

Supplemental Material for:

The Glassy Phase of Optimal Quantum Control

Alexandre G.R. Day¹, Marin Bukov², Phillip Weinberg¹, Pankaj Mehta¹ and Dries Sels^{1,3,4}

¹*Department of Physics, Boston University, 590 Commonwealth Ave., Boston, MA 02215, USA*

²*Department of Physics, University of California, Berkeley, CA 94720, USA*

³*Department of Physics, Harvard University, 17 Oxford St., Cambridge, MA 02138, USA*

⁴*Theory of quantum and complex systems, Universiteit Antwerpen, B-2610 Antwerpen, Belgium*

S1. STOCHASTIC DESCENT

Here we provide more details on the stochastic descent algorithms used to sample the local minima of the control landscape (see Algorithm ?? for a pseudocode of SD_k). As mentioned in the main text, SD_k is a k -flip stochastic descent where, starting from an initial random protocol, the algorithm proposes a new protocol chosen uniformly at random that differs by at most k -flips from the previous protocol. If the proposed protocol fidelity is higher than that of the previous protocol, the update is accepted. The SD_k algorithm halts when all possible updates with at most k -flips decrease the fidelity. The obtained protocol is said to be a local minimum of the (negative) log-fidelity landscape: all k -flip perturbations from that minimum will increase the log-fidelity. For N_T bangs and $k = 4$ (the most computationally intensive algorithm we ran), the number of fidelity evaluations needed to certify that a protocol is a local minimum is $\binom{N_T}{4} + \binom{N_T}{3} + \binom{N_T}{2} + \binom{N_T}{1}$. For $N_T = 80$ (the largest N_T we ran for SD_4), sampling a *single* local minimum required computing the fidelity of $\mathcal{O}(10^7)$ protocols (see Fig. ??). Note that evaluating the fidelity of a single bang-bang protocol with N_T time steps requires the multiplication of N_T unitaries of size 2^L . Since this is the bottleneck in the algorithm run-time we optimized the fidelity evaluation by taking into account symmetries of the qubit Hamiltonian (Eq. 1 of main text), such as translation and reflection (parity), and precomputing and storing in memory a subset of the products of unitaries.

Algorithm S1 Stochastic descent (SD_k)

```

1: Input:  $N_T, T, k$ 
2: Routines: FindAllUpdate, RandomShuffle, UpdateProtocol, Fid
3: initialize:
4:    $\mathbf{h}_{old} \leftarrow \{h(1), h(\delta t), \dots, h(N_T \delta t)\} \sim \{-4, 4\}^{N_T}$  ▷ Initialize protocol at random
5:  $\text{ListOfAllUpdates} \leftarrow \text{FindAllUpdate}(N_T, k)$  ▷ Finds the list of updates with at most  $k$ -flips
6: shuffle:
7:    $\text{ListOfAllUpdates} \leftarrow \text{RandomShuffle}(\text{ListOfAllUpdates})$  ▷ Shuffle updates in a random order
8: for update in ListOfAllUpdates do ▷ Iterate over all possible update
9:    $\mathbf{h}_{new} \leftarrow \text{UpdateProtocol}(\mathbf{h}_{old}, \text{update})$  ▷ Update protocol given the specified update
10:  if  $\text{Fid}(\mathbf{h}_{new}) > \text{Fid}(\mathbf{h}_{old})$  then ▷ Evaluates the fidelity of each protocol and compares them
11:     $\mathbf{h}_{old} \leftarrow \mathbf{h}_{new}$ 
12:    goto shuffle ▷ If update accepted, then restart for loop
return  $\mathbf{h}_{old}$ 

```

S2. ALGORITHMIC COMPLEXITY AND SCALING OF THE NUMBER OF LOCAL MINIMA

In order to verify the exponential nature of the glassy phase (i.e. verify whether there are exponentially many local minima in the glassy phase), we measure the algorithmic complexity scaling of the stochastic descent algorithms used (SD_k , $k = 1, 2, 4$), see Fig. ??). Specifically, we are interested in the computational effort required by SD_k to find the optimal protocol as a function of N_T . In order to certify the optimal fidelity protocol, we performed a brute-force search over all possible 2^{N_T} protocols. For this reason (but also because the complexity (see below) is exponentially hard to measure in the glassy phase) we were limited to effective system sizes of $N_T \leq 28$.

