

**Supplementary information**

---

**Benchmarking highly entangled states on a  
60-atom analogue quantum simulator**

---

In the format provided by the  
authors and unedited

# Supplementary Information: Benchmarking highly entangled states on a 60-atom analog quantum simulator

Adam L. Shaw,<sup>1,\*,\dagger</sup> Zhuo Chen,<sup>2,3,\*</sup> Joonhee Choi,<sup>1,4,\*</sup> Daniel K. Mark,<sup>2,\*</sup> Pascal Scholl,<sup>1</sup>  
Ran Finkelstein,<sup>1</sup> Andreas Elben,<sup>1</sup> Soonwon Choi,<sup>2,\dagger</sup> and Manuel Endres<sup>1,\dagger</sup>

<sup>1</sup>*California Institute of Technology, Pasadena, CA 91125, USA*

<sup>2</sup>*Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

<sup>3</sup>*The NSF AI Institute for Artificial Intelligence and Fundamental Interactions, Cambridge, MA 02139, USA*

<sup>4</sup>*Department of Electrical Engineering, Stanford University, Stanford, CA, USA*

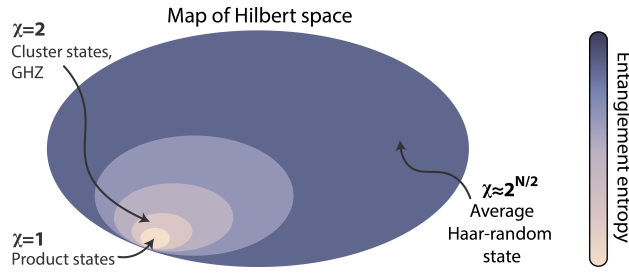
## Contents

A. Description of the Experiment	3
1. State preparation and readout	3
2. The Rydberg Hamiltonian	4
3. Maximum likelihood estimation of Hamiltonian parameters	4
4. Defining the experimental time unit	5
B. Fidelity estimation with a modified cross-entropy benchmark	6
C. Non-Markovian origins of non-exponential fidelity decay	9
1. Fidelity response to coherent Hamiltonian errors	9
2. Non-exponential decay for Rydberg quantum simulators	11
D. Fidelity extrapolation through Monte Carlo inference	13
1. Method description	13
2. Consistency checks	14
E. Scalability of fidelity estimation	18
1. Fidelity estimation with approximate algorithms for Markovian noise	18
2. Scalability of Monte Carlo inference for non-Markovian noise sources	18
3. Sample complexity	19
F. Extrapolating pure state entanglement growth	21
G. Bounds on the mixed state entanglement proxy	22
1. Negativity lower bounds and estimators	23
2. Robustness of our bounds to realistic experimental errors	25
3. Negativity of truncated MPS states	33
H. Mixed state entanglement estimation for quasi-adiabatic state preparation	36
I. Simulation methods	37
1. Description of the Caltech computing cluster	37
2. Exact simulation with a Schrieffer-Wolff transformation	37
3. Standard Trotterization-based TEBD-MPS	38
4. The Lightcone-MPS algorithm	39
5. Time-dependent variational principle	41
6. Noisy MPO-based simulation	41
7. Path integral formulation	42
8. Neural network based Hamiltonian simulation	43
J. Estimating the MPS fidelity	44
1. Approximating the MPS fidelity with truncation errors	44
2. Extrapolating the experimental $\chi^*$ value	45
3. Fitting and extrapolating the MPS fidelity	45
K. Error model simulations	47
1. Random unitary circuits	47
2. The Rydberg Hamiltonian	47
References	49

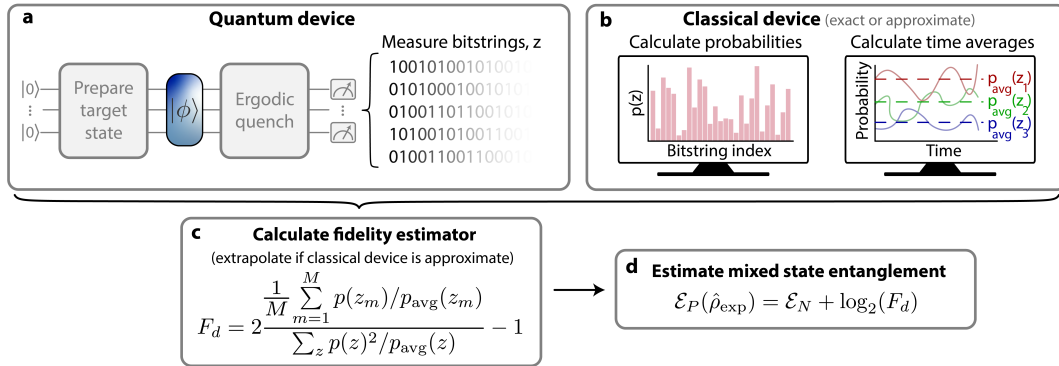
---

\* These authors contributed equally

\dagger ashaw@caltech.edu, soonwon@mit.edu, mendres@caltech.edu



**Supp. Fig. 1 | Map of the relevant regions of Hilbert space.** For any intermediate or large system size,  $N$ , an average Haar random state in the Hilbert space will have nearly maximal entanglement entropy (up to the Page correction<sup>1</sup>), and correspondingly is described by an MPS with approximately maximal bond dimension. By contrast, states such as GHZ state and cluster states – which are colloquially referred to as maximally-entangled – actually have very little extensive entanglement entropy, and can be written in an efficient MPS representation with only  $\chi = 2$ .



**Supp. Fig. 2 | Schematic for mixed state entanglement estimation of a target quantum state.** **a.** The quantum device prepares a target state<sup>2</sup> of interest,  $|\psi\rangle$ , before performing an ergodic quench, and measuring a set of resulting bitstrings. The experimental fidelity drops both during state preparation and the quench due to errors (not pictured). **b.** A classical computer then performs noise-free simulations of the entire dynamics (either exactly or approximately), calculates both the bitstring probabilities and the time-averaged bitstring probabilities,  $p_{\text{avg}}$ , and further estimates the second moment of the bitstring probability distribution. **c.** The fidelity is then estimated via a cross-entropy type quantity<sup>2-4</sup>,  $F_d$ . If the classical simulation were approximate,  $F_d$  is predicted either via Monte Carlo inference or direct extrapolation. **d.** The negativity of the target target state is estimated or assumed, and the experimental mixed state entanglement-proxy is calculated.

Experiment	Type	Date	N	Fidelity ( $\times 10^{-2}$ )	Mixed state entanglement	Reference
Sycamore	Digital	Oct 2019	53	0.22	16.7	Arute et al. <sup>[5]</sup>
Zuchongzhi 2.0	Digital	Jun 2021	56	0.066	17.0	Wu et al. <sup>[6]</sup>
Zuchongzhi 2.1	Digital	Sep 2021	60	0.037	18.1	Zhu et al. <sup>[7]</sup>
Sycamore	Digital	Apr 2023	70	0.17	25.3	Morvan et al. <sup>[8]</sup>
Quantinuum H2	Digital	May 2023	32	57	14.7	Moses et al. <sup>[9]</sup>
This work	Analog	Aug 2023	42 <sup>†</sup>	9.5	14.6	This work

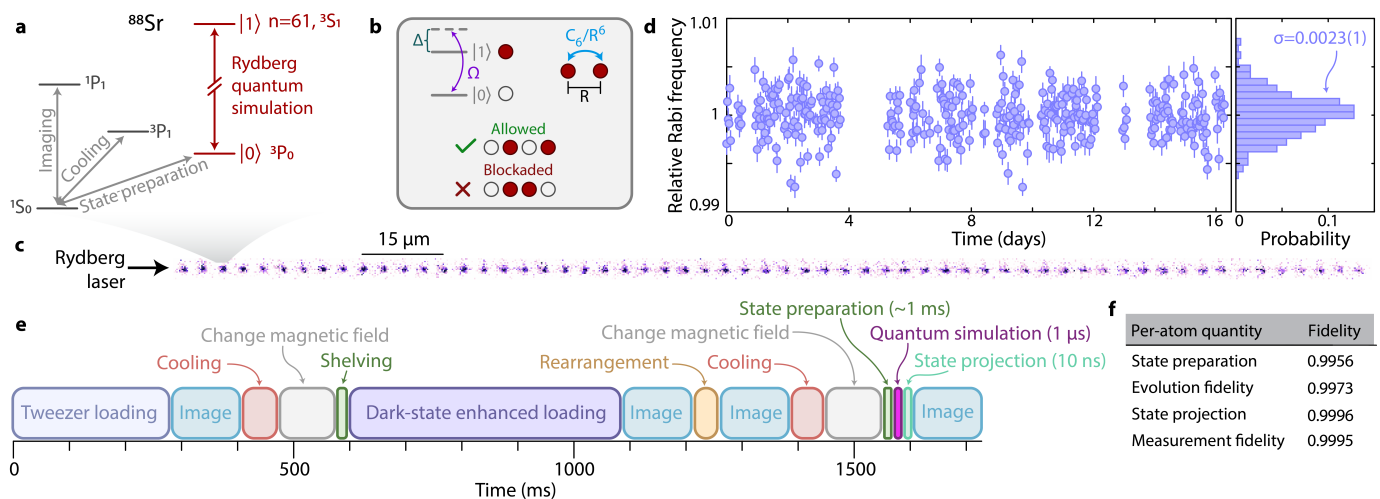
**Supp. Table I | Summary of mixed state entanglement values for various experiments.** For all literature experiments, we assume the target pure state is as entangled as a typical Haar random state, while for this work we extrapolate the entanglement directly (Supp. Fig. 16). <sup>†</sup> indicates an approximate effective system size due to the Rydberg blockade constraint, see text.

## A. Description of the Experiment

### A1. State preparation and readout

Our experiment is a Rydberg atom array quantum simulator<sup>2,10</sup> trapping individual strontium-88 atoms in optical tweezers<sup>11,12</sup>; see also Supp. Fig. 3. Up-to-date details of our apparatus may be found in previous works<sup>2,13</sup>. In brief, we use dark-state enhanced loading<sup>14</sup> to load a 68-tweezer array with  $\sim 61$  atoms, which we then rearrange<sup>15,16</sup> into defect free arrays of various sizes, with the array spacing calibrated with a laser-based ruler<sup>17</sup>. Atoms are initially in the  $5s^2 \ ^1S_0$  state, and are cooled on the narrow-line  $5s^2 \ ^1S_0 \leftrightarrow 5s5p \ ^3P_1$  transition close to their motional ground state. Atoms are prepared into the long-lived  $5s5p \ ^3P_0$  clock state with a preparation fidelity of 0.9956(1) per atom, which we then treat as a metastable ground state,  $|0\rangle$ . Atoms are then driven globally to the  $5s61s \ ^3S_1, m_J=0$  Rydberg state,  $|1\rangle$  while the traps are briefly blinked off.

Following Hamiltonian evolution, state projection is performed by autoionizing<sup>18</sup> the Rydberg atoms, leaving them dark to our fluorescent imaging, with a fidelity of  $\sim 0.9996$  per atom. Atoms in the clock state are pumped into the imaging cycle, from which we map atomic fluorescence to qubit state<sup>18,19</sup> with a detection fidelity  $\gtrsim 0.9995$  per atom<sup>13,19</sup>. For an initially defect-free array, this results in a series of bitstrings associated with the measured qubit states in the array; note that we discard any experimental bitstrings for which initial rearrangement failed. As we load many more atoms than are needed when benchmarking small system sizes, we simultaneously excite and detect multiple non-interacting subensembles when possible, and accrue several thousand bitstrings per system size and time (Supp. Fig. 15a, inset).



**Supp. Fig. 3 | Description of the experiment.** **a.** We use a Rydberg quantum simulator, based on trapping an array of single atoms of bosonic strontium-88. We treat the metastable  $5s5p \ ^3P_0$  state as an effective ground state, and consider dynamics in the Rydberg manifold. State projection is performed by autoionizing the Rydberg state<sup>18</sup> (not shown). **b.** The Rydberg Hamiltonian is Ising-like, with single particle terms characterized by a detuning,  $\Delta$ , and Rabi frequency,  $\Omega$ , between the two qubit levels. Rydberg atoms are strongly interacting, characterized by the  $C_6$  coefficient, and with a strength falling off approximately as  $1/R^6$ , where  $R$  is the interatom distance. When two atoms are close enough together, simultaneous excitation to the Rydberg state is strongly suppressed, causing the so-called Rydberg blockade effect. We operate in a regime in which the nearest neighbor interaction strength is  $\sim 13\times$  the Rabi frequency, which is strongly blockaded (see also Supp. Fig. 6d). **c.** We study one-dimensional chains of atoms (fluorescence image), with the Rydberg excitation laser aligned along the longitudinal array axis to minimize inhomogeneity. **d.** A classical control system<sup>2</sup> continuously interleaves data-taking with automated atom-based calibration of Hamiltonian parameters, resulting in highly stable operation over the course of multiple weeks. For example, the Rabi frequency is held constant with a relative standard deviation of only 0.0023(1). **e.** A single run of the experiment takes  $\sim 1.7$  s, almost entirely limited by array loading and imaging. The actual time required for quantum evolution is  $\sim 1 \mu\text{s}$ . We use dark-state enhanced loading<sup>14</sup> to reach the largest array sizes used in this work. **f.** Representative per-atom fidelities; the reported per-atom evolution fidelity is defined as in Fig. 5b of the main text.



## A2. The Rydberg Hamiltonian

The Hamiltonian of our system is well approximated by

$$\hat{H}/h = \Omega \sum_i \hat{S}_i^x - \Delta \sum_i \hat{n}_i + \frac{C_6}{a^6} \sum_{i>j} \frac{\hat{n}_i \hat{n}_j}{|i-j|^6} \quad (1)$$

describing a set of interacting two-level systems, labeled by site indices  $i$  and  $j$ , driven by a laser with Rabi frequency  $\Omega$  and detuning  $\Delta$ . The interaction strength is determined by the  $C_6$  coefficient and the lattice spacing  $a$ . Operators are  $\hat{S}_i^x = (|1\rangle_i \langle 0|_i + |0\rangle_i \langle 1|_i)/2$  and  $\hat{n}_i = |1\rangle_i \langle 1|_i$ , where  $|0\rangle_i$  and  $|1\rangle_i$  denote the electronic ground and Rydberg states at site  $i$ , respectively. Hamiltonian parameters are summarized in Supp. Fig. 4b, and are chosen to lead to high-temperature thermalization<sup>20,21</sup>, and chaotic behavior consistent with previous studies<sup>2</sup>.

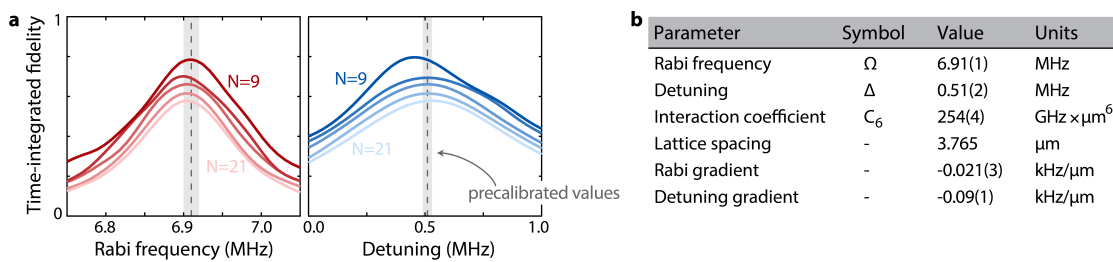
An important feature of our system arises from the Rydberg blockade<sup>10</sup> effect (Fig. 2b of the main text). Notably, the nearest neighbor interaction strength is  $13\times$  larger than the next largest energy scale (the Rabi frequency), greatly reducing the probability to have simultaneous excitation of neighboring atoms to Rydberg states. To first order, this reduces the Hilbert space dimension from  $2^N$  to  $\text{Fib}(N+2) \approx 1.62^N$ , where Fib is the Fibonacci function. This means that for a system size of  $N$  atoms, the effective Hilbert space size is only that of  $\sim 0.7N$  qubits, restricting the maximum entanglement built-up.

We account for this effect in our benchmarking protocol (discussed below), and, more concretely, in Fig. 4 of the main text we plot the experimental mixed state entanglement-proxy against the effective system size,  $\text{Fib}(N+2)$ .

## A3. Maximum likelihood estimation of Hamiltonian parameters

Hamiltonian parameters are precalibrated via *ab initio* measurements, and then refined using maximum likelihood estimation (MLE) from fidelity measurements at small system sizes ( $N \leq 21$ ), as described in Ref.<sup>2</sup>. MLE estimated parameters are insensitive to system size, so we believe they are accurate for the larger system sizes we test here. MLE results are consistent with pre-calibrated values, and tightly optimize the target Hamiltonian used in classical simulation (Supp. Fig. 4a). Note that the real-space positions of atoms are taken into account when performing MLE on parameter gradients. For instance, several small islands of atoms are benchmarked over the same spatial extent as the largest 60-atom array.

The main experimental dataset was taken continuously over the course of 17 days. Parameters such as Rabi frequency, detuning, beam alignment, state preparation, and others were automatically calibrated via our home-built control architecture<sup>2</sup>, resulting in high stability even over such a long data-taking period. The overall duty cycle of our experiment (the ratio of time spent on ‘science shots’ versus the total wall time) was roughly 36% over that period.



**Supp. Fig. 4 | Coherent Hamiltonian parameters.** **a.** We calibrate Hamiltonian parameters using a maximum-likelihood-estimation (MLE) technique based on varying parameters used during classical simulation, and finding the point of highest fidelity overlap with the experimental data; for further details, see Ref.<sup>2</sup>. Fidelities are time-integrated over all experimental measurements. MLE calibrated values are consistent for system sizes up to  $N = 21$ , as is expected as data taking for all system sizes (and times) is done in a randomized, interleaved fashion. Several smaller arrays are benchmarked simultaneously over the entire spatial extent of the largest 60-atom array to perform MLE on parameter gradients. **b.** Resultant MLE-calibrated Hamiltonian parameters (and gradients) used for all  $N$  in the classical simulation.

#### A4. Defining the experimental time unit

Throughout this manuscript, we define the time unit in terms of ‘cycles’, given by  $t_{\text{cycle}} = 1/\Omega$ , and thus concretely  $t_{\text{cycle}} \approx 145$  ns. This timescale was chosen as  $\Omega$  is the dominant energy-scale in the system (after the blockading nearest neighbor interaction energy). However, it is instructive to compare cycles versus a more natural timescale, namely the ‘unit-entanglement-time’,  $t_{\text{ebit}}$ , required to generate 1 ebit of entanglement in the early time linear growth regime (see Fig. 2c). For our system parameters, we numerically find this timescale is given by  $t_{\text{ebit}} \approx 0.83 \times t_{\text{cycle}} \approx 121$  ns. This definition is convenient, because it then allows us to directly compare our analog evolution *times* against the equivalent *depth* for digital circuit evolution.

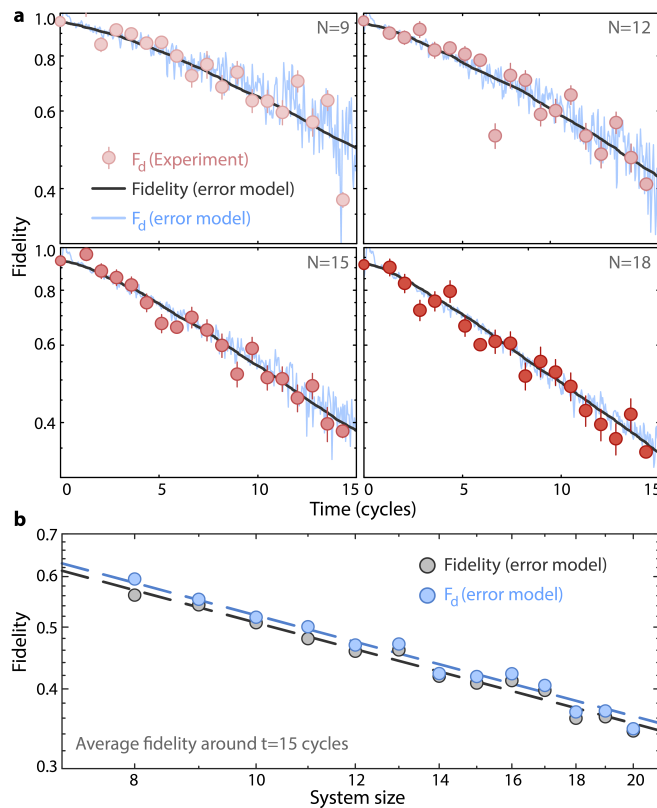
We numerically fit the early time linear entanglement growth for one-dimensional random unitary circuit (RUC) evolution, and compare against our experiment. For the gate-set of Ref.<sup>5</sup>, we find that  $t_{\text{ebit}} \approx 1.1$  layers, meaning 1 ebit of entanglement is built up in 1.1 layers. While this value is gate-set dependent, the exact numerical values are not of too much importance, but serve to demonstrate that  $t_{\text{cycle}}$  used throughout this manuscript can be treated roughly as 1 circuit depth. For further details on comparing the evolution of analog and digital devices, see Ref.<sup>2</sup>.

## B. Fidelity estimation with a modified cross-entropy benchmark

We employ a recently developed fidelity estimator<sup>3</sup>,  $F_d$ , to evaluate the many-body fidelity of ergodic quench dynamics<sup>22</sup> producing large entanglement. Importantly,  $F_d$  can be efficiently sampled using only a small number of measurements as

$$F_d = 2 \frac{\frac{1}{M} \sum_{m=1}^M p(z_m)/p_{\text{avg}}(z_m)}{\sum_z p(z)^2/p_{\text{avg}}(z)} - 1, \quad (2)$$

where  $M$  is the number of measurements,  $z_m$  is the experimentally measured bitstring at the  $m^{\text{th}}$  repetition,  $p(z)$  is the theoretically calculated probability of measuring  $z$  by a classical algorithm, and  $p_{\text{avg}}(z)$  is the time-averaged probability of measuring  $z$  over a certain time window. The denominator can be efficiently calculated by importance-sampling from the MPS<sup>23</sup> as  $\mathbb{E}_{z \sim p(z)} p(z)/p_{\text{avg}}(z)$ . For small system sizes where we accrue bitstrings from multiple non-interacting subarrays, we calculate  $F_d$  for each individual ‘island’ and then average their values, weighting by the number of measured bitstrings from each. This procedure is summarized schematically in Supp. Fig. 2.



**Supp. Fig. 5 | Accuracy of the fidelity estimator.** **a.** Comparing fidelities estimated from experiment against both the true fidelity and the estimated fidelity for our error model, showing good agreement both between experiment and theory, and between the estimator  $F_d$  and the true fidelity.  $F_d$  shows systematic typicality fluctuations around the true fidelity, but these decrease in amplitude with increasing system size, see Supp. Fig. 15b. **b.**  $F_d$  does exhibit a systematic multiplicative offset from the true fidelity in the time regime we study due to a time-delay needed for errors to scramble and become visible<sup>2,3</sup>. We find that  $F_d \approx 1.02F$ , regardless of system size around  $t = 15$  cycles.

We compute  $p_{\text{avg}}(z)$  via time-averaging the calculated bitstring probabilities. Specifically, for each time point  $t_i$  of the experiment, we estimate  $p_{\text{avg}}(z)$  as a discrete average of the probability distributions within a window of  $t_i \pm 1.4 \mu\text{s}$ , with an step size of approximately 28 ns. This means around 100 points are averaged, which limits statistical sampling fluctuations. Points included in the averaging are weighted to emphasize data with similar values of classical fidelity, namely using a weight factor of  $\min\{\frac{F_{\text{svd}}(t_i)}{F_{\text{svd}}(t_j)}, \frac{F_{\text{svd}}(t_j)}{F_{\text{svd}}(t_i)}\}$ . Here  $F_{\text{svd}}$  is the product of MPS truncation errors<sup>24,25</sup>, which

we use to approximate the classical accuracy, and defined as

$$F_{\text{svd}} = \prod_i \sum_{\alpha=1}^{\chi} s_{i,\alpha}^2, \quad (3)$$

where  $i$  runs over all MPS steps involving Schmidt value truncations, and  $s_{i,\alpha}$  are the Schmidt values at truncation step  $i$ . Here, we assume the wave function is normalized where  $\sum_{\alpha=1}^{\infty} s_{i,\alpha}^2 = 1$ . While this estimation is only approximate, it can be extremely accurate when successive truncations are independent<sup>25</sup>. We describe this approximation in more detail in Section J and Supp. Fig. 29. We weight in this way under the hypothesis that averaging classical simulations of very different accuracies (as typified by  $F_{\text{svd}}$ ) will lead to a poorer estimation of  $p_{\text{avg}}(z)$ .

To account for the Rydberg blockade mechanism, we have the option to slightly modify<sup>2</sup> the  $F_d$  formula. Indexing the different blockade sectors by the number of blockade violations in that sector,  $s$ , we can write an approximator for  $F_d$  to get an estimator for the blockade sector,  $F_{d,s=0}$ , only:

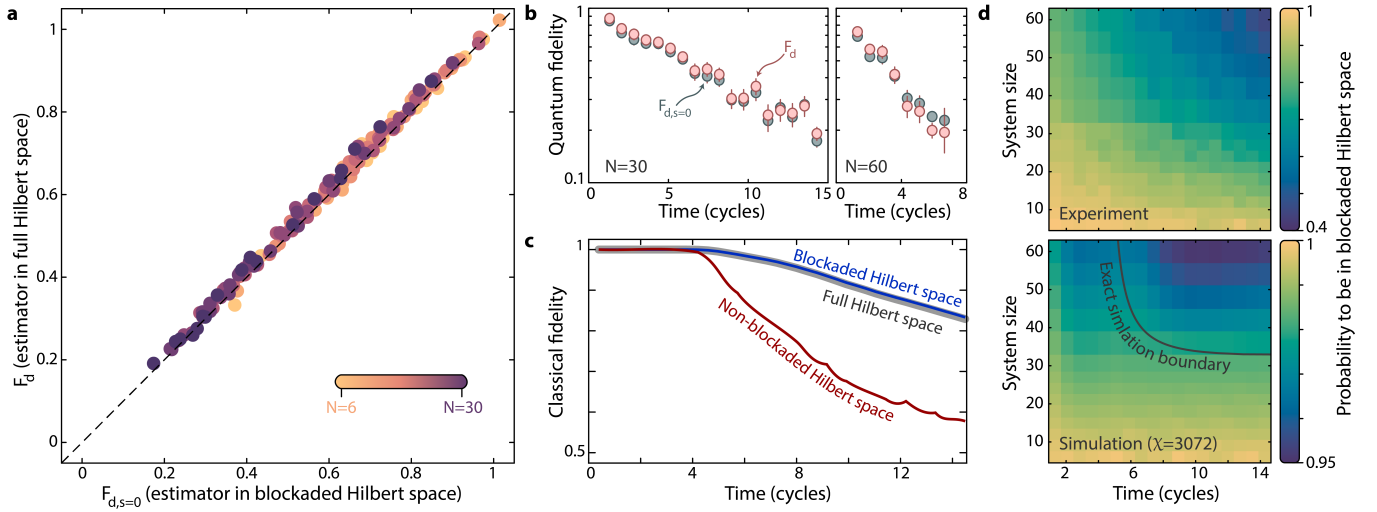
$$F_{d,s=0} = B_{\text{thy}} B_{\text{exp}} \left( 2 \frac{\frac{1}{M'} \sum_{m=1}^{M'} p'(z'_m) / p'_{\text{avg}}(z'_m)}{\sum_z p'(z)^2 / p'_{\text{avg}}(z)} - 1 \right). \quad (4)$$

Here  $B_{\text{exp}}$  ( $B_{\text{thy}}$ ) is the total probability for an experimental (simulation) bitstring to be in the blockaded Hilbert space,  $s = 0$ , which further redefines the normalized probabilities,  $p'(z) = p(z)/B_{\text{thy}}$ . Note here  $M'$  is the number of bitstrings,  $z'_m$ , measured in the blockaded Hilbert space.

