reliable [9]. The results from a first low-precision eigenvalue computation with MPS bond dimension 8 is used to select a parent Hamiltonian. The eigenvalues from a second high-precision eigenvalue computation with MPS bond dimension 24 is used to obtain the certificate.

As the certified estimate $|\psi_c^k\rangle = |\psi_{\mathrm{GS}}\rangle$ is the result of an eigenvector computation, the eigenvector computation determines the maximal bond dimension of $|\psi_{\mathrm{GS}}\rangle$: For 8 qubits, its bond dimension may reach the maximal value of 16 and for 14 qubits, its bond dimension can be up to 24. The case $k = 1$ is an exception; there, it is easy to see that a non-degenerate ground state must be a product state (bond dimension 1).

4. *Statistical analysis of the fidelity lower bound*

The fidelity lower bound in Eq. (9) relies on using the parent Hamiltonian and information about the unknown state ρ_{lab} to estimate the energy $E = \mathrm{tr}(H\rho_{\mathrm{lab}})$. The information about the lab state is obtained from a finite number of measurement outcomes distributed according to an unknown probability distribution. This leads to uncertainty in our estimate of the energy E. To determine this uncertainty in E, we construct an estimator $\epsilon(D)$ for E, where D represents the measurement data. The term *estimator* refers to a function that uses values of random variables—in our case, the measurement data D—to obtain an estimate $\epsilon(D)$ of the true value E [15]. In this section, we present an estimator $\epsilon(D)$ for E and we quantify our uncertainty about the value of E with an estimator $V_\epsilon(D)$ for the mean squared error of $\epsilon(D)$.

In order to introduce the estimator $\epsilon(D)$, we have to define the measurement data D. The measurement settings used in the experiment were given by Eq. (6) (Sec. III). Here, we describe each measurement setting as one POVM Π_j and we collect the POVMs for all measurement settings in the POVM set $\Pi_M = \{\Pi_j : j\}$. We describe the 3^k measurement settings from Sec. III with 3^k POVMs:

$$\Pi_M = \left\{ \Pi_j : \quad j = (j_1, \ldots, j_k), \quad j_a \in \{X, Y, Z\}, \quad a \in \{1, 2, \ldots, k\} \right\}. \tag{14}$$

As there are exactly 3^k different values of the POVM label j, we can equivalently treat j as integer $j \in \{1, \ldots, q\}$, $q = 3^k$.

Each POVM has 2^N distinguishable outcomes, i.e. 2^N POVM elements:

$$\Pi_j = \left\{ \Pi_{jl} : \quad l = (l_1, \ldots, l_N), \quad l_c \in \{-1, 1\}, \quad c \in \{1, \ldots, N\} \right\}. \tag{15}$$

The POVM elements are given by

$$\Pi_{jl} = P_{1,j_1,l_1} \otimes \ldots \otimes P_{k,j_k,l_k} \otimes P_{k+1,j_1,l_{k+1}} \otimes \ldots \otimes P_{2k,j_k,l_{2k}} \otimes \ldots, \tag{16}$$

where $P_{i,j_{i'},l_i} = |j_{i'} l_i\rangle_i \langle j_{i'} l_i|$ projects spin i onto the eigenvector $|j_{i'} l_i\rangle$ of the Pauli operator $j_{i'}$ (= X, Y or Z) with eigenvalue l_i ($i = 1, \ldots, N$ and $i' = 1, \ldots, k$). Note that the single-qubit measurement basis, as indicated by $j_1, \ldots j_k$, repeats after k qubits. These POVM elements describe exactly the measurement settings mentioned in Eq. (6) (Sec. III). A single measurement of one of the POVMs produces a single outcome $l = (l_1, \ldots, l_N)$ (Eq. (15)). We will refer to an outcome from a measurement of Π_j as $y_j = l = (l_1, \ldots, l_N)$.

For simplicity, we consider the case where each $\Pi_j \in \Pi_M$ has been measured exactly m times. This allows us to take one single outcome y_j from each $\Pi_j \in \Pi_M$ and store them into a vector $x = (y_1, \ldots, y_q)$. From now on, we will refer to x as a "single outcome" or as a "(single) sample". The complete dataset of

m outcomes from q POVMs is then structured as $D = (x_1, \ldots, x_m)$. A single element x_i of the dataset D is distributed according a probability density $p(x)$. The sampling distribution p_m describes the distribution of the complete dataset D and is given by $p_m(D) = p(x_1) \ldots p(x_m)$. (The explicit form of $p(x)$ and $p_m(D)$ will be provided in Sec. IV A 5.)

Our estimator will be given in terms of a real-valued function $f(x)$ of a single outcome x. To describe its properties, we will need the expectation value (often referred to as population average)

$$\mathbb{E}_p(f) = \int f(x)p(x)\mathrm{d}x. \tag{17}$$

An expectation value corresponds to an exact value which we want to obtain but cannot access directly because we do not know $p(x)$. We only have access to m samples from $p(x)$, given by $D = (x_1, \ldots, x_m)$ (which is the measurement data from the experiment). Using this data, we define the data expectation value (often referred to as sample average)

$$\mathbb{E}_D(f) = \frac{1}{m}\sum_{i=1}^{m} f(x_i). \tag{18}$$

The data expectation value is a quantity which we can compute from the samples we have, and we will use it to estimate the expectation value. The covariance and data covariance are defined as usual by

$$\mathbb{V}_p(f,g) = \mathbb{E}_p(fg) - \mathbb{E}_p(f)\mathbb{E}_p(g), \qquad \mathbb{V}_D(f,g) = \mathbb{E}_D(fg) - \mathbb{E}_D(f)\mathbb{E}_D(g), \tag{19}$$

and the variance is given by $\mathbb{V}_p(f) = \mathbb{V}_p(f,f)$. The product fg is the usual point-wise product $(fg)(x) = f(x)g(x)$. We use the following textbook relations (see Sec. IV A 6 below for a proof).

Lemma 1. *For two functions f and g of a random variable, we have*

$$\mathbb{E}_{p_m}[\mathbb{E}_D(f)] = \mathbb{E}_p(f), \tag{20}$$

$$\mathbb{V}_{p_m}[\mathbb{E}_D(f), \mathbb{E}_D(g)] = \frac{1}{m}\mathbb{V}_p(f,g), \tag{21}$$

$$\mathbb{V}_p(f,g) = \frac{m}{m-1}\mathbb{E}_{p_m}[\mathbb{V}_D(f,g)]. \tag{22}$$

In the next subsection, we will define a function f_ϵ with the property

$$\mathbb{E}_p(f_\epsilon) = E. \tag{23}$$

Indeed, f_ϵ will be a linear combination of simple functions which count how often certain partial measurement outcomes have occured in the experiment. The function $\epsilon(D)$, defined by

$$\epsilon(D) = \mathbb{E}_D(f_\epsilon), \tag{24}$$

then has the property

$$\mathbb{E}_{p_m}(\epsilon(D)) = \mathbb{E}_p(f_\epsilon) = E. \tag{25}$$

In other words, $\epsilon(D)$ is an unbiased estimator of E. This provides the equality

$$\mathbb{E}_{p_m}[(\epsilon(D) - E)^2] = \mathbb{V}_{p_m}(\epsilon(D)), \tag{26}$$

i.e., the estimator's mean squared error (left-hand side) equals its variance (right-hand side). We estimate the estimator's variance using

$$V_\epsilon(D) = \frac{1}{m}\frac{m}{m-1}\mathbb{V}_D(f_\epsilon). \tag{27}$$

Combining Eqs. (21) and (22) shows

$$\mathbb{E}_{p_m}(V_\epsilon(D)) = \mathbb{V}_{p_m}(\mathbb{E}_D(f_\epsilon)) = \mathbb{V}_{p_m}(\epsilon(D)), \tag{28}$$

i.e., $V_\epsilon(D)$ is an unbiased estimator of the variance of the estimator $\epsilon(D)$ under the sampling distribution p_m. The definition of f_ϵ, which is still missing and rather technical, is given in the next subsection.

We summarize the results from this section. Using Eq. (9), a lower bound to the fidelity between the certified estimate $|\psi_c^k\rangle = |\psi_{\mathrm{GS}}\rangle$ and the unknown lab state ρ_{lab} is obtained:

$$\langle\psi_c^k|\rho_{\mathrm{lab}}|\psi_c^k\rangle \geq F_c^k \pm \Delta F_c^k. \tag{29}$$

The value of the fidelity lower bound F_c^k and its measurement uncertainty ΔF_c^k are given by

$$F_c^k = 1 - \frac{\epsilon(D) - E_0}{E_1 - E_0}, \qquad \Delta F_c^k = \frac{\sqrt{V_\epsilon(D)}}{E_1 - E_0}, \tag{30}$$

with $\epsilon(D)$ and $V_\epsilon(D)$ from Eqs. (24) and (27) and the function $f_\epsilon(x)$ given in the next subsection. Values of the fidelity lower bound F_c^k are mentioned in the main text and shown in Fig. 3a of the main text.

The estimator $\epsilon(D)$ will be seen to be a weighted sum of functions of many independent random variables, in our case we have 27000 random variables from 27 measurement bases and 1000 measurements per measurement basis. While not all 27 measurement bases contribute equally to the weighted sum, all 1000 measurements do contribute equally and it is reasonable to expect that $\epsilon(D)$ will be distributed according to a normal distribution.

5. *Estimator for the parent Hamiltonian energy*

In the last section, we have replaced the uncertified initial estimate by a certified estimate, and we have provided the functional form of the fidelity lower bound which provides the certificate. In this section, the function f_ϵ will be defined; it is required to obtain values of the fidelity lower bound and its uncertainty and it must obey Eq. (23).

While the value of the fidelity lower bound F_c^k could be obtained from an easier computation than presented below, determining its measurement uncertainty ΔF_c^k requires the full discussion of this subsection. Incorporating the fact that the local probabilities of Eq. (8) (Sec. III) have been obtained from the global measurement bases of Eq. (6) complicates the computation of the measurement uncertainty ΔF_c^k. The complication comes from correlations which arise from two probabilities being estimated using the same set of measurement outcomes. The following minimal example illustrates how these correlations come about: The observable $\sigma_Z^{(1)} \otimes \sigma_Z^{(2)}$ is measured on the two-spin Bell state $(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle)/\sqrt{2}$ several times and the outcome probabilities $p_{\uparrow,\downarrow}^{(i)}$ of a measurement of $\sigma_Z^{(i)}$ on spin i ($i = 1, 2$) are estimated by counting how often each (partial) outcome has occured. If the estimate $p_\uparrow^{(1)}$ is larger than the true value $\frac{1}{2}$ due to statistical fluctuations, then the estimate $p_\uparrow^{(2)}$ is also be larger than its true value $\frac{1}{2}$ because the two estimates are equal. In addition, if the estimate $p_\uparrow^{(1)}$ is larger than its true value, the estimate of $p_\downarrow^{(1)}$ is smaller than its true value (this holds even for a product state). The variance estimator $V_\epsilon(D)$ from the last subsection takes all these correlations into account; it can be shown to estimate all covariances of the kind we have just discussed.

We continue with the pending definition of f_ϵ. The parent Hamiltonian H from Eq. (11) takes the form

$$H = \sum_{s=1}^{N-k+1} \mathbb{1}_{1,\ldots,s-1} \otimes h_s \otimes \mathbb{1}_{s+k,\ldots,N}, \tag{31}$$

where each term h_s acts only on k out of the total N qubits. We need to estimate the energy E, given by

$$E = \sum_{s=1}^{N-k+1} \mathrm{tr}(h_s \rho_s), \tag{32}$$

where ρ_s is the reduced state of ρ_{lab} on sites $s, s+1, \ldots, s+k-1$. Our first step is providing an expression for E in terms of the local probabilities from Eq. (8) (Sec. III). For this purpose, we define a POVM set Π_L whose outcome probabilities correspond to the named local probabilities:

$$\Pi_L = \left\{ Q_s : \quad s = 1, \ldots, N-k+1 \right\}. \tag{33}$$

The individual POVMs Q_s are given by

$$Q_s = \left\{ Q_{si} : \quad i = (a_1, \ldots, a_k, b_1, \ldots, b_k), \quad a_c \in \{X, Y, Z\}, \quad b_c \in \{-1, 1\} \right\} \tag{34}$$

with $c \in \{1, \ldots, k\}$. Their 6^k elements are given by

$$Q_{si} = P_{s,a_1,b_1} \otimes \ldots \otimes P_{s+k-1,a_k,b_k}. \tag{35}$$

As above, $P_{c,a_i,b_i} = |a_i b_i\rangle_c \langle a_i b_i|$ projects spin c onto the eigenvector $|a_i b_i\rangle$ of the Pauli operator a_i ($= X, Y$ or Z) with eigenvalue b_i. It is understood that the elements of Q_s act only on the sites $s, \ldots, s+k-1$ of the N-qubit lab state ρ_{lab}.

We will use the linear map

$$\mathcal{M}_s(\rho_s) = [\mathrm{tr}(Q_{si}\rho_s)]_{Q_{si}\in Q_s} \tag{36}$$

which maps a given k-qubit state ρ_s on the vector of POVM probabilities $p_{si} = \mathrm{tr}(Q_{si}\rho_s)$. The POVM elements of each POVM $Q_s \in \Pi_L$ span the complete space of operators. As a consequence, they are called informationally complete and the identity $\overline{\mathcal{M}_s}\mathcal{M}_s(\rho_s) = \rho_s$ holds; here, $\overline{\mathcal{M}_s}$ is the Moore–Penrose pseudoinverse of $\mathcal{M}_s$. This relation provides

$$E = \sum_{s=1}^{N-k+1} \mathrm{tr}(h_s\overline{\mathcal{M}_s}(\mathcal{M}_s(\rho_s))) = \sum_{s=1}^{N-k+1}\sum_i c_{si}p_{si}, \tag{37a}$$

$$c_{si} = \mathrm{tr}(h_s\overline{\mathcal{M}_s}(e_i)), \quad p_{si} = \mathrm{tr}(Q_{si}\rho_s), \tag{37b}$$

where e_i is the i-th standard basis vector. We have accomplished the goal of expressing the energy E in terms of the local probabilities p_{si}.