The algorithmic complexity of stochastic descent is measured by estimating the number fidelity evaluations required on *average* to locate the optimal protocol with high-probability (w.h.p.). Thus, we define the complexity as the expected number of fidelity evaluations per stochastic descent run, $\langle n_{\text{eval}} \rangle$, *multiplied* by the number of random

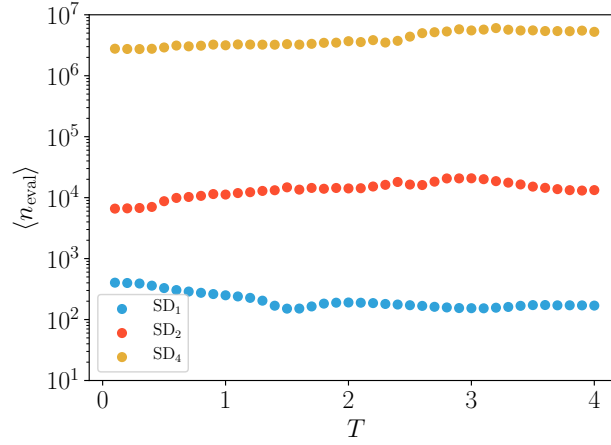


FIG. S1: Average number of protocol fidelity evaluation per stochastic descent run required to reach a local minimum of the log-fidelity landscape as a function of the protocol duration T . This is computed for $N_T = 80$ and $L = 6$ with $\mathcal{O}(10^5)$ samples for each time and SD_k .

initializations needed until on average one stochastic descent local minimum corresponds to the global minimum w.h.p.:

$$\mathcal{C} = \frac{\langle n_{\text{eval}} \rangle}{p(h(t) = h^*(t))}. \quad (\text{S1})$$

Here, $p(h(t) = h^*(t))$ is the probability that the local minimum found for a *single* stochastic descent run corresponds to the global minimum. Importantly, we expect that the scaling of the complexity *for stochastic descent* also reveals whether the number of local minima is exponential or sub-exponential in N_T (see Fig. ??). The results are shown Fig. ??) and were computed using sampling (hence the error bars) due to computational limitations. Note that the “ruggedness” exhibited by the curves in Fig. ??) is mostly due to finite size effects (as opposed to sampling noise). The “ruggedness” increases with k but also with T .

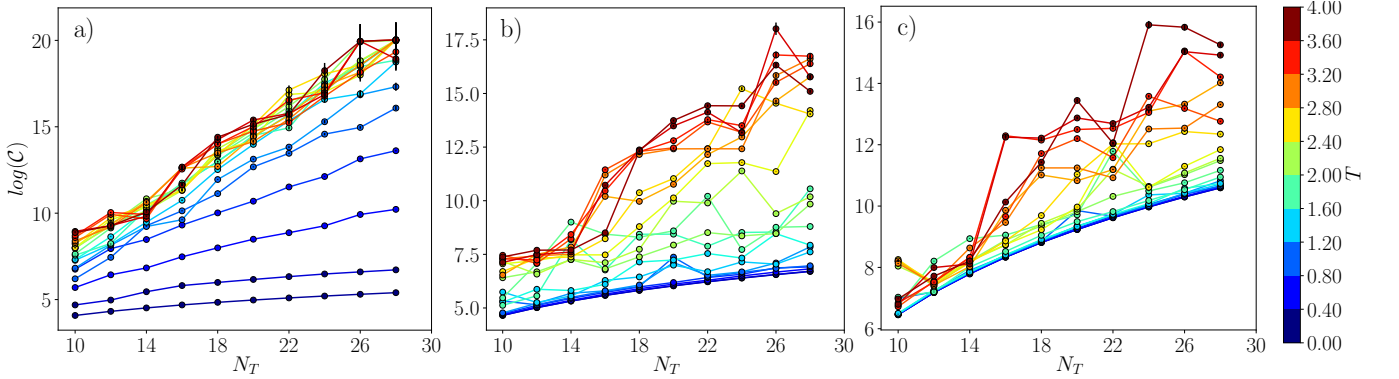


FIG. S2: The log-complexity (see Eq. (??)) of finding the optimal fidelity protocol as a function of the effective system size N_T for various protocol durations $T \in [0.1, 0.3, \dots, 3.9]$ (see color bar). Error bars are shown for every data point. (a), (b), (c) : complexity for SD_1 , SD_2 , SD_4 respectively. The value $T_c^{(k)}$ at which the logarithm of the complexity goes from a logarithmic to a linear scaling in N_T indicates the transition to a rugged landscape with exponentially many local minima. In the main text, the $T_c^{(k)}$ was defined as the point where $f_{\text{SD}_k} \approx 0.5$: we found $T_c^{(1)} \approx 0.35$, $T_c^{(2)} \approx 2.3$, $T_c^{(4)} \approx 2.5$.

S3. t -SNE: T-DISTRIBUTED STOCHASTIC NEIGHBOR EMBEDDING

t -SNE [?] is a non-parametric dimensional reduction method that can be used to effectively embed high-dimensional data in a two-dimensional space. The t -SNE embedding is obtained by minimizing a cost function which emphasizes the

conservation of local ordination (short-distance information) of the data while downplaying long-distance information. In this work, we used t -SNE with a perplexity of 60 and a Barnes-Hut angle of 0.5. We computed multiple t -SNE runs to make sure that the results were insensitive (up to global rotations) to the random seed.