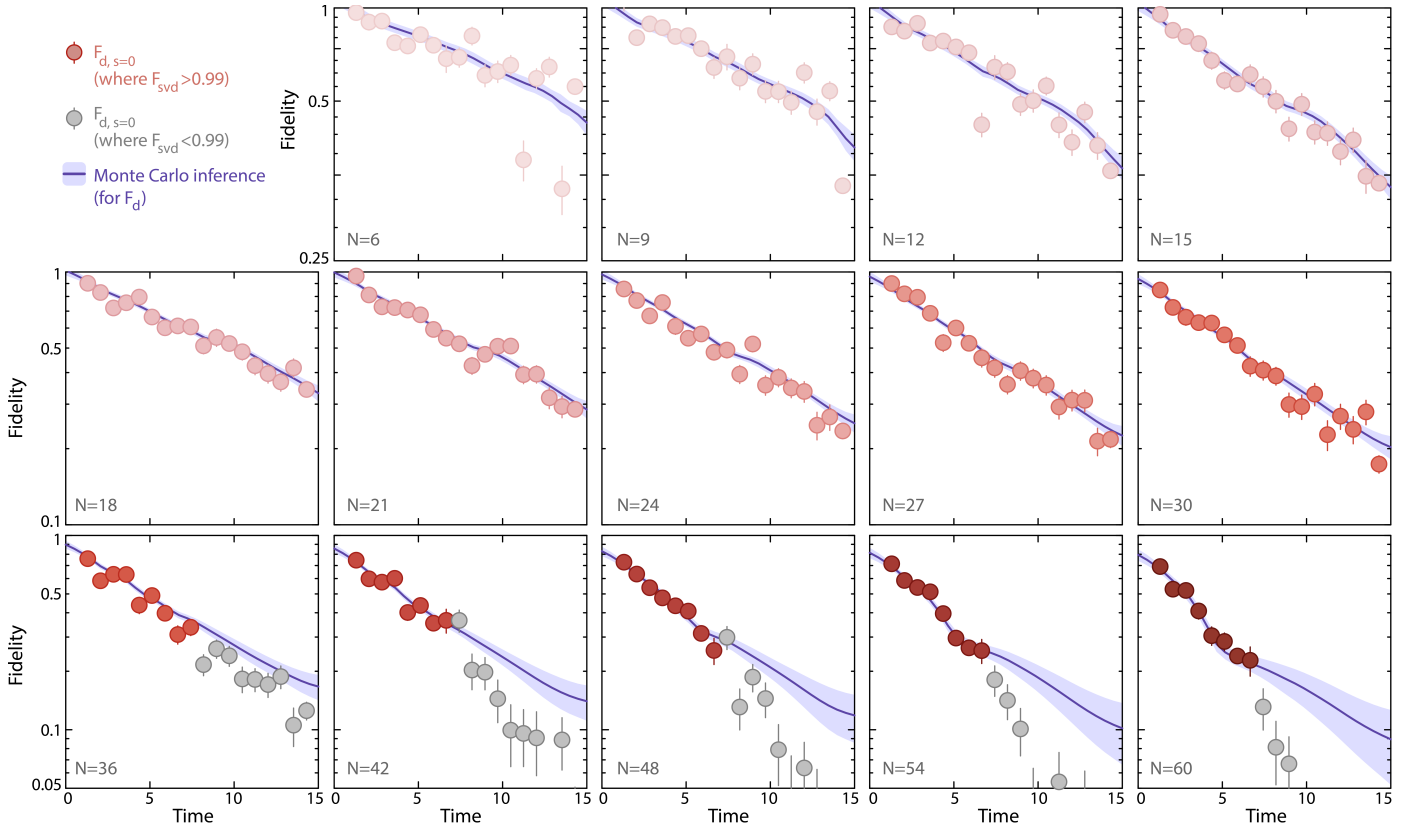
Importantly, we find  $F_{d,s=0}$  is more robust against the failure of the classical simulation algorithm, as compared to  $F_{d,s>0}$ . This is because MPS truncations more strongly affect the  $s > 0$  sectors in ways which are not necessarily visible only by looking at the MPS truncation fidelity (Supp. Fig. 6bc).

For this reason, when estimating the quantum fidelity, we employ a two-pronged approach. Where  $F_{\text{svd}} > 0.99$  we approximate  $F_d$  as  $F_{d,s=0}$ , which from measurements at small system sizes we find is always conservative (Supp. Fig. 6a). Utilizing  $F_{d,s=0}$  allows us to more confidently directly measure the fidelity out to intermediate times at the largest system sizes, and measure the according effective decay rate (Fig. 3c of the main text). However, where  $F_{\text{svd}} < 0.99$ , we estimate  $F_d$  from our Monte Carlo inference procedure (described below), which is trained directly on the values of  $F_d$ .

To be concrete, in the main text, all data shown as markers in Figs. 2 and 3 is  $F_{d,s=0}$  for  $\chi = 3072$ . We summarize all of the data, both  $F_{d,s=0}$  and the Monte Carlo fidelities, in Supp. Fig. 7.



**Supp. Fig. 6 | Benchmarking in the blocked Hilbert space.** **a.**  $F_d$  versus  $F_{d,s=0}$  (the fidelity benchmarked in only the blocked Hilbert space, see text) for system sizes up to  $N=30$ , consistent with correlation with unity slope, but with a slight  $\sim 0.01$  offset. **b.** For exactly simulatable systems,  $F_d$  typically always exceeds  $F_{d,s=0}$  (left), but for larger systems (right),  $F_d$  systematically begins underestimating  $F_{d,s=0}$ , even before the nominal exact simulation time. **c.** We understand this behavior as arising from the MPS preferentially truncating higher blockade subspaces, leading to the non-blockaded classical fidelity dropping below unity before it is visible in the full Hilbert space fidelity. This implies  $F_{d,s=0}$  stays a more accurate estimator slightly longer than  $F_d$ . **d.** Probability to be in the blocked subspace for both experiment and simulation. Note the different color-scale axes for the simulation and experiment, as many experimental errors can lead to apparently non-blockaded measurements.



**Supp. Fig. 7 | Fidelity benchmarking for all system sizes and times.** Fidelity (quantified by  $F_{d,s=0}$  with  $\chi = 3072$ ) as well as Monte Carlo inference for all system sizes. As in the main text, grayscale markers are those for which the classical fidelity is less than 0.99.

### C. Non-Markovian origins of non-exponential fidelity decay

In Fig. 3 of the main text, we experimentally observe that the many-body fidelity appears to decrease exponentially at intermediate times, before bending sub-exponentially at late times, a behavior which seems to grow stronger with increasing system size (also visible in Supp. Fig. 7). We believe this behavior originates from globally-correlated, non-Markovian Hamiltonian parameter fluctuations, which we note are often found in analog quantum simulators. In particular, a dominant noise source in our experiment is shot-to-shot variation of the Rabi frequency, for instance arising from fluctuations of the driving laser intensity, which we measure to be around  $\sim 0.5\%$ , corresponding to a shot-to-shot variation of  $\sigma \sim 40$  kHz (Supp. Fig. 8a). Here, we present analytical and numerical support that such an error source can induce the observed fidelity decay behavior.

#### C1. Fidelity response to coherent Hamiltonian errors

We discover a surprising dependence of the fidelity between states subject to coherent Hamiltonian errors.

**Theorem 1** (Gaussian fidelity response to coherent errors). *Let  $|\Psi(t, \theta)\rangle \equiv \exp(-i(\hat{H}_0 + \theta\hat{V})t)|\Psi_0\rangle$  denote the initial state  $|\Psi_0\rangle$  quench-evolved by a translationally-invariant local Hamiltonian  $\hat{H}_0 + \theta\hat{V}$  for a time  $t$ , where the perturbation  $\hat{V} = \sum_x \hat{V}^x$  is a sum of local terms and  $\theta$  quantifies the perturbation strength. We assume the Eigenstate Thermalization Hypothesis (ETH), i.e. that the expectation values  $\langle E|\hat{V}^x|E\rangle$  of each term  $\hat{V}^x$  with respect to the eigenstates  $|E\rangle$  of  $\hat{H}_0$  are independent random samples from a fixed distribution. Then in the limit of large system size  $N$  and long time  $t$ , the fidelity  $F$  between two such states  $|\Psi(t, \theta_1)\rangle$  and  $|\Psi(t, \theta_2)\rangle$  is a Gaussian function of  $(\theta_1 - \theta_2)t$ :*

$$F(\theta_1, \theta_2) \equiv |\langle \Psi(t, \theta_1) | \Psi(t, \theta_2) \rangle|^2 = \exp(-N\lambda(\theta_1 - \theta_2)^2 t^2 / 2), \quad (5)$$

for some constant  $\lambda$  defined in Eq. (14).

This result is illustrated in Supp. Fig. 8b and will be instrumental to showing our observed non-exponential fidelity decay (Corollary 1). This fidelity response is related to the Loschmidt echo, and our result is consistent with existing numerical observations in e.g. Refs.<sup>26–28</sup>.

*Proof.* Without loss of generality, it suffices to let  $\theta_1 = 0$  and  $\theta_2 = \theta$ . We compute the Taylor series of  $F(\theta) \equiv F(0, \theta)$  and show that for large  $N$ , it agrees to all orders with the Gaussian series

$$\exp\left(-\frac{\theta^2 t^2}{2\sigma^2}\right) = 1 - \frac{1}{2}\left(\frac{\theta t}{\sigma}\right)^2 + \frac{1}{8}\left(\frac{\theta t}{\sigma}\right)^4 + \dots \quad (6)$$

It is instructive to compute the first non-trivial term: the second derivative

$$\partial_\theta^2 F(\theta)|_{\theta=0} = \left[ (\partial_\theta^2 \langle \Psi(t) | \Psi(t, \theta) \rangle) \langle \Psi(t, \theta) | \Psi(t) \rangle \right. \quad (7)$$

$$+ 2(\partial_\theta \langle \Psi(t) | \Psi(t, \theta) \rangle) (\partial_\theta \langle \Psi(t, \theta) | \Psi(t) \rangle) \quad (8)$$

$$\left. + \langle \Psi(t) | \Psi(t, \theta) \rangle (\partial_\theta^2 \langle \Psi(t, \theta) | \Psi(t) \rangle) \right] \Big|_{\theta=0}$$

In order to evaluate these derivatives, we use the relation<sup>29</sup>:

$$\begin{aligned} \partial_\theta \exp(i(\hat{H}_0 + \theta\hat{V})t)|_{\theta=0} &= \int_0^t d\tau \exp(i\hat{H}_0\tau) (i\hat{V}) \exp(i\hat{H}_0(t-\tau)) \quad (9) \\ &= it \sum_E V_{EE}|E\rangle \langle E| e^{iEt} + i \sum_{E \neq E'} V_{EE'}|E\rangle \langle E'| \frac{e^{iE't} - e^{iEt}}{i(E' - E)}, \end{aligned}$$

where  $|E\rangle$  are the eigenstates of  $\hat{H}_0$  with eigenvalues  $E$ , and we have defined  $V_{EE'} \equiv \langle E' | \hat{V} | E \rangle$ . For our purposes, Eq. (9) is valid even when  $\hat{H}_0$  has degeneracies: the initial state  $|\Psi_0\rangle$  projects any degenerate subspace onto a single eigenstate  $|E\rangle$ .

The first term in Eq. (9) grows linearly with time  $t$ , while the summands in the second term oscillate and are sub-dominant. We subsequently neglect such terms. We obtain a similar expression for the second derivative:

$$\partial_\theta^2 \exp(i(\hat{H}_0 + \theta \hat{V})t) \approx -t^2 \sum_E V_{EE}^2 |E\rangle \langle E| e^{iEt} . \quad (10)$$

This gives:

$$\begin{aligned} \partial_\theta^2 F(\theta)|_{\theta=0} &\approx -2t^2 \left[ \sum_E |c_E|^2 V_{EE}^2 - \left( \sum_E |c_E|^2 V_{EE} \right)^2 \right] \\ &\equiv -2t^2 (\langle \hat{V}^2 \rangle - \langle \hat{V} \rangle^2) \\ &\equiv -2t^2 \kappa_2(\hat{V}) , \end{aligned} \quad (11)$$

where  $|c_E| \equiv |\langle E | \Psi_0 \rangle|^2$  are the populations of  $|\Psi_0\rangle$  in the energy eigenbasis. We interpret the quantities  $\sum_E |c_E|^2 V_{EE}$  and  $\sum_E |c_E|^2 V_{EE}^2$  as averages of the vector  $V_{EE}$  and  $V_{EE}^2$  with respect to the distribution  $|c_E|^2$ : we denote them  $\langle \hat{V} \rangle$  and  $\langle \hat{V}^2 \rangle$  respectively; the expression  $\langle \hat{V}^2 \rangle - \langle \hat{V} \rangle^2$  is the variance, which we denote  $\kappa_2$  for reasons made clear below.

Within the approximation made in Eq. (9),  $\partial_\theta^n F(\theta)|_{\theta=0} = 0$  for all odd  $n$ . The next term can be expressed in terms of the moments  $\langle \hat{V}^k \rangle$  as

$$\begin{aligned} \partial_\theta^4 F(\theta)|_{\theta=0} &\approx 2t^4 (\langle \hat{V}^4 \rangle - 4\langle \hat{V}^3 \rangle \langle \hat{V} \rangle + 3\langle \hat{V}^2 \rangle^2) \\ &= 2t^4 [\kappa_4(\hat{V}) + 6\kappa_2(\hat{V})^2] , \end{aligned} \quad (12)$$

where  $\kappa_4(\hat{V}) \equiv \langle (\hat{V} - \langle \hat{V} \rangle)^4 \rangle - 3\langle (\hat{V} - \langle \hat{V} \rangle)^2 \rangle^2$  is the *fourth cumulant* of  $\hat{V}$ . Crucially,  $\hat{V} = \sum_x \hat{V}^x$  is the sum of identical local terms: the Eigenstate Thermalization Hypothesis (ETH) predicts that their eigenstate overlaps  $V_{EE}^x$  fluctuate about a smooth function of energy<sup>30</sup>. Furthermore, for sufficiently high-temperature states, correlations  $\langle \hat{V}^x \hat{V}^y \rangle$  are short-ranged, and we can treat each  $\hat{V}^x$  as an effectively independent random variable.

We utilize a further property of cumulants: they are additive, obeying  $\kappa_n(\sum_x \hat{V}^x) = \sum_x \kappa_n(\hat{V}^x)$ . Therefore, all cumulants scale linearly with  $N$ . When  $N$  is large,  $\kappa_4(\hat{V}) \ll \kappa_2(\hat{V})^2$  and we can neglect the  $\kappa_4$  term in Eq. (12). This gives the expression

$$F(\theta) \approx 1 - \frac{1}{2} 2\kappa_2(\hat{V})(\theta t)^2 + \frac{1}{8} (2\kappa_2(\hat{V}))^2 (\theta t)^4 + \dots , \quad (13)$$

which agrees with Eq. (6) when we identify

$$\sigma^{-1} = 2\kappa_2(\hat{V}) = N\kappa_2(\hat{V}^x) \equiv N\lambda \quad (14)$$

In fact, all higher cumulants are subleading to the  $\kappa_2$  contribution. As long as they can be neglected, the series Eq. (6) and Eq. (13) agree *to all orders*. For example, the  $(\theta t)^6$  term of (13) is  $-2(\kappa_6(\hat{V}) + 30\kappa_4(\hat{V})\kappa_2(\hat{V}) + 60\kappa_2(\hat{V})^3)/6! \approx -(2\kappa_2(\hat{V}))^3/48$ , in agreement with the sixth-order term of Eq. (6).

We prove this equivalence to all orders using properties of *Bell polynomials*,  $B_k$  (where  $k$  is a polynomial index)<sup>31</sup>. The  $n$ -th derivative of the fidelity can be expressed as

$$\partial_\theta^n F|_{\theta=0} = (-1)^{n/2} (\theta t)^n \sum_{k=0}^n (-1)^k \binom{n}{k} \langle \hat{V}^{n-k} \rangle \langle \hat{V}^k \rangle . \quad (15)$$

We can rewrite this in terms of the cumulants  $\kappa_1, \dots, \kappa_k$  with the cumulant-moment relation<sup>32</sup>

$$\langle \hat{V}^k \rangle = B_k(\kappa_1, \kappa_2, \dots, \kappa_k) . \quad (16)$$

While the full expression is challenging, we are only interested in the coefficient of the  $\kappa_2(\hat{V})^{n/2}$  term in Eq. (15). We first note that  $\langle \hat{V}^k \rangle$  contains a  $\kappa_2(\hat{V})^{k/2}$  term only for even

$k$ . To determine its coefficient, we use the recursion relation<sup>31</sup>  $\partial_{\kappa_2} B_k = \binom{k}{2} B_{k-2}$  to conclude that  $\langle \hat{V}^k \rangle = \dots + (k!/k!!)\kappa_2(\hat{V})^{k/2} + \dots$ , where  $k!! = k(k-2)\dots 2$  is the *double factorial* of  $k$ . Returning to Eq. (15), we conclude that  $\partial_{\theta}^n F|_{\theta=0} = \dots + a_n \kappa_2 (\hat{V})^{n/2} (\theta t)^n + \dots$ , where

$$\begin{aligned} a_n &= (-1)^{n/2} \sum_{k \text{ even}} \binom{n}{k} \frac{k!(n-k)!}{k!!(n-k)!!} \kappa_2^{n/2} \\ &= (-2)^{n/2} \frac{n!}{n!!} \kappa_2^{n/2}, \end{aligned} \quad (17)$$

which is equal to the  $n$ -th derivative  $\partial_{\theta}^n \exp(-\theta^2/(2\sigma^2))|_{\theta=0} = (-1)^{n/2} \frac{n!}{n!!} \sigma^{-n}$ , thus concluding our derivation of the Gaussian fidelity response [Eq. (5)].  $\square$

## C2. Non-exponential decay for Rydberg quantum simulators

The non-exponential fidelity decay immediately follows from Theorem 1.

**Corollary 1.** *When the parameter value  $\theta$  follows a Gaussian probability distribution  $P(\theta)$  with mean  $\theta'$  and standard deviation  $\sigma$*

$$P(\theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\theta - \theta')^2}{2\sigma^2}\right), \quad (18)$$

*Theorem 1 and a straightforward integration give a universal form for the fidelity of the target state  $|\Psi(t, \theta_0)\rangle$  to the mixed state  $\rho(t) \equiv \int d\theta P(\theta) |\Psi(t, \theta)\rangle \langle \Psi(t, \theta)|$ .*

$$F = \int d\theta P(\theta) F(\theta_0, \theta) = \frac{1}{\sqrt{1+\Lambda}} \exp\left(-\frac{\delta^2}{2} \frac{\Lambda}{1+\Lambda}\right), \quad (19)$$

where  $\Lambda \equiv Nt^2\lambda\sigma^2$  and  $\delta = (\theta' - \theta_0)/\sigma$  is the overall relative parameter miscalibration.

This functional form leads to markedly different behavior as compared to simple exponential decay, as it becomes essentially like a power-law at late times (Supp. Fig. 8d). In simulation, we benchmark an  $N = 16$  atom system experiencing only global intensity fluctuations – but otherwise no noise – and see an excellent agreement between the simulated fidelity, and a fit to Eq. (19) with free  $\lambda$ . Notably, we see Eq. (19) exhibits three distinct regimes: super-exponential behavior at very early time, effectively simple-exponential behavior at intermediate times, and finally sub-exponential behavior at late times. Notably, by studying multiple system sizes, we find the divergence between Eq. (19) and an effective exponential decay model is mainly a function of fidelity, not time, and becomes visible when the fidelity reaches  $F \sim 0.2$ , which in our dataset is only the case for the highest  $N$  (Fig. 3 of the main text).

As a quantitative comparison, we define an effective model of the fidelity as

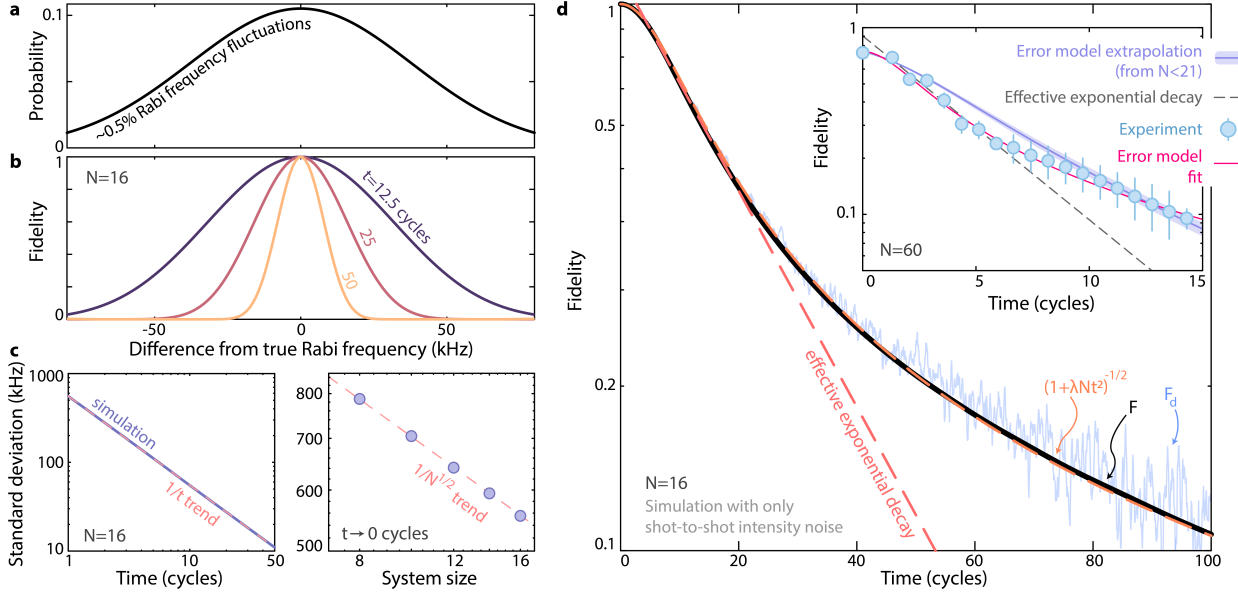
$$F \approx F_0 \exp(-\gamma t) \frac{1}{\sqrt{1+\lambda't^2}} \exp\left(-\frac{\delta^2}{2} \frac{\lambda't^2}{1+\lambda't^2}\right), \quad (20)$$

that is, Eq. (19) multiplied by an exponential decay factor to account for the other noise sources, where we have redefined the free parameter  $\lambda' \equiv \lambda N \sigma^2$ . We perform error model simulations from  $7 < N < 21$  for two cases: 1) for only global, shot-to-shot Rabi and detuning noise, 2) for all other noise sources. Then, assuming multiplicativity of the error sources (Supp. Fig. 32a), from case (1) we fit the expected behavior of  $\lambda'$ , and from case (2) we fit the expected behavior of  $\gamma$ . We extrapolate the observed parameters all the way to  $N = 60$ , and compare against the experiment (inset of Supp. Fig. 8d). Even assuming perfect calibration ( $\delta = 0$ ) we see general agreement between the fidelity decay curve of the error model prediction and the experiment.

However, the relative proportion of exponential decay versus non-exponential decay is difficult to calibrate versus experimental data at small system sizes, as this requires evolving out to very late times when other error sources (such as finite recapture probability of atoms in their traps) become noticeable. In addition, we expect that Hamiltonian parameter calibration is likely imperfect, at the very least up to our uncertainty in the Hamiltonian parameters. As a general sanity check then, we fit with  $\gamma$ ,  $\lambda'$ , and  $\delta$  free. Fitting  $N = 60$  (with  $F_0$  fixed by the separately calibrated single-atom preparation fidelity), we find  $\delta \approx 0.2$ ,  $\gamma \approx 1/4 \times \gamma_0$  and  $\lambda' \approx 4 \times \lambda'_0$ , where  $\gamma_0$  and  $\lambda'_0$  are the extrapolated predictions from error model. We note  $\delta \approx 0.2$  corresponds to a potential error of  $\approx 10$  kHz in the calibrated Rabi frequency, which is within our measured error bar (Supp. Fig. 4b).



We also note that other error sources, in particular atomic position fluctuations, also likely have some contribution to the non-exponential decay behavior, which may improve the agreement above. However, the analytical expression for their effect becomes quite involved, and so we disregard them here for simplicity, and leave a more in-depth analytic and numerical study of the various error sources to future work.



**Supp. Fig. 8 | Origins of the non-exponential fidelity decay.** **a.** A major noise source for our platform is Rabi frequency fluctuations (for instance arising from variation in the driving laser intensity), which have a shot-to-shot standard deviation of  $\sim 0.5\%$  ( $\sim 40$  kHz). **b.** For global, correlated intensity noise, the fidelity is Gaussian with respect to error in the Rabi frequency. **c.** The standard deviation of this Gaussian dependence decreases as  $1/t\sqrt{N}$ . **d.** This behavior leads to an analytic solution for the fidelity decay of  $F \sim (1 + \lambda N t^2)^{-1/2}$ , for free parameter  $\lambda$ . This decay is characterized by super-exponential curvature at early times, and sub-exponential curvature at late times. In the inset, we use our error model with all error sources enabled, and extrapolate to the expected fidelity behavior at  $N = 60$ , seeing generally good agreement with the experiment, which is improved by directly fitting and allowing unity-level changes in the extrapolated error model parameters.

## D. Fidelity extrapolation through Monte Carlo inference

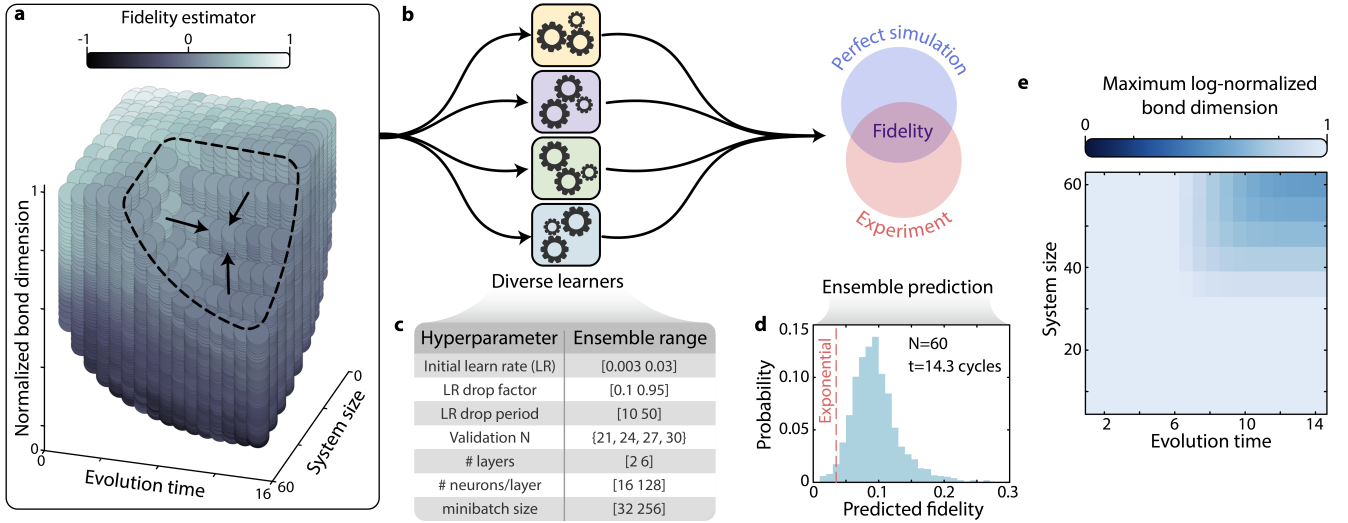
To extrapolate fidelities into the high-entanglement (i.e. late-time, large system-size) regime, we use a Monte Carlo inference approach. More specifically, we train an ensemble<sup>33</sup> of fully-connected regression neural networks, implemented in MATLAB, with three inputs (system size, evolution time, and bond dimension), and one output (estimated fidelity). The neural networks are instantiated with randomized hyperparameters (including the initial seed for the weights and biases). We emphasize that this approach is not trying to learn directly from experimental observations<sup>34–37</sup>, and instead we are simply extrapolating a smoothly varying 3-to-1 function. Fundamentally, the goal of this approach is to utilize all aspects of the data simultaneously, rather than extrapolating along only a single axis at a time, for instance either via time-extrapolation or system-size scaling alone.

### D1. Method description

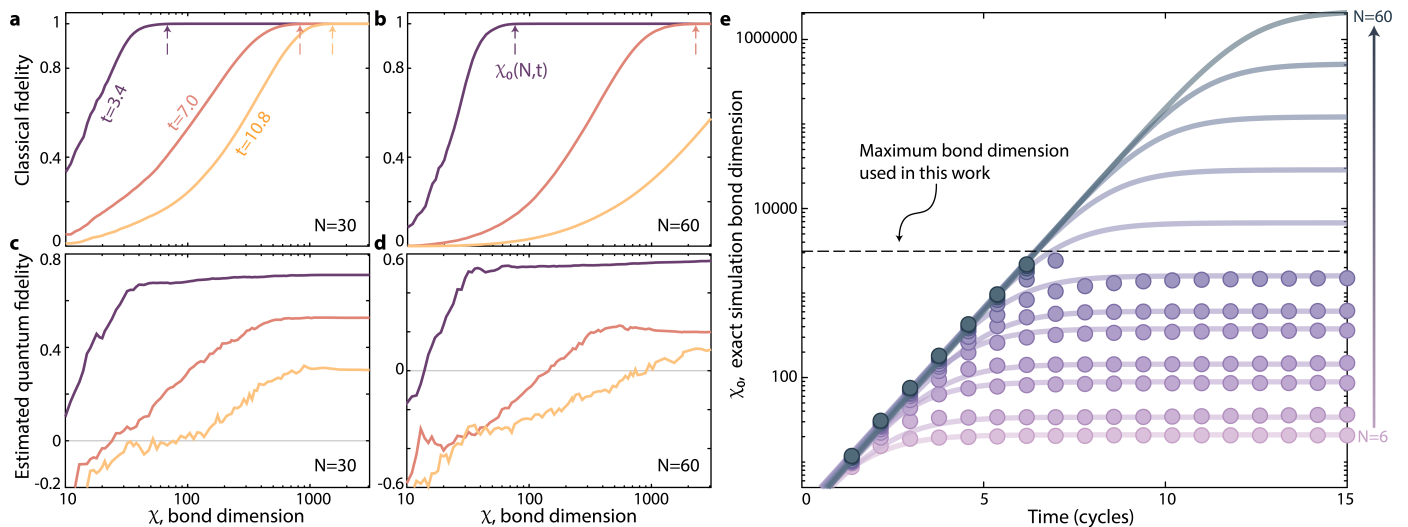
For training data, we use the fidelities from all combinations of the 14 system sizes, 18 evolution times, and 91 simulated bond dimensions (from  $\chi = 2$  to 3072), resulting in a total dataset of  $\sim 23000$  points (Supp. Fig. 9a). The network is then trained to learn the functional dependence of  $F(N, \chi, t)$ .

To improve generalization and avoid fine-tuning, we employ a Monte Carlo approach through ensemble learning<sup>33</sup>, where the training is repeated 1500 times. Each iteration is instantiated with a randomized seed, and every three iterations we randomly select from a range of hyperparameters, the domain of which are specified in Supp. Fig. 9c. The fidelity for a given set of inputs is then taken to be the unweighted ensemble average, while the error bar is taken as the ensemble standard error of the mean added in quadrature with the sampling error at a given time.