If we had measurement data for the POVM set Π_L available, we could use simple counting functions

$$\theta_{si}(x) = \begin{cases} 1, & \text{if } y_s = i \text{ with } x = (y_1, \ldots, y_{N-k+1}) \\ 0, & \text{otherwise} \end{cases}, \tag{38}$$

where we have combined single outcomes y_s from $Q_s \in \Pi_L$ into vectors $x = (y_s)_s$ (as has been done in the last subsection for $\Pi_j \in \Pi_M$). It is simple to see that such counting functions estimate probabilities (see below for an explicit example):

$$\mathbb{E}_{p_m}[\mathbb{E}_D(\theta_{si})] = \mathbb{E}_p(\theta_{si}) = p_{si}. \tag{39}$$

In this case, we could obtain the function f_ϵ by replacing p_{si} with θ_{si} in Eq. (37a).

However, our measurement data is data for the POVM set Π_M defined above in Eqs. (14)–(16) and we must estimate the local probabilities p_{si} from that data. We need to establish a relation between the two POVM sets Π_M and Π_L. Because of the particular choice we made for these two sets, it is easy to find non-negative coefficients $c_{si,jl'}$ such that [22]

$$Q_{si} = \sum_{jl'} c_{si,jl'}\Pi^{(s)}_{jl'}, \quad Q_{si} \in Q_s, \quad Q_s \in \Pi_L, \tag{40}$$

where the sum over l' is over the 2^k different partial traces

$$\Pi^{(s)}_{jl'} = \mathrm{tr}_{1,\ldots,s-1,s+1,\ldots,N}(\Pi_{jl}) \tag{41}$$

of the 2^N elements $\Pi_{jl} \in \Pi_j$. (As before, $\Pi_j \in \Pi_M$.) The partial traces $\Pi^{(s)}_{jl'}$ are rank-1 projectors onto a particular k-fold tensor product of eigenvectors of Pauli matrices. Therefore, we enumerate them with an index vector $l' = (l'_1, \ldots, l'_k)$, $l'_a \in \{+1, -1\}$, $a \in \{1, \ldots, k\}$ where l'_a specifies whether the a-th eigenvector is spin up or spin down in some direction (X, Y or Z).

In order to estimate the local probabilities p_{si} from data of the global POVM set Π_M, we define counting functions for occurences of a local part $(l_s, \ldots, l_{s+k-1})$ of a global outcome $(l_1, \ldots, l_N)$:

$$\theta_{jsl'}(x) = \begin{cases} 1, & \text{if } (l_s, \ldots, l_{s+k-1}) = (l'_1, \ldots, l'_k) \\ & \text{with } x = (y_1, \ldots, y_q) \text{ and } y_j = (l_1, \ldots, l_N), \\ 0, & \text{otherwise.} \end{cases} \tag{42}$$

To show that $\theta_{jsl'}$ can be used to estimate $\mathrm{tr}(\Pi^{(s)}_{jl'}\rho_s)$, we have to complete some definitions. As has been mentioned above, the sampling distribution $p_m(D) = p(x_1)\ldots p(x_m)$ describes the probability distribution of the complete dataset $D = (x_1, \ldots, x_m)$. Single outcomes $x = (y_1, \ldots, y_q)$ contain one outcome y_j for each POVM $\Pi_j \in \Pi_M$. The single-outcome probability density therefore is $p(x) = p_1(y_1)\ldots p_q(y_q)$. Each POVM Π_j gives rise to to a discrete probability distribution $p^{(j)}_l = \mathrm{tr}(\Pi_{jl}\rho_{\mathrm{lab}})$, $\Pi_{jl} \in \Pi_j$. We embed it into a probability density via

$$p_j(y_j) = \sum_l \delta(y_j - l)\,\mathrm{tr}(\Pi_{jl}\rho_{\mathrm{lab}}) \tag{43}$$

where we have used l as an integer from $\{1, 2, \ldots, 2^N\}$. The counting functions defined above then have the property

$$\mathbb{E}_p(\theta_{jsl'}) = \mathrm{tr}(\Pi^{(s)}_{jl'}\rho_{\mathrm{lab}}) = \mathrm{tr}(\Pi^{(s)}_{jl'}\rho_s). \tag{44}$$

Finally, we can put everything together to obtain the final function f_ϵ, which will provide the estimator $\epsilon(D) = \mathbb{E}_D(f_\epsilon)$ of E. First, we define

$$f_{si}(x) = \sum_{jl'} c_{si,jl'} \theta_{jsl'}(x) \tag{45}$$

and observe

$$\mathbb{E}_p(f_{si}) = \sum_{jl'} c_{si,jl'} \operatorname{tr}(\Pi^{(s)}_{jl'} \rho_s) = \operatorname{tr}(Q_{si}\rho_s) = p_{si}. \tag{46}$$

We define f_ϵ by

$$f_\epsilon(x) = \sum_s \sum_i c_{si} f_{si}(x), \quad c_{si} = \operatorname{tr}(h_s \overline{M_s}(e_i)). \tag{47}$$

Using what we have learned so far, we obtain

$$\mathbb{E}_{p_m}(\epsilon(D)) = \mathbb{E}_{p_m}(\mathbb{E}_D(f_\epsilon)) = \mathbb{E}_p(f_\epsilon) = \sum_{si} c_{si} \mathbb{E}_p(f_{si}) = E. \tag{48}$$

This shows that the function $\epsilon(D)$ is an unbiased estimator of the energy $E = \operatorname{tr}(H\rho_{\text{lab}})$. This scheme has been implemented as part of the Python library mpnum [14, function `mpnum.povm.MPPovmList.est_lfun_from`].

6. *Proof for basic variance relations*

In this section, we proof three basic equalities used in Sec. IV A 4. Their proof is included for completeness.

Let $p(x)$ be some probability density function, $D = (x_1, \ldots, x_m)$ a dataset of m samples from p, and let $p_m(D) = p(x_1) \ldots p(x_m)$ the sampling distribution which describes the probability density of the complete dataset D. We will also use the definitions from Eqs. (17)– (19) on page 10. The equalities which we proof here provide relations between expectation values, sampling distribution expectation values and data expectation values. Above, they have been used to estimate the estimator's variance from data. Lemma 1 states: For two functions f and g of a random variable, we have

$$\begin{aligned} \mathbb{E}_{p_m}[\mathbb{E}_D(f)] &= \mathbb{E}_p(f), \\ \mathbb{V}_{p_m}[\mathbb{E}_D(f), \mathbb{E}_D(g)] &= \frac{1}{m}\mathbb{V}_p(f,g), \\ \mathbb{V}_p(f,g) &= \frac{m}{m-1}\mathbb{E}_{p_m}[\mathbb{V}_D(f,g)] \end{aligned}$$

Proof. First equation:

$$\mathbb{E}_{p_m}[\mathbb{E}_D(f)] = \int \frac{1}{m} \sum_{i=1}^m f(x_i) p(x_1) \ldots p(x_m) \mathrm{d}x_1 \ldots \mathrm{d}x_m = \frac{m}{m}\mathbb{E}_p(f). \tag{49}$$

For the second and third equation, we first compute

$$\begin{aligned} \mathbb{E}_{p_m}[\mathbb{E}_D(f), \mathbb{E}_D(g)] &= \int \frac{1}{m^2} \sum_{i=1}^m \sum_{j=1}^m f(x_i) g(x_j) p(x_1) \ldots p(x_m) \mathrm{d}x_1 \ldots \mathrm{d}x_m \\ &= \frac{m}{m^2}\mathbb{E}_p(fg) + \frac{m^2-m}{m^2}\mathbb{E}_p(f)\mathbb{E}_p(g) \\ &= \frac{1}{m}\mathbb{V}_p(f,g) + \mathbb{E}_p(f)\mathbb{E}_p(g) \end{aligned} \tag{50}$$

This provides

$$\begin{aligned} \mathbb{V}_{p_m}[\mathbb{E}_D(f), \mathbb{E}_D(g)] &= \mathbb{E}_{p_m}[\mathbb{E}_D(f), \mathbb{E}_D(g)] - \mathbb{E}_{p_m}[\mathbb{E}_D(f)]\, \mathbb{E}_{p_m}[\mathbb{E}_D(g)] \\ &= \frac{1}{m}\mathbb{V}_p(f,g) \end{aligned} \tag{51}$$

and

$$\mathbb{E}_{p_m}[\mathbb{V}_D(f,g)] = \mathbb{E}_p(fg) - \mathbb{E}_{p_m}[\mathbb{E}_D(f), \mathbb{E}_D(g)] = (1 - \frac{1}{m})\mathbb{V}_p(f,g), \tag{52}$$

which is the required relation. □

B. Simulations of MPS tomography and certification

In Figure 3a of the main text, fidelity lower bounds F_c^k based on an ideal model of the tomographic process are presented. Here we explain how they were obtained.

The ideal model of the tomographic process assumes a perfect initial state $|\phi(0)\rangle = |\uparrow\downarrow\uparrow\downarrow \ldots\rangle$ in the σ^z basis and an idealised time evolution of the quantum simulator, described by the Hamiltonian of eq. 1 The transverse fields B, B_i and coupling matrix elements J_{ij} have been described above. Note that in the limit $B \gg |J_{ij}|$, which is upheld in the experiments, the above Hamiltonian is equivalent to an XY model in a transverse field, as described in Sec. II. For more details see the supplementary material of [3]. The ideal time evolution $|\phi(t)\rangle = \exp(-iH_{\mathrm{Ising}}t)\,|\phi(0)\rangle$ is computed using the library function `scipy.sparse.linalg.expm_multiply` [13]. (For the simulation with a mixed initial state discussed in Sec. IV D below, the same library function has been used to propagate 15 different pure initial states in time. It was convenient to convert the resulting state to a purified MPS representation with a single ancilla site of dimension 15.)

The simulated MPS tomography and certification can be performed with exact knowledge of local probabilities or data from a finite number of simulated measurements (simulation mentioned in Sec. IV D below).

Exact knowledge of local observables. Assuming exact knowledge of local observables simplifies the simulation. The MPS tomography algorithms are run with the exact values of the 6^k probabilities describing the measurement outcomes of the 3^k k-fold tensor products of the Pauli X, Y and Z matrices for each of the $N - k + 1$ local blocks. Computing the energy $E = \mathrm{tr}(H\rho_{\mathrm{lab}})$ of the (now known) ideal lab state $|\phi(t)\rangle$ in terms of the parent Hamiltonian H is simplified considerably as we can compute the exact local reductions of ρ_{lab}. As a consequence, the resulting fidelity lower bound is known without uncertainty as well.

This numerical simulation represents the expected performance of certified MPS tomography in a case where the perfect initial state is prepared, the simulator produces ideal unitary dynamics and an infinite number of perfect measurements are performed. This is shown in Fig. 3a in the main text.

Finite number of simulated measurements of global observables. While the measurement of a global observable such as $X^{\otimes N}$ can yield one of 2^N outcomes, it is simple to draw a sample from the probability distribution of measurement outcomes if a matrix product description of the state on which measurements shall be simulated is available: One can simply compute the local reduced state on the first qubit, simulate a single-qubit measurement there and continue by computing the reduced state state of the second qubit conditioned on the outcome of the first measurement, etc. [23]. A more direct way to implement the same procedure uses a matrix product description of the POVM in question to obtain a matrix product representation of the measurement outcome probabilities. Marginal and conditional distributions of the full outcome probability distribution can be efficiently obtained. This procedure has been implemented as part of the mpnum library [14] (function `mpnum.povm.MPPovm.sample(method='cond')`). However, with our code for 14 qubits, it was still faster to convert the matrix product representation of the outcome probability distribution to a full array with 2^{14} elements and sample from the full description. This simulation represents the expected performance of certified MPS tomography on our system, assuming everything is perfect but allowing for a finite number of measurements (1000) per basis. The outcome of these simulations are shown in FIG. S2.

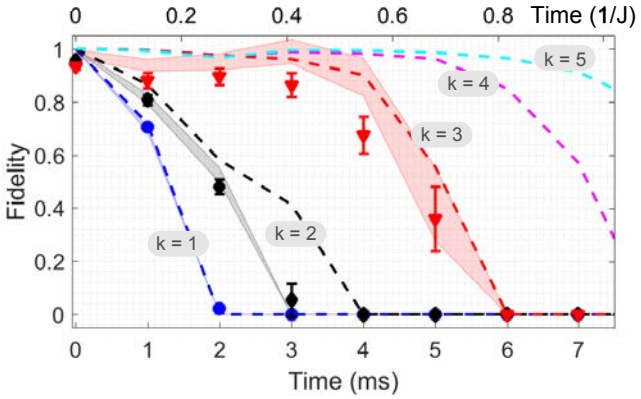


FIG. S2: **Fidelity lower bounds for the 8-spin quench: comparison with theory.** Certified lower bounds on the fidelity between MPS estimates $|\psi_c^k\rangle$, reconstructed from measurements over k sites, and the quantum simulator state ρ_{lab}. Shapes: data points with errors, 1 standard deviation. Error bars reflect uncertainty due to finite measurement number and can be calculated directly from raw measurement statistics through the certification process (Methods). Data is compared to two theoretical models (dashed lines and shaded areas). Both models consider ideal simulator states $|\phi(t)\rangle$. Dashed lines: using exact knowledge of k-site local reductions derived from $|\phi(t)\rangle$. Shaded areas: using imperfect knowledge of k-site local reductions derived from the ideal states by simulating the outcomes of the finite number of measurements per measurement basis used in experiments (1000). See Methods. Ones sees that the effect of imperfect knowledge of local reductions accounts for many of the differences between data and exact theory, and that there isn't much to gain by taking more measurements. A possible reason for additional discrepancies between data and model is mixture in the laboratory state, which will reduce the quality of the pure-state MPS description.

C. Resource cost for a constant estimation error in MPS tomography.

In the final section in the Methods we discuss the resource cost for a constant estimation error in MPS tomography. We performed extensive numerical simulations (see FIG. S3) in order to demonstrate the practicality of the scheme and its efficient scalability to large systems in comparison to full state tomography. We found that for a constant reconstruction fidelity the required number of measurements scales as the third power in the number of subsystems N. The post-processing time to obtain the state increases with $\approx N^{1.2}$. Hence MPS tomography involves an experimental and computational effort that scales polynomially in N while standard tomography scales exponentially.