From the t -SNE results, we used density clustering (similar to DBSCAN) in order to accurately identify the clusters. Density clustering does not require to specify the number of clusters as would be the case using K -means for instance. The explicit clusters found using density clustering are shown in Fig. ?? . In order to further corroborate the presence of clusters, we computed the mean inter-distance (in the original space of protocols) for every pair of protocols in between the clusters found (see Fig. ??).

We verified that the clusters found by t -SNE are indeed separated by extensive barriers by computing the Hamming distance matrix of the sampled protocols (see Fig. 3 of the main text and Fig. ??). Note that the scales (axis values) in t -SNE maps can sometimes be misleading and one should NOT compare the scales *in between* t -SNE plots. In the main text, Fig. 3.c, while it may seem that all protocols (local minima) are closeby, computing the exact pairwise distances (see Fig. 3.d-f) reveals that in fact all protocols are almost maximally distant (the distribution is peaked at 0.5). Indeed, if two protocols are drawn at uniformly at random, they will on average be separated by a (N_T normalized) Hamming distance of 0.5. This can be seen by inspecting the Hamming distance matrix in Fig. 3.f. In Fig. ?? we plot the distribution of the pairwise distances for the distance matrices presented in the main text.

Last, note that many software packages that implement t -SNE in different programming languages have been written and we refer the enthusiastic reader to consult <https://lvdmaaten.github.io> for more information and for the version that we used. We implemented an easy-to-use Python wrapper available on https://github.com/alexandreday/tsne_visual. For a more detailed discussion about the use of t -SNE we encourage the reader to consult Ref. [?] .

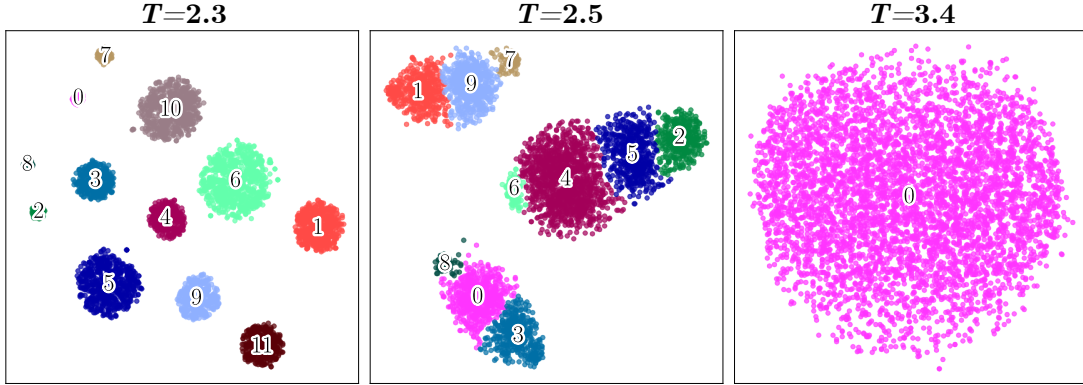


FIG. S3: Density clustering of the t -SNE embedding. We used a modern density clustering based approach based on [?] and available on https://github.com/alexandreday/fast_density_clustering. Each protocol corresponds to a point on the 2D t -SNE embedding and is assigned a cluster label by the clustering algorithm. Protocols belonging to the same cluster were grouped together in Fig. 3 of the main text. $N_T = 200, L = 6$ and we sampled 5000 unique protocols using SD_2 .

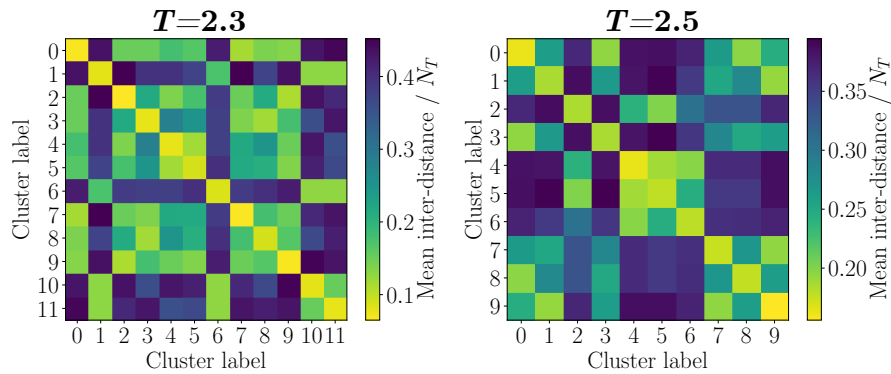


FIG. S4: Mean inter-distance for each pair of clusters found and labelled in Fig. ?? : for every pair of protocols within two distinct clusters, we compute the Hamming distance between. The mean inter-distance is then obtained by averaging over all protocol pairs. $N_T = 200, L = 6$ and we sampled 5000 unique protocols using SD_2 .