To preprocess the data, we log-normalize the bond dimensions as  $\tilde{\chi} = \log(\chi)/\log(\chi_0)$ , where  $\chi_0$  is the size- and time-dependent saturation bond dimension required for the simulation to be exact (Supp. Fig. 10e). This makes the learned function have a more consistent domain for different system sizes and evolution times. We also test if we train using simply  $\log(\chi)$ . We stress that in doing so, the largest trained  $\log(\chi) \approx 8$ , but for instance for  $N = 60$  we predict the fidelity at  $\log(\chi) \approx 14.5$ , with no training data in the interim regime. We find consistent mean predictions between the two methods, but the standard deviation of Monte Carlo instances is reduced by training on  $\tilde{\chi}$ .



**Supp. Fig. 9 | Monte Carlo fidelity inference.** **a.** All the data in this study, mapping the three inputs (system size,  $N$ , evolution time,  $t$ , and log-normalized bond dimension,  $\tilde{\chi} = \log(\chi)/\log(\chi_0)$ ) to an estimated fidelity. Here  $\chi_0$  is the time- and size-dependent saturation bond dimension (Supp. Fig. 10e). **b.** To improve generalization into the unknown regime, we use a Monte Carlo inference approach. We train an ensemble of 1500 fully-connected neural networks, each instantiated and trained with random choices of parameters within reasonably specified ranges (defined in **c**); the Monte Carlo predicted fidelity is then taken as the ensemble average for a given input. **d.** Monte Carlo prediction for the fidelity of the  $N = 60$  system at the entanglement saturation time. Over  $\sim 98\%$  of the individual Monte Carlo instance predictions exceed the expectation from the early time exponential decay (see Fig. 3 of the main text), resulting in an average of 0.095(11), where the error bar represents the standard error on the mean, added in quadrature with the intrinsic sampling error of the underlying data. **e.** Maximum log-normalized bond dimension available for training as a function of system size and time, for  $\chi = 3072$ .



**Supp. Fig. 10 | Determining the log-normalized bond dimension.** **a, b.** As a function of bond dimension, the classical simulation fidelity rises before saturating to 1; we define  $\chi_0$  (the saturation bond dimension), as the minimum value for which the classical fidelity is greater than 0.999. **c, d.** By  $\chi_0$ , the experimentally benchmarked fidelities have also saturated. **e.** We fit the dependence of  $\chi_0$  as a function of system size and time (solid lines) in order to extrapolate beyond the range of  $\chi$  values which we directly simulate in this work.

We determine  $\chi_0$  by finding the minimum  $\chi$ , at a given system size and evolution time, where the MPS fidelity is greater than 0.999. We extrapolate the behavior of  $\chi_0$  into the high-entanglement regime of large system sizes and late evolution times (Supp. Fig. 10e). We further normalize the input (subtracting off the mean and dividing by the standard deviation) to improve learnability.

To avoid overfitting, we randomly select a few  $N$  from 21 to 30, for which we excise from the training any data with  $\tilde{\chi} > 2/3$  and  $t > 6.6$  cycles. We then use the data for these system sizes and times with  $\tilde{\chi} \sim 1$  as the validation dataset, and stop the training when the validation loss – defined by the prediction root mean square error (RMSE) – does not decrease for 50 epochs in a row (see for instance Supp. Fig. 12b).

We summarize a partial subset of the data showing the Monte Carlo inference as a function of  $N$ ,  $t$ , and  $\chi$  in Supp. Fig. 11.

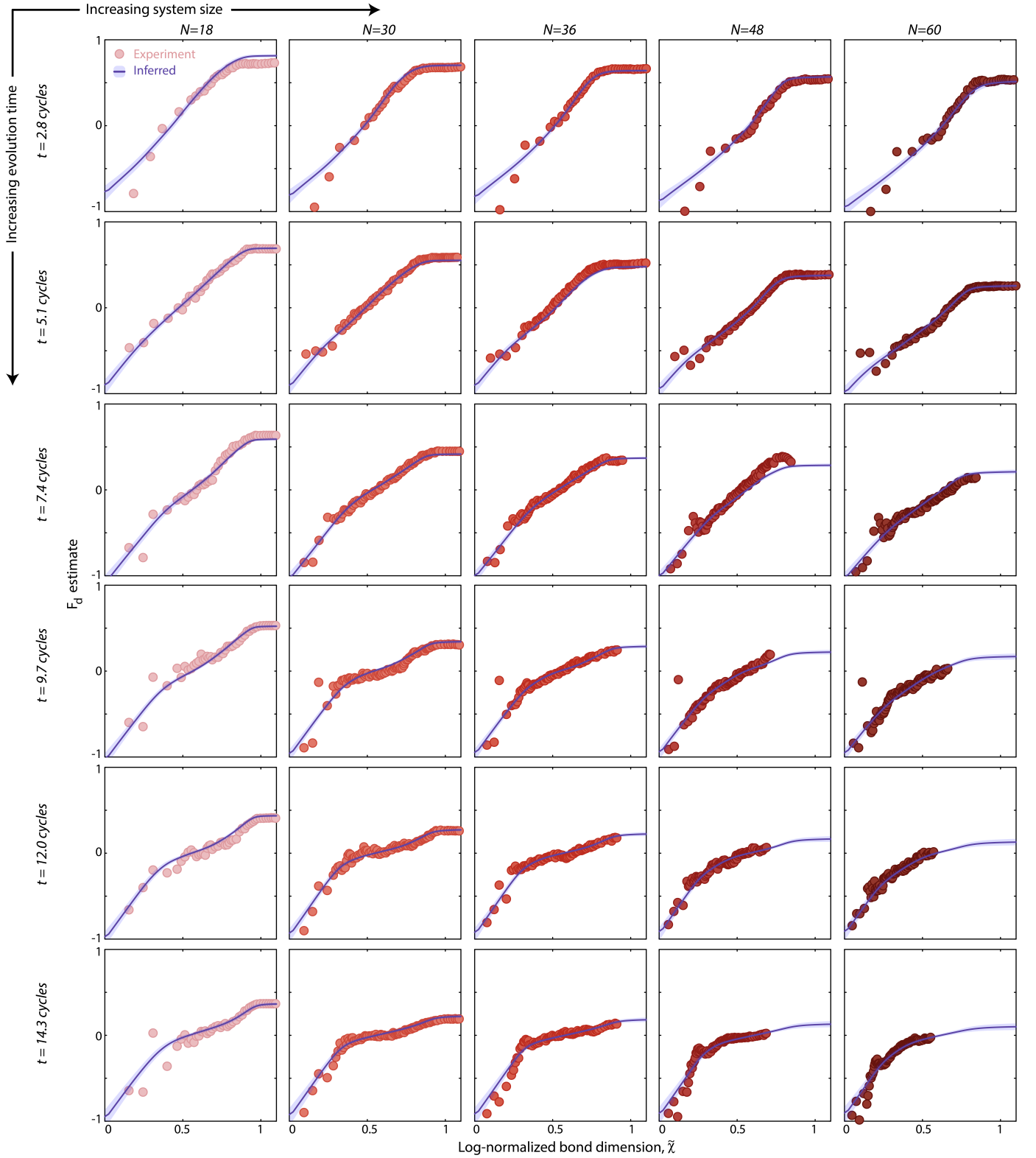
## D2. Consistency checks

In order to test that this approach is accurate in estimating the fidelity, we perform several consistency checks, ranging from varying the number of learnable parameters, testing the result as a function of the amount of experimental training data, checking against small system sizes in experiment and error model where we have exact ground truth data, and more. Several of these tests are described below.

*Small system sizes, experimental*— For the largest system sizes we do not have ground truth data for  $F_d$  in the late-time regime, and so cannot directly verify our model’s efficacy. However, we can emulate the effect in this regime at smaller system sizes where we do have ground truth data in all regimes.

We employ the same training methodology as on the full dataset, but only consider  $N \leq 30$  where we have the full ground truth. For  $t > 6.6$  cycles and  $N \geq 24$ , we excise from the training data all bond dimensions for which  $\tilde{\chi} > 2/3$ . This value is chosen as it is approximately the maximum normalized bond dimension available at  $N = 60$  (Supp. Fig. 9e). Concretely, this approach is essentially pretending as if  $N \geq 24$  is beyond the exactly simulable regime. In reality, of course, we have the ground truth data to compare against, and can for instance use it to study the behavior of the loss functions (Supp. Fig. 12ab).

Averaging over all Monte Carlo instantiations, we find the training loss decreases consistently as a function of training epoch, but the validation loss (where the ground truth  $N \geq 24$  results are used as the validation data set) reaches a minimum after  $\sim 50$  epochs, after which it proceeds into an overfitting regime. As with the full dataset, we thus stop the training for each instantiation at the point where the validation loss is non-decreasing for  $\sim 50$  epochs in a row, and then take the result from the best epoch. We note that both losses drop most dramatically after only a single epoch of training from initially randomized configurations, meaning the gradient descent quickly finds a nearly



**Supp. Fig. 11 | Monte Carlo inference for all data. a.** Fidelity-estimator,  $F_d$ , for various system sizes and evolution times as a function of the log-normalized bond dimension,  $\tilde{\chi} = \log(\chi)/\log(\chi_0)$ . The fidelity is estimated at  $\tilde{\chi} = 1$ . A key feature to note is that as system size or evolution time increases (i.e. as  $\chi_0$  increases), the data points shift left: the classical simulation breakdown regime is when no data is available at  $\tilde{\chi} = 1$ .

optimal regime, implying that the loss function is both smooth and non-flat.

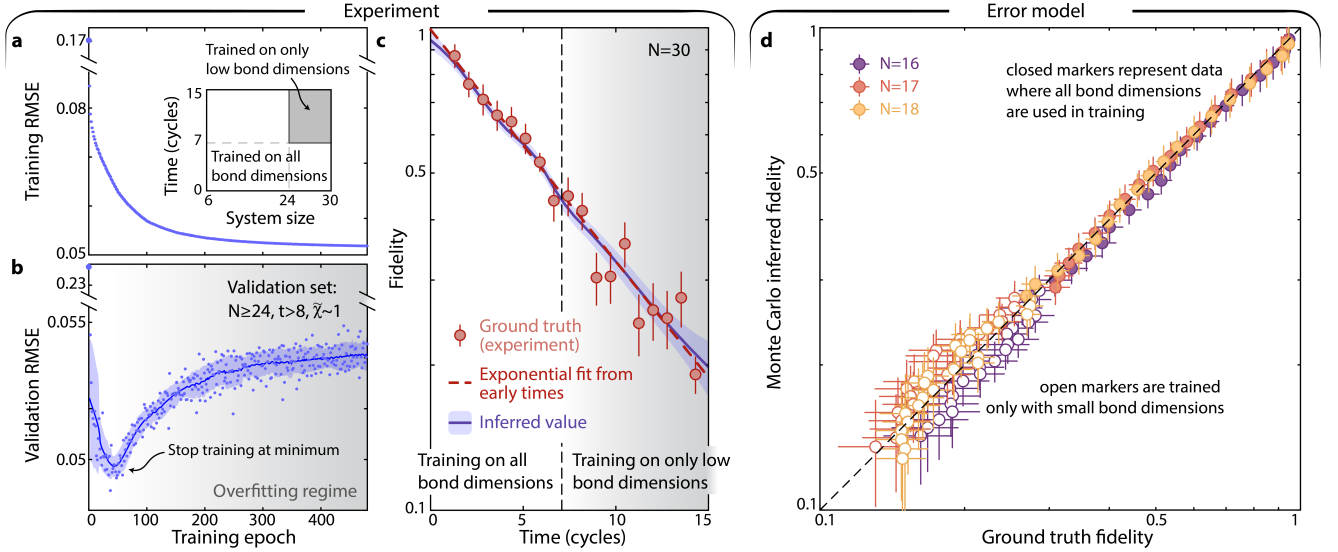
We then compare directly between the Monte Carlo predicted  $F_d$  and the ground truth value for  $N = 30$  across all times. We see good agreement even in the emulated breakdown regime at long times (Supp. Fig. 12c), with no indication that the learned fidelities are misestimating the true values, despite not having access to the underlying ground truth data at large bond dimensions.

*Small system sizes, error model*— As described, we can check our model’s efficacy at small system sizes directly on experimental data, but the non-exponential fidelity decay behavior is not yet strongly visible. In the main text, we reach this regime using Monte Carlo fidelity inference on error model simulations (Fig. 3 of the main text). Here we further describe this analysis.

We perform error model simulations for  $N = 8 - 18$  up to the maximum experimental time, but artificially increase the shot-to-shot Rabi frequency variation by a factor of 4 in order to roughly emulate the predicted experimental fidelities at large  $N$ . We choose to do this, rather than simulate for much longer timescales where the behavior would become naturally evident (Supp. Fig. 8d), because we want to replicate as much as possible the fact that for the experimental dataset the entanglement is still growing for most times for the largest system sizes. Performing error model simulations out to  $\sim 100$  cycles would dominate the training by a regime where the entanglement had already saturated for these small system sizes.

We perform MPS simulations for all bond dimensions up to  $\chi_0(N = 18, t \rightarrow \infty) = 89$ , and benchmark the error model with each. Though we run the error model with a finer timestep of  $\sim 10$  ns, we only select data at the experimentally measured times for training. We then emulate the experimental dataset by treating all  $N < 16$  as “small  $N$ ” where we have full ground truth data. We excise from the training any data for which  $\tilde{\chi} > 2/3$ ,  $N > 15$ , and  $t > 6.6$  cycles, to emulate the available data for the full experimental dataset. We use  $N = 13 - 15$  as the validation dataset to prevent overfitting. We then train in the same way as for the experiment, and compare Monte Carlo predictions versus ground truth output (Supp. Fig. 12d). We find the two are consistent, both in terms of the values obtained, and the non-exponential fidelity decay, indicating the Monte Carlo inference is able to learn this behavior.

*Number of learnable parameters*— For the full dataset, a total of  $\sim 23000$  training data points are generated, while depending on the choice of hyperparameters, the number of learnable parameters can vary from  $\sim 350 - 83000$ . For

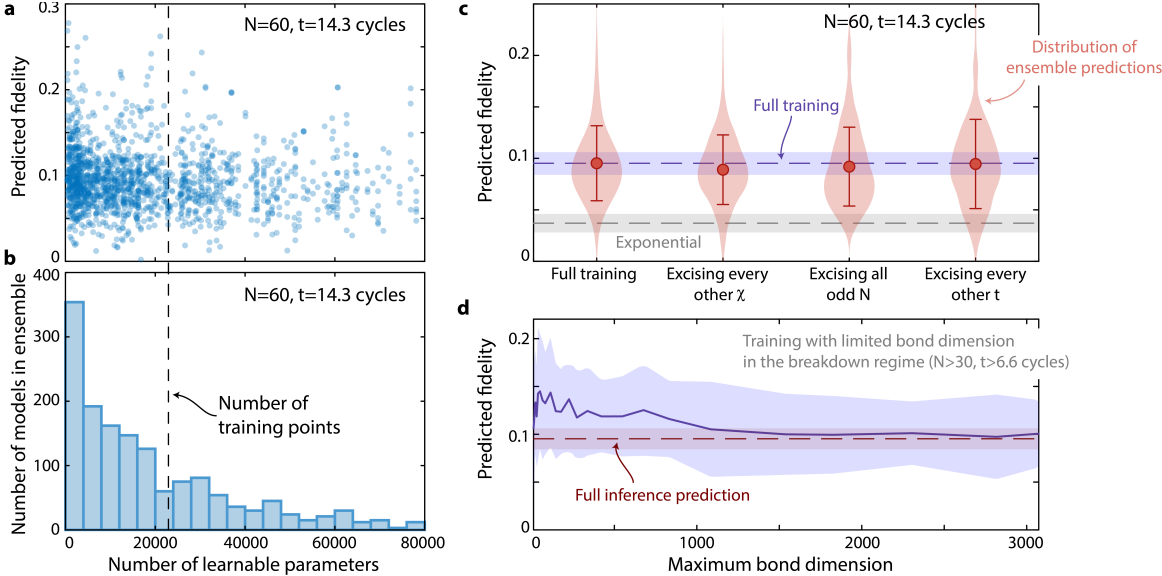


**Supp. Fig. 12 | Evaluating the Monte Carlo inference at small system sizes.** **a, b.** As a cross-check of our approach, we use the same training methodology employed on the full dataset, but truncate to only include  $N \leq 30$  where we have the full ground truth. We bipartition the data (inset), and for the large system sizes and late evolution times we only directly train with  $\tilde{\chi} < 2/3$ . Averaging over 1500 instantiations, we find the training loss decreases consistently, but validation loss at  $\tilde{\chi} \sim 1$  reaches a minimum after  $\sim 50$  epochs, and as expected proceeds into an over-fitting regime after. **c.** Model predictions show the Monte Carlo inference accurately predicts fidelities at late times, despite not having access to the underlying ground truth data at large bond dimensions in this regime. **d.** We perform a similar test using error model simulations of our experiment, as for instance shown in Fig. 3b of the main text. We train with  $N = 8$  to 18, excising  $\tilde{\chi} > 2/3$  for  $N > 15$  and  $t > 6.6$  cycles. The resultant Monte Carlo inference predictions are well correlated with the ground truth, across the whole range of fidelities. X-error bars represent estimated typicality fluctuations, y-error bars represent the Monte Carlo inference standard deviation.

~70% of the Monte Carlo instances, the number of learnable parameters is lower than the number of training points (Supp. Fig. 13ab). Further we do not see any strong correlation between the predicted fidelity and the number of learnable parameters, or more generally between the fidelity and either network depth or width.

*Reducing the experimental dataset*— We perform the same training procedure on several variants of reduced experimental datasets, where we excise from the training half of the bond dimensions, system sizes, or evolution times used. We find the resulting Monte Carlo mean is always consistent with the prediction from training on the full dataset (Supp. Fig. 13c).

*Maximum bond dimension used*— The maximum bond dimension used in the training is constrained by our computational resources, but we can still study if the Monte Carlo prediction appears to be converged as a function of the bond dimension. In Supp. Fig. 13d we systematically scale the training bond dimension used in the breakdown regime of large system sizes and long evolution times, and find that the results appear converged with a bond dimension of ~1500, about a factor of 2 lower than the maximum bond dimension we consider in this work. However, we note that by supplementing the experimental data with numerical data obtained from our error model simulations, this required bond dimension is dramatically reduced (see Section G, Supp. Fig. 14).



**Supp. Fig. 13 | Robustness of Monte Carlo inference.** **a, b.** For the majority (~70%) of hyperparameter combinations sampled during the Monte Carlo inference (Supp. Fig. 9c), the number of learnable parameters is smaller than the number of training points (~23000). Further, we see no correlation between number of learnable parameters (nor any individual specific choice of hyperparameter) with the predicted fidelity. **c.** The Monte Carlo mean prediction is insensitive to the choice of data: removing half of the data along any axis leads to the same predicted fidelity. Note that while the “Full training” is composed of 1500 models, the “excised” ensembles only feature 192 trained models. Markers show mean and standard deviation for each distribution. **d.** To check the convergence of our protocol, we repeat the Monte Carlo inference procedure with varying maximum bond dimension allowed in the breakdown regime of large system sizes and long evolution times. We find the fidelity (at  $N = 60$  and the latest experimental time) is generally consistent over the entire range, with convergence achieved around a bond dimension of 1500.



## E. Scalability of fidelity estimation

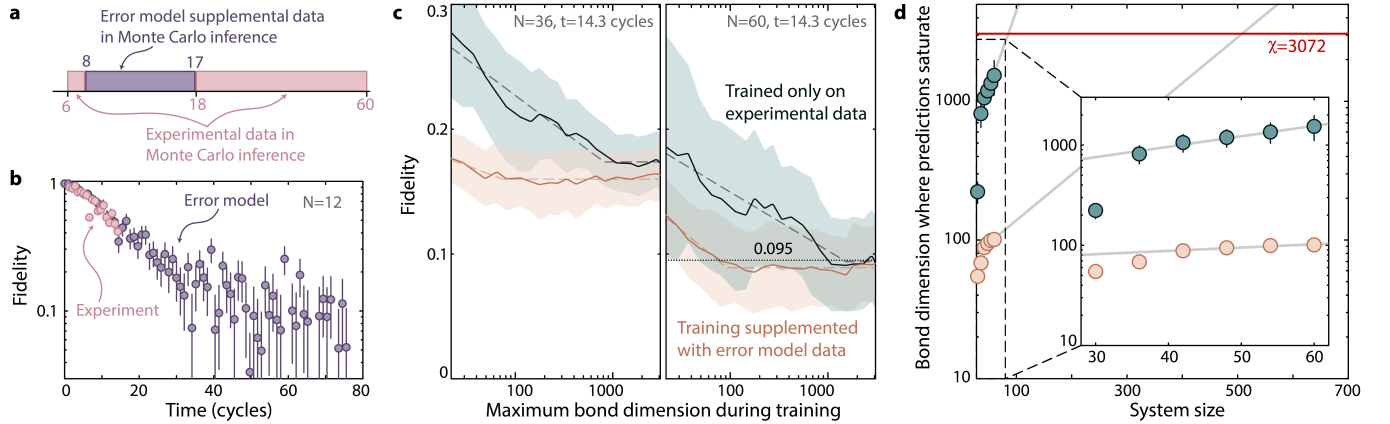
### E1. Fidelity estimation with approximate algorithms for Markovian noise

A major advantage of the fidelity estimator which we use,  $F_d$ , is its accuracy at early times<sup>2,3</sup>. This is in contrast to the original linear cross-entropy which only becomes accurate for deep circuits<sup>5</sup>. As we establish in the main text, using tensor network simulation methods such as MPS, we are able to compute exact classical simulation references for essentially arbitrarily large systems up to constant time. Assuming the noise affecting the system is purely Markovian (i.e. the fidelity decays exponentially<sup>38</sup>), then to estimate the fidelity in the late-time, high-entanglement regime it is sufficient to measure the fidelity with high-precision at early times, and then fit the resulting exponential decay rate.

This gives another axis for fidelity extrapolation which may be particularly useful for digital quantum devices implementing random unitary circuits, which typically use a form of system size scaling to estimate the fidelity in the high-entanglement regime<sup>8</sup>. Such an approach should work for fidelity estimation of systems with largely arbitrary size and dimension, but does require the exact functional form of the fidelity decay be assumed *a priori*, and thus would not apply for systems with more non-trivial noise sources like those affecting our present experiment (or potentially also some digital quantum computers<sup>5</sup>).

### E2. Scalability of Monte Carlo inference for non-Markovian noise sources

In Supp. Fig. 13d, we investigated the minimum bond dimension of the classical (approximate) simulation,  $\chi_{\text{extrap}}$ , required for the Monte Carlo inference procedure to reach a reliable, saturated prediction for the fidelity at  $N = 60$  at the latest experimental time (14.5 cycles). There, we found  $\chi_{\text{extrap}}$  saturated for a bond dimension around  $\chi \sim 1500$ . However, an important question concerns the applicability at the Monte Carlo inference technique at larger system sizes, i.e. the scaling of  $\chi_{\text{extrap}}$  with  $N$ . If the required bond dimension scales poorly with system size then it becomes difficult to trust the Monte Carlo inference results at larger system sizes, assuming there is a cap in the maximum bond dimension usable by the classical simulation.



**Supp. Fig. 14 | Scalability of Monte Carlo inference.** **a.** In order to improve scalability of our Monte Carlo inference protocol, we supplement training with error model data evaluated out to late evolution times for system sizes from  $N = 8$  and  $N = 17$ . **b.** Error model and experimental  $F_d$  values are in good agreement at early times where we have data (see also Ext Data. Fig. 5). Error bars in the late time regime are due to typicality error which reduces with increasing system size (see Supp. Fig. 15). **c.** We plot the predicted fidelity for  $N = 36$  and  $N = 60$  at the latest experimental time for two possible sets of training data: 1) considering only the experimental data (from  $N = 6$  to  $N = 60$  up to 15 cycles), and 2) supplementing with error model data out to late times (75 cycles). In both cases we vary the maximum bond dimension used during training, and then fit to find the bond dimension where predictions saturate,  $\chi_{\text{extrap}}$ . Fits are proportional to  $\log(\chi)$  and then constant after  $\chi_{\text{extrap}}$  (dashed lines). **d.**  $\chi_{\text{extrap}}$  as a function of system size for the predicted fidelity to saturate at the latest experimental time. The  $\chi_{\text{extrap}}$  is improved by roughly an order of magnitude by supplementing with error model data. Conservatively extrapolating with  $\chi_{\text{extrap}} \propto \exp(N)$  (grey lines) and setting a maximum  $\chi = 3072$  shows maximum system sizes of  $\sim 500$  and  $\sim 90$  with and without the error model supplement.

We consider two versions of the training procedure (Supp. Fig. 14). In the first, we perform Monte Carlo inference using just the experimental data, as done previously. In the second approach we replace the training data from  $N = 8$

to 17 with data obtained from error model simulations up to 75 evolution cycles (sampled every 1 cycle).  $F_d$  values are calculated for error model data for  $\chi$  ranging from 1 to 90. We replace with error model results, rather than simply augmenting, so that the training data does not have multiple outputs defined for the same set of inputs.

We have extensively checked the validity of our error model both in terms of fidelity and KL divergence (also see Refs.<sup>2,13</sup> and Supp. Fig. 5), and using error model allows us to generate training data for  $F_d$  as a function of bond dimension in the low- $N$ , long-time regime where the non-exponential fidelity decay becomes more clearly visible. We emphasize that access to an error model is not a general requirement, but is instead simply a tool we are employing in this case as the original experimental dataset was planned and taken without accounting for the non-exponential fidelity decay, as it was a novel observation of our present analysis. Using the error model alone (with no experimental data) for inference is not expected to give good fidelity extrapolation at large system sizes given the large difference between the largest simulable error model size ( $N \approx 20$ ) and the largest experimental size.

In both cases (purely experimental data, or supplemented with error model data), we plot  $\chi_{\text{extrap}}$  as a function of system size. Here  $\chi_{\text{extrap}}$  is found from a phenomenological fit of the predicted fidelity,  $F_p$ , as a function of the maximum bond dimension used in training,  $\chi$ ,

$$F_p = \begin{cases} F + A \log(\chi) & \chi < \chi_{\text{extrap}} \\ F & \chi \geq \chi_{\text{extrap}}, \end{cases} \quad (21)$$

where  $F$  and  $A$  are free fit parameters (Supp. Fig. 14c). In general, we see that supplementing with error model data produces consistent fidelity predictions as our original, all-experimental inferences, but saturates far earlier. This builds additional confidence in both the error model and the Monte Carlo inference procedure.

We then study how  $\chi_{\text{extrap}}$  scales as a function of  $N$ , and set a cap of  $\chi \sim 3072$ , which represents a significant, but still achievable bond dimension. We choose a conservative phenomenological fit model of

$$\chi_{\text{extrap}} = A \exp(BN), \quad (22)$$

where  $A$  and  $B$  are fit constants.

We find the all-experimental inference exceeds this cap for  $N \sim 90$ , but the supplemented dataset goes as high as  $N \sim 500$ . We emphasize that this value represents a lower bound that could be improved either through use of larger classical resources, or possibly through an improved classical algorithm. This indicates that as long as care is taken to acquire experimental data in a diverse regime of system sizes and evolution times, we conservatively believe the Monte Carlo inference procedure we show here should provide scalable predictions for an order of magnitude more atoms than we currently study.

### E3. Sample complexity

To estimate error bars on the raw data,  $\sigma_F$ , we propagate from separately calculating the sampling error of the numerator and denominator of  $F_d$  as

$$\sigma = \sqrt{(\mathbb{E}(\tilde{p}(z)^2) - \mathbb{E}(\tilde{p}(z))^2) / M}, \quad (23)$$

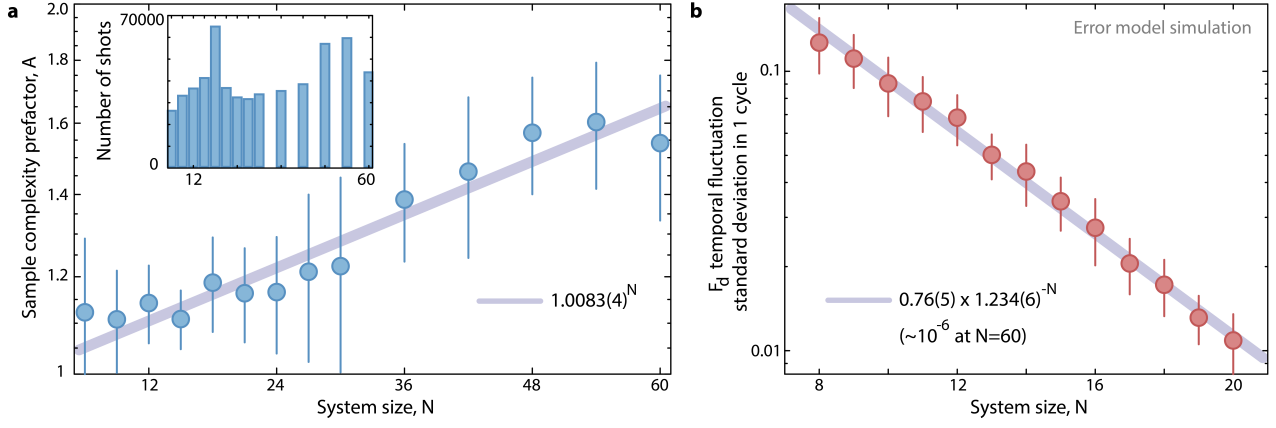
where  $\tilde{p}$  is the rescaled bitstring probability distribution obtained from MPS simulation,  $M$  is the number of samples, and the expectation value is taken over the set of experimentally (numerically) sampled bitstrings for the numerator (denominator). Enough samples are taken numerically that the dominant overall sampling error is from the experiment. See Section B for more details.