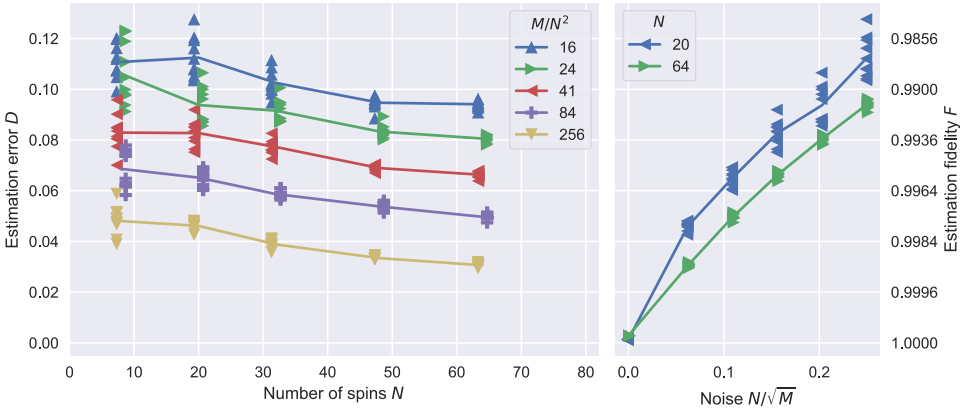


FIG. S3: **Estimation error of MPS tomography** with a finite number of measurements. The estimation error D and estimation fidelity $F = |\langle\psi_{\text{ideal}}|\psi_{\text{est}}\rangle|^2$ are related by $D = \sqrt{1-F}$. M measurements per basis were simulated (numerically on a classical computer) on an N-spin ideal state $|\psi_{\text{ideal}}\rangle$ and the estimate $|\psi_{\text{est}}\rangle$ was obtained with MPS tomography; this procedure was repeated 10 times. Markers show errors obtained in 10 individual reconstructions and lines connect average values. MPS tomography was applied to a state from a nearest-neighbour quench (see Methods in the main text). Left: Estimation error as function of number of spins N. Right: Estimation error as function of $N/\sqrt{M}$; M is the number of measurements per basis. Markers are shifted slightly horizontally to enhance visibility and straight lines serve as guide to the eye.

D. Modelling initial Néel state errors for the 14-spin experiments

In the main text we state that the differences, between the experimentally-obtained and ideal-simulator model fidelity bounds F_c^3 at $t_{14} = 4$ ms are largely explained by errors in the initial Néel state preparation for 14 spins. Here we aim to convince the reader of that.

In section I B of this supplementary material we showed that the initial 14-spin Néel state was prepared with a (directly-measured) fidelity of 0.89 ± 0.01, compared to 0.967 ± 0.006 for the 8-spin case. This corresponds to a significantly larger error-per-particle (e.g. quantified by $\log(F)/N$) for the 14-spin case: $-\log(0.89 \pm 0.01)/14 = 0.012 \pm 0.01$, while $\log(0.967 \pm 0.006)/8 = 0.007 \pm 0.001$

We performed numerical simulations to determine the effect of the errors in the 14-spin initial state on the MPS reconstruction of the 14-spin quench experiment. Specifically, we asked how detrimental the initial state errors are expected to be to the performance of the MPS tomography for 3-site measurements after 4 ms of evolution (the time at which direct state fidelity estimation was carried out).

To do this, we modelled the initial state error in the following way. Analysis of the direct measurement results for the 14-spin initial state in the lab show that, out of 1000 times that we prepared and measured the state in the z-basis, 893 times we observed the Néel state (hence the fidelity of 0.893). 93 times we observed a state with one spin flip error. For the remaining 12 times, we observed two spin flip errors. The errors are most likely caused by fluctuations in the intensity of our addressed laser beam, meaning that these erroneous states are added in mixture with the ideal Neel state. We built a noisy model for the initial lab state as an appropriately weighted mixture of the ideal Neel state and single spin flip errors. We call this model of the noisy initial 14 spin state $\rho^{14}_{noisysim}$.

In the next step, we numerically simulate outcomes of the same measurements as have been performed in the actual experiments in the lab (1000 measurements of each setting described in Sec. III). We insert measurement outcomes from numerical simulation into the MPS tomography search algorithm. The MPS reconstruction then proceeds as usual, as described above. The output is a certified estimate for the fidelity lower bound $F^3_{c,noisy}$, that we would expect to achieve when measuring such a noisy state in the lab. The result for the 14 spin noisy model and $t = 0$ ms is the lower bound $F^3_{c,noisysim} = 0.90 \pm 0.03$. This is to be compared with the direct (exact) fidelity measurement of 0.89 ± 0.01 in the lab and the MPS-tomography lower fidelity bound, for measurements on $k = 1$-site, of 0.90 ± 0.04. Clearly all agree well and the lower bounds are tight.

The result for the 14 spin noisy model at $t_{14} = 4$ ms is $F^3_{c,noisysim} = 0.49 \pm 0.07$. This is to be compared with the lower bound obtained in the experiment $F_c^3 = 0.39 \pm 0.08$. Of course, the described errors in the initial state are not the only errors in the experiment, however, we conclude that they are largely responsible for the difference between the fidelity lower bounds obtained from experimental data and from an idealised model of a perfect simulator. So, noise adding mixture to the initial state preparation explains why the experimentally obtained fidelity lower bounds for 14 spins are lower than the ideal case. In the future, we should work on keeping the error-per-particle constant when scaling up our system. This should be relatively straightforward for up to several tens of spins, by rebuilding the optical setup used to deliver the addressed laser beam. Other sources of error that we considered, and found to play minor roles by equivalent numerical modelling, are: the finite lifetime of the excited spin (atomic) state; correlated dephasing due to correlated fast fluctuations in real magnetic fields around the ion string; small errors in the analysis pulses that determine the measurement bases.

V. EXTENDED EXPERIMENTAL ANALYSIS AND RESULTS

A. Single site magnetisation dynamics for 8 and 14 spin quenches

Figures 2a and 4a in the main text present the measured single-site magnetisation dynamics for 8 and 14-spin quenches, for the Néel initial state. In order to calibrate our experimental system, and compare the model dynamics with the results in the lab, we run quenches starting with a spin state that contains a single spin excitation (local quench [2]). The subsequent dynamics reveal the spreading of correlations from a single site and provide a useful visualisation of the approximate light-like cones (approximate maximum group velocity for the spread of information). FIG. S4 compares the measured and model single-site magnetisation dynamics for a local quench and the Néel initial state for 8 spins, showing how the single excitation spreads out in the system. Two kinds of light-like cones are presented, as described in the caption. The faster of which is the same as those presented in the main text. The maximum rate of information spreading should not depend on the initial state, but only the spin-spin interaction Hamiltonian. FIG. S5 presents the same but for 14 spins. The experimental results are the same as those in figures 2a and 4a in the main text. In all cases, the match between data and model is excellent.

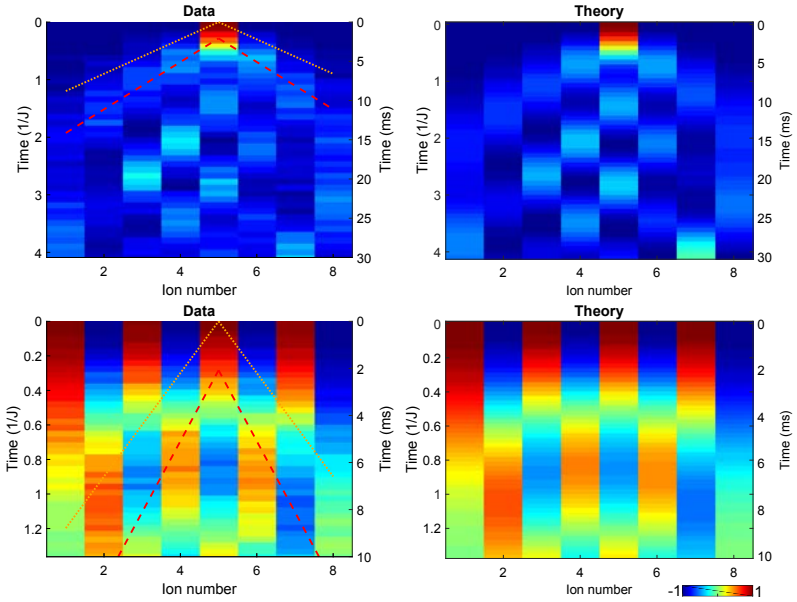


FIG. S4: **Magnetisation dynamics during an 8 spin quench.** Time evolution of magnetisation $\langle \sigma_i^z(t) \rangle$ for an initial single spin excitation (top row) and an initial Néel state (bottom row). Left column: Experimental data. Right column: predictions from idealised model ($|\phi(t)\rangle$). The two time axes distinguish between the real time passed in the laboratory (right axis) and time renormalised by the mean nearest-neighbour spin-spin interaction (Methods: Normalised time units $1/J$). The colours identify the spin state: Dark blue indicates a $|\downarrow\rangle$ state, red a $|\uparrow\rangle$ state. In both data panels, red dashed lines are fits to the peaks of the observed magnetisation (information) wavefronts seen in the *single* excitation dynamics (top row). The spreading out (dispersion) of the initially localised spin-excitations can be approximated bounded by light-like cones: Orange dotted lines show the maximum expected velocity at which correlations spread out, estimated by renormalising the mean nearest-neighbour interaction strength by the algebraic tail (see Methods: Light-like cones and interaction ranges). The orange dotted lines are the same for both data panels and the method to derive them is identical (it does not depend on the initial state). From the top data panel one sees that the orange light-like cones contain the vast majority of the information wavefronts. **The orange dotted lines are the ones shown as black dotted lines in Figure 2 of the main text.** The light-like cones presented in this paper serve as a useful guide to interpret correlation spreading in our system and when one should expect measurements over larger sites to become necessary for successful MPS tomography.

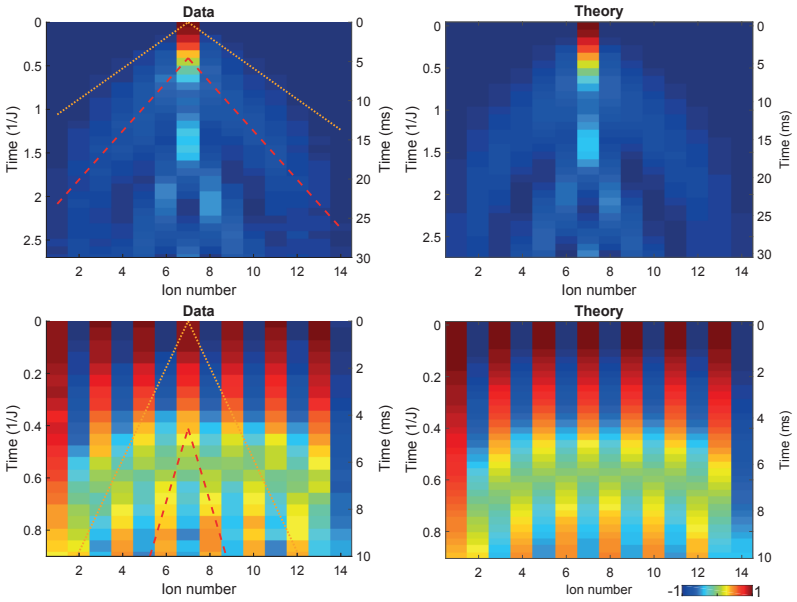
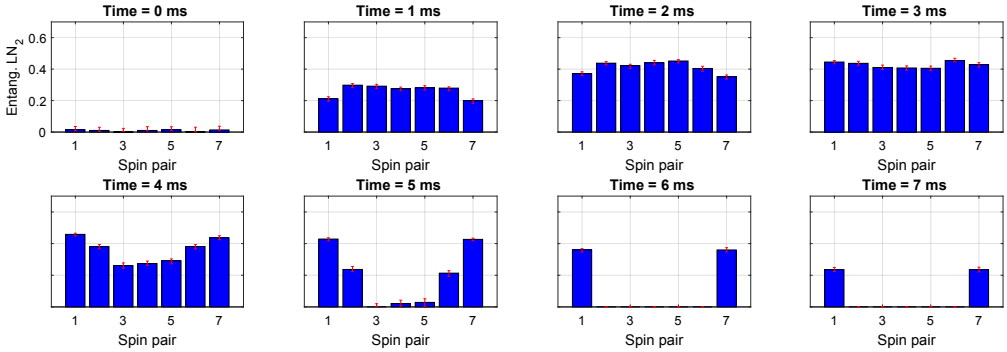


FIG. S5: **Magnetisation dynamics during an 14 spin quench.** Time evolution of magnetisation $\langle \sigma_i^z(t) \rangle$ for initial single spin excitation (top row) and initial Néel state (bottom row). Left column: Experimental data. Right column: Predictions from idealised model right column ($|\phi(t)\rangle$). For normalised time units see Methods: Normalised time units $1/J$. In both data panels, red dashed lines are fits to the peaks of the observed magnetisation wavefronts seen in the *single* excitation dynamics (top row): the maxima of the leading information wavefronts. The spreading out (dispersion) of the initially localised spin-excitations can be approximated bounded by light-like cones: Orange dotted lines show the maximum expected velocity at which correlations spread out, estimated by renormalising the mean nearest-neighbour interaction strength by the algebraic tail (see Methods: Light-like cones and interaction ranges). The Orange dotted lines are the same for both Data panels and the method to derive them is identical (it does not depend on the initial state). From the top Data panel one sees that the orange light-like cones contain the vast majority of the information wavefronts. **The orange dotted lines are the ones shown as black dotted lines in Figure 4 of the main text**. The light-like cones presented in this paper serve as a useful guide to interpret correlation spreading in our system and when one should expect measurements over larger sites to become necessary for successful MPS tomography.

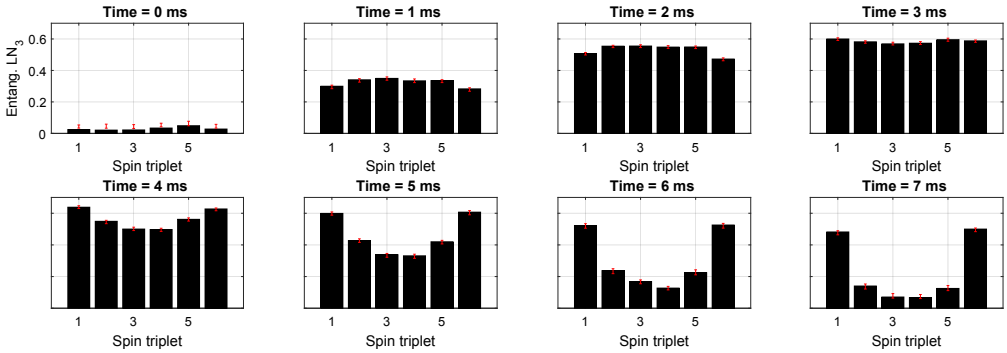
B. Local reductions and correlation matrices for the 8-spin quench

We now present dynamical properties of the experimentally reconstructed local reductions of single spins, neighbouring spin pairs and neighbouring spin triplets, during the 8-spin quench experiments. These local reductions are reconstructed directly from the local measurements, using full quantum state tomography (not MPS tomography), which searches over all possible physical states (pure and mixed) to find an optimum fit with the data.