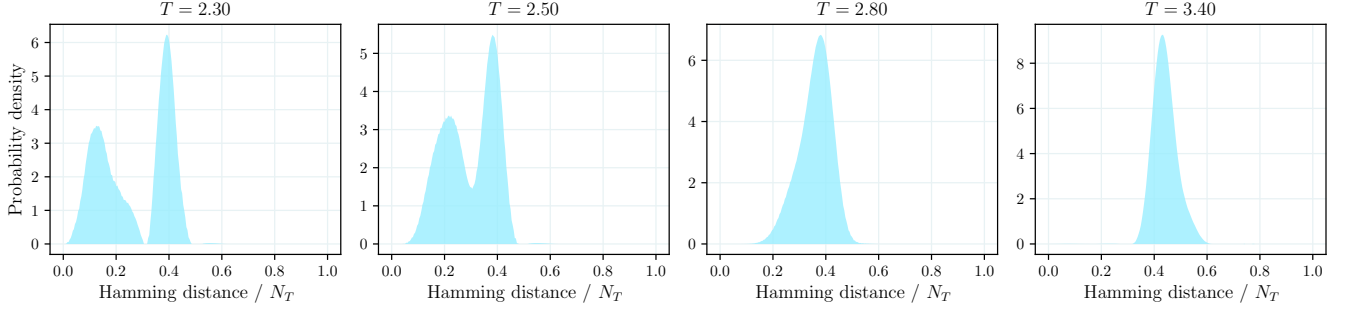


FIG. S5: Distribution of the pairwise Hamming distances for $M > 10000$ sampled local minima protocols using SD_2 at protocol durations $T = 2.3, 2.5, 2.8, 3.4$ with $N_T = 200$ and $L = 6$. At $2.3 \approx T_c^{(2)} \lesssim T \lesssim 2.55$, the distribution is bi-modal, indicating the presence of well-separated clusters : there is a small number of clusters with high entropy of protocols that are connected by small barriers. At longer protocol duration times, tightly concentrated clusters fracture into an exponential number of clusters (with a low-entropy) separated by extensive barriers. The latter is seen from the fact that the mode of the distribution approaches 0.5.

S4. FINITE-SIZE SCALING OF THE DENSITY OF STATES AND ELEMENTARY EXCITATIONS

Phase transitions appear in the TD limit, which we can also define for $\mathcal{H}_{\text{eff}}(T)$: to add more degrees of freedom to the classical model, we increase the number of bangs $N_T \rightarrow \infty$, which requires sending $\delta t \rightarrow 0$ to keep the protocol duration $T = N_T \delta t$ fixed. We refer to the thermodynamic limit of the effective model $\mathcal{H}_{\text{eff}}(T)$ as the ‘continuum limit’, to distinguish it from the TD limit of the physical quantum many-body system $L \rightarrow \infty$.

Figure ?? shows the finite-size scaling of the normalized density of states and the elementary excitations on top of the optimal protocol h_j^* as we vary the physical system size L . Note that the log-fidelity $\log F_h(T) \sim L$ is an extensive quantity. Therefore, to carry out the scaling, we consider

$$C_h(T) = -\frac{1}{L} \log F_h(T) \quad (\text{S2})$$

which remains finite in the limit $L \rightarrow \infty$. The plots were generated by computing the fidelity for all 2^{N_T} protocols for $N_T = 28$. The collapse of the curves suggests that the physics of the optimization problem is close to the TD limit. This is consistent with similar results obtained in Ref. [?].

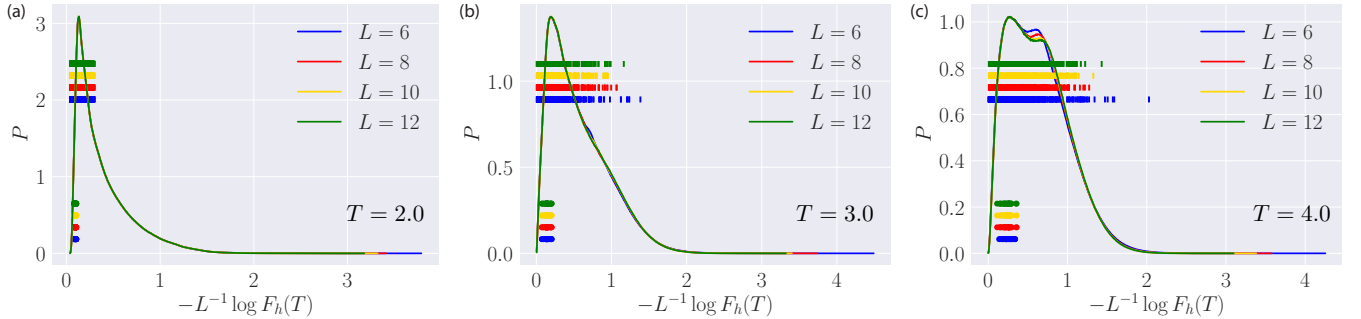


FIG. S6: Finite-size scaling of the normalized DOS as a function of the number of qubits L , shows the system is close to the thermodynamic limit $L \rightarrow \infty$. The positions of the local elementary excitations on top of the optimal protocol w.r.t. the fidelity axis are marked in circles for 1-flip excitations (lower batch) and vertical bars for 2-flip excitations (upper batch), respectively. The position of the excitations along the y -axis carries no meaning. All plots are based on the exact computation of the fidelity for all 2^{N_T} protocols for $N_T = 28$.