We expect the statistical error of our benchmarking protocol<sup>2,3</sup> to scale as  $A/\sqrt{M}$ , where  $A$  is the *sample complexity* which scales exponentially in system size with a base which depends on the fidelity<sup>3</sup>. We measure  $A$  at an approximately fixed fidelity of  $F_d = 0.5$  by performing an average of all the experimentally measured error bars from Eq. (23), weighted by their inverse distance from  $F_d = 0.5$ , such that

$$A(N) = \frac{\sum_t \sigma_F(N, t) \sqrt{M(N, t)} w(N, t)}{\sum_t w(N, t)}, \quad (24)$$

with  $w(N, t) = 1/|F_d(N, t) - 0.5|$ . With this approach, we find  $A = 1.0083(4)^N$  (Supp. Fig. 15a).





**Supp. Fig. 15 | Uncertainty sources of the benchmarking protocol.** There are two fundamental sources of noise in calculating  $F_d$ : statistical error, and systematic error arising from so-called typicality fluctuations<sup>2</sup>. **a.** We first evaluate the statistical error behavior. At a fixed fidelity, the statistical error bars of our protocol are given<sup>3</sup> by  $A/\sqrt{M}$ , where  $M$  is the number of samples and  $A$  is the sample complexity. At approximately a fixed fidelity of  $F_d = 0.5$  (see text), we find  $A$  grows weakly with system size, indicating favorable scalability of our protocol from a statistical perspective. For this work,  $M \sim 40000$  for all system sizes, summing over all times (inset). **b.** Typicality error manifests as temporal fluctuations of the estimator  $F_d$ , with respect to the true fidelity  $F$ , due to benchmarking in a finite-sized Hilbert space. The standard deviation of these fluctuations scales<sup>2</sup> as  $1/\sqrt{D}$ , where  $D$  is the Hilbert space dimension. From error model simulations, we calculate the standard deviation of temporal fluctuations of  $F_d$  in a 1-cycle period at late times, which we find decreases as  $1.23^{-N}$  (or  $\sim 1.2/\sqrt{D}$ ), implying a negligible typicality error of  $\sim 10^{-6}$  at  $N = 60$ .

## F. Extrapolating pure state entanglement growth

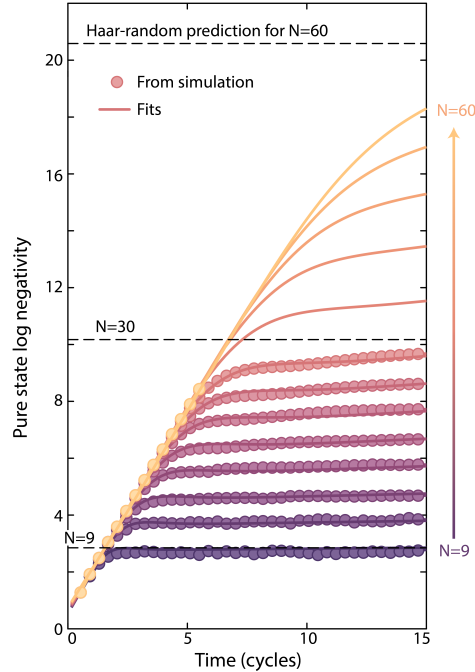
We target the generation of high entanglement entropy states through effectively infinite-temperature quench dynamics. In order to estimate the mixed state negativity, we first must estimate the pure state entanglement, even for system sizes for which we can no longer exactly simulate the dynamics. To do this, we extrapolate the scaling behavior from small system sizes. Specifically, we fit entanglement growth,  $S$ , with the following functional form as a function of time and system size:

$$S = S_0 + m_1 t + \frac{(m_2 - m_1)(t - t_{\text{ent}} - \sigma)}{1 + \exp(-\frac{t - t_{\text{ent}}}{\sigma})}. \quad (25)$$

To respect the physical constraints of the entanglement growth, we fix  $m_1$  – the early time linear entanglement growth – to be system size independent. The Rydberg Hamiltonian induces fast entanglement growth over the first  $\sim 0.5$  cycles of evolution which makes fitting the linear growth erroneous without also including a system-size independent offset,  $S_0$ . We then expect  $t_{\text{ent}}$ , which we identify with the approximate entanglement saturation time, to scale linearly with  $N$ , which we empirically find is true also for  $\sigma$  (a smoothness parameter). We further empirically find that  $m_2$  (the shallow entanglement growth slope after  $t_{\text{ent}}$ ) scales approximately quadratically with  $N$ .

We show the results of these fits in Supp. Fig. 16, where markers are measured directly from simulation, and lines are fits. Though here we show log negativity, the von Neumann entropy scales and acts in the same way.

It is important to note that as system size increases, we gradually reach relatively less entangled states at  $t_{\text{ent}}$  as compared to the expectation for a Haar random state (dashed lines in Supp. Fig. 16). Given the Rydberg blockade constraint, we define the Haar random prediction as  $S_{\text{Haar}} \approx \log_2(D_0) - 0.47$ , where  $D_0$  is the blockaded Hilbert space dimension of the half-cut chain, and the numerical offset of -0.47 is a type of Page correction<sup>1</sup>, which is actually<sup>39,40</sup>  $\approx \log_2(\frac{64}{9\pi^2})$ , as explained further below.



**Supp. Fig. 16 | Pure state entanglement growth dynamics for the Rydberg Hamiltonian.** We directly calculate the entanglement (here log negativity) for system sizes up to  $N = 60$ , and fit the resulting data (see text) to extrapolate the entanglement dynamics for large  $N$  at late times. We also compare against the prediction for Haar random states at various system sizes (black dashed lines), and find that as system size increases we gradually undershoot this expectation more, which we attribute potentially to imperfect parameter selection making the dynamics effectively less than infinite temperature at large  $N$ .

## G. Bounds on the mixed state entanglement proxy

In this work, we have introduced the entanglement-proxy,  $\mathcal{E}_P$ , which estimates the mixed state entanglement,  $\mathcal{E}_N(\hat{\rho})$ , in terms of the pure state entanglement,  $\mathcal{E}_N(|\psi\rangle)$ , and the fidelity,  $F \equiv \langle \psi | \hat{\rho} | \psi \rangle$  (see Eq. (2) of the main text).

Explicitly, consider a system bipartitioned into two subsystems,  $A$  and  $B$ , and described by density matrix,  $\hat{\rho}_{AB}$ . The log negativity<sup>41</sup> is given by

$$\mathcal{E}_N = \log_2 \|\hat{\rho}_{AB}^{TA}\|_1,$$

where  $\hat{\rho}_{AB}^{TA}$  is the partially transposed density matrix, and  $\|\cdot\|_1$  is the trace norm. More plainly, the log negativity measures the log absolute sum of all eigenvalues of  $\hat{\rho}_{AB}^{TA}$ . If  $\hat{\rho}_{AB}$  is separable, the log negativity is zero, but for entangled states it generically serves as an upper bound to the distillable entanglement of the system<sup>42</sup>, and is an entanglement monotone<sup>42</sup>. For pure states the log negativity is equivalent to the Rényi-1/2 entropy.

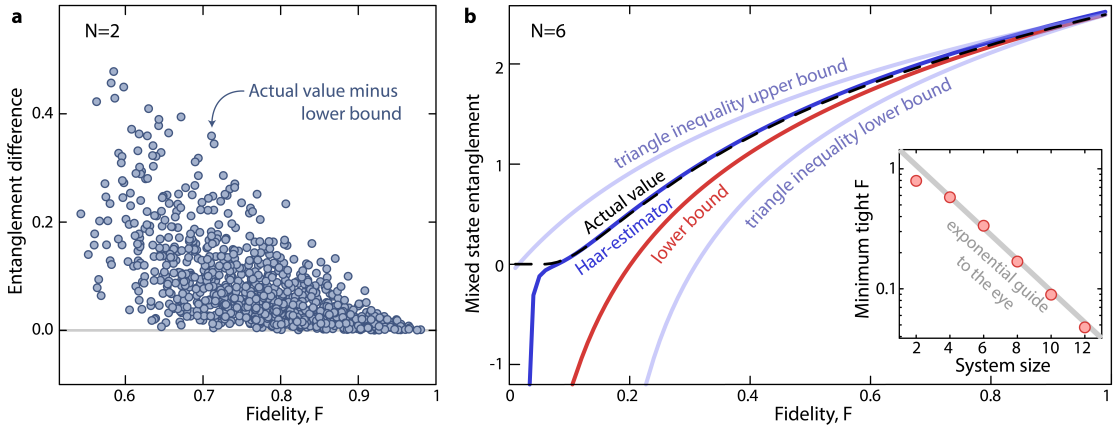
The following discussion is generally split into three subsections:

In the first, we study lower bounds and estimators of the mixed state entanglement. We prove the validity of our entanglement-proxy,  $\mathcal{E}_P$ , for the case when the target pure state,  $|\psi\rangle$  is an eigenstate of the mixed state,  $\hat{\rho}$ , as for instance is the case if  $|\psi\rangle$  is the highest fidelity pure state to  $\hat{\rho}$ . We further improve our result for the case where  $|\psi\rangle$  is a Haar-random state. Finally, we prove more general unconditional upper and lower bounds to argue that our entanglement-proxy becomes exponentially tighter as the system size increases.

Next, we discuss possible violations of the lower bound given by  $\mathcal{E}_P$ , if we allow  $|\psi\rangle$  to no longer be an eigenstate of  $\hat{\rho}$ . We consider two physically realistic possibilities: 1) globally correlated coherent errors, and 2) incoherent local errors. For each case we provide numeric or analytic evidence showing that at worst our entanglement-proxy lower bound is potentially violated by an  $\mathcal{O}(1)$  amount.

Finally, we make a specific connection between the entanglement content of an erroneous mixed state, and of a truncated MPS state. We show that at equal fidelity, the truncated MPS is less entangled than the depolarized Haar random state, which explains an observed discrepancy between the experimental mixed state entanglement and  $\chi^*$  found in the main text.

As a brief aside, we note an important point, arising from the fact that in the main text we have used either  $F_d$  as the fidelity for experiment, or the linear cross-entropy,  $F_{\text{XEB}}$ , as the fidelity for the literature examples. However,  $F_d$  and  $F_{\text{XEB}}$  are both lowered by measurement error, which strictly speaking does not affect the mixed state entanglement. However, both  $F_d$  and  $F_{\text{XEB}}$  are expected to decrease due to measurement errors, and so this only serves to make our



**Supp. Fig. 17 | Estimation of log negativity from fidelity.** **a.** We demonstrate the validity of the lower bound Eq. (2) of the main text, by plotting the difference  $\mathcal{E}_N(\hat{\rho}) - \mathcal{E}_P$ , with  $|\psi\rangle$  the pure state with highest fidelity to  $\hat{\rho}$ . Here, we generate 1000 random two-qubit mixed states as uniform incoherent mixtures of two Haar random states. The difference is always above zero, indicating the validity of the lower bound. **b.** Dependence between log negativity  $\mathcal{E}_N(\hat{\rho})$  and fidelity  $F$  for globally depolarized Haar random states, Eq. (34). For small values of  $F$ , the negativity  $\mathcal{E}_N(\hat{\rho})$  (black dashed) deviates from the lower bound  $\mathcal{E}_P$  (red), illustrated here for the half-chain negativity of six-qubit depolarized Haar random states. Remarkably, for such states,  $\mathcal{E}_N(\hat{\rho})$  is uniquely determined by the fidelity  $F$  and Hilbert space dimension  $D$ . We find an analytic series expression for this dependence, Eq. (36), with first-order correction illustrated in blue. We also illustrate general lower and upper bounds for the log negativity, Eq. (41), in light blue. Inset: The lower bound (red) improves exponentially as a function of system size, quantified here as the minimum fidelity for which the lower bound is 90% of the actual value.

lower bound less tight. In principle the mixed state entanglement-proxy could be measurement-corrected to account for this effect, but we choose to include measurement errors as a more general quality-factor of the experiments.

### G1. Negativity lower bounds and estimators

The validity of our entanglement-proxy has been shown in the past<sup>43</sup> for the case of *isotropic states*; here we extend this result and validate our entanglement-proxy when the target pure state,  $|\psi\rangle$ , is an eigenstate of the mixed state,  $\hat{\rho}$  (Theorem 2). We then improve our result for the case of depolarized Haar-random pure states (Lemma 1), and then show more general lower bounds which allow us to argue our entanglement-proxy should tighten exponentially with system size (Lemma 2). Numerical support of these analytical claims is shown in Supp. Fig. 17.

We begin with the case of  $|\psi\rangle$  being an eigenstate of  $\hat{\rho}$

**Theorem 2.** *For any mixed state  $\hat{\rho}$  and pure state  $|\psi\rangle$  which is an eigenstate of  $\hat{\rho}$ , the logarithmic negativities  $\mathcal{E}_N(\hat{\rho})$  and  $\mathcal{E}_N(|\psi\rangle)$  are related by*

$$\mathcal{E}_N(\hat{\rho}) \geq \mathcal{E}_P \equiv \mathcal{E}_N(|\psi\rangle) + \log_2(F) , \quad (26)$$

where  $F \equiv \langle \psi | \hat{\rho} | \psi \rangle$  is their fidelity.

*Proof.* The proof of this Theorem follows in three steps. It is equivalent to derive the bound

$$\|\hat{\rho}^{TA}\|_1 \geq F \|\psi\rangle\langle\psi|^{TA}\|_1 . \quad (27)$$

First, we express the state

$$\hat{\rho} = F|\psi\rangle\langle\psi| + (1 - F)\hat{\rho}_\perp \quad (28)$$

for some positive semi-definite state  $\hat{\rho}_\perp$ , which is guaranteed because  $|\psi\rangle$  is an eigenstate of  $\hat{\rho}$ .

Next, we use the variational definition of the trace-norm as a maximum over projectors  $\hat{P}$ , i.e. operators with eigenvalues either 0 or 1.

$$\|\hat{\rho}^{TA}\|_1 = 2 \max_P \text{tr}(\hat{P}\hat{\rho}^{TA}) - 1 . \quad (29)$$

This maximum is attained when  $\hat{P}$  is the projector onto the positive-eigenvalue subspace of  $\hat{\rho}^{TA}$ . This variational definition gives the lower bound:

$$\|\hat{\rho}^{TA}\|_1 \geq F \|\psi\rangle\langle\psi|^{TA}\|_1 + (1 - F)(2\text{tr}(\hat{P}_\psi\hat{\rho}_\perp^{TA}) - 1) , \quad (30)$$

where  $\hat{P}_\psi$  is the projector onto the positive-eigenvalue subspace of  $|\psi\rangle\langle\psi|^{TA}$ .

Finally, we utilize an explicit construction of  $\hat{P}_\psi$  to show that  $\text{tr}(\hat{P}_\psi\hat{\rho}_\perp^{TA}) \geq 1/2$ , giving the desired bound Eq. (27). This step is the most involved, which we show below.

The eigenvalues  $\lambda$  and eigenvectors  $|v\rangle$  of  $|\psi\rangle\langle\psi|^{TA}$  can be expressed in terms of the Schmidt values  $s_j$  and Schmidt basis  $|a_j\rangle, |b_j\rangle$  of  $|\psi\rangle$ , with respect to the same bipartition<sup>43</sup>. With  $|\psi\rangle = \sum_j s_j |a_j\rangle |b_j\rangle$ , the eigenvalues  $\lambda$  are  $s_j^2$  or  $\pm s_i s_j$  (for  $i \neq j$ ), and their corresponding eigenvectors are respectively  $|v_{jj}\rangle \equiv |a_j\rangle |b_j\rangle$  and  $|v_{ij}^\pm\rangle \equiv \frac{1}{2}(|a_i\rangle |b_j\rangle \pm |a_j\rangle |b_i\rangle)$ . Then  $\hat{P}_\psi$  has explicit expression:

$$\hat{P}_\psi = \sum_j |v_{jj}\rangle\langle v_{jj}| + \sum_{i < j} |v_{ij}^+\rangle\langle v_{ij}^+| . \quad (31)$$

Next, we use the relation  $\text{tr}(\hat{M}^{TA}\hat{N}^{TA}) = \text{tr}(\hat{M}\hat{N})$ , valid for all operators  $\hat{M}$  and  $\hat{N}$  defined on  $AB$ . Thus  $\text{tr}(\hat{P}_\psi\hat{\rho}_\perp^{TA}) = \text{tr}(\hat{P}_\psi^{TA}\hat{\rho}_\perp)$ .  $\hat{P}_\psi^{TA}$  can be expressed as

$$\hat{P}_\psi^{TA} = \frac{1}{2} (I + |\Phi\rangle\langle\Phi|) , \quad (32)$$

where  $|\Phi\rangle \equiv \sum_j |a_j\rangle|b_j\rangle$  is the *unnormalized* maximally entangled state in the Schmidt bases  $\{|a_j\rangle\}$  and  $\{|b_j\rangle\}$ . Since  $\hat{\rho}_\perp$  is a positive semidefinite state, we have

$$\text{tr}(\hat{P}_\psi^{TA} \hat{\rho}_\perp) = \frac{1}{2} \text{tr}(\hat{\rho}_\perp) + \frac{1}{2} \langle \Phi | \hat{\rho}_\perp | \Phi \rangle \geq \frac{1}{2}. \quad (33)$$

This proves the desired bound Eq. (27).  $\square$

This Theorem validates the entanglement proxy in the simplest case. We next move to improve it, first specifically in the case where  $|\psi\rangle$  is taken to be a Haar random state,  $|\psi_{\text{Haar}}\rangle$ .

**Lemma 1.** *For depolarized Haar random states, i.e. states of the form*

$$\hat{\rho} = F' |\psi_{\text{Haar}}\rangle \langle \psi_{\text{Haar}}| + (1 - F') \frac{I}{D}, \quad (34)$$

where  $F' = F - (1 - F)/(D - 1)$  (chosen so the fidelity  $\langle \psi_{\text{Haar}} | \hat{\rho} | \psi_{\text{Haar}} \rangle$  equals  $F$ ), the mixed state entanglement has the form

$$\mathcal{E}_N(\hat{\rho}) \approx \mathcal{E}_N(|\psi\rangle) + \log_2(\mathfrak{f}(F)), \quad (35)$$

where the more general function  $\mathfrak{f}$  can be expressed to low order as

$$\mathfrak{f}(F) = F' + \frac{9}{64} (F' D_A)^{-1} (2 \ln(F' D_A) + 8 \ln 2 - 1) + \mathcal{O}((F' D_A)^{-3}), \quad (36)$$

where  $D_A$  is the Hilbert space dimension of subsystem  $A$ , here assumed to be for a half-chain bipartitions.

*Proof.* Here, we restrict our discussion to equal bipartitions, although our results can be generalized to any bipartition.

First, with Haar random pure states, with high probability we have<sup>39,40</sup>

$$\mathcal{E}_N(|\psi\rangle) = \frac{N}{2} + \log_2 \frac{64}{9\pi^2} \approx \frac{N}{2} - 0.472 \quad (37)$$

This expression follows from random matrix theory (RMT), which states that the Schmidt values  $s_j$  of a Haar random state follow the *quarter-circle law*, while  $s_j^2$  follow the *Marchenko-Pastur distribution*.

This enables us to compute the distribution of eigenvalues  $\lambda$  of  $|\psi\rangle\langle\psi|^{TA}$ . The eigenvalues of the form  $\lambda = s_j^2$  contribute a constant amount to  $\|\hat{\rho}^{TA}\|_1$ . The dominant (exponentially larger) contribution to the log negativity comes from the eigenvalues of the form  $\lambda = \pm s_i s_j$  which follow the distribution<sup>44</sup>

$$P(\tilde{\lambda}) = \frac{2}{\pi^2} \left[ \left( 1 + \frac{\tilde{\lambda}^2}{16} \right) K \left( 1 - \frac{\tilde{\lambda}^2}{16} \right) - 2E \left( 1 - \frac{\tilde{\lambda}^2}{16} \right) \right], \quad (38)$$

where  $\tilde{\lambda} \equiv D_A \lambda$ ,  $P(\tilde{\lambda})$  is supported in  $\tilde{\lambda} \in [-4, 4]$ , and  $K$  and  $E$  are complete elliptic integrals.

The negativity of the depolarized state  $\hat{\rho}$  can be calculated using this distribution: the eigenvalues  $\lambda'$  of  $\hat{\rho}^{TA}$  are given by  $\lambda' = F' \lambda + (1 - F')/D$ . Therefore, the trace norm can be calculated from  $P(\tilde{\lambda})$  as

$$\|\hat{\rho}^{TA}\|_1 \approx 2F' D_A \int_0^\infty \tilde{\lambda} P(\tilde{\lambda}) d\tilde{\lambda} + 2F' D_A \int_0^{(F' D_A)^{-1}} \left[ (F' D_A)^{-1} - \tilde{\lambda} \right] P(\tilde{\lambda}) d\tilde{\lambda} \quad (39)$$

The first term evaluates to  $\frac{64}{9\pi^2} F' D_A$ , while the second term can be systematically computed by using the expansion  $P(\tilde{\lambda}) \approx -\frac{2}{\pi^2} \ln |\tilde{\lambda}| + \frac{8 \ln 2 - 4}{\pi^2}$ , valid for small  $\tilde{\lambda}$ , giving Eq. (36).  $\square$

Importantly, if more terms are added in the expansion of Eq. (36) Lemma 1 is an exact estimator of the mixed state entanglement. Further, if  $D_A$  is large, as is the case for large system sizes, then the original simple proxy is restored as  $\mathfrak{f} \sim F' \sim F$ . We show the efficacy of this estimator, with  $\mathfrak{f}(F)$  evaluated to first order, in Supp. Fig. 17b.

Finally we prove an even more general set of upper and lower bounds. While not amenable to be computed experimentally, they allow us to argue our entanglement-proxy likely becomes exponentially tight as the system size increases.

**Lemma 2.** For any state pure state  $|\psi\rangle$  and mixed state  $\hat{\rho}$  with decomposition

$$\hat{\rho} = F|\psi\rangle\langle\psi| + (1 - F)\hat{\rho}_\perp, \quad (40)$$

where  $\hat{\rho}_\perp$  is the (not necessarily positive semi-definite) remainder of the noisy state, we can bound the mixed state log negativity by

$$\begin{aligned} \mathcal{E}_N(\hat{\rho}) &\geq \log_2 \left( F 2^{\mathcal{E}_N(|\psi\rangle)} - (1 - F) \|\hat{\rho}_\perp^{TA}\|_1 \right) \\ \mathcal{E}_N(\hat{\rho}) &\leq \log_2 \left( F 2^{\mathcal{E}_N(|\psi\rangle)} + (1 - F) \|\hat{\rho}_\perp^{TA}\|_1 \right), \end{aligned} \quad (41)$$

where  $F$  is a lower bound for the fidelity  $\langle\psi|\hat{\rho}|\psi\rangle$ .

*Proof.* Unlike above,  $\hat{\rho}_\perp$  does not necessarily commute with  $|\psi\rangle\langle\psi|$ .

To bound the negativity for such a state, we simply apply the triangle inequality on the trace norm  $\|\cdot\|_1$ , yielding Eqs. (41).  $\square$

This bound is particularly useful in settings such as ours. Here, we can expect the noisy state to be weakly entangled, i.e.  $\|\hat{\rho}_\perp^{TA}\|_1 = \mathcal{O}(1)$ , while the pure state is highly entangled:  $\| |\psi\rangle\langle\psi|^{TA} \|_1 = \mathcal{O}(\exp(N))$ , where  $N$  is the system size. Then the above bounds imply that our entanglement-proxy is in fact exponentially close to the true mixed state negativity:  $\mathcal{E}_N(\hat{\rho}) = \mathcal{E}_P + \mathcal{O}(\exp(-N))$ , as observed in the inset of Ext. Data Fig. 17b.

## G2. Robustness of our bounds to realistic experimental errors

In the above section, we showed several proofs which let us argue our entanglement-proxy,  $\mathcal{E}_P$  is a tight lower bound for the mixed state entanglement under specific assumptions. Here, we relax these assumptions, but show that under realistic noise sources, maximum possible violations of our bound are at most  $\mathcal{O}(1)$ .

**Error model simulations**— To start, we reiterate and emphasize the results of Fig. 4a of the main text. There, we showed numerically that the entanglement-proxy appears to be a genuine lower bound of the mixed state entanglement either for a random unitary circuit (RUC) undergoing incoherent local noise, or for our experiment undergoing the full set of error sources (Supp. Fig. 32).

Now, our aim is to provide further analytical and numerical evidence for the robustness of the entanglement-proxy under: 1) global coherent errors, or 2) incoherent local errors.

**Global coherent errors**— In this section, we consider potential violations of our bound due to global coherent errors. To study this, we first prove a Theorem for a more general bound of the mixed state entanglement, based on studying the norm of the commutator between the mixed state  $\hat{\rho}$  and pure state  $|\psi\rangle$ . This bound is equivalent to our original entanglement-proxy up to a correction term, which we then show analytically is small if the normalized entanglement of the target state is large. Finally, we present numerical support of this discussion.

**Theorem 3** (Commutator based bound of the mixed state entanglement). *For any mixed state  $\hat{\rho}$  and pure state  $|\psi\rangle$  whose commutator has Frobenius norm  $\mathcal{C} \equiv \|\hat{\rho}, |\psi\rangle\langle\psi|\|_F$ , the logarithmic negativities  $\mathcal{E}_N(\hat{\rho})$  and  $\mathcal{E}_N(|\psi\rangle)$  are related by*

$$\mathcal{E}_N(\hat{\rho}) \geq \mathcal{E}_N(|\psi\rangle) + \log_2(F) + 2 \log_2(1 - (\mathcal{C}/F) \sqrt{(1 - \alpha)/(2\alpha)}) . \quad (42)$$

where  $F \equiv \langle\psi|\hat{\rho}|\psi\rangle$  is the fidelity,  $\alpha \equiv 2^{\mathcal{E}_N(|\psi\rangle)}/D_A \in [0, 1]$  is a normalized entanglement of  $|\psi\rangle$ , and  $D_A$  is the Hilbert space dimension of the subsystem  $A$  (assumed to be smaller than its complement).

*Proof.* We write  $\hat{\rho}$  in a suitable basis as:

$$\hat{\rho} = \begin{pmatrix} F & B^\dagger \\ B & \hat{\rho}' \end{pmatrix}, \quad (43)$$

where the first row/column denotes the basis vector  $|\psi\rangle$ , and the second denotes the subspace perpendicular to  $|\psi\rangle$ . With  $P_\perp$  the projector onto this subspace, the matrix  $\hat{\rho}' \equiv P_\perp \hat{\rho} P_\perp$  is

positive semi-definite. The off-diagonal element  $B \equiv P_{\perp} \hat{\rho} |\psi\rangle$  is a  $(d-1) \times 1$  vector, which we take to be an unnormalized vector  $|B\rangle$  in the full Hilbert space (with first entry equal to zero).

Using the variational definition of the trace-norm used in the proof of Theorem 2, we write

$$\|\hat{\rho}^{TA}\|_1 \geq \text{tr}(\hat{P}_{\psi} \hat{\rho}^{TA}) = \langle \Phi | \hat{\rho} | \Phi \rangle \quad (44)$$

$$= F |\langle \psi | \Phi \rangle|^2 + \langle \Phi | \hat{\rho}' | \Phi \rangle + \langle \Phi | \psi \rangle \langle B | \Phi \rangle + \langle \Phi | B \rangle \langle \psi | \Phi \rangle, \quad (45)$$

where  $|\Phi\rangle$  is the *unnormalized* maximally entangled state in the full Hilbert space, in the Schmidt bases  $\{|a_j\rangle\}$  and  $\{|b_j\rangle\}$  of  $|\psi\rangle$ . We then bound the terms of this expression.

First, the Frobenius norm of the commutator  $\mathcal{C} = \|\llbracket \hat{\rho}, |\psi\rangle\langle\psi| \rrbracket\|_F$  has the expression

$$\mathcal{C}^2 = 2\langle \psi | \hat{\rho}^2 | \psi \rangle - 2\langle \psi | \hat{\rho} | \psi \rangle^2 = 2\langle B | B \rangle. \quad (46)$$

Therefore,  $\langle B | B \rangle \leq \mathcal{C}^2/2$ . We then use the *Schur complement* condition for positive semi-definite matrices. With our decomposition, the fact that  $\hat{\rho}$  is positive semi-definite implies that the Schur complement  $\hat{\rho}' - F^{-1}|B\rangle\langle B|$  is positive semi-definite. We conclude that

$$\langle \Phi | B \rangle \langle B | \Phi \rangle \leq F \langle \Phi | \hat{\rho}' | \Phi \rangle, \quad (47)$$

Using the fact that,  $|B\rangle \perp |\psi\rangle$  and that  $|\Phi\rangle$  is an unnormalized state with norm  $D_A$  and overlap  $|\langle \Phi | \psi \rangle|^2 \equiv \alpha D_A$ , then accounting for the norms of  $|\Phi\rangle$  and  $|B\rangle$ , the maximum possible value of  $|\langle B | \Phi \rangle|$  is  $|\langle B | \Phi \rangle| \leq \sqrt{(1-\alpha)D_A \mathcal{C}}/\sqrt{2}$ . Eq. (47) also implies that  $\langle \Phi | \hat{\rho}' | \Phi \rangle \geq (1-\alpha)D_A \mathcal{C}^2/(2F)$ .

Combining these ingredients with Eq. (45), we then have

$$\|\hat{\rho}^{TA}\|_1 \geq \langle \Phi | \hat{\rho} | \Phi \rangle \geq F\alpha D_A + (1-\alpha)D_A \mathcal{C}^2/(2F) - 2\sqrt{\alpha D_A^2(1-\alpha)\mathcal{C}^2/2} \quad (48)$$

$$= \left( \sqrt{F\alpha} - \sqrt{(1-\alpha)\mathcal{C}^2/(2F)} \right)^2 D_A \quad (49)$$

$$= F\alpha D_A \left( 1 - (\mathcal{C}/F)\sqrt{(1-\alpha)/(2\alpha)} \right)^2 \quad (50)$$

Taking the logarithm of both sides, we arrive at Theorem 3.  $\square$

With this Theorem 3 in hand, to validate our entanglement proxy we need only bound the size of the quantity  $\mathcal{C}/F$ , and show it is  $\mathcal{O}(1)$ . Here, we show this is the case under the physically realistic scenario of the system undergoing coherent global errors, which are a dominant error source in our experiment (Supp. Fig. 32). First, we prove a useful Lemma rewriting the coefficient to the correction term,  $\mathcal{C}/F$ , in terms of only  $F$ .