FIG. S6 presents the **dynamics of entanglement**, quantified by the logarithmic negativity, in the experimentally-reconstructed local reductions. Spin pair entanglement maximises at 2 ms and spin triplet entanglement between 3 and 4 ms. As the simulator evolves further, entanglement reduces, first in pairs then in triplets agreeing with the spread of correlations in the system. The measured results closely fit an ideal model of the simulator (not shown for clarity).



(a) Bipartite logarithmic negativity LN_2



(b) Tripartite logarithmic negativity LN_3

FIG. S6: **Evolution of entanglement in local reductions during the 8 spin quench.** (a) Bipartite logarithmic negativity LN_2 for neighbouring spin pairs and (b) tripartite logarithmic negativity LN_3 for neighbouring spin triplets. The entanglement measure LN_3 is defined as the geometric mean of all three bipartite logarithmic negativity splittings of a triplet (Methods). The simulator evolution time for each panel is shown in its title. The values are extracted directly from the experimentally-reconstructed k-spin local reductions (density matrices). Error bars, representing 1 standard deviation confidence, are derived with the standard Monte Carlo method to simulate quantum project noise (finite measurement number effects). For clarity, values from an ideal simulator model are not shown: the data is largely indistinguishable from the ideal model for all times.

In Figures 3b,c of the main text, **correlation matrices** are presented showing correlations in various bases between spin pairs across the 8-spin system. FIG. S7 presents correlations measured in additional bases and compares them with those captured in the measured MPS reconstructions and those from an ideal model of the simulator.

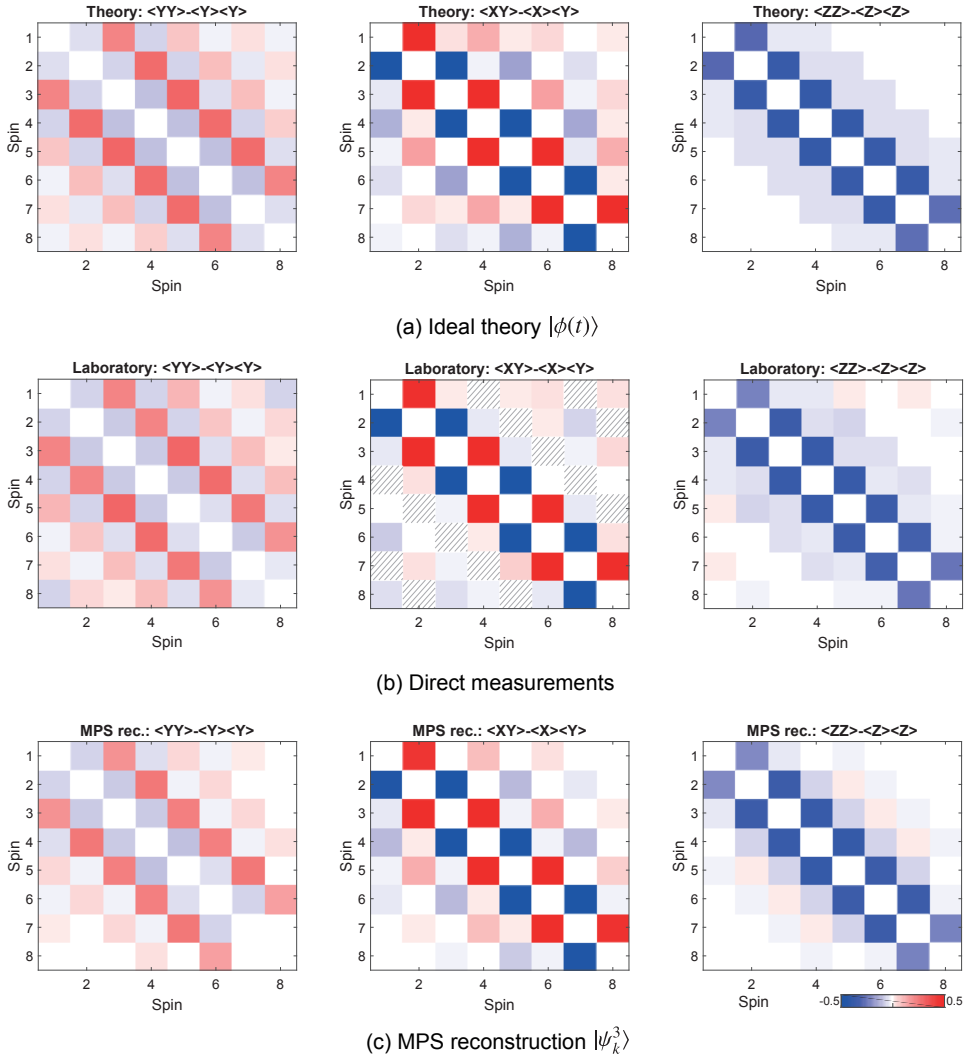


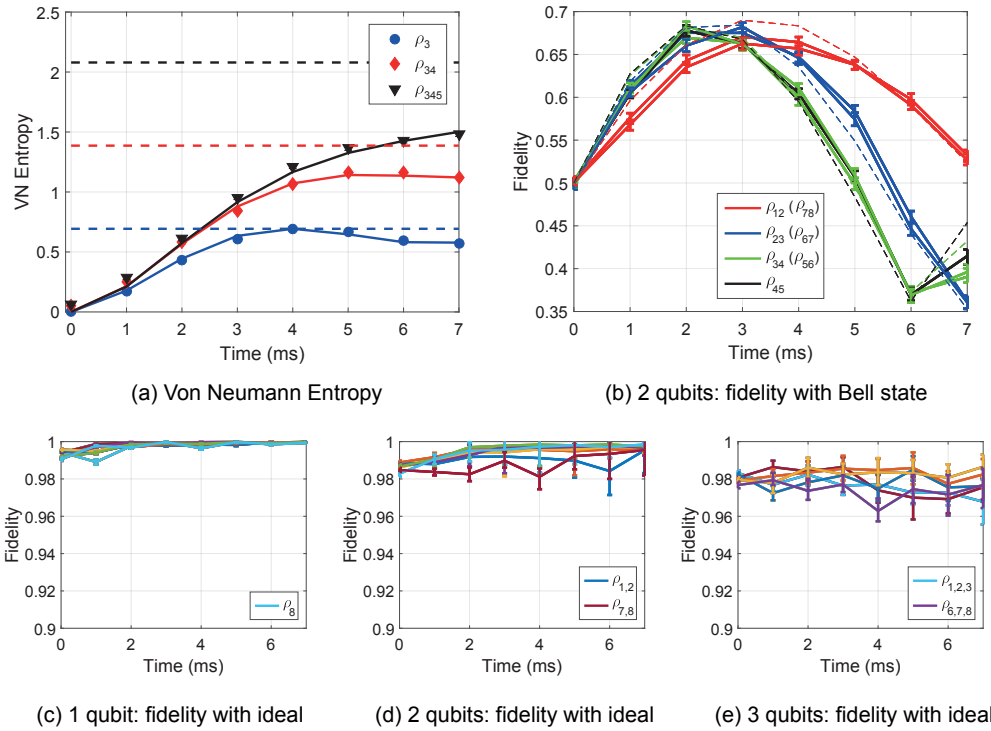
FIG. S7: **Two-spin correlation matrices for 8 spins.** $\langle Y_i(t)Y_j(t)\rangle - \langle Y_i(t)\rangle\langle Y_j(t)\rangle$ (left column) and $\langle X_i(t)Y_j(t)\rangle - \langle X_i(t)\rangle\langle Y_j(t)\rangle$ (middle column) and $\langle Z_i(t)Z_j(t)\rangle - \langle Z_i(t)\rangle\langle Z_j(t)\rangle$ (right column) at $t = 3$ ms evolution time (see sub panel titles). Top row (a): Calculated from theoretical ideal state $|\phi(t)\rangle$. Middle row (b) outcome of direct measurements on the state in the laboratory (requiring more measurements than used for MPS tomography). Bottom row (c): Calculated from the experimentally-reconstructed certified MPS $|\psi_c^3\rangle$ for the 8 ion Néel state after 3 ms simulator evolution (using measurements on $k = 3$ sites). The hatched squares denote correlations that were not measured. A subset of the above matrices are presented in Figure 3 in the main text. Note: X,Y and Z are the standard Pauli spin operators.

The **Von Neumann entropy** ('quantum entropy') of example local reductions during the 8 spin quench are presented in FIG. S8(a). The maximal value for Von Neumann entropy of an N spin (qubit) state is $\ln(2^N)$, corresponding to a maximally mixed state (shown as horizontal dashed lines in the figure).

FIG. S8(b) presents the **fidelity** between the experimentally reconstructed neighbouring spin-pair states and a **maximally entangled two-spin state** ($|\Psi^+\rangle$ Bell state). The entanglement between neighbours reaches a maximum between 2 and 3 ms. Quantifying entanglement in terms of the fidelity with a maximally entangled two-qubit state has operational meaning: states with fidelities above 50% are distillable. That is, given many copies of states with fidelities above this threshold, fewer states with higher quality entanglement can be distilled [16].

The **fidelities** of the directly reconstructed local reductions with the **ideal simulator model** are presented in FIG. S8(c-e). For fidelity F of an N-spin state, a measure of error-per particle is $E_N = \log_2(F)/N$. For the initial states (time $t = 0$) in FIG. S8(c-e), we find E_1, E_2 and E_3 all agree to within measurement uncertainty. That is, the error in the initial states of single spins, pairs and triplets is statistically indistinguishable. The single spin fidelities reach unity after a few ms of evolution. This can be

understood by considering that single spin states rapidly become fully mixed due to quantum correlations. The fully mixed state is unchanged under any unitary rotations and many physical noise sources.



(a) Von Neumann Entropy (b) 2 qubits: fidelity with Bell state

(c) 1 qubit: fidelity with ideal (d) 2 qubits: fidelity with ideal (e) 3 qubits: fidelity with ideal

FIG. S8: **Properties of the local reductions measured during the 8-spin quench**. Measurements over k-sites enables direct reconstruction of the k-site local reductions (density matrices) via standard Maximum Likelihood reconstruction. From these experimentally-reconstructed local reductions, various properties can be derived, with error bars representing one standard deviation obtained via the Monte Carlo simulation of quantum projection noise. (a) Von Neumann entropy in example local reductions of a single spin (blue), two spins (red), three spins (black). Shapes: data, entropy of experimentally-reconstructed local reductions via full QST. Solid lines: model based on ideal quantum simulator states. Dashed lines: the maximum entropy for fully mixed state of N spins is N (qubits). Error bars in data are almost smaller than the symbol size. (b) Overlap of the maximally entangled $|\Psi^+\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$ Bell state with the absolute value of the experimentally reconstructed neighbouring 2-spin density matrices. Spin pairs symmetrically distributed around the centre of the string are shown in the same color. Solid lines connecting points with error bars: data. Dashed lines: values from ideal model of the simulator.
(c - e) Overlap between the 8-spin state in the laboratory and the ideal state. Time dependent overlap of the measured reduced single-qubit (c), two-qubit (d) and three-qubit (e) density matrices with the theoretical ideal density matrices.

C. Correlation matrices for the 14-spin quench

In Figure 4c-e of the main text, correlation matrices are presented showing correlations in various bases between spin pairs across the 14-spin system. FIG. S9 presents correlations measured in additional bases and compares them with those captured in the measured MPS reconstructions and those from an ideal model of the simulator.

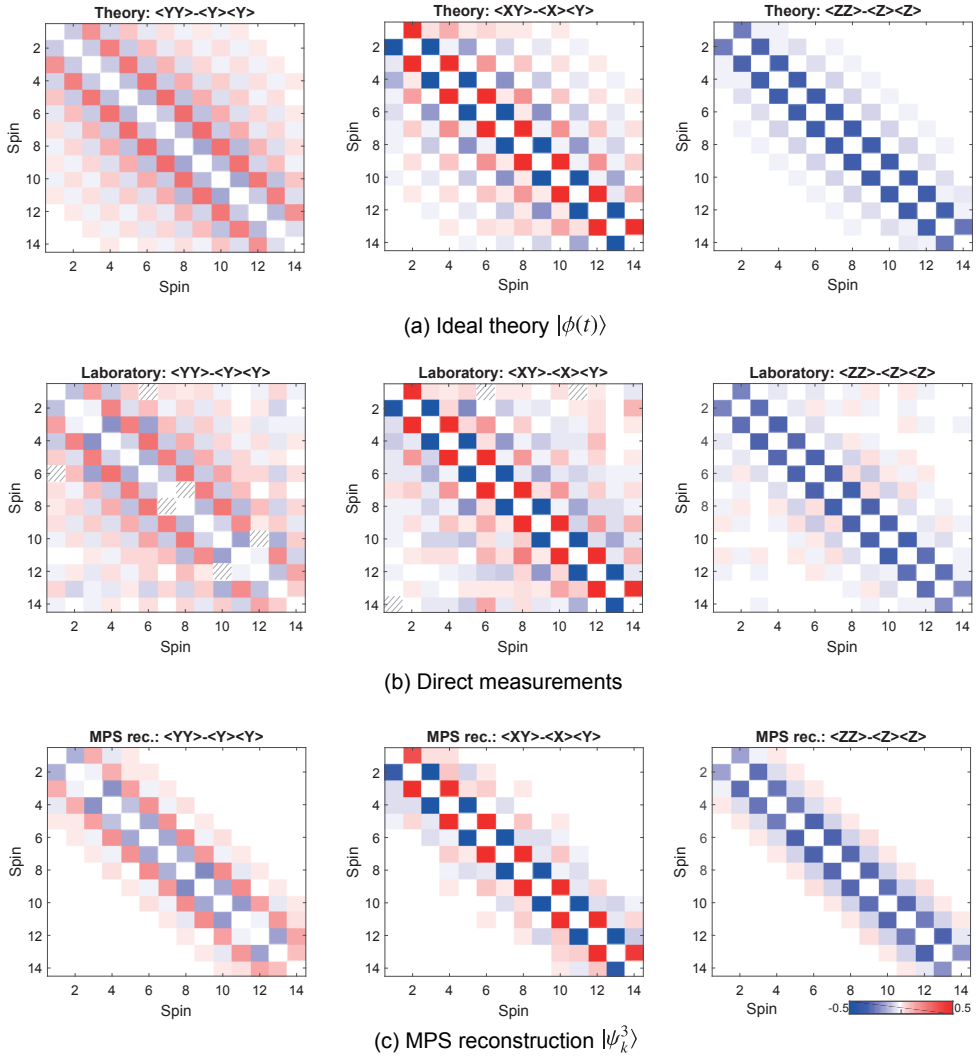


FIG. S9: **Two-spin correlation matrices for the 14 spin quench.** $\langle Y_i(t)Y_j(t)\rangle - \langle Y_i(t)\rangle\langle Y_j(t)\rangle$ (left column), $\langle X_i(t)Y_j(t)\rangle - \langle X_i(t)\rangle\langle Y_j(t)\rangle$ (middle column) and $\langle Z_i(t)Z_j(t)\rangle - \langle Z_i(t)\rangle\langle Z_j(t)\rangle$ (right column) at time $t = 4$ ms. At this time, the fidelity between the $|\psi_c^3\rangle$ MPS estimate and lab state is 0.74±0.05 (via direct fidelity estimation). Top row (a): Calculated from theoretical ideal state $|\phi(t)\rangle$. Middle row (b) outcome of direct measurements on the state in the laboratory (requiring more measurements than used for MPS tomography). Bottom row (c): Calculated from the experimentally-reconstructed certified MPS $|\psi_c^3\rangle$ for the 14 ion Néel state after 4 ms simulator evolution (using measurements on $k = 3$ sites). The hatched squares denote correlations that were not measured. A subset of the above matrices are presented in Figure 4 in the main text. Note: X,Y and Z are the standard Pauli spin operators.