Figure ?? shows the scaling of the normalized density of states against the number of bangs N_T , i.e. the system size of the effective model $\mathcal{H}_{\text{eff}}(T)$. These results are computed using the full set of 2^{N_T} protocols for $N_T \leq 28$. While the results presented are exact for these value of N_T , because we are limited by computational bottlenecks to $N_T \leq 28$, the results display some finite-size effects in the log-fidelity of the elementary excitations. In order to access greater system sizes (in N_T), we used stochastic descent (to find good fidelity protocols) along with uniform sampling of the

protocols (to compute the black curve in Fig. 4 of the main text). Finally, we remark that the shape of the DOS responds only weakly to increasing N_T .

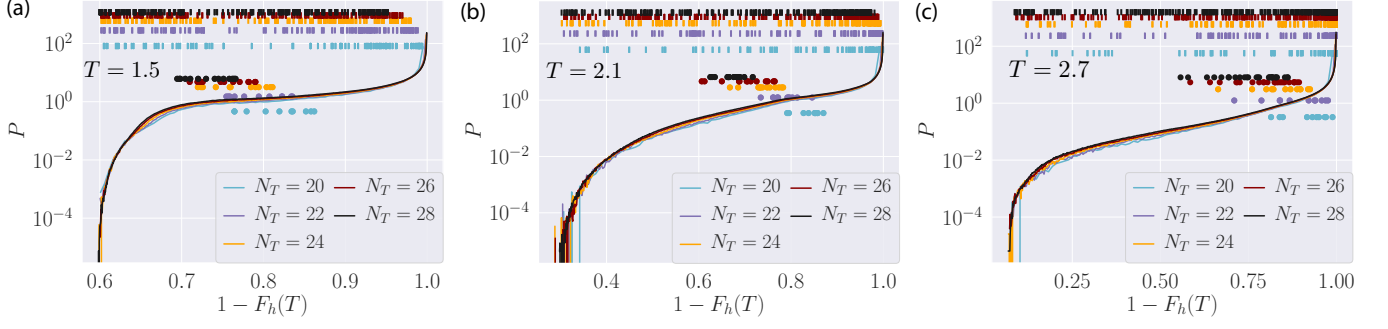


FIG. S7: Finite-size scaling of the normalized DOS as a function of the number of bangs N_T . The positions of the local elementary excitations on top of the optimal protocol w.r.t. the fidelity axis are marked in circles (1-flip) and vertical bars (2-flip), respectively. The position of the excitations along the y -axis carries no meaning. The DOS data is obtained from fits of histograms, hence the noise.

S5. FINITE-SIZE SCALING OF THE ORDER PARAMETERS $q(T)$ AND $f(T)$.

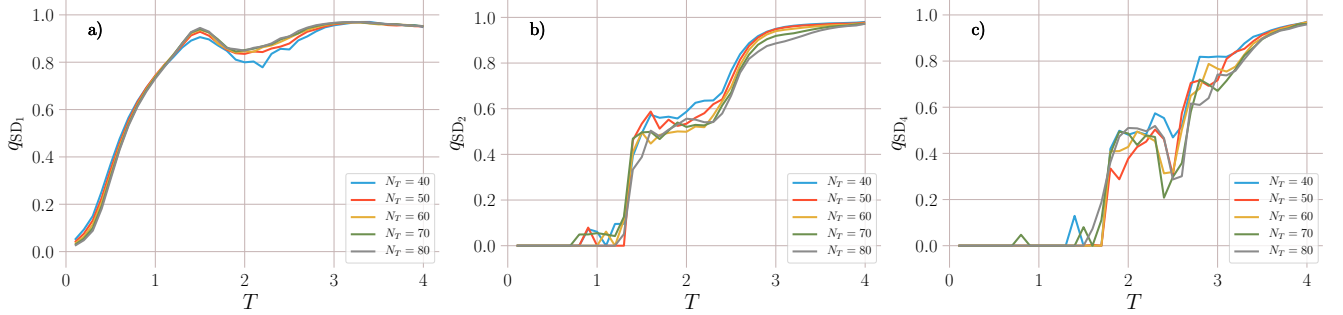


FIG. S8: Correlator ($q(T)$) scaling with the effective system size $N_T \in \{40, 50, 60, 70, 80\}$ as a function of the protocol duration T . (a)-(b)-(c) : the correlator for SD_1 , SD_2 and SD_4 respectively. Here we used $L = 6$. The correlation is computed from the sampled protocols that have a fidelity greater than 95% of the best encountered protocol fidelity over the whole sampling at a fixed T . We sampled at least 10000 protocols for every N_T and protocol duration T .