**Lemma 3.** *For normally distributed global coherent quasistatic parameter errors, the coefficient to the correction term in Theorem 3,  $\mathcal{C}/F$ , can be written entirely in terms of the fidelity as*

$$\frac{\mathcal{C}}{F\sqrt{2}} = \sqrt{\frac{2}{\sqrt{3+2F^2-F^4}} - 1}. \quad (51)$$

*Proof.* To begin, consider evolution under a fixed parameter value, which results in the pure state  $|\Psi(t, \theta)\rangle \equiv e^{-i(\hat{H}_0 + \theta \hat{V})t} |\Psi_0\rangle$ . In Theorem 1, we found a Gaussian fidelity dependence between states corresponding to any two parameter values  $\theta_1$  and  $\theta_2$ :

$$F(\theta_1, \theta_2) \equiv |\langle \Psi(t, \theta_1) | \Psi(t, \theta_2) \rangle|^2 = \exp\left(-N\lambda t^2 \frac{(\theta_1 - \theta_2)^2}{2}\right), \quad (52)$$

For some constant  $\lambda$  [Eq. (5)]. Assuming a normal distribution  $P(\theta)$  of parameter values  $\theta$  with mean  $\theta_0$  and standard deviation  $\sigma$ , we obtain the mixed state  $\hat{\rho}(t) = \int d\theta P(\theta) |\Psi(t, \theta)\rangle\langle\Psi(t, \theta)|$ . In Eq. (19), we computed the fidelity of  $\rho(t)$  to the pure state  $|\Psi(t, \theta_0)\rangle$  and found it to be  $F = 1/\sqrt{\Lambda + 1}$ , where  $\Lambda \equiv Nt^2\lambda\sigma^2$ .

The Frobenius norm of the commutator,  $\mathcal{C}$ , then has expression

$$\mathcal{C}^2 = \|\hat{\rho}(t), |\Psi(t, \theta_0)\rangle\langle\Psi(t, \theta_0)|\|_F^2 \quad (53)$$

$$= 2\langle\Psi(t, \theta_0)|\hat{\rho}(t)^2|\Psi(t, \theta_0)\rangle - 2\langle\Psi(t, \theta_0)|\hat{\rho}(t)|\Psi(t, \theta_0)\rangle^2 \quad (54)$$

$$= 2 \int d\theta_1 P(\theta_1) \int d\theta_2 P(\theta_2) \langle\Psi(t, \theta_0)|\Psi(t, \theta_1)\rangle \langle\Psi(t, \theta_1)|\Psi(t, \theta_2)\rangle \langle\Psi(t, \theta_2)|\Psi(t, \theta_0)\rangle - 2F^2 \quad (55)$$

$$= \frac{4}{\sqrt{3\Lambda^2 + 8\Lambda + 4}} - \frac{2}{\Lambda + 1}. \quad (56)$$

In this computation, the unknown phases of the inner products  $\langle\Psi(t, \theta_i)|\Psi(t, \theta_j)\rangle$  cancel, allowing the use of Theorem 1. Rearranging, it is straightforward to then arrive at Eq. (51).  $\square$

We can then combine Theorem 3 and Lemma 3, to arrive at the simplified result

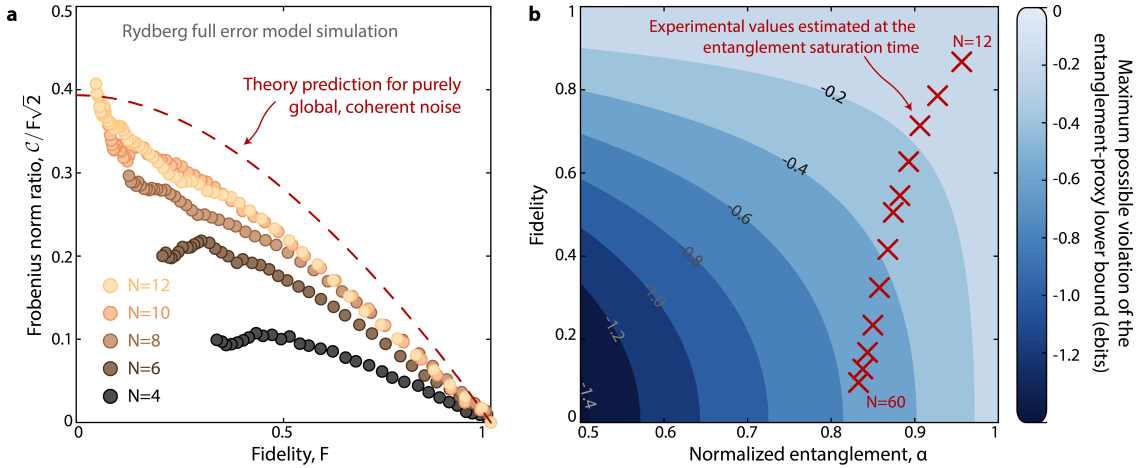
$$\mathcal{E}_N(\hat{\rho}) \geq \mathcal{E}_N(|\psi\rangle) + \log_2(F) + 2 \log_2 \left( 1 - \sqrt{\left( \frac{2}{\sqrt{3 + 2F^2} - F^4} - 1 \right) \frac{1 - \alpha}{\alpha}} \right) \quad (57)$$

$$\approx \begin{cases} \mathcal{E}_N(|\psi\rangle) + \log_2(F) & F \approx 1 \\ \mathcal{E}_N(|\psi\rangle) + \log_2(F) + 2 \log_2 \left( 1 - 0.39 \sqrt{\frac{1 - \alpha}{\alpha}} \right) & F \ll 1 \end{cases} \quad (58)$$

which shows that our entanglement-proxy is robust in the presence of quasistatic parameter fluctuations, up to a  $\mathcal{O}(1)$  deviation which decreases with the normalized entanglement  $\alpha \equiv 2^{\mathcal{E}_N(|\psi\rangle)}/D_A$ .

More concretely, let us assume a normalized entanglement of  $\alpha = 0.8$  (as is approximately the expectation at  $N = 60$  and  $t = 14.3$  cycles, see Supp. Fig. 16). In that case, Eq. (58) reduces to

$$\mathcal{E}_N(\hat{\rho}) \gtrsim \mathcal{E}_P(\hat{\rho}) - 0.6. \quad (59)$$



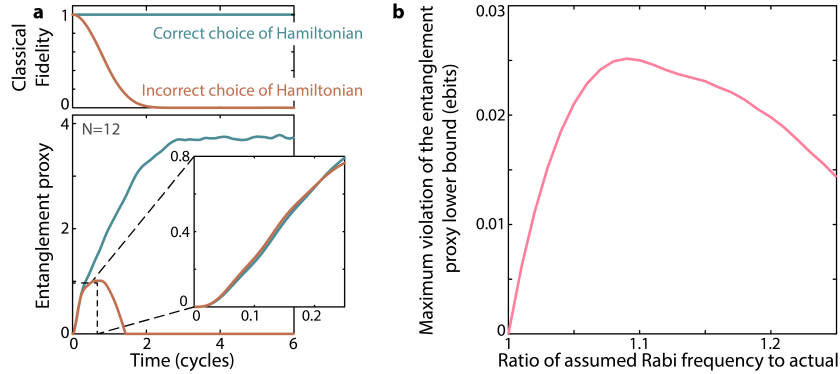
**Supp. Fig. 18 | Bounding the entanglement-proxy in the presence of errors.** **a.** We consider full error model simulations of the Rydberg dynamics studied in the main text, i.e. simultaneously under both global and local errors which each can be either Markovian or non-Markovian. We study the ratio of the commutator norm to the fidelity versus fidelity, truncating to only show data for which  $F > 1/D_A$ , where  $D_A$  is the half chain Hilbert space dimension. The data appears to converge as a function of system size, approaching the theory prediction obtained for the case of purely global, coherent Hamiltonian fluctuations, see Eq. (57). **b.** Here we plot the expected maximum violation of our entanglement-proxy lower bound as a function of the fidelity and the normalized entanglement; more specifically, we plot the correction term from Eq. (57). We find that in the reasonable parameter regime shown, the violation is always  $\lesssim 1$  ebit. From analysis shown in Supp. Fig. 16, we plot the expected experimental values from  $N = 12$  to 60, and find the maximum violation appears to saturate around  $\approx 0.6$  ebits. Together these results imply that our entanglement-proxy is robust up to  $\mathcal{O}(1)$  corrections in the face of realistic noise.



In other words, we expect the maximum possible violation of our bound is at most around half an ebit of entanglement. We emphasize that this does not mean the bound is in fact violated by this amount - only that it *potentially could be*.

In order to validate that the coefficient to the correction term is small, in Supp. Fig. 18a we show  $\mathcal{C}/(F\sqrt{2})$  as a function of  $F$  for various system sizes, calculated with our full Rydberg error model (which includes both Markovian and non-Markovian terms, and both global and local errors). Despite the error model including more noise terms than just global parameter fluctuations, we see curves as a function of  $N$  increase but appear to converge near the prediction of Eq. (51). Then, assuming the prediction of Eq. (51), we show the maximum possible violation of the lower bound provided by  $\mathcal{E}_P$  (i.e. the correction term in Eq. (57)) as a function of normalized entanglement and fidelity. We further include estimated experimental values up to  $N = 60$  (Supp. Fig. 18b). We find the experimental values appears to saturate around a maximum possible violation of  $\approx 0.6$  ebits.

**Systematic miscalibration errors**— Next, we comment on the case where the target state  $|\psi\rangle$  is set incorrectly, for instance because of a miscalibration of the Hamiltonian parameters. While this type of error can lead to a violation of our lower bound, the violation is small and the period for which it occurs is ephemeral, because the incorrect choice of  $|\psi\rangle$  will lead to a fidelity decay which is stronger than any typical increase in pure state entanglement. To showcase this, we consider quench dynamics resulting in linear growth of negativity followed by saturation (Supp. Fig. 19). We study the two entanglement regimes (saturated or growing linearly) separately.



**Supp. Fig. 19 | Sensitivity to incorrect pure state.** **a.** We study a possible error source in our entanglement proxy, that the pure state negativity could be assumed incorrectly, for instance if the wrong Hamiltonian parameters are chosen in simulation. We consider error-free quench evolution with  $\Omega/2\pi = 6.9$  MHz, leading to normal linear growth and then saturation of pure state negativity (which is equal to the entanglement proxy, as  $F = 1$ ). If we calculated the entanglement proxy assuming the Rabi frequency was instead  $\Omega/2\pi = 7.2$  MHz, then at very early times the calculated entanglement proxy can actually slightly exceed the true entanglement, showcasing an adversarial example where our entanglement proxy fails to be a lower bound (inset). However, loss of fidelity due to the Hamiltonian parameter mismatch will eventually dominate (top), restoring the validity of the lower bound for large system sizes or late evolution times. **b.** Even for a wide range of possible Rabi frequency miscalibrations (represented by the ratio of the miscalibrated Rabi frequency to the actual Rabi frequency), the maximum possible violation of the proxy lower bound is only  $\sim 0.02$  ebits.

First, we consider the case where the entanglement has saturated. We imagine the true evolution was performed with a Rabi frequency  $\Omega_0$ , resulting in a state  $|\psi_0\rangle$  at time  $t$ . However, we consider the situation where the parameters in the Hamiltonian have been mischosen, for instance with Rabi frequency  $\Omega_1 = \Omega_0 + \omega$ . From Theorem 1 we know the resulting fidelity overlap between the two states is given by

$$F = |\langle\psi_0|\psi_1\rangle|^2 = 2^{-N\lambda\omega^2 t^2/2}. \quad (60)$$

Because we are studying the entanglement saturated regime, the pure state negativities of  $|\psi_0\rangle$  and  $|\psi_1\rangle$  are approximately equal, but the mixed state negativity will now be calculated as

$$\mathcal{E}_P(|\psi_1\rangle) = \mathcal{E}_N(|\psi_1\rangle) + \log_2(F) \quad (61)$$

$$\approx \mathcal{E}_P(|\psi_0\rangle) - N\lambda\omega^2 t^2/2. \quad (62)$$

Therefore, any change in the assumed Rabi frequency will lead to a quadratic loss in the mixed state entanglement proxy, and the proxy will remain a lower bound.

The situation is largely the same if we now work in the linear growth regime, where we assume the pure state entanglement grows as  $\mathcal{E}(|\psi_0\rangle) = \alpha\Omega_0 t$ , and  $\mathcal{E}(|\psi_1\rangle) = \alpha\Omega_1 t$ , where  $\alpha$  is the entanglement velocity related to Lieb-Robinson bounds. In this case, the entanglement proxy can be written as

$$\mathcal{E}_P(|\psi_1\rangle) = \mathcal{E}_N(|\psi_1\rangle) + \log_2(F) \quad (63)$$

$$= \alpha\Omega_1 t - N\lambda\omega^2 t^2/2 \quad (64)$$

$$= \mathcal{E}_P(|\psi_0\rangle) + \alpha\omega t - N\lambda\omega^2 t^2/2 \quad (65)$$

$$= \mathcal{E}_P(|\psi_0\rangle) + \omega t(\alpha - N\lambda\omega t/2). \quad (66)$$

Thus, as long as  $N\lambda\omega t/2 > \alpha$ , the erroneous choice of target state will still result in a lower overall mixed state entanglement proxy, thus maintaining the lower bound. This will then become increasingly tight for larger system sizes and later evolution times. In Supp. Fig. 19 we show simulations demonstrating this, indicating that Hamiltonian parameter miscalibration should not substantially affect the veracity of our mixed state entanglement lower bound past very early times.

**Incoherent local errors**— Next, we turn to prove our entanglement proxy is a valid lower bound for the case of random unitary circuits (RUCs) undergoing realistic noise, namely local depolarization error. To do so, we will introduce a new theorem, which bounds the violation in our entanglement proxy by a value related to the fidelity overlap with eigenstates of  $\hat{\rho}$ , and then will show that this correction term is small.

We proceed first by proving a more general Lemma which will assist us; the proof has similarities to the proof of Theorem 3, but we will present it again here for completeness.

**Lemma 4.** *For any mixed state  $\hat{\rho}$  and pure state  $|\psi\rangle$*

$$\mathcal{E}_N(\hat{\rho}) \geq \mathcal{E}_N(|\psi\rangle) + \log_2(F) + 2\log_2(1 - \sqrt{(1-F)(1-\alpha)/(F\alpha)}), \quad (67)$$

where  $F \equiv \langle\psi|\hat{\rho}|\psi\rangle$  is their fidelity,  $\alpha \equiv 2^{\mathcal{E}_N(|\psi\rangle)}/D_A \in [0, 1]$  is a normalized entanglement of  $|\psi\rangle$ , and  $D_A$  is the Hilbert space dimension of the subsystem  $A$  (assumed to be smaller than its complement).

*Proof.* In deriving Theorem 2, we utilized the fact that when  $|\psi\rangle$  is an eigenstate of  $\hat{\rho}$ , the remainder state  $\hat{\rho}_\perp$  is positive semi-definite, which allowed us to bound  $\langle\Phi|\hat{\rho}_\perp|\Phi\rangle \geq 0$ . To show Lemma 4, we use constraints that arise from the fact that  $\hat{\rho}$  is positive semidefinite. We write  $\hat{\rho}$  in a suitable basis as:

$$\hat{\rho} = \begin{pmatrix} F & B^\dagger \\ B & \hat{\rho}' \end{pmatrix}, \quad (68)$$

where the first row/column denotes the basis vector  $|\psi\rangle$ , and the second denotes the subspace perpendicular to  $|\psi\rangle$ . With  $P_\perp$  the projector onto this subspace, the matrix  $\hat{\rho}' \equiv P_\perp \hat{\rho} P_\perp$  is positive semi-definite. The off-diagonal element  $B \equiv P_\perp \hat{\rho} |\psi\rangle$  is a  $(d-1) \times 1$  vector, which we take to be an unnormalized vector  $|B\rangle$  in the full Hilbert space (with first entry equal to zero).

Using Eq. (30), the negativity can be obtained from

$$\|\hat{\rho}^{TA}\|_1 \geq \langle\Phi|\hat{\rho}|\Phi\rangle = F|\langle\psi|\Phi\rangle|^2 + \langle\Phi|\hat{\rho}'|\Phi\rangle + \langle\Phi|\psi\rangle\langle B|\Phi\rangle + \langle\Phi|B\rangle\langle\psi|\Phi\rangle. \quad (69)$$

Only the last two terms above may be negative. We bound their magnitude using the *Schur complement* condition for positive semi-definite matrices. With our decomposition, the fact that  $\rho$  is positive semi-definite implies that the Schur complement  $\hat{\rho}' - F^{-1}|B\rangle\langle B|$  is positive semi-definite. We conclude that

$$\langle\Phi|B\rangle\langle B|\Phi\rangle \leq F\langle\Phi|\hat{\rho}'|\Phi\rangle, \quad (70)$$

which allows us to further conclude that

$$\langle\Phi|\hat{\rho}|\Phi\rangle \geq (\sqrt{F}|\langle\Phi|\psi\rangle| - \sqrt{\langle\Phi|\hat{\rho}'|\Phi\rangle})^2. \quad (71)$$

We then use the fact that  $|\Phi\rangle/\sqrt{D_A}$  is a normalized state. Since  $|\langle\Phi|\psi\rangle|^2 = \alpha D_A = \|\langle\psi|\hat{\rho}^{TA}\|_1$ ,  $\text{tr}\hat{\rho}' = 1 - F$  and  $\langle\psi|\hat{\rho}'|\psi\rangle = 0$ , we conclude that

$$\langle\Phi|\hat{\rho}'|\Phi\rangle \leq (1-F)(1-\alpha)D_A. \quad (72)$$

Then when  $F \geq 1 - \alpha$ , we obtain the claimed Eq. (67).  $\square$

We then simply generalize Lemma 4 to the case where  $|\psi\rangle$  has fidelity overlap with a particular eigenstate of  $\hat{\rho}$

**Theorem 4** (Eigenstate-fidelity based bound). *For any mixed state  $\hat{\rho}$  and pure state  $|\psi\rangle$  which has a fidelity  $f \equiv |\langle\psi|\lambda\rangle|^2$  with an eigenstate  $|\lambda\rangle$  of  $\hat{\rho}$ , the logarithmic negativities  $\mathcal{E}_N(\hat{\rho})$  and  $\mathcal{E}_N(|\psi\rangle)$  are related by*

$$\mathcal{E}_N(\hat{\rho}) \geq \mathcal{E}_N(|\psi\rangle) + \log_2(F) + 2 \log_2(1 - \sqrt{(1-f)(1-\alpha)/(f\alpha)}), \quad (73)$$

where  $F \equiv \langle\psi|\hat{\rho}|\psi\rangle$  is the fidelity,  $\alpha \equiv 2^{\mathcal{E}_N(|\psi\rangle)}/D_A \in [0, 1]$  is a normalized entanglement of  $|\psi\rangle$ , and  $D_A$  is the Hilbert space dimension of the subsystem  $A$  (assumed to be smaller than its complement).

*Proof.* We first write  $\hat{\rho} = F_t|v\rangle\langle v| + (1-F_t)\hat{\rho}_\perp$ , where  $|v\rangle$  is an eigenstate of  $\hat{\rho}$  with eigenvalue  $F_t$ , and  $\hat{\rho}_\perp$  is the state projected onto the orthogonal subspace to  $|v\rangle$ . We write the state  $|\psi\rangle$  that we benchmark against  $\hat{\rho}$  in terms of  $|v\rangle$  as  $|\psi\rangle = \sqrt{f}|v\rangle + \sqrt{1-f}|\psi_\perp\rangle$ . Then writing  $\hat{\rho}$  in the same basis as Eq. (68) gives

$$\hat{\rho} = \begin{pmatrix} F_t f & F_t \sqrt{f(1-f)}|\psi\rangle\langle\psi_\perp| \\ F_t \sqrt{f(1-f)}|\psi_\perp\rangle\langle\psi| & \hat{\rho}_\perp + F_t(1-f)|\psi_\perp\rangle\langle\psi_\perp| \end{pmatrix}, \quad (74)$$

As with the above, the only negative contribution to  $\langle\Phi|\hat{\rho}|\Phi\rangle$  arises from the off-diagonal terms of Eq. (74). This allows us to improve our bound to the claimed Theorem 4.  $\square$

We now show that under incoherent local errors, the correction term presented in Theorem 4 is small. We study the case of random unitary circuits (RUCs) in the presence of local depolarization in order to treat the problem analytically and make connection to the digital quantum circuits we compare against in the main text.

We study the fidelity and purity and bound the fidelity  $f \equiv |\langle\psi(t)|v\rangle|^2$  between the ideal evolved state  $|\psi(t)\rangle$  with the largest eigenstate  $|v\rangle$  of the noisy evolved state  $\hat{\rho}(t)$ , by computing both the purity and fidelity of  $|\psi(t)\rangle$  and  $\hat{\rho}(t)$ . The relationship between these quantities in fact bounds the fidelity of  $|\psi(t)\rangle$  to the largest eigenstate of  $\hat{\rho}$ , allowing us to apply Theorem 4 to evaluate the robustness of our negativity proxy.

Physically, our result means that the state  $\hat{\rho}$  is a mixture of a particular pure state  $|v_1\rangle$  with large population  $F = \lambda_1$  and many other orthogonal states with smaller populations. For a circuit with spacetime volume  $V = Nt$ , the largest eigenvalue decays exponentially  $\lambda_1 \approx (1-p)^V$ , yet remains larger than all other eigenvalues. The other states correspond to trajectories with errors. The average number of errors is  $k = pV$ , and there are roughly  $\frac{V!}{k!(V-k)!}$  different spacetime locations of  $k$  errors. When the number of trajectories are much less than the Hilbert space dimension  $D$ , we expect the resultant wavefunctions resultant to be orthogonal to one another. This is true if the  $k$  errors are sparsely spread over the system in spacetime (such that their separations are sufficiently large) and the effects of errors are scrambled quickly. When  $p$  is sufficiently small and when the quantum dynamics is sufficiently chaotic, these assumptions are expected to hold.

We proceed to justify this intuition by averaging over random circuits. At every time step, we apply random unitary two-qubit gates in a brickwork circuit geometry, and subject it to local depolarization, which locally depolarizes the state  $\rho$  at every site with probability  $p$ , equivalent to applying the local channel  $\rho \mapsto (1-p)\rho + p(I_j/d) \otimes \text{tr}_j(\rho)$  independently to every site  $j$ , where  $d = 2$  is the local Hilbert space dimension (Supp. Fig. 20a). For simplicity, we also take the initial state  $|\psi_0\rangle$  to be the product state  $|0\rangle^{\otimes N}$ . After averaging over RUCs, the purity  $P \equiv \text{tr}[\rho(t)^2]$  and fidelity  $F \equiv \langle\psi(t)|\rho(t)|\psi(t)\rangle$  can be expressed as the partition functions of two different Ising ferromagnets, with the Ising spins  $\sigma = \pm 1$  respectively denoting the local identity and swap permutations (Supp. Fig. 20b).

The derivation of the mapping between RUCs and classical spin models is described in detail in many works including Refs.<sup>45,46</sup>. Here we simply quote the result:

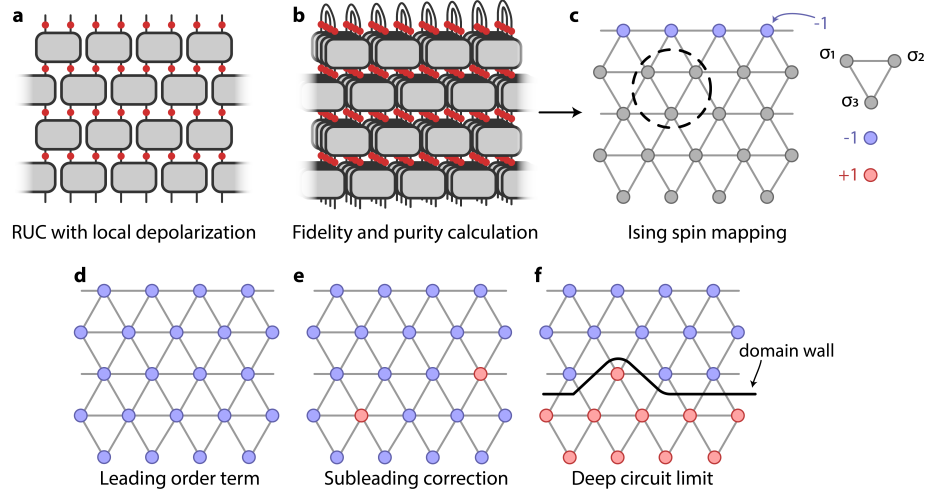
$$P = \sum_{\{\sigma\}} \prod_{\nabla} W_P(\sigma_1, \sigma_2, \sigma_3; p), \quad (75)$$

$$F = \sum_{\{\sigma\}} \prod_{\nabla} W_F(\sigma_1, \sigma_2, \sigma_3; p), \quad (76)$$

where the summation is over all possible Ising configurations of classical spin variables  $\sigma = \pm 1$  and the product is over all downward-facing triangles in the resulting triangular lattice (Supp. Fig. 20c).  $W_{P/F}(\sigma_1, \sigma_2, \sigma_3; p)$  act as local ‘‘Boltzmann weights,’’ parameterized by the error probability  $p$ . Their values are enumerated in Supp. Table II. The

$\sigma_1, \sigma_2, \sigma_3$	$W_P(\sigma_1, \sigma_2, \sigma_3; p)$	$W_F(\sigma_1, \sigma_2, \sigma_3; p)$
+, +, +	1	1
+, +, -	0	0
+, -, +	$\frac{d}{d^2+1} + \frac{2d}{d^4-1}p + \frac{4d}{3(d^4-1)}p^2$	$\frac{d}{d^2+1} + \frac{d}{d^4-1}p$
+, -, -	$\frac{d}{d^2+1} + \frac{-2d^3}{d^4-1}p + \frac{4d^3}{3(d^4-1)}p^2$	$\frac{d}{d^2+1} + \frac{-d^3}{d^4-1}p$
-, +, +	$\frac{d}{d^2+1} + \frac{2d}{d^4-1}p + \frac{4d}{3(d^4-1)}p^2$	$\frac{d}{d^2+1} + \frac{d}{d^4-1}p$
-, +, -	$\frac{d}{d^2+1} + \frac{-2d^3}{d^4-1}p + \frac{4d^3}{3(d^4-1)}p^2$	$\frac{d}{d^2+1} + \frac{-d^3}{d^4-1}p$
-, -, +	$\frac{4d^2}{d^4-1}p + \frac{-20d^2}{3(d^4-1)}p^2 + \frac{16d^2}{3(d^4-1)}p^3 + \frac{-16d^2}{9(d^4-1)}p^4$	$\frac{2d^2}{d^4-1}p + \frac{-d^2}{d^4-1}p^2$
-, -, -	$1 + \frac{-4d^4}{d^4-1}p + \frac{20d^4}{3(d^4-1)}p^2 + \frac{-16d^4}{3(d^4-1)}p^3 + \frac{16d^4}{9(d^4-1)}p^4$	$1 + \frac{-2d^4}{d^4-1}p + \frac{d^4}{d^4-1}p^2$

**Supp. Table II** | Ising Boltzmann weights of local downward-facing triangle configurations for the effective statistical mechanical models. Their partition functions equal the purity  $P$  [Eq. (75)] and fidelity  $F$  [Eq. (76)] of random unitary circuits (RUCs) with local depolarization. The geometric arrangement of  $\sigma_1, \sigma_2, \sigma_3$  is indicated in Fig. 20c. In this work, we consider RUCs of qubits, with local Hilbert space dimension  $d = 2$ , and only consider the  $\mathcal{O}(1)$  and  $\mathcal{O}(p)$  terms. In this table, however, we state the full expressions in terms of general  $d$  for future reference.

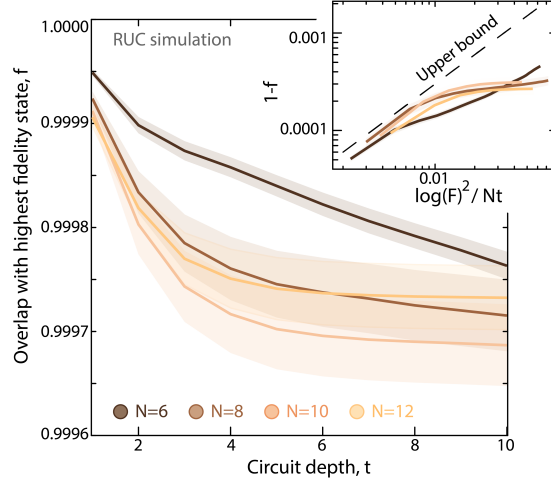


**Supp. Fig. 20** | Purity and fidelity in random unitary circuits (RUCs) with local depolarization. **a.** We consider a brickwork RUC, where local depolarization with strength  $p$  is applied after every layer and at every site (red dots). **b.** Typical values of purity or fidelity can be computed by random circuit averages. In both quantities, swap boundary conditions are applied on four copies of the circuit, representing two copies of the states. For purity, dephasing is applied to both copies (depicted here), while for fidelity, dephasing is only applied to one copy. **c.** After averaging over random circuits, the purity and fidelity can be expressed as the partition functions of effective statistical mechanical Ising models, with three-spin Boltzmann weights on every downward-facing triangle (listed in Table II), the top boundary fixed to be ‘-’ (blue circles) and the bottom boundary free. **d.** The dominant Ising spin configuration for both quantities is where all Ising spins are in the ‘-’ (swap) state (blue circles). **e.** Subleading corrections arise from configurations with sparse, isolated “bubbles” of ‘+’ Ising spins (red circles). **f.** For sufficiently deep circuits, the dominant configurations are those with a boundary domain of ‘-’ spins and a bulk domain of ‘+’ spins. However, this is only relevant in a regime where the fidelity is trivially small.