D. Certified MPS reconstructions for 8 spin quench

In Figure 3a of the main text, the fidelity bounds for the 8-spin quench experiment are presented. In that figure, the data are compared with a theoretical model (numerical simulation) which shows the MPS tomography fidelity bounds that would be obtained for an idealised model of the simulator. Specifically, the exact local reductions of the model state $|\phi(t)\rangle = \exp(-iH_{\text{Ising}}t)\,|\phi(0)\rangle$ are used as input to the MPS tomography algorithm and certification process. This idealised simulation is described in section IV B. We would only expect to achieve these results in the laboratory if first, our system exactly implemented the XY model Hamiltonian and, second, we carried out an infinite number of perfect measurements to determine the local reductions. Clearly we do neither of these things. FIG. S2 compares the same experimental data with a more realistic model (Shaded areas). This model again uses the ideal simulator states but considers the effect of carrying out a 1000 (perfect) measurements per basis to identify local reductions, as done in the experiments. This model is described in more detail in section IV B. From this we conclude that: 1. the differences between data and the original perfect model are largely explained

by the finite number of measurements used in experiments and: 2. there is not much to be gained from doing more measurements in the lab.

E. Reconstructed MPS non-separable across all bipartitions

We calculated the Von Neumann entropies S of the pure reconstructed MPS $|\psi_c^3\rangle$ for all bipartitions over the string and for each time step, both from experimental 8-spin data and theoretical simulations.

$$S = -tr(\rho \log_2 \rho) \text{ or also } S = -\sum_j \eta_j \log_2 \eta_j \quad ,$$

with the reduced state ρ and the eigenvalues η_j. We find that the entropy of the reconstructed MPS agrees with the expected values derived from theoretical simulations. The entropy is observed to increase in time and reaches values between $0.7 < S < 1$ for all bipartitions, after 5 ms. This means that—across almost all bipartitions—the same amount of entanglement is present as in a maximally entangled state of two qubits ($S = 1$). The entropies may grow beyond $S = 1$ at later times, but at those times, we do not have a positive fidelity lower bound and no certified estimate.

VI. DIRECT FIDELITY ESTIMATION

The fidelity lower bounds returned by the certification procedure described in the main text are merely lower bounds. That is, the actual overlap (fidelity) between the two states could take any value between the certificate and unity. Which value does the fidelity actually take is a natural question to ask. To estimate the overlap, we implement the method of direct fidelity estimation (DFE) [17, 18]. The DFE method uses measurements on a lab state to determine an estimate of the fidelity between the lab state (generally mixed) and a given pure state, which we set as the output MPS from the efficient tomography procedure. In this section, we provide an overview of DFE with emphasis on the current experiment.

A. Overview of direct fidelity estimation procedure

Before describing the employed DFE method, we recollect relevant notation. Consider the state ρ_{lab} implemented in the laboratory and let $\left|\psi_c^3\right\rangle$ be the pure-state estimate obtained from MPS tomography on ρ_{lab}. The fidelity of the estimate with respect to the lab state is defined as

$$F\left(\left|\psi_c^3\right\rangle, \rho_{\text{lab}}\right) \stackrel{\text{def}}{=} \left\langle \psi_c^3 \right| \rho_{\text{lab}} \left| \psi_c^3 \right\rangle. \tag{53}$$

Define $\{P^k : k = 1, 2, \ldots 4^N\}$ as the orthonormal Pauli operators $\text{tr}\left(P^\ell P^k\right) \stackrel{\text{def}}{=} \delta_{\ell,k}$. The Pauli operators form a basis for Hermitian operators acting on the system Hilbert space. In this basis, the lab state and the MPS estimate can be represented by their characteristic functions

$$\rho_{\text{lab}}^k = \text{tr}\left(P^k \rho_{\text{lab}}\right), \quad \sigma^k = \text{tr}\left(P^k \left|\psi_c^3\right\rangle \left\langle \psi_c^3 \right|\right) = \left\langle \psi_c^3 \right| P^k \left| \psi_c^3 \right\rangle \tag{54}$$

respectively. The fidelity (53) is expressed in terms of the characteristic functions as

$$F\left(\left|\psi_c^3\right\rangle, \rho_{\text{lab}}\right) = \sum_{k=1}^{4^N} \rho_{\text{lab}}^k \sigma^k. \tag{55}$$

Estimating the fidelity by a straightforward application of Equation (55) requires measuring 4^N observables, each of which requires exponentially many (in N) measurements, and is thus infeasible. This implies that over two hundred million observables need to be measured in our setting of fourteen spins, which is clearly infeasible.

The DFE method leverages the knowledge of the MPS estimate to overcome this infeasibility. Specifically, the full summation of Equation (55) is replaced by a preferential summation over those values of k for which MPS-estimate components σ^k are likely to be large. In other words, more measurements are made in those basis elements P^k for which the MPS estimate is known to have a large expectation value.

The preferential summation to obtain the fidelity estimate is performed as follows. First, the fidelity is expressed as the expectation value

$$F\left(\left|\psi_c^3\right\rangle, \rho\right) = \sum_{k=1}^{4^N} \rho_{\text{lab}}^k \sigma^k = \sum_{k=1}^{4^N} q^k \frac{\rho_{\text{lab}}^k}{\sigma^k}, \tag{56}$$

of a random variable $\rho_{\text{lab}}^k/\sigma^k$ over probability distribution $\left\{q^k \stackrel{\text{def}}{=} (\sigma^k)^2 : k = 1, 2, \ldots, 4^N\right\}$. Next, this expectation value is estimated using a Monte Carlo approach. That is, M random indices $\left\{k_1, k_2, \ldots, k_M; k_m \in \{1, 2, \ldots, 4^N\}\right\}$ are generated according to the probability distribution q^k where M is chosen based on a desired threshold of error. In the experiment we set $M = 250$. In other words, the number of observables required to estimate the fidelity are reduced by six orders of magnitude.

The fidelity is obtained from the estimator

$$\overline{F} \stackrel{\text{def}}{=} \frac{1}{M}\sum_{i=1}^{M}\frac{\overline{\rho}_{\text{lab}}^{k_i}}{\sigma^{k_i}} \approx F\left(\left|\psi_c^3\right\rangle, \rho_{\text{lab}}\right), \tag{57}$$

where $\overline{\rho}_{\text{lab}}^{k_m}$ estimates the lab-state expectation value (54). These estimates are obtained from measuring $M = 250$ observables using many copies of the state for each observable, where a total of 5×10^5 copies of the state were used. The number of copies N_k spent to measure a particular Pauli operator P_k was chosen to be proportional to the inverse square of its calculated expectation value σ^{k_i} for ψ_c^3 in order to prevent the error in the estimator $\overline{F}$ to be dominated by those terms of the sum in Eq. 57 for which σ^{k_i} is very small.

The fidelity estimate (57) requires sampling the indices $\{k_1, k_2, \ldots, k_M\}$; this sampling is efficiently performed using classical algorithms outlined in the supplementary material of Reference [18]. Finally, the values $\{\sigma^{k_1}, \sigma^{k_2}, \ldots, \sigma^{k_M}\}$ are determined by efficient MPS-based classical algorithms. This completes a summary of the direct fidelity estimation procedure.

The values of $\overline{\rho}_{\text{lab}}^{k_m}$ obtained in the experiment and the corresponding calculated σ^{k_m} are depicted in FIG. S10(a), for the $M = 250$ different observables, which are indexed by m. The distribution of $\overline{\rho}_{\text{lab}}^{k_m}/\sigma^{k_m}$ for different m is presented in FIG. S10(b). Based on this distribution, we infer a fidelity estimate of 0.74. We present the procedure for calculating the error bars on this estimate in the next section.

B. Mean-square error and bias of DFE estimates

The fidelity estimate (57) is amenable to random error, which arises from (i) choosing a smaller number M of indices than the maximum possible 4^N and from (ii) using a finite number of measurements to estimate the expectation values $\{\rho_{\text{lab}}^{k_1}, \rho_{\text{lab}}^{k_2}, \ldots, \rho_{\text{lab}}^{k_M}\}$. This random error is quantified by the variance estimator

$$\overline{\text{var}}[F] \stackrel{\text{def}}{=} \frac{1}{M(M-1)}\sum_{m=1}^{M}\left(\frac{\overline{\rho}_{\text{lab}}^{k_m}}{\sigma^{k_m}} - \overline{F}\right)^2, \tag{58}$$

where $\overline{F}$ is determined using Equation 57. In the remainder of this section, we justify that $\overline{F}\left(\left|\psi_c^3\right\rangle, \rho_{\text{lab}}\right)$ (57) and $\overline{\text{var}}[F]$ (58) are unbiased estimators of the fidelity and the variance of the fidelity. In other words, the expectation value of the random variables $\overline{F}$ and $\overline{\text{var}}[F]$ are respectively equal to the true value of the fidelity and the fidelity variance.

In order to account for random error from the experiment, we connect the fidelity estimator (57) and variance estimator (58) with the measurement outcomes from the experiment. We account for random error in estimates $\overline{\rho}_{\text{lab}}^{k_m}$ by expressing $\rho_{\text{lab}}^{k_m}$ in terms of measurement outcomes:

$$\rho_{\text{lab}}^k = \text{tr}\left(P^k\rho_{\text{lab}}\right) = \sum_{i=1}^{2^N}\lambda_i^k\,\text{tr}\left(\left|\psi_i^k\right\rangle\left\langle\psi_i^k\right|\rho_{\text{lab}}\right) = \sum_{i=1}^{2^N}\lambda_i^k\left\langle\psi_i^k\right|\rho_{\text{lab}}\left|\psi_i^k\right\rangle = \sum_{i=1}^{2^N}\lambda_i^k p_i^k, \tag{59}$$

where $\{\lambda_i^k : k = 1, 2, \ldots, 2^N\}$ are the eigenvalues of Pauli operators

$$P^k = \sum_{i=1}^{2^N}\lambda_i^k\left|\psi_i^k\right\rangle\left\langle\psi_i^k\right|, \quad \lambda_i^k = \pm\frac{1}{\sqrt{2^N}}, \tag{60}$$

and $\left\{p_i^k \stackrel{\text{def}}{=} \left\langle\psi_i^k\right|\rho_{\text{lab}}\left|\psi_i^k\right\rangle : k = 1, 2, \ldots, 2^N\right\}$ is a probability distribution. Finally, the fidelity can be expressed as the expectation value

$$\begin{aligned} F\left(\left|\psi_c^3\right\rangle, \rho_{\text{lab}}\right) &= \sum_{k=1}^{4^N} q^k\frac{\rho_{\text{lab}}^k}{\sigma^k} = \sum_{k=1}^{4^N} q^k\frac{\sum_{i=1}^{2^N}\lambda_i^k p_i^k}{\sigma^k} \\ &= \sum_{k=1}^{4^N}\sum_{i=1}^{2^N} q^k p_i^k\frac{\lambda_i^k}{\sigma^k} = \sum_{k=1}^{4^N}\sum_{i=1}^{2^N} u_i^k\frac{\lambda_i^k}{\sigma^k}, \end{aligned} \tag{61}$$

of the random variable $\frac{\lambda_i^k}{\sigma^k}$ over the probability distribution $\left\{u_i^k \stackrel{\text{def}}{=} q^k p_i^k : k = 1, 2, \ldots, 4^N, i = 1, 2, \ldots, 2^N\right\}$.