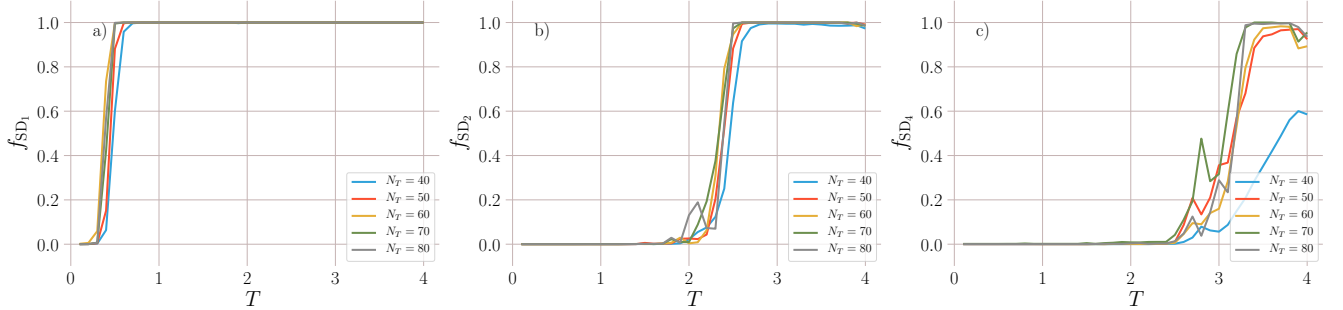


FIG. S9: Local minima fraction ($f(T)$) scaling with the effective system size $N_T \in \{40, 50, 60, 70, 80\}$ as a function of the protocol duration T . (a)-(b)-(c) : the correlator for SD_1 , SD_2 and SD_4 respectively. Here we used $L = 6$. The correlation is computed from the sampled protocols that have a fidelity greater than 95% of the best encountered protocol fidelity over the whole sampling at a fixed T . We sampled at least 10000 protocols for every N_T and protocol duration T .

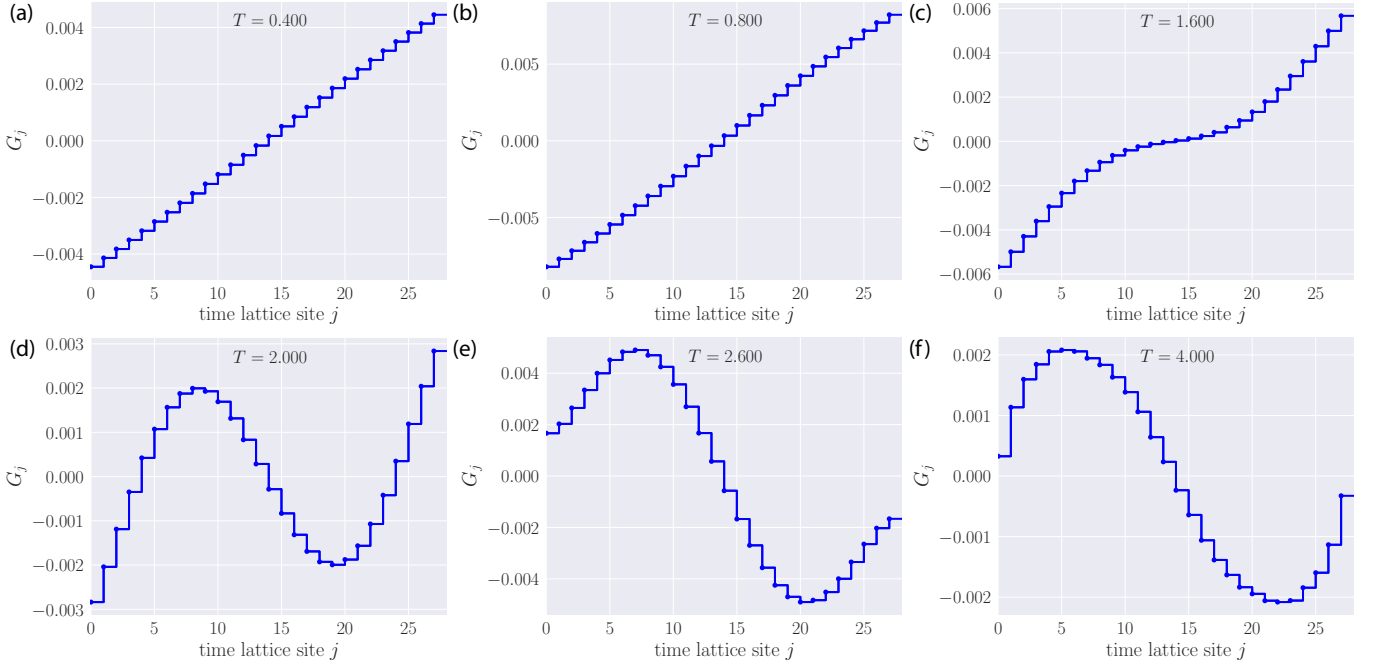


FIG. S10: Spatial dependence of the single-spin (‘onsite magnetic field’) term in the exact effective Hamiltonian $H_{\text{eff}}(T)$. The time lattice sites correspond to the bangs of the bang-bang protocols used to prepare the state. The parameters are $N_T = 28$, $L = 6$.