Ising spins at the bottom are determined by the initial condition and are effectively unconstrained (open boundary condition), while the purity and fidelity expressions result in an additional layer of Ising spins at the top, which are pinned to be the swap permutation ‘-’.

When  $p = 0$ , notice that  $W_{P/F}(-, -, +) = 0$ . Combined with the fact that all top boundary spins have  $\sigma = -$ , the only non-vanishing contribution to  $P$  or  $F$  is the one where all  $\sigma = -$  (Supp. Fig. 20d). In this case,  $P = \prod_{\sigma} W_P(-, -, -) = 1$  and similarly  $F = 1$ . This is expected since in the absence of errors,  $\rho$  is a pure state with perfect fidelity.

Now consider small, nonzero  $p$ .  $P$  and  $F$  have contributions from configurations where some of the spin variables are flipped to the ‘+’ state, with weight  $\mathcal{O}(p)$  (Supp. Fig. 20e). We call such spin flips “bubbles,” because such bubbles are penalized from growing by an effective line-tension term, which weights larger bubbles by a factor of  $(2/5)^l$ , where  $l$  is the length of the boundary of the bubble. These small bubbles will effectively have renormalized weights, but will remain  $\mathcal{O}(p)$  and whose precise values are not consequential to our conclusions. Therefore, in this regime, the



**Supp. Fig. 21 | Bounding the entanglement-proxy in the presence of errors.** We consider the case of incoherent local errors, specifically for random unitary circuits undergoing local depolarization. The overlap,  $f$ , between the target pure state and the highest fidelity state to  $\hat{\rho}$  remains near unity. Further,  $1 - f$  is linearly bounded by the quantity  $\log(F)^2/Nt$  with a small proportionality constant 0.03 (inset), consistent with our RUC analysis leading up to Eq. (82), and giving the overall bound Eq. (83).

number of such spin flips will be rare and sparse. To this end, we approximate  $P$  and  $F$  by simply summing over the contributions from sparse, disjoint bubbles. However, this approximation breaks down at larger values of  $p$ , discussed below. Within this approximation,

$$P \approx \left(1 - \frac{4d^2}{d^4 - 1}p + \mathcal{O}(p^2)\right)^V \exp(A_P V) \equiv \exp(-B_P V), \quad (77)$$

$$F \approx \left(1 - \frac{2d^2}{d^4 - 1}p + \mathcal{O}(p^2)\right)^V \exp(A_F V) \equiv \exp(-B_F V), \quad (78)$$

where  $V$  is the spacetime volume (number of Ising spins) and  $A_F$  and  $A_P$  are the renormalized relative weights of a single bubble.

We are interested in the ratio  $P/F^2 \approx e^{(2B_F - B_P)V} \geq 1$ . When  $p \ll 1$ ,

$$\begin{aligned} 2B_F - B_P &= -2A_F + A_P - 2 \ln \left(1 - \frac{2d^2}{d^4 - 1}p + \mathcal{O}(p^2)\right) + \ln \left(1 - \frac{4d^2}{d^4 - 1}p + \mathcal{O}(p^2)\right) \\ &= A_P - 2A_F + \mathcal{O}(p^2). \end{aligned} \quad (79)$$

Because each bubble already has  $\mathcal{O}(p)$  weight, the leading order contribution to  $A_P - 2A_F$  is determined by that of  $W_P(-, -, +)$  and  $W_F(-, -, +)$ , which have ratio precisely 2. Therefore, we conclude that all  $\mathcal{O}(p)$  terms vanish and

$$P/F^2 = e^{Cp^2V}, \quad (80)$$

for some  $\mathcal{O}(1)$  constant  $C$ .

In order to translate this into a bound on the eigenstate-fidelity  $f$ , we denote the fidelities between  $|\psi\rangle$  and the eigenstates  $|v_i\rangle$  of  $\hat{\rho}$  as  $\mu_i \equiv |\langle v_i | \psi \rangle|^2$ , with  $f \equiv \mu_1$ . With  $\lambda_i$  the eigenvalues of  $\hat{\rho}$ , we have  $F = \sum_i \lambda_i \mu_i$  and  $P = \sum_i \lambda_i^2$ . The Cauchy-Schwarz inequality implies that

$$\mu_1^2 + (1 - \mu_1)^2 \geq \sum_i \mu_i^2 \geq F^2/P \Rightarrow 1 - \mu_1 \leq \frac{P/F^2 - 1}{2P/F^2}. \quad (81)$$

Using Eq. (80), we obtain

$$1 - f \equiv 1 - \mu_1 \leq (1 - \exp(-Cp^2V))/2 = \mathcal{O}(p^2V). \quad (82)$$

Using Theorem 4, we obtain

$$\mathcal{E}_N(\hat{\rho}) \gtrsim \mathcal{E}_N(|\psi\rangle) + \log_2(F) + 2 \log_2 \left(1 - \sqrt{\frac{1 - \alpha}{\alpha}} C_2 \frac{\ln(F)^2}{V}\right), \quad (83)$$

for another  $\mathcal{O}(1)$  constant  $C_2$ , indicating that the entanglement-proxy  $E_P$  is valid up to a small correction which is negligible as long as  $F \gg \exp(-\sqrt{V}) = \exp(-\sqrt{Nt}) \sim \exp(-N)$ , i.e.  $F$  is not trivially small. We give numerical support for this in Supp. Fig. 21.

Finally, we remark that additional configurations may come to dominate the partition function sums Eqs. (75), (76) when the circuit is sufficiently deep. This is easiest to understand in the statistical mechanical picture. For both the purity and fidelity, the bulk has a small preference for the '+' state, because depolarization breaks the Ising symmetry and favors the identity element. Meanwhile, the boundary condition (associated with evaluating purity or fidelity) pins the boundary to the '-' state, acting as a strong boundary field. When the circuit is shallow, the boundary term dominates, giving rise to the global '-' domain assumed above. However, when the circuit is sufficiently deep, it becomes energetically favorable for a large domain wall to form, dividing the spins into a '-' domain near the boundary, and a '+' domain in the bulk (Supp. Fig. 20f), at the expense of a domain wall with free-energy scaling with the spatial size. In our above calculations, we neglect this phenomenon because it is only relevant when the fidelity is trivially small,  $\mathcal{O}(D^{-1})$ , where  $D$  is the total Hilbert space dimension. In this regime, the purity and fidelity decouple from the bulk dynamics and assume this constant value which arises from the boundary Ising '-' domain.

### G3. Negativity of truncated MPS states

Finally, we investigate an interesting observation in the main text, that at  $N = 60$  the entanglement-proxy found for the experiment in Fig. 4 is  $\sim 3$  ebits higher than  $\log_2(\chi^*)$  in Fig. 5. In other words, the fidelity-equivalent MPS simulation is less entangled than the experiment. Here we study the fidelity and entanglement (negativity) of truncated pure state MPS representations to explain this behavior.

For a theoretical treatment, we consider a simplified setup where we start with an ideal pure state and its equal bipartite Schmidt decomposition  $|\psi\rangle = \sum_{j=1}^{D_A} s_j |a_j\rangle |b_j\rangle$  where  $D_A$  is the Hilbert space dimension of the half system and the Schmidt values  $s_j$  are assumed to be in descending order. Then the Schmidt rank is truncated to a maximum bond dimension  $\chi \leq D_A$ . Denote the truncated state and its Schmidt decomposition as  $|\psi_\chi\rangle = \sum_{j=1}^{\chi} s'_j |a_j\rangle |b_j\rangle$ , with  $s'_j = s_j / \sqrt{\sum_{k=1}^{\chi} s_k^2}$  (normalization condition). It is straightforward to see that

$$F = |\langle \psi_\chi | \psi \rangle|^2 = \sum_{j=1}^{\chi} s_j^2. \quad (84)$$

To derive the entanglement negativity, the eigenvalues  $\lambda$  of the partial transposed state  $|\psi\rangle\langle\psi|^{T_A}$  is given by  $s_j^2$  or  $\pm s_i s_j$  as mentioned above. This gives

$$\| |\psi_\chi\rangle\langle\psi_\chi|^{T_A} \|_1 = \frac{\left(\sum_{j=1}^{\chi} s_j\right)^2}{\sum_{j=1}^{\chi} s_j^2}, \quad (85)$$

and therefore

$$\mathcal{E}(|\psi_\chi\rangle) = \log_2 \| |\psi_\chi\rangle\langle\psi_\chi|^{T_A} \|_1 = 2 \log_2 \sum_{j=1}^{\chi} s_j - \log_2 \sum_{j=1}^{\chi} s_j^2. \quad (86)$$

For generic states with unknown  $s_j$ , it is hard to draw any meaningful relationship between fidelity and entanglement negativity. Therefore, we consider two cases: maximally entangled states and Haar random states.

**Maximally entangled states**—For a maximally entangled state, the Schmidt values are given by  $s_j = 1/\sqrt{D_A}$ . It follows directly that

$$F = \frac{\chi}{D_A}; \quad (87)$$

$$\mathcal{E}(|\psi_\chi\rangle) = \log_2 \chi = \log_2 D_A + \log_2 F, \quad (88)$$

with  $\log_2 D_A$  the log negativity of the ideal maximally entangled pure state. Thus, the log negativity for a fidelity-equivalent truncated maximally entangled state is equal to that of an isotropic mixed state.

**Haar random states**—The case for Haar random states is more interesting, as such states arise from random unitary circuit evolution, and are closer to the output of our Rydberg experiment as well. As mentioned previously, the Schmidt values of Haar random states follow the *quarter-circle law* as

$$P(\tilde{s}) = \frac{1}{\pi} \sqrt{4 - \tilde{s}^2}, \quad (89)$$

where  $\tilde{s} \equiv \sqrt{D_A} s$ . Setting a maximum Schmidt rank can be translated to setting a minimum cutoff  $s_{\min}$  (and the corresponding  $0 \leq \tilde{s}_{\min} \leq 2$ ) of the Schmidt values (keeping only Schmidt values above  $s_{\min}$ ), where the relation between  $\chi$  and  $\tilde{s}_{\min}$  is given by

$$\chi = D_A \int_{\tilde{s}_{\min}}^2 P(\tilde{s}) d\tilde{s} = D_A \left[ \frac{4}{\pi} \sin^{-1} \left( \frac{\sqrt{2 - \tilde{s}_{\min}}}{2} \right) - \frac{1}{2\pi} \tilde{s}_{\min} \sqrt{4 - \tilde{s}_{\min}^2} \right]. \quad (90)$$

Depending on the cutoff,

$$\sum_{j=1}^{\chi(\tilde{s}_{\min})} s_j = \sqrt{D_A} \int_{\tilde{s}_{\min}}^2 \tilde{s} P(\tilde{s}) d\tilde{s} = \frac{\sqrt{D_A}}{3\pi} (4 - \tilde{s}_{\min}^2)^{\frac{3}{2}}; \quad (91)$$

$$\sum_{j=1}^{\chi(\tilde{s}_{\min})} s_j^2 = \int_{\tilde{s}_{\min}}^2 \tilde{s}^2 P(\tilde{s}) d\tilde{s} = \frac{4}{\pi} \sin^{-1} \left( \frac{\sqrt{2 - \tilde{s}_{\min}}}{2} \right) + \frac{1}{4\pi} \tilde{s}_{\min} (2 - \tilde{s}_{\min}^2) \sqrt{4 - \tilde{s}_{\min}^2}. \quad (92)$$

Therefore, the fidelity and log negativity can be written in parametric forms as

$$F(\tilde{s}_{\min}) = \frac{4}{\pi} \sin^{-1} \left( \frac{\sqrt{2 - \tilde{s}_{\min}}}{2} \right) + \frac{1}{4\pi} \tilde{s}_{\min} (2 - \tilde{s}_{\min}^2) \sqrt{4 - \tilde{s}_{\min}^2}; \quad (93)$$

$$\mathcal{E}(|\psi_{\chi(\tilde{s}_{\min})}\rangle) = \log_2 D_A + \log_2 \frac{(4 - \tilde{s}_{\min}^2)^3}{9\pi^2} - \log_2 F(\tilde{s}_{\min}). \quad (94)$$

It is easy to verify that when  $\tilde{s}_{\min} = 0$ , the equations reduce to the case of ideal Haar random states, with  $\chi = D_A$ ,  $F = 1$ , and  $\mathcal{E} = \log_2 D_A + \log_2 64/9\pi^2$ , where  $64/9\pi^2 \approx 0.72$  is the Page correction<sup>1</sup>.

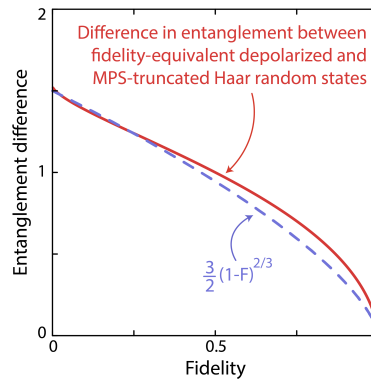
Importantly, we can then compare this prediction against the prediction for a depolarized Haar random state, where we use the first-order approximation from Eq. (36), in other words taking  $D_A \rightarrow \infty$ . At a given fidelity for each, the difference in estimated entanglement,  $\delta$ , is given as

$$\delta = \mathcal{E}_N(\hat{\rho}_{\text{Haar}}) - \mathcal{E}_N(|\psi_{\chi(\tilde{s}_{\min})}\rangle) \quad (95)$$

$$= 3 \log_2 \left( \frac{4}{4 - \tilde{s}_{\min}^2} \right) + 2 \log_2(F) \quad (96)$$

$$\approx \left( 2 + \log_2 \left( \frac{64}{9\pi^2} \right) \right) (1 - F)^{2/3} \approx \frac{3}{2} (1 - F)^{2/3} \quad (97)$$

In the last line we have performed series expansions of both  $F$  and  $\delta$  in terms of  $\tilde{s}_{\min}$  to arrive at a non-parametric form. In Supp. Fig. 22 we plot  $\delta$  (as well as the non-parametric approximation) as a function of  $F$ , which is always positive, indicating that for all  $F$ , the MPS representation of a truncated Haar random state is less entangled (by  $\mathcal{O}(1)$  ebits) than a equivalent-fidelity depolarized Haar random mixed state. Given that the experiment does not produce a perfect Haar random state, as well as the fact that a real MPS simulation performs many truncations between potentially correlated timesteps, these results are in good agreement with the experimentally found discrepancy of  $\sim 3$  ebits.



**Supp. Fig. 22 | Entanglement difference between truncated and depolarized Haar random states.** At fixed fidelity, MPS truncation and depolarization affect the entanglement content of a pure state differently. For a Haar random state, we can solve this relation analytically, finding the depolarized state has a higher entanglement (log negativity) content than the MPS-truncated state at equal fidelity.



## H. Mixed state entanglement estimation for quasi-adiabatic state preparation

The mixed state entanglement proxy we have introduced (Eq. 2 of the main text) is a general lower bound in generic situations for any target pure state of interest (see Section F). We have demonstrated its applicability in quantum quench dynamics, and we anticipate its direct applicability in a much broader range of quantum simulation experiments. To exemplify this, we consider in this section the estimation of mixed state entanglement during quasi-adiabatic state preparation and at quantum critical points.

To estimate the mixed state entanglement with our proxy, we require knowledge of the entanglement of the pure target state and the fidelity of experimental state to the pure target state. At quantum critical points, the scaling of entanglement is well studied<sup>47</sup>, and the pure state entanglement can be extrapolated from exact simulations at small system size.

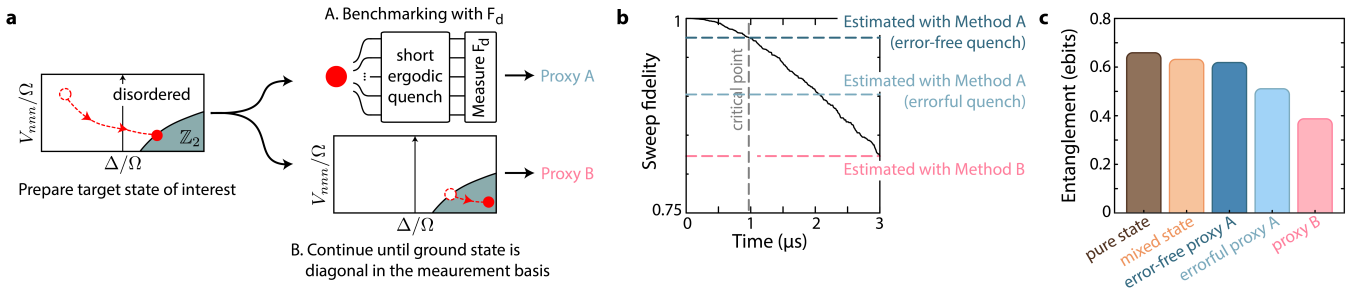
It remains then to estimate the fidelity, for which we identify two possible methods (Supp. Fig. 23a):

Method A) For particular states of interest,  $F_d$  may not work directly. However,  $F_d$  can still be applied by following the quasi-adiabatic sweep preparation with a short, ergodic quench<sup>2,3</sup>. Importantly,  $F_d$  is accurate for short quenches, meaning there is little fidelity loss in this process, and that tensor network methods can be used to generate exact classical simulation references even for large system sizes where critical states have area law entanglement with potentially logarithmic corrections<sup>47</sup>. Further, the performance of this method could be improved by calibrating the fidelity loss due to the quench via measurements at multiple times.

Method B) Continue the adiabatic sweep into a phase with a ground state for which the fidelity is easily measured (such as a product state). The fidelity of preparing the critical state can then be approximated as the fidelity at the end of the sweep.

We demonstrate both fidelity estimation methods using our standard error model (with no preparation or measurement errors, see Section J) to simulate a quasi-adiabatic sweep to prepare the ground state of the quantum phase transition between the disordered and classical  $\mathbb{Z}_2$  ordered phases in the 1D Rydberg phase diagram (Supp. Fig. 23b) with  $N = 11$  atoms. The sweep in question has a tangent detuning profile with maximum detuning range of  $\pm 2\pi \times 40$  MHz, and Rabi frequency of  $2\pi \times 5$  MHz. The sweep duration is  $3 \mu\text{s}$ .

For Method A, we simulate the preparation of the target state at the critical point. We then use the error model to simulate a quench for 4 cycles with a time-independent Hamiltonian with  $\Omega = 2\pi \times 6.9$  MHz, and with  $2\pi \times \pm 2$  MHz random on-site disorder drawn from a uniform distribution. Such a technique in a similar parameter regime has been well studied theoretically in our previous works<sup>2,3</sup>. For Method B, we measure the population of the classical  $\mathbb{Z}_2$  ordered state ( $|10101\dots 0101\rangle$ ) at the end of the entire quench. Similar techniques can also be used for estimating fidelities of two-dimensional arrays<sup>48,49</sup>. For both methods the mixed state entanglement proxy remains a lower bound to the actual mixed state negativity at the quantum phase transition, with Method A providing a tighter bound. For Method A we also provide the value assuming the measurement quench is error-free (or where the fidelity loss due to the quench is calibrated away), which closely estimates the actual mixed state negativity (Supp. Fig. 23c).



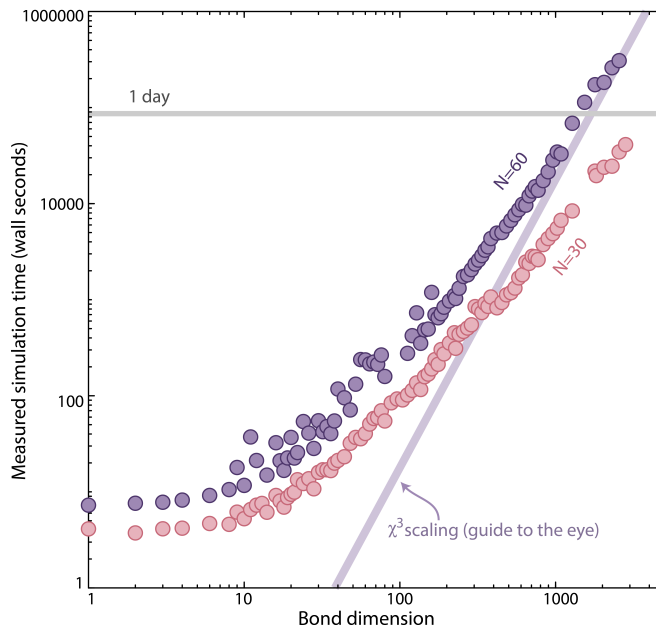
**Supp. Fig. 23 | Estimating mixed state entanglement of adiabatic state preparation.** **a.** The mixed state entanglement proxy presented in the main text can be applied to the case of adiabatic state preparation. For instance, a quasi-adiabatic sweep may be performed to prepare the critical state at a quantum phase transition. To estimate the fidelity, there are two possible methods: A) use the  $F_d$  formula presented here (and as further discussed in this context of *target state benchmarking* in Ref.<sup>2</sup>), or B) continue the sweep into a classical phase for which the ground state fidelity can be directly measured. **b, c.** We use our error model to simulate a sweep to the quantum phase transition between the disordered and  $\mathbb{Z}_2$  phases of the ground state Rydberg phase diagram in 1D. For method A, after reaching the critical point we performed a disordered quench for 4 cycles and calculate  $F_d$  to estimate the fidelity. For method B, we use the  $\mathbb{Z}_2$  population at the end of the entire sweep to estimate the fidelity. While both options work, for the particular sweep simulated here, method A yields a tighter lower bound on the actual mixed state negativity at the critical point, particularly if the measurement quench is error-free.

## I. Simulation methods

### II. Description of the Caltech computing cluster

This work extensively utilizes the Caltech Resnick High Performance Computing (HPC) Center, running CentOS Linux 7. In our work, each simulation is assigned to a single node, using 16 cores of Intel Cascade Lake CPU (at 2.2 or 2.4 GHz), with a total node memory of 386 GB (at 2933 MT/s). We note that many classical simulation times are reported as core-time, which we simplify as the wall time multiplied by the number of cores.

In terms of raw wall time, the largest simulation we performed (for  $N = 60$  and  $\chi = 3072$ ) took  $\sim 8.3$  days ( $\sim 11.1$  days if including the additional time points past the last experimental time in order to calculate  $p_{\text{avg}}$ ). Representative times to simulate up to the last experimental time point are shown as a function of bond dimension in Supp. Fig. 24; extrapolating the expected  $\mathcal{O}(\chi^3)$  scaling we predict simulation for  $N = 60$  and  $\chi^* = 3400$  would entail a simulation time of 11.3 wall-days (180 core-days, as reported in the main text). In total, we generated roughly  $\sim 100$  TB of data in order to create and store the MPS tensors at various bond dimensions used in this work.

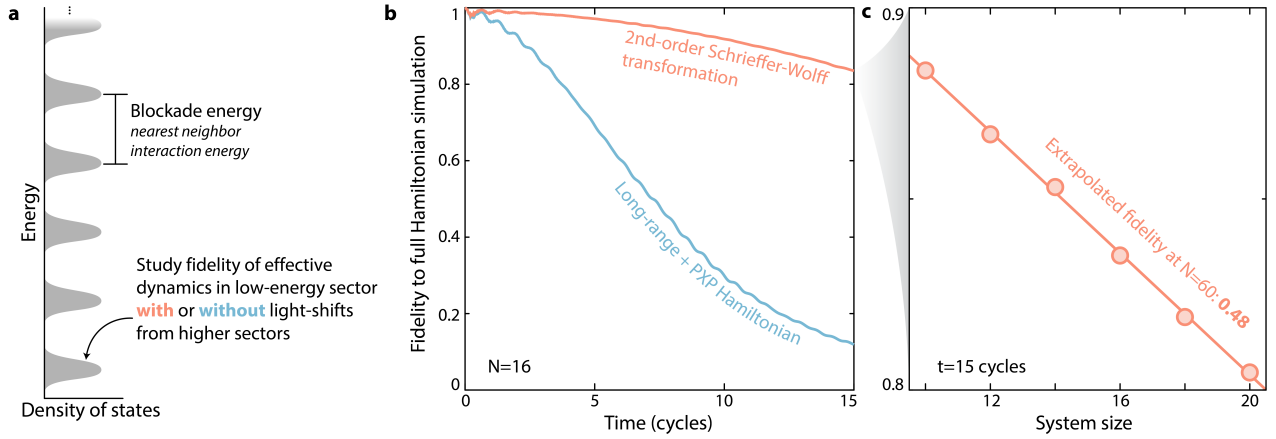


**Supp. Fig. 24 | Simulation time as a function of bond dimension.** After an initial overhead plateau, the Lightcone-MPS simulation time (measured in raw wall time on the Caltech central cluster) to reach the last experimental time increases as a function of  $N$ , approaching the expected asymptotic  $\chi^3$  scaling.

### II. Exact simulation with a Schrieffer-Wolff transformation

We first consider the possibility of exact brute-force simulation of our system. The current state-of-the-art in exact RUC simulation is 45 qubits<sup>50</sup>, while the state-of-the-art for continuous Hamiltonian time evolution is 38 qubits<sup>51</sup>. While clearly neither of these approaches would work natively for the largest experimental system sizes of  $N = 60$ , it is possible they could become applicable through a Schrieffer-Wolff transformation by rewriting the Rydberg Hamiltonian only in the blockaded subspace, applying appropriate level shifts to mimic the effect of higher blockade sectors, as described in the supplement of Ref.<sup>52</sup>. Because of the Rydberg blockade constraint, the effective system size is given by  $\tilde{N} = \log_2(\text{Fib}(N + 2)) \approx 0.7N$ , where Fib is the Fibonacci function. For the 60-atom system,  $\tilde{N} \sim 42$  qubits, so such brute-force techniques could allow for potentially pseudo-exact simulation.

However, here we discount such approaches for multiple reasons. First they both rely on specialized hardware – a supercomputing system<sup>50</sup> with 100s of TB of RAM, or tensor processing units<sup>51</sup> which are not widely available. Further, even with this hardware it is not clear that our system would be easily simulable, because of difficulties associated with treating the blockaded subspace. In particular, applying local gate operations becomes more complicated to implement efficiently, because the blockaded subspace is not factorizable into the product of local bases.



**Supp. Fig. 25 | Using an effective Hamiltonian simulation.** **a.** The Rydberg blockade constraint stratifies the energy levels of the Rydberg Hamiltonian, enabling approximating the dynamics in the full Hilbert space by an effective set of dynamics in the lowest energy Hilbert space (relative size of energy sectors are not to scale). **b.** While a lowest-order approximation (blue) quickly causes the fidelity of the reduced-Hilbert-space simulation to drop compared to simulation in the full space, adding the second-order term from a Schrieffer-Wolff Hamiltonian transformation<sup>52</sup> (orange) improves the fidelity. **c.** Even with the second-order correction, the simulation fidelity decreases as a function of system size. It is possible that numerically solving for higher order terms in the Schrieffer-Wolff transformation will improve this fidelity, but it likely can never achieve unity given the finite blockade violation probability (Supp. Fig. 6d).

Finally, applying this transformation is still an approximation to the true dynamics, for which it is not clear if the resulting fidelities will actually be competitive at these large system sizes and late evolution times. In Supp. Fig. 25 we test the accuracy of the Schrieffer-Wolff transformation at second order for system sizes up to  $N = 20$  ( $\tilde{N} \sim 14$ ), where we see the fidelity falls by the latest experimental time. Importantly, the fidelity decay increases with system size. A simple linear extrapolation implies a simulation accuracy of  $\sim 0.48$  for  $N = 60$ . Further, employing this approach will always have some fundamental inaccuracy as the blockade constraint is imperfect, and we expect the ideal target state to always have some finite population in the blockade-violating sectors, a symptom which grows stronger with increasing system size and evolution time (Supp. Fig. 6d). However, we do not discount the possibility that by adding successively higher-order terms in the Schrieffer-Wolff transformation, the accuracy of this method could be at least improved.

We strongly emphasize that storing even a single copy of the wavefunction at  $\tilde{N} \sim 42$  still requires  $\sim 65$  TB of memory, and that  $\tilde{N} \sim 42$  exceeds the current state-of-the-art for Hamiltonian simulation<sup>51</sup>,  $N = 38$ . Still, we welcome collaboration on studying the boundaries of nearly-exact classical simulation with the most advanced hardware.

### 13. Standard Trotterization-based TEBD-MPS

A standard approach for simulating large one-dimensional (1D) quantum systems is to use a matrix product state (MPS) representation<sup>53</sup>. In this representation, the quantum state is decomposed into a tensor network in the form of a 1D chain through repeated Schmidt decompositions. The key insight of such schemes is that if particles are weakly entangled at any bipartition of the system, then the on-site tensors will be low rank, and so can be efficiently *truncated* to a smaller size while preserving most of the information. The size to which the tensor is truncated is user-controllable, and is known as the *bond dimension*,  $\chi$ . Thus, larger bond dimensions equate to truncating less information when converting a state into its MPS representation. For a more general review of MPS, see Ref.<sup>24</sup>.

An important aspect of the tensor decomposition of the MPS is that unitary evolution can be efficiently applied to just part of the tensor network. This enables the standard time-evolving block decimation (TEBD) algorithm<sup>54</sup>, in which Hamiltonian evolution is discretized (or *Trotterized*) using the Suzuki-Trotter formula

$$e^{-i(\sum_i \hat{H}_i)\delta} = \prod_i e^{-i\hat{H}_i\delta} + \mathcal{O}(\delta^2). \quad (98)$$

For a many-body Hamiltonian with (geometrically) local interactions  $\hat{H}_i$ , each term is applied sequentially onto the MPS. With suitable implementation, Trotterization results in a second-order decomposition of the time evolution

operator.