In the experiment, we choose M observables $\{P^1, P^2, \ldots, P^M\}$. Each observable P^m is measured N_m times, with each measurement returning outcome value $\lambda_{i_n}^{k_m}$. The returned measurement outcomes are used to obtain expectation values as

$$\overline{\rho}_{\text{lab}}^{k_m} = \frac{1}{N_m} \sum_{n=1}^{N_m} \lambda_{i_n}^{k_m}. \tag{62}$$

Thus, the fidelity estimate (57) is obtained from measurement outcomes as

$$\overline{F}\left(\left|\psi_c^3\right\rangle, \rho_{\text{lab}}\right) = \frac{1}{M} \sum_{m=1}^{M} \frac{\frac{1}{N_m} \sum_{n=1}^{N_m} \lambda_{i_n}^{k_m}}{\sigma^{k_m}}. \tag{63}$$

The variance of $\overline{F}\left(\left|\psi_c^3\right\rangle, \rho_{\text{lab}}\right)$ is obtained using estimator (58), which we express in terms of measurement outcomes by the following simplification. Consider

$$\overline{\text{var}}[F] \stackrel{\text{def}}{=} \frac{1}{M(M-1)} \sum_{m=1}^{M} \left(\frac{\overline{\rho}_{\text{lab}}^{k_m}}{\sigma^{k_m}} - \overline{F}\right)^2, \tag{64}$$

$$= \frac{1}{M(M-1)} \sum_{m=1}^{M} \left(\frac{\left(\overline{\rho}_{\text{lab}}^{k_m}\right)^2}{\left(\sigma^{k_m}\right)^2} - 2\frac{\overline{\rho}_{\text{lab}}^{k_m}}{\sigma^{k_m}}\overline{F} + \left(\overline{F}\right)^2\right), \tag{65}$$

$$= \frac{1}{M(M-1)} \sum_{m=1}^{M} \frac{\left(\overline{\rho}_{\text{lab}}^{k_m}\right)^2}{\left(\sigma^{k_m}\right)^2} - \frac{M}{M(M-1)}\left(\overline{F}\right)^2, \tag{66}$$

where we have used the definition (57) to obtain (66) from (65). Substituting the expression for estimators $\overline{F}$ and $\overline{\rho}_{\text{lab}}^k$, we obtain

$$\overline{\text{var}}[F] = \frac{1}{M(M-1)} \sum_{m=1}^{M} \frac{\left(\frac{1}{N_m} \sum_{n=1}^{N_m} \lambda_{i_n}^{k_m}\right)^2}{\left(\sigma^{k_m}\right)^2} - \frac{M}{M(M-1)} \left(\frac{1}{M} \sum_{m=1}^{M} \frac{\frac{1}{N_m} \sum_{n=1}^{N_m} \lambda_{i_n}^{k_m}}{\sigma^{k_m}}\right)^2 \tag{67}$$

$$= \frac{1}{M(M-1)} \sum_{m=1}^{M} \frac{1}{N_m^2} \sum_{n,n'=1}^{N_m} \frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}} \frac{\lambda_{i_{n'}}^{k_m}}{\sigma^{k_m}} - \frac{1}{M^2(M-1)} \sum_{m,m'=1}^{M} \frac{1}{N_m N_{m'}} \sum_{n,n'=1}^{N_m} \frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}} \frac{\lambda_{i_{n'}}^{k_{m'}}}{\sigma^{k_{m'}}} \tag{68}$$

$$= \frac{1}{M^2} \sum_{m=1}^{M} \frac{1}{N_m^2} \sum_{n,n'=1}^{N_m} \frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}} \frac{\lambda_{i_{n'}}^{k_m}}{\sigma^{k_m}} - \frac{1}{M^2(M-1)} \sum_{\substack{m,m'=1 \\ m \neq m'}}^{M} \frac{1}{N_m N_{m'}} \sum_{n,n'=1}^{N_m} \frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}} \frac{\lambda_{i_{n'}}^{k_{m'}}}{\sigma^{k_{m'}}}, \tag{69}$$

which is the variance estimator in terms of measurement outcomes

Now we show that the fidelity estimator (63) is an unbiased estimator. Consider the expectation value of the fidelity

$$\mathbb{E}[\overline{F}] = \mathbb{E}\left[\frac{1}{M} \sum_{m=1}^{M} \frac{\frac{1}{N_m} \sum_{n=1}^{N_m} \lambda_{i_n}^{k_m}}{\sigma^{k_m}}\right] \tag{70}$$

$$= \frac{1}{M} \sum_{m=1}^{M} \frac{1}{N_m} \sum_{n=1}^{N_m} \mathbb{E}\left[\frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}}\right]. \tag{71}$$

We note that the expectation value of $\frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}}$ is equal to the true fidelity

$$\mathbb{E}\left[\frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}}\right] = \sum_{k_m=}^{4^N} \sum_{i_n=1}^{2^N} u_{i_n}^{k_m} \frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}} = F\left(\left|\psi_c^3\right\rangle, \rho_{\text{lab}}\right) \tag{72}$$

because each of the $\lambda_{i_n}^{k_m}$ values are drawn from the same distribution for each value of m and n. Substituting Equation (72) in the fidelity expectation value (71), we obtain

$$\mathbb{E}[\overline{F}] = F\left(\left|\psi_c^3\right\rangle, \rho\right), \tag{73}$$

which implies that $\overline{F}$ is an unbiased estimator of the fidelity.

Finally, we show $\overline{\text{var}}[F]$ is an unbiased estimator, i.e., that the expectation value of $\overline{\text{var}}[F]$ is the same as the true variance

$$\text{var}[F] \stackrel{\text{def}}{=} \mathbb{E}\left[\overline{F}^2\right] - \left(\mathbb{E}\left[\overline{F}\right]\right)^2 \tag{74}$$

of the estimator $\overline{F}$. From Equation (69), we have the expectation value,

$$\mathbb{E}\left(\overline{\text{var}}[F]\right) = \mathbb{E}\left(\frac{1}{M^2}\sum_{m=1}^{M}\frac{1}{N_m^2}\sum_{n,n'=1}^{N_m}\frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}}\frac{\lambda_{i_{n'}}^{k_m}}{\sigma^{k_m}} - \frac{1}{M^2(M-1)}\sum_{\substack{m,m'=1\\ m\neq m'}}^{M}\frac{1}{N_m N_{m'}}\sum_{n,n'=1}^{N_m}\frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}}\frac{\lambda_{i_{n'}}^{k_{m'}}}{\sigma^{k_{m'}}}\right), \tag{75}$$

which we simplify as

$$\begin{aligned}
\mathbb{E}\left(\overline{\text{var}}[F]\right) &= \mathbb{E}\left(\frac{1}{M^2}\sum_{m=1}^{M}\frac{1}{N_m^2}\sum_{n,n'=1}^{N_m}\frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}}\frac{\lambda_{i_{n'}}^{k_m}}{\sigma^{k_m}}\right) \\
&\quad - \mathbb{E}\left(\frac{1}{M^2(M-1)}\sum_{\substack{m,m'=1\\ m\neq m'}}^{M}\frac{1}{N_m N_{m'}}\sum_{n,n'=1}^{N_m}\frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}}\frac{\lambda_{i_{n'}}^{k_{m'}}}{\sigma^{k_{m'}}}\right) && (76)\\
&= \mathbb{E}\left(\frac{1}{M^2}\sum_{m=1}^{M}\frac{1}{N_m^2}\sum_{n,n'=1}^{N_m}\frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}}\frac{\lambda_{i_{n'}}^{k_m}}{\sigma^{k_m}}\right) \\
&\quad - \frac{1}{M^2(M-1)}\sum_{\substack{m,m'=1\\ m\neq m'}}^{M}\mathbb{E}\left(\sum_{n=1}^{N_m}\frac{1}{N_m}\frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}}\right)\mathbb{E}\left(\sum_{n'=1}^{N_m}\frac{1}{N_{m'}}\frac{\lambda_{i_{n'}}^{k_{m'}}}{\sigma^{k_{m'}}}\right) && (77)\\
&= \mathbb{E}\left(\frac{1}{M^2}\sum_{m=1}^{M}\frac{1}{N_m^2}\sum_{n,n'=1}^{N_m}\frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}}\frac{\lambda_{i_{n'}}^{k_m}}{\sigma^{k_m}}\right) - \frac{1}{M^2(M-1)}\sum_{\substack{m,m'=1\\ m\neq m'}}^{M}\left(\mathbb{E}\left[\overline{F}\right]\right)^2, && (78)
\end{aligned}$$

where in the last step we have used Equation (63) and that each $\lambda_{i_n}^{k_m}$ is drawn from the same distribution. We add and subtract $\left(\mathbb{E}\left[\overline{F}\right]\right)^2$ to obtain

$$\begin{aligned}
\mathbb{E}\left(\overline{\text{var}}[F]\right) &= \mathbb{E}\left(\frac{1}{M^2}\sum_{m=1}^{M}\frac{1}{N_m^2}\sum_{n,n'=1}^{N_m}\frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}}\frac{\lambda_{i_{n'}}^{k_m}}{\sigma^{k_m}}\right) - \\
&\quad \frac{1}{M^2(M-1)}\sum_{\substack{m,m'=1\\ m\neq m'}}^{M}\left(\mathbb{E}\left[\overline{F}\right]\right)^2 + \left(\mathbb{E}\left[\overline{F}\right]\right)^2 - \left(\mathbb{E}\left[\overline{F}\right]\right)^2 && (79)\\
&= \mathbb{E}\left(\frac{1}{M^2}\sum_{m=1}^{M}\frac{1}{N_m^2}\sum_{n,n'=1}^{N_m}\frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}}\frac{\lambda_{i_{n'}}^{k_m}}{\sigma^{k_m}}\right) - \frac{1}{M^2(M-1)}\sum_{\substack{m,m'=1\\ m\neq m'}}^{M}\left(\mathbb{E}\left[\overline{F}\right]\right)^2 \\
&\quad + \frac{1}{M(M-1)}\sum_{\substack{m,m'=1\\ m\neq m'}}^{M}\left(\mathbb{E}\left[\overline{F}\right]\right)^2 - \left(\mathbb{E}\left[\overline{F}\right]\right)^2 && (80)\\
&= \mathbb{E}\left(\frac{1}{M^2}\sum_{m=1}^{M}\frac{1}{N_m^2}\sum_{n,n'=1}^{N_m}\frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}}\frac{\lambda_{i_{n'}}^{k_m}}{\sigma^{k_m}}\right) \\
&\quad + \frac{1}{M^2}\sum_{\substack{m,m'=1\\ m\neq m'}}^{M}\left(\mathbb{E}\left[\overline{F}\right]\right)^2 - \left(\mathbb{E}\left[\overline{F}\right]\right)^2. && (81)
\end{aligned}$$

Performing the simplification of Equation (77)–(78), we combine the summations of the first two terms as

$$\mathbb{E}\left(\overline{\mathrm{var}}[F]\right) = \mathbb{E}\left(\frac{1}{M^2}\sum_{m=1}^{M}\frac{1}{N_m^2}\sum_{n,n'=1}^{N_m}\frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}}\frac{\lambda_{i_{n'}}^{k_m}}{\sigma^{k_m}}\right) + \frac{1}{M^2}\sum_{\substack{m,m'=1\\ m\neq m'}}^{M}\mathbb{E}\left(\sum_{n=1}^{N_m}\frac{1}{N_m}\frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}}\right)\mathbb{E}\left(\sum_{n'=1}^{N_m}\frac{1}{N_{m'}}\frac{\lambda_{i_{n'}}^{k_{m'}}}{\sigma^{k_{m'}}}\right) - \left(\mathbb{E}\left[\overline{F}\right]\right)^2 \tag{82}$$

$$= \frac{1}{M^2}\sum_{m,m'=1}^{M}\mathbb{E}\left(\sum_{n=1}^{N_m}\frac{1}{N_m}\frac{\lambda_{i_n}^{k_m}}{\sigma^{k_m}}\sum_{n'=1}^{N_{m'}}\frac{1}{N_{m'}}\frac{\lambda_{i_{n'}}^{k_{m'}}}{\sigma^{k_{m'}}}\right) - \left(\mathbb{E}\left[\overline{F}\right]\right)^2 \tag{83}$$

$$= \mathbb{E}\left[\overline{F}^2\right] - \left(\mathbb{E}\left[\overline{F}\right]\right)^2, \tag{84}$$

which is the same as the variance (74) of the fidelity estimate. Thus, we conclude that the variance estimator $\overline{\mathrm{var}}[F]$ is an unbiased estimator.

In summary, we have presented a procedure for estimating error bars on the DFE and have shown that the procedure returns an unbiased estimator of the variance. Using this procedure and the DFE procedure described above we obtain the estimate 0.74 ± 0.05 for the fidelity of the experimental 14-ion state at $t_{14} = 4$ ms with our MPS estimate.

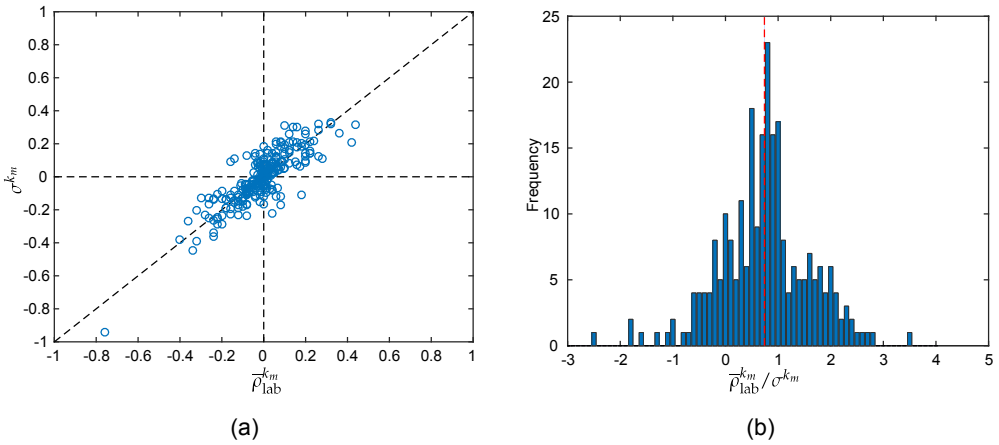


FIG. S10: **Measured expectation values used for direct fidelity estimation of the quenched 14-spin MPS state.** Outcomes are shown for measurements of the 14-spin quench state $|\psi_k^3\rangle$ after simulator evolution time of $t = 4$ ms. For explanation of direct fidelity estimation (DFE) and definition of algebra used below, see Methods. (a) Scatter plot of the expected (for MPS state) and observed (for lab state) expectation values, σ^{k_m} and $\overline{\rho}_{\mathrm{lab}}^{k_m}$ respectively, for the chosen $M = 250$ observables. Positive correlation is apparent. (b) The distribution of the random variable $\overline{\rho}_{\mathrm{lab}}^{k_m}/\sigma^{k_m}$ for the different observables. The mean (vertical red dashed line) and standard deviation of this distribution are the respective estimators of the fidelity estimate and its error. For our experiment, the obtained fidelity estimate is 0.74 ± 0.05.