S6. THE EFFECTIVE CLASSICAL SPIN MODEL

The quantum state preparation problem is an optimisation task which, over the space of bang-bang protocols, is equivalent to finding the ground state of a classical spin model. As we discussed in the main text, each bang-bang protocol can be thought of as a classical Ising spin state, while the discrete time points are mapped onto lattice sites. The TD limit of the effective classical spin model coincides with the continuum limit of the discrete time evolution: $N_T \rightarrow \infty$, $\delta t \rightarrow 0$ with $\delta t N_T = T = \text{const.}$

An intriguing and natural question arises as to what the underlying classical spin model $\mathcal{H}_{\text{eff}}(T)$ actually looks like. This classical spin energy function governs the phase transitions of the quantum control problem, and our interpretation of the latter in terms of spin phases can potentially benefit if one is able to extract some useful information about the properties of the different terms in $\mathcal{H}_{\text{eff}}(T)$. For instance, information about the form of its couplings is contained in the (negative) log-fidelity spectrum, i.e. the log-fidelities corresponding to all possible bang-bang configurations. In this section, we work with the log-fidelity $C_h(T) = -L^{-1} \log F_h(T)$, which depends only weakly on the system size L .

When it comes to studying control phases, of particular interest is to determine the locality properties of the dominant spin couplings. For this purpose, we make the following ansatz for the most general form of the energy function of a two-state classical spin degree of freedom $h_j \in \{\pm 1\}$:

$$\mathcal{H}_{\text{eff}}(T) = C_0(T) + \sum_j G_j(T) h_j + \frac{1}{N_T} \sum_{i \neq j} J_{ij}(T) h_i h_j + \frac{1}{N_T^2} \sum_{i \neq j \neq k} K_{ijk}(T) h_i h_j h_k + \dots \quad (\text{S3})$$

A. Exact Coupling Strengths

Consider first a generic k -body interaction term. If $k = 0$, this is just the constant log-fidelity offset C_0 ; for $k = 1$, we have an effective magnetic field G_j , while the $k = 2$ case can be interpreted as a two-body interaction J_{ij} , and so forth. If we consider a protocol of N_T time steps (bangs), there are a total of $\{h_{j;s}\}_{s=1}^{2^{N_T}}$ different protocol configurations, each of which comes with its own log-fidelity $C_s(T)$. Here the index s runs over all 2^{N_T} protocol configurations. Suppose we know the entire exact log-fidelity spectrum, but not the underlying effective energy function \mathcal{H}_{eff} . One