Thus, the basic algorithmic loop of the TEBD approach is: 1) perform Schmidt decompositions to write the state as a tensor network, 2) *truncate* the tensor ranks, 3) *Trotterize* the Hamiltonian, 4) and apply it to the relevant tensors, after which we repeat step 1. The two fundamental sources of error in this approach come from steps 2 (the truncation) and 3 (the Trotterization) (Supp. Fig. 26).

The error of the Trotterization decomposition depends on the product of the commutator of the terms in the Hamiltonian and the time step  $dt$  (here  $dt$  is defined as a full cycle of application of local evolution from the leftmost site to the rightmost site and then back to the leftmost site). For the Rydberg system, the commutator is on the order of the nearest neighbor interaction  $V_{nn}$ . Therefore, even with the second-order Suzuki–Trotter decomposition, the perturbative expansion converges only for  $V_{nn} \cdot dt \lesssim 1$ . Due to the Rydberg blockade, the nearest neighbor interaction for Rydberg Hamiltonian is strong, putting a significant constraint on the maximum  $dt$  possible for the simulation. In practice, with a second-order decomposition of the time-evolution operator, we found that a maximum of  $dt \approx 1/100$  Rabi cycle is necessary to have a 0.5% Trotterization error.

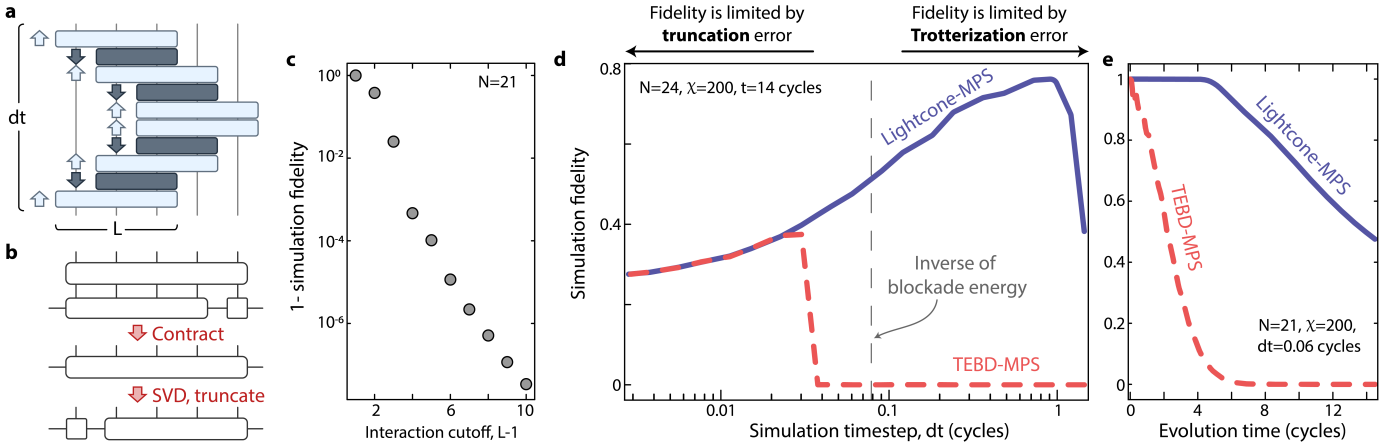
Importantly, however, for too small of a timestep, the TEBD algorithm will perform many more truncations, each time lowering the fidelity. This leads to a fundamental tension, where an optimal  $dt$  must be chosen which balances these competing effects of truncation and Trotterization errors (Supp. Fig. 26d).

Moreover, a truncation of the long-range interaction must be chosen to efficiently use the TEBD algorithm. In practice, we found that keeping an interaction length of 5 results in the desired error of less than 0.5% (Supp. Fig. 26c). This results in a local evolution involving at most 6 atoms.

We emphasize that we do not believe this performance could be strongly improved through the use of a Schrieffer–Wolff transformation as described above. This is because in performing the Schmidt decompositions, MPS-based methods already find the optimal working basis, which is nearly entirely within the blockaded Hilbert space (Supp. Fig. 6d). In fact, it is likely such a transformation will *decrease* the MPS performance, because of the added complexity in local gate application due to the non-locally-factorizable structure of the blockaded subspace.

#### I4. The Lightcone-MPS algorithm

One of the challenges of the Trotterization TEBD algorithm is the small  $dt$  required to faithfully simulate the quantum dynamics. However, this difficulty can be circumvented by efficiently utilizing the lightcone dynamics of quantum systems, allowing for simulation with much larger  $dt$ , and thus higher fidelity as truncation errors are minimized. In particular, in Ref.<sup>55</sup> it was shown that a global time evolution operator could be decomposed into



**Supp. Fig. 26 | Comparing Lightcone-MPS and TEBD algorithms.** **a.** Illustration of the Lightcone-MPS Hamiltonian decomposition over a single time step  $dt$ . The long range Rydberg interaction is truncated after a cutoff distance of  $L$ . **b.** Illustration of the MPS update procedure. **c.** We choose the Lightcone-MPS block size to be  $L = 6$ , which we estimate as introducing an error of  $< 0.5\%$  at  $N = 60$ . **d.** and **e.** For MPS algorithms there exists an optimal simulation timestep; for too large a timestep, the Hamiltonian Trotterization introduces simulation errors. For too small a timestep, the MPS is contracted and truncated too often, leading to loss of accuracy. For the TEBD algorithm, the maximum timestep is set by the inverse blockade energy, but the Lightcone-MPS algorithm can be used accurately with much larger timesteps. This ultimately means that at equal bond dimension,  $\chi$ , the Lightcone-MPS algorithm can be optimized to higher fidelities than the TEBD.

small blocks of local evolution by combining forward and backward time evolution. In this proposed non-perturbative decomposition, the time step is only limited by the Lieb-Robinson velocity  $v$  and the block size  $L$  of the local time evolution, not the blockade interaction strength<sup>56</sup>.

Here, we propose a modified version of the Hamiltonian decomposition in Ref.<sup>55</sup>, which is especially amenable to efficient integration with MPS state representations, and which we call Lightcone-MPS. This modification allows for a smaller block size  $L$  with the same cutoff of long-range interactions. This algorithm consists of alternating forward-time and backward-time evolution, where the forward-time evolution has a block size  $L$ , and the backward evolution has a block size of  $L - 1$ . Each block performs the time evolution for all Hamiltonian terms strictly inside the block. Crucially, in this implementation,  $dt$  is no longer limited by blockade interactions at the edges of the local evolution blocks (Supp. Fig. 26d).

Specifically, for the Rydberg Hamiltonian on  $N$  atoms

$$\hat{H}/h = \Omega \sum_{i=1}^N \hat{S}_i^x - \Delta \sum_{i=1}^N \hat{n}_i + \frac{C_6}{a^6} \sum_{i=1}^N \sum_{j=i+1}^N \frac{\hat{n}_i \hat{n}_j}{|i-j|^6}, \quad (99)$$

we consider the Hamiltonian defined locally from  $k$  to  $k + L - 1$  sites as

$$\hat{H}_{k,k+L-1}/h = \Omega \sum_{i=k}^{k+L-1} \hat{S}_i^x - \Delta \sum_{i=k}^{k+L-1} \hat{n}_i + \frac{C_6}{a^6} \sum_{i=k}^{k+L-1} \sum_{j=i+1}^{k+L-1} \frac{\hat{n}_i \hat{n}_j}{|i-j|^6}, \quad (100)$$

with the forward evolution operator on  $L$  sites and backward evolution operator on  $L - 1$  sites defined as

$$\hat{U}_{k,k+L-1}^{\text{forward}} = \exp\left(-i \frac{\hat{H}_{k,k+L-1}}{h} \frac{dt}{2}\right); \quad (101)$$

$$\hat{U}_{k,k+L-2}^{\text{backward}} = \exp\left(+i \frac{\hat{H}_{k,k+L-2}}{h} \frac{dt}{2}\right) \quad (102)$$

Then, the evolution within one timestep is decomposed (Supp. Fig. 26a) as

$$\exp\left(-i \frac{\hat{H}}{h} dt\right) \approx \hat{U}_{1,L}^{\text{forward}} \prod_{i=2}^{N-L+1} \left(\hat{U}_{i,i+L-2}^{\text{backward}} \hat{U}_{i,i+L-1}^{\text{forward}}\right) \prod_{i=N-L+1}^2 \left(\hat{U}_{i,i+L-1}^{\text{forward}} \hat{U}_{i,i+L-2}^{\text{backward}}\right) \hat{U}_{1,L}^{\text{forward}} \quad (103)$$

Here, the order of applying the time evolution operator is reversed in every other layer (i.e. left to right for odd layers and right to left for even layers), and each two layers are defined as one timestep. In practice, each forward evolution can be combined with the backward evolution (each pair of operators in Eq. (103)) into a single operator, reducing the simulation cost.

The simulation cost can be further reduced by always keeping a tensor with  $L - 1$  atoms fused together, without performing addition decompositions of the MPS. (Supp. Fig. 26b). More specifically, we can keep an MPS in the form

$$\langle z_1, z_2, \dots, z_N | \psi_{MPS} \rangle = \sum_{\alpha_1, \alpha_2, \dots, \alpha_{N-1}} A_{z_1}^{\alpha_1} A_{z_2}^{\alpha_1 \alpha_2} \dots M_{z_j, \dots, z_{j+L-2}}^{\alpha_{j-1} \alpha_j + L-2} \dots B_{z_N}^{\alpha_{N-1}}, \quad (104)$$

without decomposing  $M_{z_j, \dots, z_{j+L-2}}^{\alpha_{j-1} \alpha_j + L-2}$  into  $L - 1$  tensors. Here,  $z_i$  denotes the physical indices and  $\alpha_i$  denotes the bond dimensions. In addition, we use  $A$  to denote the tensors in the left canonical form,  $B$  to denote tensors in the right canonical form, and  $M$  to denote tensors in the center canonical form. Basically, all tensors to the left of  $M_{z_j, \dots, z_{j+L-2}}^{\alpha_{j-1} \alpha_j + L-2}$  should be in the left canonical form and all the tensors to the right of  $M_{z_j, \dots, z_{j+L-2}}^{\alpha_{j-1} \alpha_j + L-2}$  should be in the right canonical form. When applying the operator of length  $L$  between site  $j$  and site  $j + L - 1$ , we can first combine  $M_{z_j, \dots, z_{j+L-2}}^{\alpha_{j-1} \alpha_j + L-2}$  with  $B_{z_{j+L-1}}^{\alpha_{j+L-2} \alpha_j + L-1}$  into a tensor of length  $L$

$$M_{z_j, \dots, z_{j+L-1}}^{\alpha_{j-1} \alpha_j + L-1} = \sum_{\alpha_{j+L-2}} M_{z_j, \dots, z_{j+L-2}}^{\alpha_{j-1} \alpha_j + L-2} B_{z_{j+L-1}}^{\alpha_{j+L-2} \alpha_j + L-1}. \quad (105)$$

Then, we update the tensor by applying the operator of length  $L$

$$M_{z_j, \dots, z_{j+L-1}}^{\alpha_{j-1} \alpha_j + L-1} = \sum_{z'_j, \dots, z'_{j+L-1}} U_{z_j, \dots, z_{j+L-1}; z'_j, \dots, z'_{j+L-1}} M_{z'_j, \dots, z'_{j+L-1}}^{\alpha_{j-1} \alpha_j + L-1}, \quad (106)$$

and decompose the updated tensor back into two tensors of length 1 and  $L - 1$

$$M'_{z_j, \dots, z_{j+L-1}}^{\alpha_{j-1} \alpha_{j+L-1}} \xrightarrow{SVD} A'_{z_j}^{\alpha_{j-1} \alpha_j} M'_{z_{j+1}, \dots, z_{j+L-1}}^{\alpha_j \alpha_{j+L-1}}. \quad (107)$$

Here, the process of truncating the Schmidt coefficient and combining it into  $M'_{z_{j+1}, \dots, z_{j+L-1}}^{\alpha_j \alpha_{j+L-1}}$  is omitted for simplicity. The same process can be carried on over the rest of the system, and reversed if the operators are applied in the reverse order.

The Lightcone-MPS algorithm allows for a maximum  $dt \sim L/v$ . Although computing the Lieb-Robinson velocity is difficult, based on empirical results, we find in practice that  $dt \approx 0.5$  Rabi cycle is sufficient for a 0.5% decomposition error with a block size of 6 (long-range interaction cutoff at 5). Again, the  $dt$  is defined as a full cycle of application of local evolution from the left-most site to the right-most site and then back to the left-most site.

Crucially, there is still a tension on  $dt$  from the competing effects of truncation and Trotterization errors, but for the Lightcone-MPS algorithm the Trotterization errors are minimized up to much larger  $dt$ . This means at equal bond dimension, if selecting the optimal  $dt$ , the Lightcone-MPS can reach higher fidelities than TEBD (Supp. Fig. 26de). Equivalently, to reach the same fidelity necessitates the TEBD to use a larger bond dimension, as is visible in Fig. 5 of the main text.

Additionally, even if a larger bond dimension is used for the TEBD (where truncation error is small), the optimal  $dt$  in TEBD is empirically  $\sim 50\times$  smaller than that for the Lightcone-MPS, increasing the raw simulation time further. In practice, to benchmark the experiment, we need to choose  $dt$  such that we can match the time step chosen in the experiment (around 0.77 Rabi cycle). Therefore, we use half the experimental time step as the  $dt$  for the Lightcone-MPS algorithm, and use 1/80 the experimental time step as the  $dt$  for the TEBD algorithm. This still results in a 1/40 reduction in number of simulation steps thanks to the Lightcone-MPS algorithm.

## 15. Time-dependent variational principle

The Time-Dependent Variational Principle<sup>57</sup> (TDVP) algorithm is another MPS time evolution algorithm. In this algorithm, the Schrödinger equation is projected onto the variational manifold of MPS, where an effective (nonlinear) partial differential equation (PDE) is obtained to update the MPS parameters. This method is better than TEBD in certain situations such as simulating emergent hydrodynamics, due to its ability to preserve all symmetries and conservation laws. In addition, TDVP can naturally handle long-range interactions.

However, the TEBD algorithm (and the Lightcone-MPS algorithm) is expected to be better suited in our setup. First, the TEBD algorithm is optimal for minimizing the infidelity in each evolution time step. This is done by truncating the Schmidt coefficients in a way that the infidelity is minimized. In the TDVP algorithm, on the other hand, the infidelity is not necessarily minimized at the cost of preserving conserved quantities. Second, when combined with MPS representation, the TDVP implementation is nontrivial. This is due to the existence of the gauge degree of freedom in MPS representation. More specifically, the representation of MPS is not unique due to gauge transformation. This makes it difficult to utilize MPS-based TDVP by leveraging existing PDE solvers. Ref.<sup>58</sup> presents a solution to this problem, which shows that TDVP can be implemented by using the canonical form of MPS, but it relies on evolving the system locally in small time steps, akin to the Trotterization of the Hamiltonian. In addition, it turns out that this solution is a minor modification of the original TEBD algorithm, where a backward time evolution is added. This backward evolution is already naturally incorporated in our Lightcone-MPS algorithm.

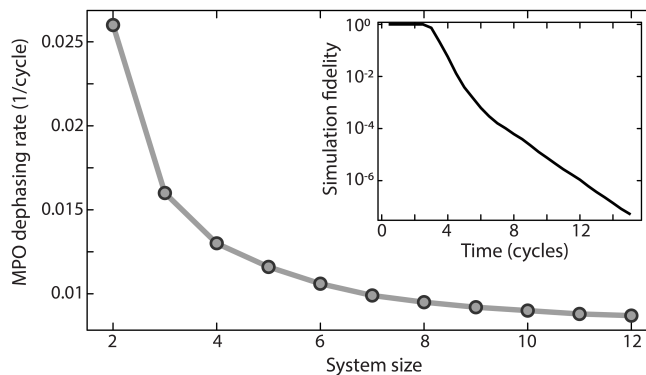
Moreover, regardless of the implementation of the TDVP algorithm, the PDE needs to be solved numerically, with the maximum time step limited by the largest term of the Hamiltonian or the largest term in the PDE. Therefore, the TDVP algorithm has the same bottleneck as the TEBD algorithm —the time step must be small compared to the blockade interaction strengths, limiting its utility. In our Lightcone-MPS algorithm, however, this limitation is lifted by utilizing extra physical properties such as Lieb-Robinson bound and our knowledge about the geometry of the system (Supp. Fig. 26). Given this, we do not explicitly test a TDVP implementation in this work.

## 16. Noisy MPO-based simulation

Instead of simulating perfect quantum dynamics with limited fidelity, an alternative method of benchmarking with the experiment is to (almost) exactly simulate noisy quantum dynamics  $\hat{\rho}(t)$  with similar fidelity. We explore this possibility using the recently developed matrix product operator (MPO) based simulation<sup>59</sup>. In MPO simulation, each local tensor has an extended Hilbert space dimension of 4 instead of 2. The increased local dimension significantly increases the cost of simulation using the efficient decomposition of Ref.<sup>55</sup>. Therefore, we fall back to simulating the Trotterized Hamiltonian dynamics and ignore the long-range interactions. In addition, we choose  $dt$  to be half the Rabi cycle, and ignore Trotterization error, in order to consider this method in the most favorable light.

For simplicity, we consider local dephasing error as the only error source and find the effective dephasing rate  $\gamma$  per qubit as a function of system size such that the fidelity of the simulation matches the fidelity of the experiment (Supp. Fig. 27), extrapolated using the exponential decay fit. As shown in the figure, the effective dephasing rate decreases as the system size increases, which means the dephasing rate at  $N = 12$  overestimates the dephasing rates at larger system sizes. Since the MPO entanglement decreases as the decoherence rate increases, this leads to an underestimation of the MPO simulation complexity. To make the case more favorable for the MPO method, we thus use the dephasing rate at  $N = 12$  for the simulation at  $N = 60$ .

Assuming the early time exponential fit of the experimental fidelity, the Lightcone-MPS algorithm achieves similar fidelity with the experiment at  $N = 60$  using  $\chi = 913$  (Ext Data Fig. 31). This motivates us to test if the MPO method with a similar or larger bond dimension can achieve the same fidelity. We simulate the noisy quantum dynamics at  $N = 60$  using  $\chi = 1100$  and the dephasing rate upper bound, yielding the un-normalized simulation mixed state  $\hat{\rho}_{trun}$ , where we don't explicitly normalize  $\hat{\rho}_{trun}$  at each time step. Here,  $\text{tr}(\hat{\rho}_{trun})$  serves as a measure of the simulation accuracy, where it equals 1 when the simulation is perfect and decreases as we truncate the density matrix. We measure  $\text{tr}(\hat{\rho}_{trun})$  as a function of time (Supp. Fig. 27, inset) and find that it drops to  $\sim 10^{-7}$ . This shows that the MPO algorithm is far from an exact simulation of the noisy quantum dynamics. We note that the  $\text{tr}(\hat{\rho}_{trun})$  measures the accuracy of the algorithm compared to the noisy dynamics. It does not measure the fidelity of the simulation compared to the ideal dynamics, which can be even smaller. Thus, this study, in combination with the Trotterization errors not considered here, implies the MPO-based approach will achieve far lower fidelity than the experiment and the lightcone-MPS method for reasonable simulation parameters.



**Supp. Fig. 27 | Matrix product operator (MPO) simulation.** Effective dephasing rate per qubit to obtain noisy dynamics with a fidelity similar to the exponential decay fit of the experiment. Inset) Truncation accuracy of the MPO method with  $\chi = 1100$  at  $N = 60$ . The effective dephasing rate is taken as the dephasing rate at  $N = 12$ , which should serve as an upper bound of the effective dephasing rate at  $N = 60$ .

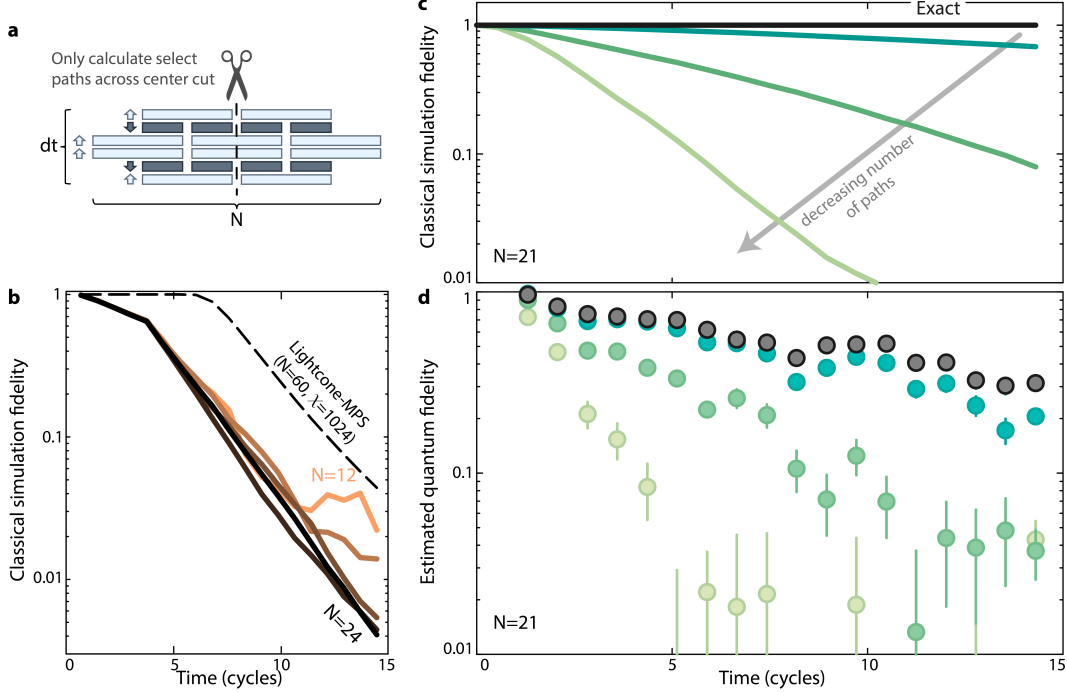
## 17. Path integral formulation

A hybrid path integral algorithm was introduced in Ref.<sup>5</sup> for quantum circuit simulation. In this algorithm, the system is cut into two patches. When no gate is applied between the two patches, each patch evolves independently under the Schrödinger equation. When a gate is applied across the two patches, the gate is decomposed into sums of gates that apply independently on the two patches, where one path was sampled out of this decomposition. More formally, we write the gate as  $\hat{U} = \sum_i \hat{V}_i \otimes \hat{W}_i$ . For each path, we only choose a particular  $\hat{V}_i \otimes \hat{W}_i$ , so that the two patches stay in a product state. The wave function of each path can be then summed over to obtain the desired result. Suppose each gate across the cut can be minimally decomposed as  $r$  product gates and there are  $m$  gates across the cut in total,  $r^m$  paths are required to fully describe the system.

In this work, we adapt the hybrid path integral algorithm to the long-range Hamiltonian dynamics. We first decompose<sup>55</sup> the global time evolution operator into local gates with block size  $L = 10$ . We further reverse the order of the gate application at every other layer so the gates can be fused between the two layers (Supp. Fig. 28). This reduces the number of gates that need to be decomposed at the cut. This allows us to have a minimum of 19 cuts throughout the simulation to match the experiment time steps. Similar to the original hybrid path integral algorithm, we cut the system into two halves. For gates applied across the two halves of the system, we use Schmidt decomposition to write the gate as  $\hat{U} = \sum_i \sigma_i \hat{V}_i \otimes \hat{W}_i$ , where  $\sigma_i$  can be then absorbed into either of the halves. However, a naïve decomposition results in a large interaction between the blockade and non-blockade sectors, which results in poor



convergence. Therefore, we project the gate into the blockade sector before performing the decomposition.



**Supp. Fig. 28 | A hybrid path integral simulation.** **a.** Illustration of the hybrid lightcone-path integral (Lightcone-PI) method. **b.** Simulation fidelity of the Lightcone-PI with  $4 \times 10^6$  paths, showing decreasing fidelity as a function of  $N$ . We compare against  $F_{\text{svd}}$  for the Lightcone-MPS algorithm at  $\chi = 1024$  and  $N = 60$ , which outperforms the L-PI method. **c, d.** At a fixed  $N$ , the infidelity in the Lightcone-PI method from using too few paths is visible when benchmarking directly against experimental data, as was the case for the Lightcone-MPS algorithm (Fig. 2 of the main text).

We test this algorithm for up to 24 atoms where we truncate the rank of the gates across the cut to  $r_i$  to mimic actually simulating the path integral with  $\prod_i r_i$  paths. We find that the resulting fidelity mostly depends on the number of paths and only weakly depends on the system size, where the dependency becomes weaker as the system size increases. This is because, for all system sizes, we have the same number of gates across the cut, which we expect to be the only factor affecting the fidelity for large enough system sizes. Extrapolating to 60 atoms, we find that we need at least  $4 \times 10^6$  paths to obtain a fidelity of  $\sim 0.01$ , where *each path* requires an exact simulation of two 30-atom systems. Even using our Lightcone-MPS method, simulating each 30-atom system takes 60 core-hours (with  $\chi = 1536$ , which is nearly exact). This adds up to a formidable amount of computational resources. In comparison, a direct application of the Lightcone-MPS algorithm for 60 atoms achieves better fidelity with a bond dimension of  $\chi = 1024$  (Supp. Fig. 28b).

Interestingly, we can still benchmark the experimental fidelity using the hybrid path integral approach, which we see shows qualitatively the same behavior as the Lightcone-MPS algorithm in Fig. 2 of the main text. That is, when the number of paths is too low, the classical simulation fidelity drops, and this drop is visible directly on the benchmarked experimental fidelity (Supp. Fig. 28cd).

This raises the interesting potential prospect of using the experimental benchmarking protocol in reverse to determine the fidelity of the classical algorithm, which we leave to future work.

## 18. Neural network based Hamiltonian simulation

Neural networks have shown great potential in simulating quantum dynamics<sup>60,61</sup>. However, such protocols are still in an early age of development and their application to the Rydberg Hamiltonian with ergodic, long-time dynamics is still challenging. We considered two different neural network algorithms used in Refs<sup>62,63</sup>, but neither results in a fidelity higher than the Lightcone-MPS algorithm, even for comparing both at system sizes as low as  $N = 12$ . While it is possible that neural network based algorithms require additional fine-tuning and hyper-optimization, as it is not the main focus of this work, we leave them for future study. We also note that neural network algorithms are under active development, and we welcome readers to benchmark their algorithms on our data.



## J. Estimating the MPS fidelity

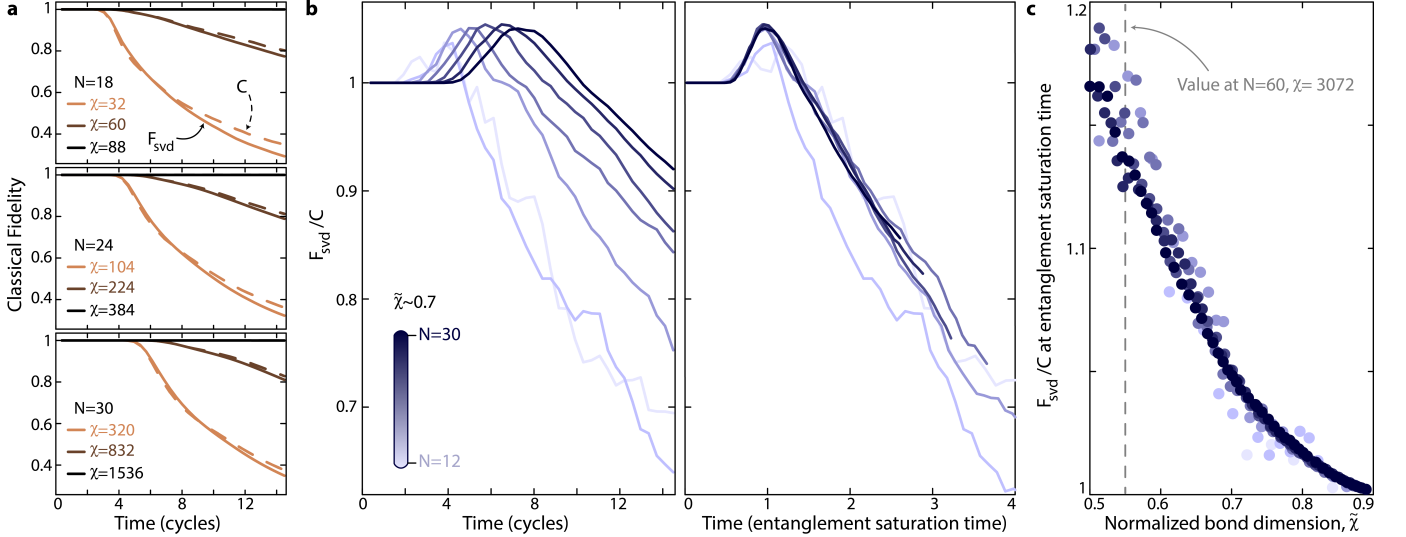
### J1. Approximating the MPS fidelity with truncation errors

While MPS gives a controlled approximation of the exact wave function, an exact calculation of the MPS fidelity,  $C$  – defined as the overlap between the MPS state and the ideal target state – is not possible over multiple time steps for system sizes greater than  $N \sim 30$ . However, it is possible to estimate the MPS fidelity from the truncation fidelity

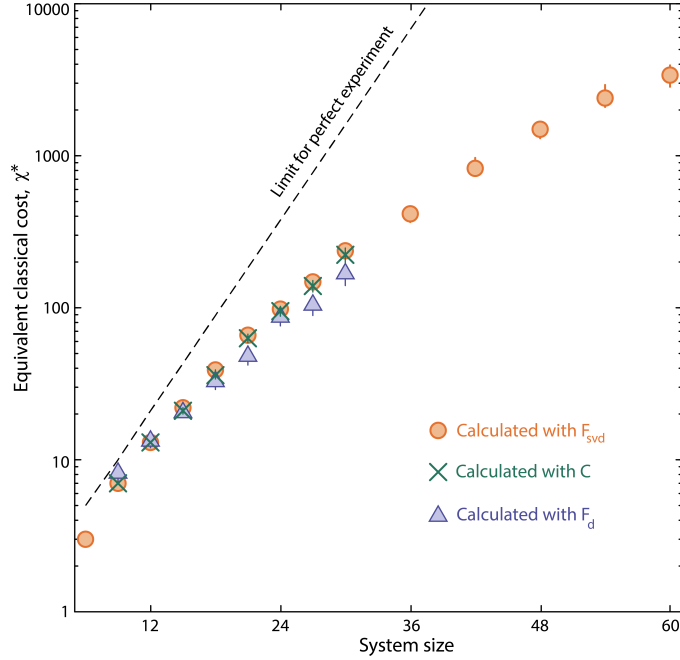
$$F_{\text{svd}} = \prod_i \sum_{\alpha=1}^{\chi} s_{i,\alpha}^2, \quad (108)$$

where  $i$  runs over all steps involving Schmidt value truncations, and  $s_{i,\alpha}$  are the Schmidt values at truncation step  $i$ . Here, we assume the wave function is normalized where  $\sum_{\alpha=1}^{\infty} s_{i,\alpha}^2 = 1$ . While this estimation is only approximate, it can be extremely accurate when successive truncations are independent<sup>25</sup>. For  $N \leq 30$ , a bond dimension of  $\chi = 3072$  almost exactly captures the dynamics, therefore, it is possible to compute the MPS fidelity,  $C$ , for various bond dimensions. In Supp. Fig. 29a, we plot  $F_{\text{svd}}$  versus the actual MPS fidelity,  $C$ , to show the general agreement when using the Lightcone-MPS algorithm. This is expected due to the large  $dt$  of the algorithm, which makes successive truncations approximately independent.