VII. CERTIFIED MPS TOMOGRAPHY IS EFFICIENT FOR 1D LOCAL QUENCH DYNAMICS

A. Summary of the results

In this section, we show that certified MPS tomography can be used to characterise local quench dynamics with resources that scale efficiently in system size. Specifically, we provide upper bounds to the resources, both experimental and computational, required to characterise a state obtained by evolving a pure product state under a 1D nearest-neighbour Hamiltonian [24]. These required resources grow no faster than polynomially in the size N of the system, inverse polynomially with the tolerated infidelity $\mathcal{I}$ of characterisation, and exponentially in the time t of evolution. That is, at any given time t during the quench dynamics, the resources to characterise the state scale efficiently (polynomially) in system size.

To show that certified MPS tomography is efficient for quenched states, a necessary condition is that these states admit an efficient (in N) MPS representation. This follows from simple arguments in addition to Corollary 3 of Reference [19]. However, the existence of the MPS is not sufficient to guarantee the

existence of a parent Hamiltonian with suitable spectral properties, which is essential in our certification procedure. In this section, we show that such a Hamiltonian for quenched states indeed exists, and this existence enables the use of the certified MPS tomography procedure for these states.

Our argument is structured as follows. First, we show that pure product states have parent Hamiltonians that have a unit gap above a non-degenerate ground state and that comprise local terms acting on single sites only. Next, we generalise the pure product state result to quenched states, i.e., states that start out as pure product states and undergo a time evolution under a nearest-neighbour Hamiltonian. We show that such states too can be well approximated by states which have a gapped parent Hamiltonian. These parent Hamiltonians comprise local terms that act on subsystems whose sizes 2Ω scale linearly in time and logarithmically in N and in $1/\mathcal{I}$.

Physically, the existence of gapped parent Hamiltonians implies that the quenched states can be uniquely identified using only their local reductions because these Hamiltonians are local and have a unique ground state. Furthermore, by showing the existence of 2Ω-sized gapped parent Hamiltonians, we impose upper bounds on the required resources (number of measurements and computational time) required to characterise the state. Characterising a ground state with 2Ω-sized parent Hamiltonian requires measuring and classically processing a linear (in N) number of 2Ω-sized local reductions on a 1D chain. This characterisation task requires resources (number of measurements and classical post-processing time) that scale exponentially in Ω and linearly in N via certified MPS tomography [6, 7]. From this, and the scaling of Ω in the parameters N, $\mathcal{I}$ and t, we obtain the mentioned scaling of the experimental and computational cost in terms of these parameters. In the next subsection, we recall results regarding gapped parent Hamiltonians of pure quantum states.

B. Background: parent Hamiltonian certificates

Here we recall briefly relevant notation regarding the parent Hamiltonian of a pure quantum state of N qubits on a linear chain as introduced in Section IV A 3. The parent Hamiltonian of a pure state $|\psi\rangle$ is any Hermitian linear operator H such that $|\psi\rangle$ is the ground state of H. We assume that $|\psi\rangle$ is normalized, we set the ground state energy, i.e., the lowest eigenvalue $\langle\psi|H|\psi\rangle$, of H to zero and we denote by $E_1 \geq 0$ the second smallest eigenvalue in which eigenvalues are repeated according to their multiplicities. The Hamiltonian is said to have a gap of size E_1. The gap is non-zero (i.e., $E_1 > 0$) if and only if the ground state is non-degenerate.

Now we consider the energy of any arbitrary, possibly mixed, state ρ with respect to H and H is assumed to have a non-degenerate ground state. Then, the fidelity $F(|\psi\rangle, \rho) = \langle\psi|\rho|\psi\rangle$ of ρ with respect to $|\psi\rangle$ is bounded below according to Eq. (9) with $E_0 = 0$. That is,

$$F(|\psi\rangle, \rho) \geq 1 - \frac{\mathrm{tr}(\rho H)}{E_1}. \tag{85}$$

Thus, the energy of ρ in terms of H provides a lower bound to the fidelity between $|\psi\rangle$ and ρ; we call a lower bound to the fidelity a certificate.

The certification of the lab state using H is efficient, that is, requires number of measurements and computation-time that scale polynomially in the number of qudits if H is of the following form. Suppose that H is a sum

$$H = \sum_{i=1}^{N-k+1} h_i \tag{86}$$

of terms which act non-trivially

$$h_i = \mathbb{1}_1 \otimes \mathbb{1}_2 \otimes \cdots \otimes \mathbb{1}_{i-1} \otimes h^{\ell}_{i,i+1,\ldots,i+k-1} \otimes \mathbb{1}_{i+k} \otimes \cdots \otimes \mathbb{1}_N \tag{87}$$

only on small (i.e., size $k \in O(\log N)$) subsets of the complete system. Then, only the matching local reductions of ρ are necessary to obtain the energy $\mathrm{tr}(\rho H)$:

$$\mathrm{tr}(\rho H) = \sum_i \mathrm{tr}(\rho h_i) = \sum_i \mathrm{tr}(\rho_i h_i), \tag{88}$$

where ρ_i are the reduced density matrices that act on subsystem $\{i, i+1, \ldots, i+k-1\}$. In this case the certificate can be obtained from a number of measurements that scales linearly in the number of particles (and exponentially in the subsystem size k). This is an exponential improvement over the number of measurements required for estimating fidelity by performing standard quantum state tomography. Furthermore, the summation of Eq. (88) can be performed in linear (in N) computational time as compared to the exponentially large computation time required if the output from full tomography is used to obtain fidelity. In summary, determining the fidelity certificate is efficient with respect to measurement and computation time.

C. Product states have simple parent Hamiltonians

In this section, we show that pure product states admit a parent Hamiltonian that has unit gap and only single-site local terms (Lemma 2). This result is a simple special case of prior work involving matrix product states [6, 10, 11].

Lemma 2. *Let* $|\varphi\rangle = |\varphi_1\rangle \otimes |\varphi_2\rangle \otimes \cdots \otimes |\varphi_N\rangle$ *be a product state on* N *qudits of dimension* $d_i \geq 2$, $i \in \{1, \ldots, N\}$. *Let* $\langle\varphi_i|\varphi_i\rangle = 1$ *for all site indices* i. *Define*

$$H = \sum_{i=1}^{N} h_i, \quad h_i = \mathbb{1}_{1,\ldots,i-1} \otimes P_{\ker(\rho_i)} \otimes \mathbb{1}_{i+1,\ldots,N} \tag{89}$$

where $P_{\ker(\rho_i)}$ *is the orthogonal projection onto the null space of the reduced density operator* $\rho_i = |\varphi_i\rangle\langle\varphi_i|$ *of* $|\varphi\rangle$ *on site* i. *Then, the eigenvalues of* H *are given by* $\{0, 1, 2, \ldots, N\}$, *the smallest eigenvalue zero is non-degenerate and* $|\varphi\rangle$ *is an eigenvector of eigenvalue zero.*

Proof. Expand the null space projectors h_i in terms of the single-site pure states $|\varphi_i\rangle$ as

$$\begin{aligned} h_i &= \mathbb{1}_{1,\ldots,i-1} \otimes P_{\ker(\rho_i)} \otimes \mathbb{1}_{i+1,\ldots,N} \\ &= \mathbb{1}_{1,\ldots,i-1} \otimes (\mathbb{1}_i - |\varphi_i\rangle\langle\varphi_i|) \otimes \mathbb{1}_{i+1,\ldots,N}. \end{aligned} \tag{90}$$

We obtain a complete eigendecomposition of H as follows. For each site i, construct an orthonormal basis $|\mu_{i,1}\rangle, \ldots, |\mu_{i,d_i}\rangle$ such that $|\mu_{i,1}\rangle = |\varphi_i\rangle$. Consider the corresponding orthonormal product basis of the complete system given by

$$|\mu_L\rangle = |\mu_{1,\ell_1}\rangle \otimes |\mu_{2,\ell_2}\rangle \otimes \cdots \otimes |\mu_{N,\ell_N}\rangle, \tag{91a}$$

$$L = (\ell_1, \ell_2, \ldots, \ell_N), \quad \ell_i \in \{1, \ldots, d_i\}. \tag{91b}$$

We have

$$h_i |\mu_L\rangle = \begin{cases} 0, & \text{if } \ell_i = 1, \\ 1, & \text{if } \ell_i > 1. \end{cases} \tag{92}$$

and thus

$$H |\mu_L\rangle = \lambda_L |\mu_L\rangle, \quad \lambda_L = \left|\left\{i \colon \ell_i > 1, i \in \{1, \ldots, N\}\right\}\right|. \tag{93}$$

The basis vectors $|\mu_L\rangle$ are seen to constitute an orthonormal eigenbasis of H and each eigenvalue λ_L is given by the number of tensor product factors $|\mu_{i,\ell_i}\rangle$ with $\ell_i > 1$, i.e., with $\langle\varphi_i|\mu_{i,\ell_i}\rangle = 0$. Because we have required $d_i \geq 2$, this already shows everything stated in the theorem: The eigenvalues of H are given by $\{0, 1, 2, \ldots, N\}$, the smallest eigenvalue zero is non-degenerate and $|\varphi\rangle$ is an eigenvector of eigenvalue zero. □

D. Parent Hamiltonian for a locally time evolved state

This section presents the main results regarding the parent Hamiltonians of quenched states. We show that the experimental and computational cost of characterising quenched states scales no faster than polynomially in $N/\mathcal{I}$ and exponentially in the quench time t.

Our proof is in two parts. First, we follow [20] to construct a state which closely approximates our time-evolved state. This approximate state is obtained by starting with a tensor product of pure states on at most Ω neighbouring sites and applying a single unitary operation that is a tensor product of unitaries on at most Ω sites (FIG. S11). Here, Ω depends on t, N and $\mathcal{I}$ but is seen to scale moderately.

Next, in Theorem 4 we show that this approximate state is the non-degenerate ground state of a suitable parent Hamiltonian. From Lemma 2, we know that the pure product state has a parent Hamiltonian with terms acting on at most Ω sites. The unitary operation does not increase the range of the terms in the pure-state parent Hamiltonian from Ω to more than 2Ω, which is small, i.e., $O(\log N)$, for suitably low quench time $O(\log N)$. This parent Hamiltonian enables the usual certification procedure described in Sec. VII B and the main text.

We construct the approximate state using the following theorem from Reference [20].

Theorem 3 (ϵ-QCA decomposition, [20])**.** *Let* H *be a nearest-neighbour Hamiltonian on* N *qubits in a linear chain, i.e.,* $H = \sum_{i=1}^{N-1} h_{i,i+1}$. *We fix a positive integer* Ω *and partition the linear chain into*

$\mathcal{N} = 2N/\Omega$ contiguous blocks each containing at most $\Omega/2$ qubits (FIG. S11). There is an approximation of the time evolution operator $U = \mathrm{e}^{-\mathrm{i}Ht}$ of the form

$$U' = [U_{12} \otimes U_{34} \otimes \cdots \otimes U_{\mathcal{N}-1,\mathcal{N}}][V_1 \otimes V_{23} \otimes V_{45} \otimes \cdots \otimes V_{\mathcal{N}-2,\mathcal{N}-1} \otimes V_{\mathcal{N}}] \tag{94}$$

where $U_{j,j+1}$, $V_{j,j+1}$ and V_j are unitaries acting on the blocks specified by the subscripts. This approximation satisfies $\|U - U'\| \leq \epsilon$ under the restriction that

$$\Omega \geq c_0|t| + c_1 \log(N/\epsilon) \tag{95}$$

where $\|\cdot\|$ is the operator norm and c_0 and c_1 are constants.

Let $|\psi(0)\rangle = |\psi_1\rangle \otimes \cdots \otimes |\psi_N\rangle$ be an initial product state [25]. The state $|\psi'\rangle = U'|\psi(0)\rangle$ is an approximation of the time-evolved state $|\psi(t)\rangle = U|\psi(0)\rangle$ with $\||\psi'\rangle - |\psi(t)\rangle\| \leq \epsilon$. The approximation $|\psi'\rangle$ has a matrix product state representation with bond dimension no more than 2^{Ω}.

Thus, the state $|\psi'\rangle$ closely approximates our time-evolved state. Specifically, if Ω is required to scale with $c' \log(N)$ where c' is a suitable constant, then the norm-distance between $|\psi\rangle$ and $|\psi'\rangle$ scales as as inverse polynomial in N. Now we prove the existence of the parent Hamiltonian of $|\psi'\rangle$ and find an upper bound to the number of sites that the parent Hamiltonian terms act on.

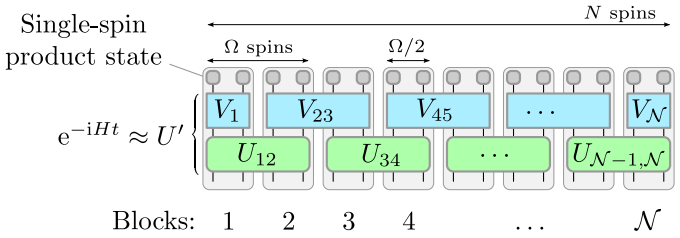


FIG. S11: **The ϵ-QCA decomposition** of a unitary time evolution under a 1D nearest-neighbour Hamiltonian. A linear chain of N spins is divided into $\mathcal{N} = 2N/\Omega$ blocks, such that each block contains at most $\Omega/2$ spins. The decomposition and the relation between N, Ω, approximation error ϵ and evolution time t have been introduced in Ref. [20]. We use the decomposition in Supplementary Material Sec. VII D to prove efficient certification of MPS tomography for 1D local quench dynamics starting from a pure product state. This is accomplished with a parent Hamiltonian G whose local terms act on at most four blocks, i.e., at most 2Ω spins (Theorems 4 and 7 in Supplementary Material Sec. VII D).