In addition, we notice that the agreement becomes better when the system size increases, indicating that  $F_{\text{svd}}$  can be a good estimator of the MPS fidelity for large systems (Supp. Fig. 29b). If anything, it appears that in the parameter regime of evolution times shorter than the entanglement saturation time,  $F_{\text{svd}}$  slightly over-estimates  $C$ , making it a conservative quantity to study (Supp. Fig. 29bc). Further, we emphasize that for  $N < 30$ , where we have access to both  $C$  and  $F_{\text{svd}}$ , the values of  $\chi^*$  are consistent (Supp. Fig. 30).



**Supp. Fig. 29 | Estimating the classical fidelity.** **a.** MPS truncation accuracy,  $F_{\text{svd}}$ , and MPS fidelity,  $C$ , for various bond dimensions and system sizes. The MPS fidelity is measured against another MPS calculation with  $\chi = 3072$ , which is almost exact for  $N \leq 30$ . **b.** For a fixed log-normalized bond dimension,  $\tilde{\chi} \sim 0.7$ ,  $F_{\text{svd}}$  first over-estimates  $C$ , before under-estimating it at late time (left). However, rescaling the time by the entanglement saturation time (which is linearly proportional to system size, see Fig. 3e of the main text) causes the data to collapse (right). This implies that before  $t_{\text{ent}}$ ,  $F_{\text{svd}}$  will over-estimate  $C$ , so in the parameter regimes we study in this work, using  $F_{\text{svd}}$  is the conservative choice when comparing quantum and classical fidelities. **c.** The ratio of  $F_{\text{svd}}/C$  around the entanglement saturation time, showing the over-estimation increases with decreasing  $\tilde{\chi}$ . Given the data collapse, we expect that  $F_{\text{svd}}$  overestimates by a factor of 1.15 at the entanglement saturation time for  $N = 60$ . Taking this effect into account implies the actual  $\chi^*(N=60) \approx 4000$ , higher than the value of 3400 quoted in the main text. However, we choose to present the lower value as a more conservative estimate. We note that this relationship between  $F_{\text{svd}}$  and  $C$  is only observed to hold for the case of the Lightcone-MPS algorithm, because the timestep is long enough that truncations are relatively independent. The agreement is not observed with a TEBD MPS implementation (which requires much finer timesteps, see text).



**Supp. Fig. 30 | Estimating  $\chi^*$  from  $F_d$  and true classical fidelity.** For system sizes larger than  $N > 30$ , we are not easily able to calculate the exact classical MPS fidelity (as it is impractical to simulate the exact reference state), so in the main text we used  $F_{\text{svd}}$  to compare against the experimental quantum fidelity. However, for  $N \leq 30$ , we can directly compare against the classical MPS fidelity,  $C$ , which we find gives consistent values of  $\chi^*$  in the applicable range. Finally, for these small  $N$ , we can directly calculate  $F_d$  between an approximate MPS simulation and an exact simulation. We find the corresponding  $\chi^*$  is also consistent with the values based solely on fidelity.

## J2. Extrapolating the experimental $\chi^*$ value

We only directly perform Lightcone-MPS simulations up to  $\chi = 3072$ . At this highest  $\chi$ , the MPS fidelity for  $N = 60$  at the latest experimental time is 0.088, as quantified by  $F_{\text{svd}}$ . Given that this is lower than the experimental fidelity of 0.095(11), we need to extrapolate in order to find the predicted value of  $\chi^*$ . The fidelities are very close, so we make a simple linear extrapolation approximation, from which we find  $\chi^* = 3392 \approx 3400$  as quoted in the main text (Supp. Fig. 31a).

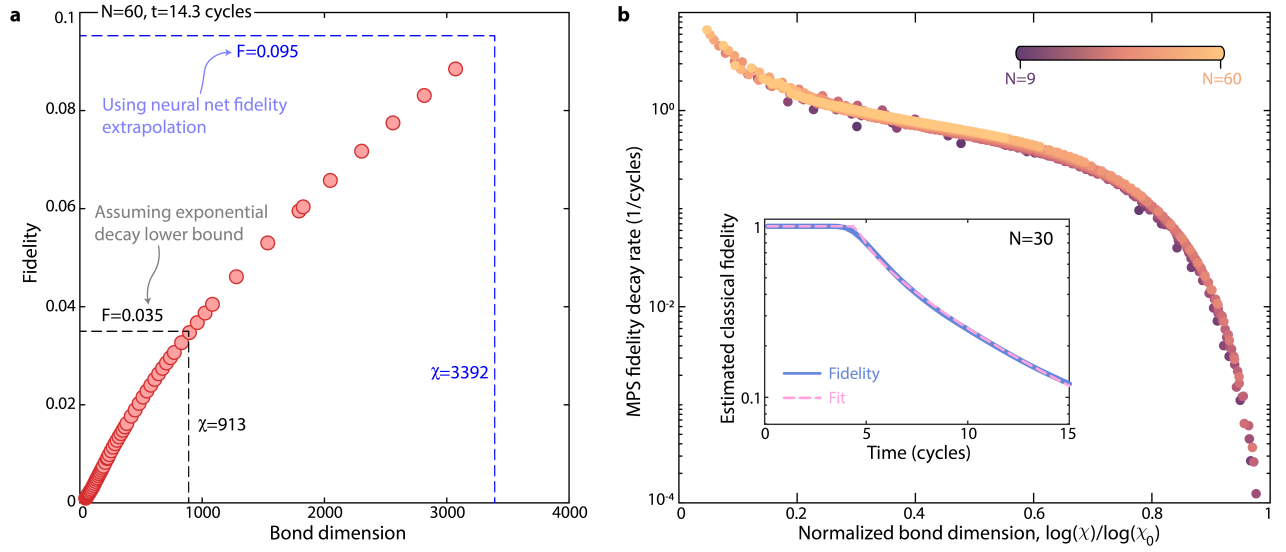
## J3. Fitting and extrapolating the MPS fidelity

In the previous subsection, we used a linear extrapolation of MPS fidelity as a function of bond dimension in order to predict  $\chi^*$  beyond the regime we directly simulated. However, in Fig. 5b of the main text, we compare the experimental and MPS fidelity over a much larger hypothetical fidelity range, where the linear approximation breaks down (see for instance Supp. Fig. 10ab).

In order to make the comparisons in Fig. 5b, we perform fits of the MPS fidelity, approximated by  $F_{\text{svd}}$ . For RUCs,  $F_{\text{svd}}$  is essentially exponentially decaying past  $t_{\text{ex}}$ , as the truncations are independent<sup>25</sup>. In our case this behavior is not perfect, and so we fit the MPS fidelity as a stretched exponential decay after  $t_{\text{ex}}$ .

We are able to fit the exponential parameters, and in particular observe an apparent universality in the decay rate as a function of the log-normalized bond dimension (Supp. Fig. 31b), an interesting phenomena which we leave to future study. We ultimately use these fits to predict the MPS fidelity at arbitrary bond dimension to compare against experiment; we find the fits generally agree to within a few percent with the true fidelities.

To generate Fig. 5b, we assume an effective per-atom error rate,  $\mathcal{F}$ , and find the minimum  $\chi^*$  such that the fitted MPS fidelity is always greater than  $F = \mathcal{F}^{Nt}$ . The experimental  $\mathcal{F}$  is determined as the value required for exponential decay of the experiment to match the fidelity value of 0.095(11) for the experiment at  $t = 14.3$  cycles and  $N = 60$ . We then convert  $\chi^*$  into runtime using the known conversion of  $\chi = 3072 \rightarrow 136$  core-days, measured empirically on the Caltech cluster, and assuming the runtime scales as  $\mathcal{O}(\chi^3)$  (Supp. Fig. 24).



**Supp. Fig. 31 | Extrapolating  $\chi^*$ .** **a.** MPS fidelity (as quantified by the MPS truncation,  $F_{\text{svd}}$ ) for bond dimensions up to  $\chi = 3072$ , the largest directly simulated. None of the MPS fidelities (markers) exceeded the experimental value (blue dashed lines), but given the functional smoothness in this regime, we use a linear extrapolation of the last few data points to predict the  $\chi^*$  for the experimental fidelity, yielding  $\chi^* = 3392 \approx 3400$ . **b.** For estimating the  $\chi$  required to reach a broader range of fidelities (as for instance is used in Fig. 5b of the main text), we fit the Lightcone-MPS fidelity as a stretched exponential decay past the exact simulation time (inset). We observe a universal collapse of the decay rate as a function of the log-normalized bond dimension, for all system sizes and bond dimension up to  $N = 60$  and  $\chi = 3072$ .

## K. Error model simulations

### K1. Random unitary circuits

The dynamics of the one-dimensional random unitary circuit (RUC) shown in Fig. 4a are simulated using randomly sampled two-qubit  $SU(4)$  unitary gates from the Haar measure. Specifically, the time evolution of an  $N$ -qubit system starting from the initial state  $|0\rangle^{\otimes N}$  can be described as follows:

$$|\psi(t)\rangle = \hat{U}_t \hat{U}_{t-1} \cdots \hat{U}_2 \hat{U}_1 |0\rangle^{\otimes N}, \quad (109)$$

where  $\hat{U}_{\text{odd}} = \{\hat{U}_1, \hat{U}_3, \dots\}$  and  $\hat{U}_{\text{even}} = \{\hat{U}_2, \hat{U}_4, \dots\}$  are the odd- and even-time unitaries composed of local two-qubit unitaries as

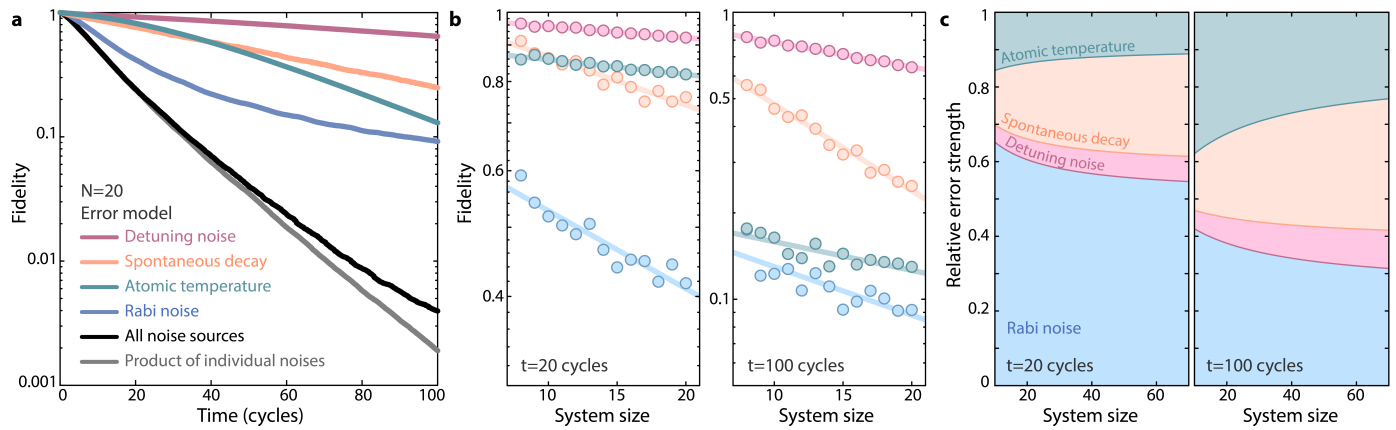
$$\hat{U}_{\text{odd}} = \prod_{i=1}^{N/2} \hat{U}_{2i-1, 2i}, \quad \hat{U}_{\text{even}} = \prod_{i=1}^{N/2} \hat{U}_{2i, 2i+1} \quad (110)$$

with open boundary condition, and  $t$  is the circuit depth and  $\hat{U}_{\mu, \nu}$  is the randomly-sampled  $SU(4)$  gate acting on two qubits at site  $\mu$  and  $\nu$ . Note that at each circuit depth  $t$ , we randomly sample two-qubit random unitaries to generate a many-body unitary  $\hat{U}$ , which leads to chaotic dynamics. To emulate noisy quantum dynamics, we utilize a stochastic evolution method<sup>64</sup> incorporating local errors represented by the Pauli operators  $\hat{S}^{x, y, z}$ . These local errors are stochastically applied to individual qubits at a single-qubit error rate of  $p_{\text{err}} = 0.007$  per circuit layer. By repeating the simulations of the noisy dynamics more than  $\sim 2000$  times with a fixed  $p_{\text{err}}$  for 10 different circuit realizations, we obtain good approximations of the density matrices,  $\hat{\rho}_{\text{RUC}}$ , from which we extract the logarithmic negativity as a measure of entanglement in the mixed state.

### K2. The Rydberg Hamiltonian

The chaotic dynamics of a one-dimensional Rydberg-atom array presented in this study can be described by time evolution governed by the Rydberg Hamiltonian given in Eq. (1) of the Supp. Info. To simulate the open quantum dynamics of the Rydberg atom array, we employ an *ab initio* error model that incorporates realistic error sources observed in our experiment. For detailed information about the error model, please refer to Refs.<sup>2,13</sup>. Similar to the case of the RUC, we simulate the noisy quantum evolution using the stochastic wavefunction method<sup>64</sup> to obtain the density matrix. This allows us to calculate the logarithmic negativity (Fig. 4a of the main text). We generally find excellent agreement between our error model and the experimental fidelities across a range of system sizes (Supp. Fig. 5), and note that this is the same error model employed in our recent study of record high-fidelity two-qubit Bell state generation<sup>13</sup>.

Of present interest is understanding the main limitations to our experimental fidelity, to best improve the system in the future. To this end, we perform error budget simulations, grouping error sources generally into four main categories: noise on the detuning, noise on the Rabi frequency, spontaneous decay, and atomic temperature effects (Supp. Fig. 32). We generally find that Rabi frequency noise (in particular, as arising from shot-to-shot laser intensity fluctuations) is a dominant noise source across all times. We extrapolate from error model simulations at small system sizes to larger systems, and find that likely spontaneous decay can become of similar strength for large systems (Supp. Fig. 32). We note that in these simulations, collective Rydberg decay effects are not accounted for<sup>65</sup>, but we do not yet observe evidence of such effects in the  $t < 15$  cycles time window in which we experimentally operate.



**Supp. Fig. 32 | Impact of different error sources on fidelity.** **a.** Simulated fidelity decay for a variety of choices of noise sources affecting our Rydberg quantum simulator. **b.** Fidelities for error sources in **a** at short (left) and long (right) times. Solid lines are exponential fits as a function of system size. **c.** We use the fits in **b** to estimate the relative contribution of different error sources for increasing system size. Here relative error contribution is defined as  $\log(F_i)/\log(F)$ , where the subscript,  $i$ , indexes the error source in question, and  $F = \prod_i F_i$ , the product of all individual errors. Note that collective Rydberg decay processes are not accounted for here, and may further lower the fidelity measured for spontaneous decay noise at late times.

- 
- [1] D. N. Page, Average entropy of a subsystem, *Physical Review Letters* **71**, 1291 (1993).
- [2] J. Choi, A. L. Shaw, I. S. Madjarov, X. Xie, R. Finkelstein, J. P. Covey, J. S. Cotler, D. K. Mark, H.-Y. Huang, A. Kale, et al., Preparing random states and benchmarking with many-body quantum chaos, *Nature* **613**, 468 (2023).
- [3] D. K. Mark, J. Choi, A. L. Shaw, M. Endres, and S. Choi, Benchmarking Quantum Simulators Using Ergodic Quantum Dynamics, *Physical Review Letters* **131**, 110601 (2023).
- [4] X. Zhang, E. Kim, D. K. Mark, S. Choi, and O. Painter, A superconducting quantum simulator based on a photonic-bandgap metamaterial, *Science* **379**, 278 (2023).
- [5] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, et al., Quantum supremacy using a programmable superconducting processor, *Nature* **574**, 505 (2019).
- [6] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, et al., Strong Quantum Computational Advantage Using a Superconducting Quantum Processor, *Phys. Rev. Lett.* **127**, 180501 (2021).
- [7] Q. Zhu, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, M. Gong, et al., Quantum computational advantage via 60-qubit 24-cycle random circuit sampling, *Science Bulletin* **67**, 240 (2022).
- [8] J. Morvan, B. Villalonga, X. Mi, S. Mandrà, A. Bengtsson, P. V. Klimov, Z. Chen, S. Hong, C. Erickson, I. K. Drozdov, et al., Phase transition in Random Circuit Sampling, [arXiv:2304.11119](https://arxiv.org/abs/2304.11119) (2023).
- [9] S. A. Moses, C. H. Baldwin, M. S. Allman, R. Ancona, L. Ascarrunz, C. Barnes, J. Bartolotta, B. Bjork, P. Blanchard, M. Bohn, et al., A Race Track Trapped-Ion Quantum Processor, [arXiv:2305.03828](https://arxiv.org/abs/2305.03828) (2023).
- [10] A. Browaeys and T. Lahaye, Many-body physics with individually controlled Rydberg atoms, *Nature Physics* **16**, 132 (2020).
- [11] A. Cooper, J. P. Covey, I. S. Madjarov, S. G. Porsev, M. S. Safronova, and M. Endres, Alkaline-Earth Atoms in Optical Tweezers, *Physical Review X* **8**, 41055 (2018).
- [12] M. A. Norcia, A. W. Young, and A. M. Kaufman, Microscopic Control and Detection of Ultracold Strontium in Optical-Tweezer Arrays, *Physical Review X* **8**, 41054 (2018).
- [13] P. Scholl, A. L. Shaw, R. B.-S. Tsai, R. Finkelstein, J. Choi, and M. Endres, Erasure conversion in a high-fidelity Rydberg quantum simulator, *Nature* **622**, 273 (2023).
- [14] A. L. Shaw, P. Scholl, R. Finkelstein, I. S. Madjarov, B. Grinkemeyer, and M. Endres, Dark-State Enhanced Loading of an Optical Tweezer Array, *Phys. Rev. Lett.* **130**, 193402 (2023).
- [15] M. Endres, H. Bernien, A. Keesling, H. Levine, E. R. Anschuetz, A. Krajenbrink, C. Senko, V. Vuletic, M. Greiner, and M. D. Lukin, Atom-by-atom assembly of defect-free one-dimensional cold atom arrays, *Science* **354**, 1024 (2016).
- [16] D. Barredo, S. de Leseleuc, V. Lienhard, T. Lahaye, and A. Browaeys, An atom-by-atom assembler of defect-free arbitrary two-dimensional atomic arrays, *Science* **354**, 1021 (2016).
- [17] A. L. Shaw, R. Finkelstein, R. B.-S. Tsai, P. Scholl, T. H. Yoon, J. Choi, and M. Endres, Multi-ensemble metrology by programming local rotations with atom movements, *Nature Physics* (2024).
- [18] I. S. Madjarov, J. P. Covey, A. L. Shaw, J. Choi, A. Kale, A. Cooper, H. Pichler, V. Schkolnik, J. R. Williams, and M. Endres, High-fidelity entanglement and detection of alkaline-earth Rydberg atoms, *Nature Physics* **16**, 857 (2020).
- [19] J. P. Covey, I. S. Madjarov, A. Cooper, and M. Endres, 2000-Times Repeated Imaging of Strontium Atoms in Clock-Magic Tweezer Arrays, *Physical Review Letters* **122**, 173201 (2019).
- [20] A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, and M. Greiner, Quantum thermalization through entanglement in an isolated many-body system, *Science* **353**, 794 (2016).
- [21] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, Colloquium : Many-body localization, thermalization, and entanglement, *Reviews of Modern Physics* **91**, 21001 (2019).
- [22] J. S. Cotler, D. K. Mark, H.-Y. Huang, F. Hernández, J. Choi, A. L. Shaw, M. Endres, and S. Choi, Emergent Quantum State Designs from Individual Many-Body Wave Functions, *PRX Quantum* **4**, 10311 (2023).
- [23] A. J. Ferris and G. Vidal, Perfect sampling with unitary tensor networks, *Physical Review B* **85**, 165146 (2012).
- [24] J. C. Bridgeman and C. T. Chubb, Hand-waving and interpretive dance: an introductory course on tensor networks, *Journal of Physics A: Mathematical and Theoretical* **50**, 223001 (2017).
- [25] Y. Zhou, E. M. Stoudenmire, and X. Waintal, What Limits the Simulation of Quantum Computers?, *Physical Review X* **10**, 41038 (2020).
- [26] T. Gorin, T. Prosen, T. H. Seligman, and M. Žnidarič, Dynamics of Loschmidt echoes and fidelity decay, *Physics Reports* **435**, 33 (2006).
- [27] W. H. Zurek, F. M. Cucchietti, and J. P. Paz, Gaussian Decoherence and Gaussian Echo from Spin Environments (2007).
- [28] P. R. Zangara and H. M. Pastawski, Loschmidt echo in many-spin systems: a quest for intrinsic decoherence and emergent irreversibility, *Physica Scripta* **92**, 33001 (2017).
- [29] I. Najfeld and T. F. Havel, Derivatives of the matrix exponential and their computation, *Advances in applied mathematics* **16**, 321 (1995).
- [30] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics, *Advances in Physics* **65**, 239 (2016).
- [31] E. T. Bell, Exponential polynomials, *Annals of Mathematics* pp. 258–277 (1934).
- [32] C. S. Withers and S. Nadarajah, Moments from cumulants and vice versa, *International Journal of Mathematical Education in Science and Technology* **40**, 842 (2009).
- [33] M. A. Ganaie, M. Hu, A. K. Malik, M. Tanveer, and P. N. Suganthan, Ensemble deep learning: A review, *Engineering*

Applications of Artificial Intelligence **115**, 105151 (2022).

- [34] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Neural-network quantum state tomography, *Nature Physics* **14**, 447 (2018).
- [35] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita, Reconstructing quantum states with generative models, *Nature Machine Intelligence* **1**, 155 (2019).
- [36] X. Zhang, M. Luo, Z. Wen, Q. Feng, S. Pang, W. Luo, and X. Zhou, Direct Fidelity Estimation of Quantum States Using Machine Learning, *Physical Review Letters* **127**, 130503 (2021).
- [37] H.-Y. Huang, R. Kueng, G. Torlai, V. V. Albert, and J. Preskill, Provably efficient machine learning for quantum many-body problems, *Science* **377**, eabk3333 (2023).
- [38] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, Random quantum circuits transform local noise into global white noise, *arXiv:2111.14907* (2021).
- [39] U. T. Bhosale, S. Tomsovic, and A. Lakshminarayanan, Entanglement between two subsystems, the Wigner semicircle and extreme-value statistics, *Phys. Rev. A* **85**, 62331 (2012).
- [40] A. Datta, Negativity of random pure states, *Phys. Rev. A* **81**, 52312 (2010).
- [41] G. Vidal and R. F. Werner, Computable measure of entanglement, *Phys. Rev. A* **65**, 32314 (2002).
- [42] M. B. Plenio, Logarithmic Negativity: A Full Entanglement Monotone That is not Convex, *Phys. Rev. Lett.* **95**, 90503 (2005).
- [43] S. Lee, D. P. Chi, S. D. Oh, and J. Kim, Convex-roof extended negativity as an entanglement measure for bipartite quantum systems, *Phys. Rev. A* **68**, 62304 (2003).
- [44] M. Žnidarič, T. Prosen, G. Benenti, and G. Casati, Detecting entanglement of random states with an entanglement witness, *Journal of Physics A: Mathematical and Theoretical* **40**, 13787 (2007).
- [45] M. P. A. Fisher, V. Khemani, A. Nahum, and S. Vijay, Random Quantum Circuits, *Annual Review of Condensed Matter Physics* **14**, 335 (2023).
- [46] Y. Bao, S. Choi, and E. Altman, Theory of the phase transition in random unitary circuits with measurements, *Phys. Rev. B* **101**, 104301 (2020).
- [47] J. Eisert, M. Cramer, and M. B. Plenio, Colloquium: Area laws for the entanglement entropy, *Rev. Mod. Phys.* **82**, 277 (2010).
- [48] P. Scholl, M. Schuler, H. J. Williams, A. A. Eberharter, D. Barredo, K.-N. Schymik, V. Lienhard, L.-P. Henry, T. C. Lang, T. Lahaye, et al., Quantum simulation of 2D antiferromagnets with hundreds of Rydberg atoms, *Nature* **595**, 233 (2021).
- [49] S. Ebadi, T. T. Wang, H. Levine, A. Keesling, G. Semeghini, A. Omran, D. Bluvstein, R. Samaajdar, H. Pichler, W. W. Ho, et al., Quantum phases of matter on a 256-atom programmable quantum simulator, *Nature* **595**, 227 (2021).
- [50] T. Häner and D. S. Steiger, 0.5 Petabyte Simulation of a 45-Qubit Quantum Circuit, in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis* (Association for Computing Machinery, New York, NY, USA, 2017), SC '17, ISBN 9781450351140.
- [51] M. Hauru, A. Morningstar, J. Beall, M. Ganahl, A. Lewis, and G. Vidal, Simulation of quantum physics with Tensor Processing Units: brute-force computation of ground states and time evolution, *arXiv:2111.10466* (2021).
- [52] D. Bluvstein, A. Omran, H. Levine, A. Keesling, G. Semeghini, S. Ebadi, T. T. Wang, A. A. Michailidis, N. Maskara, W. W. Ho, et al., Controlling quantum many-body dynamics in driven Rydberg atom arrays, *Science* **371**, 1355 (2021).
- [53] G. Vidal, Efficient Classical Simulation of Slightly Entangled Quantum Computations, *Phys. Rev. Lett.* **91**, 147902 (2003).
- [54] G. Vidal, Efficient Simulation of One-Dimensional Quantum Many-Body Systems, *Phys. Rev. Lett.* **93**, 40502 (2004).
- [55] J. Haah, M. B. Hastings, R. Kothari, and G. H. Low, Quantum Algorithm for Simulating Real Time Evolution of Lattice Hamiltonians, *SIAM Journal on Computing* **0**, 18 (0).
- [56] M. C. Tran, A. Y. Guo, Y. Su, J. R. Garrison, Z. Eldredge, M. Foss-Feig, A. M. Childs, and A. V. Gorshkov, Locality and Digital Quantum Simulation of Power-Law Interactions, *Physical Review X* **9**, 31006 (2019).
- [57] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, Time-Dependent Variational Principle for Quantum Lattices, *Physical Review Letters* **107**, 70601 (2011).
- [58] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, Unifying time evolution and optimization with matrix product states, *Physical Review B* **94**, 165116 (2016).
- [59] K. Noh, L. Jiang, and B. Fefferman, Efficient classical simulation of noisy random quantum circuits in one dimension, *Quantum* **4**, 318 (2020).
- [60] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, *Science* **355**, 602 (2017).
- [61] I. L. Gutiérrez and C. B. Mendl, Real time evolution with neural-network quantum states, *Quantum* **6**, 627 (2022).
- [62] G. Carleo, K. Choo, D. Hofmann, J. E. T. Smith, T. Westerhout, F. Alet, E. J. Davis, S. Efthymiou, I. Glasser, S.-H. Lin, et al., NetKet: A machine learning toolkit for many-body quantum systems, *SoftwareX* **10**, 100311 (2019).
- [63] D. Luo, Z. Chen, K. Hu, Z. Zhao, V. M. Hur, and B. K. Clark, Gauge-invariant and anyonic-symmetric autoregressive neural network for quantum lattice models, *Phys. Rev. Res.* **5**, 13216 (2023).
- [64] K. Mølmer, Y. Castin, and J. Dalibard, Monte Carlo wave-function method in quantum optics, *Journal of the Optical Society of America B* **10**, 524 (1993).
- [65] C. Nill, K. Brandner, B. Olmos, F. Carollo, and I. Lesanovsky, Many-Body Radiative Decay in Strongly Interacting Rydberg Ensembles, *Physical Review Letters* **129**, 243202 (2022).