Theorem 4. *For $|\psi'\rangle$ as described in Theorem 3, there is a Hermitian linear operator $G = \sum_{j=1}^{N/\Omega+1} g_j$ with non-degenerate smallest eigenvalue zero and eigenvector $|\psi'\rangle$, second smallest eigenvalue one and largest eigenvalue $N/\Omega + 1$. Each local term g_j acts on no more than 2Ω consecutive qubits.*

Proof. We define the intermediate product state

$$|\phi\rangle = [V_1 \otimes V_{23} \otimes \cdots \otimes V_{\mathcal{N}}]|\psi(0)\rangle \tag{96}$$

$$=: |\phi_{01}\rangle \otimes |\phi_{23}\rangle \otimes \cdots \otimes |\phi_{\mathcal{N},\mathcal{N}+1}\rangle\,, \tag{97}$$

i.e., $|\phi_{j,j+1}\rangle$ is a state on blocks $j, j+1$. Blocks 0 and $\mathcal{N}+1$ are empty and have been introduced for notational convenience [26]. From Lemma 2, we have a parent Hamiltonian

$$F = f_{01} + f_{23} + \cdots + f_{\mathcal{N},\mathcal{N}+1} \tag{98}$$

of $|\phi\rangle$ with

$$f_{j,j+1} = \mathbb{1}_{1,\ldots,j-1} \otimes P_{\ker(|\phi_{j,j+1}\rangle\langle\phi_{j,j+1}|)} \otimes \mathbb{1}_{j+2,\ldots,\mathcal{N}}. \tag{99}$$

Furthermore, F has a non-degenerate smallest eigenvalue zero with eigenvector $|\phi\rangle$, second smallest eigenvalue one and largest eigenvalue $\mathcal{N}/2 + 1 = N/\Omega + 1$.

We define

$$\tilde{U} = U_{12} \otimes U_{34} \otimes \cdots \otimes U_{\mathcal{N}-1,\mathcal{N}} \tag{100}$$

and we define $|\psi'\rangle = \tilde{U}|\phi\rangle$. Because $\tilde{U}$ is unitary, the operator $G = \tilde{U}F\tilde{U}^{\dagger}$ has the same eigenvalue spectrum as F. That is, G has a non-degenerate smallest eigenvalue zero with eigenvector $|\psi'\rangle$, a second smallest eigenvalue one and a largest eigenvalue $N/\Omega + 1$. We obtain the following representation of G

$$G = g_{12} + g_{1234} + g_{3456} + \cdots + g_{\mathcal{N}-1,\mathcal{N}}, \tag{101}$$

where the identity operators are omitted and

$$g_{jklm} = [U_{jk} \otimes U_{lm}]\, f_{kl}\, [U_{jk} \otimes U_{lm}]^\dagger. \tag{102}$$

The border terms are given by $g_{12} = U_{12} f_{01} U_{12}^\dagger$ and $g_{N-1,N} = U_{N-1,N} f_{N,N+1} U_{N-1,N}^\dagger$. There are $N/2 + 1$ terms in the sum and each term acts on at most four blocks, i.e., at most 2Ω qubits. □

This completes our proof regarding the parent Hamiltonian of the approximate time-evolved state. In the following corollary, we use the parent Hamiltonian to obtain a fidelity certificate (Section VII B) for the lab state ρ with respect to the approximate state.

Corrolary 5. *Consider $|\psi'\rangle$ and G as described in Theorem 4 and define $\psi' \stackrel{\text{def}}{=} |\psi'\rangle\langle\psi'|$. Denote by $\|\cdot\|_1$ the trace norm of an operator. Let ρ be an arbitrary quantum state and let $\delta = \|\rho - \psi'\|_1$. Then*

$$\langle\psi'|\,\rho\,|\psi'\rangle \geq 1 - \mathrm{tr}(\rho G) \geq 1 - (N/\Omega + 1)\delta. \tag{103}$$

Proof. As G has unit gap the fidelity lower bound (cf. Eq. (85)) becomes

$$\begin{aligned} \langle\psi'|\,\rho\,|\psi'\rangle &\geq 1 - \mathrm{tr}(\rho G)/1 \\ &= 1 - \mathrm{tr}(\rho G), \end{aligned} \tag{104}$$

which is the first inequality of (103). Consider the energy $\mathrm{tr}(\rho G)$ of ρ with respect to G. Using $\mathrm{tr}(\psi' G) = 0$, we have

$$\mathrm{tr}(\rho G) = \left|\mathrm{tr}(\rho G) - \mathrm{tr}(\psi' G)\right| = \left|\mathrm{tr}([\rho - \psi']G)\right|. \tag{105}$$

Thus, the energy

$$\begin{aligned} \mathrm{tr}(\rho G) &\leq \|\rho - \psi'\|_1 \|G\|_\infty \\ &= (N/\Omega + 1)\delta \end{aligned} \tag{106}$$

is at most $\mathrm{tr}(\rho G) \leq (N/\Omega + 1)\delta$, where $\|\cdot\| = \|\cdot\|_\infty$ denotes the operator norm. Combining Equations (104) and (106), we obtain the required inequalities. □

The final Theorem requires the following Lemma:

Lemma 6. *Let $\|\cdot\|_1$ be the trace norm, $\psi = |\psi\rangle\langle\psi|$ and $\psi' = |\psi'\rangle\langle\psi'|$. If $\||\psi\rangle - |\psi'\rangle\| \leq \epsilon \leq \sqrt{2}$, then $\|\psi - \psi'\|_1 \leq 2\epsilon$.*

Proof. Assume that $\||\psi\rangle - |\psi'\rangle\| \leq \epsilon$ holds. This gives us

$$\epsilon^2 \geq 2(1 - \Re(\langle\psi|\psi'\rangle)) \geq 2(1 - \sqrt{F}) \tag{107}$$

where $F = |\langle\psi|\psi'\rangle|^2 = F(|\psi\rangle, |\psi'\rangle)$. This gives $\sqrt{F} \geq 1 - \epsilon^2/2$ and $1 - F = \epsilon^2 - \epsilon^4/4 \leq \epsilon^2$. The equality $\|\psi - \psi'\|_1 = 2\sqrt{1 - F}$ completes the proof [12, Eqs. 9.11, 9.60, 9.99]. □

Our final theorem considers a state ρ close to a locally time-evolved state $|\psi(t)\rangle$; as before, $|\psi'\rangle$ is an approximation of $|\psi(t)\rangle$. The theorem states the conditions under which the fidelity $\langle\psi'|\,\rho\,|\psi'\rangle$ can be lower bounded by at least $1 - \mathcal{I}$, for some infidelity $\mathcal{I}$:

Theorem 7. *Let H be a nearest-neighbour Hamiltonian on N qubits in a linear chain, i.e., $H = \sum_{i=1}^{N-1} h_{i,i+1}$. Let $|\psi(0)\rangle = |\psi_1\rangle \otimes \ldots |\psi_N\rangle$ be an initial product state [27] and let $|\psi(t)\rangle = \mathrm{e}^{-iHt}|\psi(0)\rangle$ be the time-evolved state. Define $\psi(t) = |\psi(t)\rangle\langle\psi(t)|$.*

Let $\gamma = \|\rho - \psi(t)\|_1$ be the trace distance between an unknown state ρ and the time-evolved state. Fix an infidelity $\mathcal{I}$ that satisfies $\mathcal{I} > 2N\gamma$. Choose an integer $\Omega \geq 1$ such that

$$\Omega \geq c_0|t| + c_1 \log\left(\frac{2N}{\mathcal{I}/2N - \gamma}\right), \tag{108}$$

with c_0, c_1 from Theorem 3. $\psi' = |\psi'\rangle\langle\psi'|$ is the approximation of $|\psi(t)\rangle$ from the same theorem. We also use the parent Hamiltonian G from Theorem 4.

Then, the fidelity lower bound between $|\psi'\rangle$ and ρ will be at least

$$\langle\psi'|\,\rho\,|\psi'\rangle \geq 1 - \mathrm{tr}(\rho G) \geq 1 - \mathcal{I} \tag{109}$$

Proof. Theorem 3 applies for

$$\epsilon = \frac{1}{2}\left(\frac{\mathcal{I}}{2N} - \gamma\right) \tag{110}$$

and it guarantees $\||\psi(t)\rangle - |\psi'\rangle\| \leq \epsilon$. As a consequence, we have $\|\psi(t) - \psi'\|_1 \leq 2\epsilon$ (Lemma 6) and

$$\|\rho - \psi'\|_1 \leq \|\rho - \psi(t)\|_1 + \|\psi(t) - \psi'\|_1 \leq \gamma + 2\epsilon. \tag{111}$$

In addition, we observe (using $N \geq 1$ and $\Omega \geq 1$)

$$\mathcal{I} = 2N(2\epsilon + \gamma) \geq (N+1)(2\epsilon + \gamma) \geq \left(\frac{N}{\Omega} + 1\right)(2\epsilon + \gamma). \tag{112}$$

We use Corollary 5 with $\delta \leq 2\epsilon + \gamma$. It provides

$$\langle\psi'|\,\rho\,|\psi'\rangle \geq 1 - \mathrm{tr}(\rho G) \geq 1 - (N/\Omega + 1)(2\epsilon + \gamma) \geq 1 - \mathcal{I}, \tag{113}$$

which is the required result.

□

If the experimental state ρ is the same as the quenched state $|\psi(t)\rangle$, then the requirement (108) for Ω changes to

$$\Omega \geq c_0|t| + c_2 \log(N) + c_3 \log\left(\frac{1}{\mathcal{I}}\right) + c_4 \tag{114}$$

which enables us to quantify the resources required for certification.

E. Conclusion

In summary, the time-evolved state $|\psi(t)\rangle$ from Theorem 7 can be certified up to infidelity $\mathcal{I}$ with respect to a state $|\psi'\rangle$, which has a parent Hamiltonian G with non-degenerate ground state and unit gap. The local terms in G act on at most 2Ω sites. Ω can be chosen as the lowest integer that satisfies Equation (114) depending on the evolution time t, number of qubits N and infidelity $\mathcal{I}$. Note that Ω grows no faster than linearly with time and logarithmically with $N/\mathcal{I}$.

The existence of gapped parent Hamiltonians with 2Ω-sized terms means that the quenched states can be uniquely identified, in principle, using only 2Ω-sized local reductions. Whether such a state can actually be obtained using existing numerical algorithms is discussed in References [6, 7]. Although, no formal proofs for the convergence of these algorithms are available, we observe (main text) that these algorithms perform well in practice. Theorem 7 complements these discussions by ensuring that the fidelity of any reconstructed state with respect to the lab state can always be bounded from below. If this lower bound is smaller than desired, then the numerical algorithms can be run again, perhaps with more measurements to reduce quantum projection noise or with measurements on larger-sized subsystems to capture all correlations.

Theorem 7 also imposes upper bounds on the required number of measurements and the required computational time for characterising the state. Specifically, the experimental and computational costs for performing certified MPS tomography of quenched states scale no faster than polynomially in N, inverse polynomially in $\mathcal{I}$ and exponentially in the quench time t. Thus, certified MPS tomography is efficient in the size of the system and in the inverse infidelity tolerance for quenched states.

[1] Schindler, P. *et al.* A quantum information processor with trapped ions. *New J. Phys.* **15**, 123012 (2013).
[2] Jurcevic, P. *et al.* Quasiparticle engineering and entanglement propagation in a quantum many-body system. *Nature* **511**, 202–205 (2014).
[3] Jurcevic, P. *et al.* Spectroscopy of interacting quasiparticles in trapped ions. *Phys. Rev. Lett.* **115**, 100501 (2015).
[4] Sidje, R. B. A software package for computing matrix exponentials. *ACM Trans. Math. Softw.* **24**, 130–156 (1998).
[5] Ježek, M., Fiurášek, J. & Hradil, Z. Quantum inference of states and processes. *Phys. Rev. A* **68**, 012305 (2003).
[6] Cramer, M. *et al.* Efficient quantum state tomography. *Nat. Commun.* **1**, 149 (2010).
[7] Baumgratz, T., Nüßeler, A., Cramer, M. & Plenio, M. B. A scalable maximum likelihood method for quantum state tomography. *New J. Phys.* **15**, 125004 (2013).
[8] Holzäpfel, M., Baumgratz, T., Cramer, M. & Plenio, M. B. Scalable reconstruction of unitary processes and Hamiltonians. *Phys. Rev. A* **91**, 042129 (2015). 1411.6379.
[9] Schollwöck, U. The density-matrix renormalization group in the age of matrix product states. *Ann. Phys.* **326**, 96–192 (2011).
[10] Perez-Garcia, D., Verstraete, F., Wolf, M. M. & Cirac, J. I. Matrix product state representations. *Quantum Inf. Comput.* **7**, 401 (2007). quant-ph/0608197.
[11] Baumgratz, T. *Efficient system identification and characterization for quantum many-body systems*. Ph.D. thesis, Ulm University (2014).
[12] Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2007), 9 edn.
[13] Jones, E., Oliphant, T., Peterson, P. *et al.* SciPy: Open source scientific tools for Python (2001). URL `http://www.scipy.org/`.
[14] Suess, D. & Holzäpfel, M. mpnum: A matrix product representation library for python (2016). URL `https://github.com/dseuss/mpnum`.
[15] Jaynes, E. T. *Probability Theory: The Logic of Science* (Cambridge University Press, Cambridge, 2003).
[16] Bennett, C. H., Bernstein, H. J., Popescu, S. & Schumacher, B. Concentrating partial entanglement by local operations. *Phys. Rev. A* **53**, 2046–2052 (1996).
[17] Flammia, S. T. & Liu, Y.-K. Direct Fidelity Estimation from Few Pauli Measurements. *Phys. Rev. Lett.* **106**, 230501 (2011).
[18] da Silva, M. P., Landon-Cardinal, O. & Poulin, D. Practical Characterization of Quantum Devices without Tomography. *Phys. Rev. Lett.* **107**, 210404 (2011).
[19] Brandão, F. G. S. L. & Horodecki, M. Exponential decay of correlations implies area law. *Commun. Math. Phys.* **333**, 761–798 (2014).
[20] Osborne, T. J. Efficient approximation of the dynamics of one-dimensional quantum spin systems. *Phys. Rev. Lett.* **97**, 157202 (2006). quant-ph/0508031.
[21] One can also make a direct connection from the POVM description of the 6^k outcome probabilities to the ability to reconstruct the corresponding (local k-spin reduced) density matrix. See the paragraph following Eq. (36) on page 12.
[22] Whenever there is some freedom in choosing the $c_{si,jl'}$, we choose them with equal magnitude.
[23] A similar problem, where probabilities are a quadratic instead of a linear function of a pure density matrix ρ, has been discussed in the supplementary material of Ref. [18].
[24] The results in this section hold for nearest-neighbour Hamiltonians. Strictly speaking, the interaction in the lab is not of this kind but similar results are expected to hold because the lab interaction is finite ranged.
[25] One could also allow a product state on the $\mathcal{N}$ blocks instead of product states on N qubits.
[26] For the application of Lemma 2 in the next sentence, the empty blocks 0 and $N+1$ shall be dropped.
[27] One could also allow a product state on the $\mathcal{N}$ blocks instead of product states on N qubits.