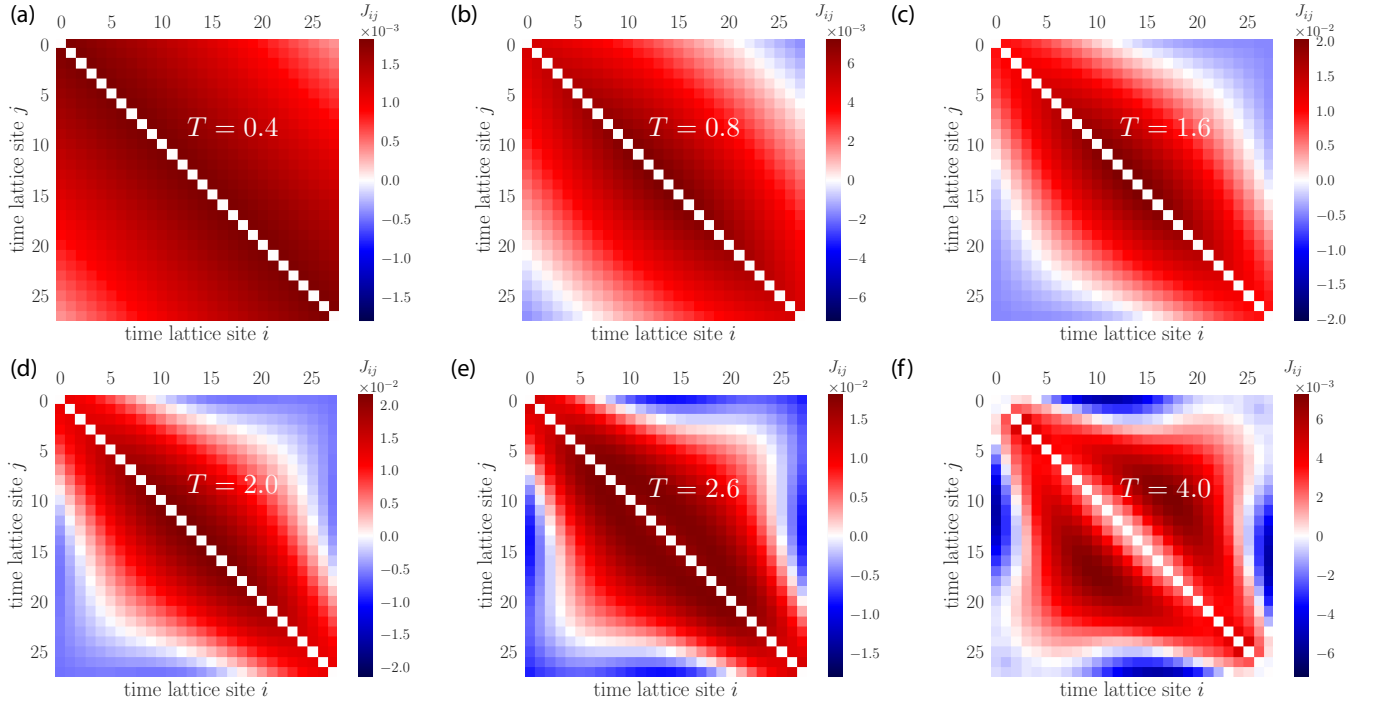


FIG. S11: Spatial dependence of the two-body interaction term in the exact effective Hamiltonian $H_{\text{eff}}(T)$. The time lattice sites correspond to the bangs of the bang-bang protocols used to prepare the state. The parameters are $N_T = 28$, $L = 6$.

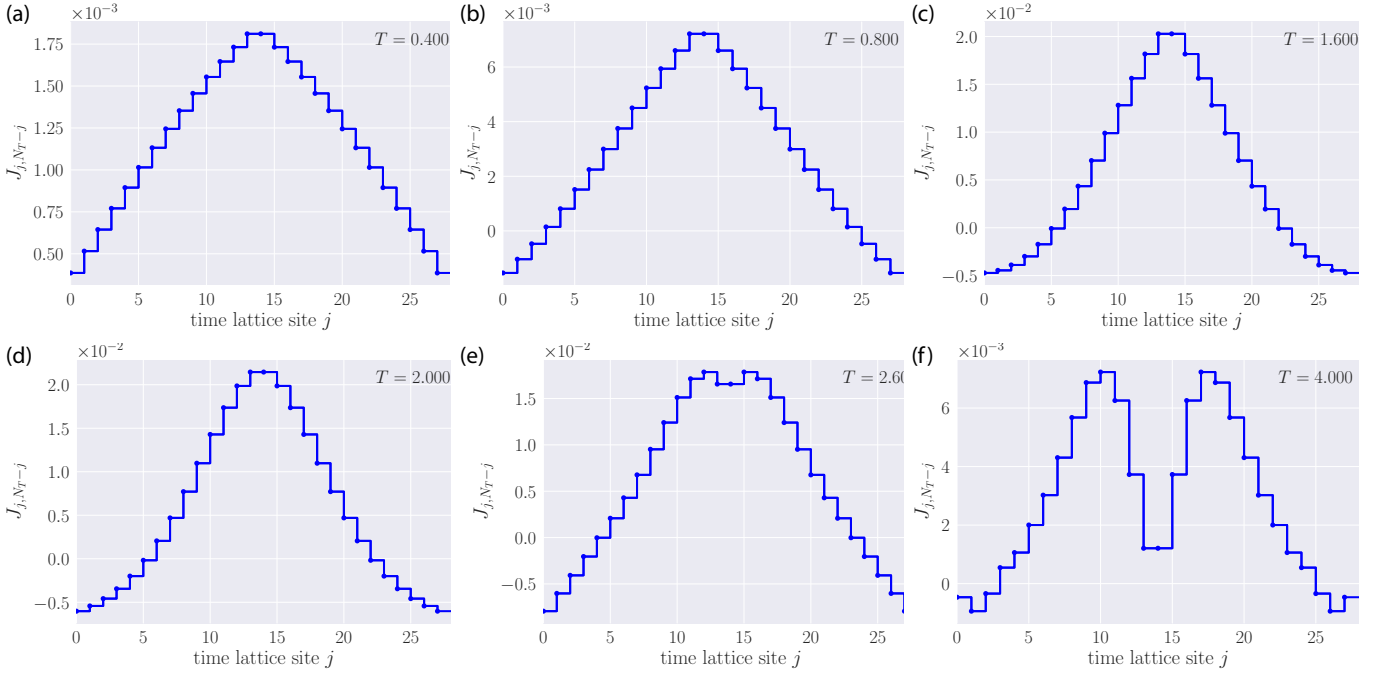


FIG. S12: Decay of the two-body interaction terms J_{ij} along the anti-diagonal, c.f. Fig. ???. The time lattice sites correspond to the bangs of the bang-bang protocols used to prepare the state. The parameters are $N_T = 28$, $L = 6$.

can compute all protocols and their log-fidelities numerically for up to $N_T = 28$ bangs. Then, one can convince oneself

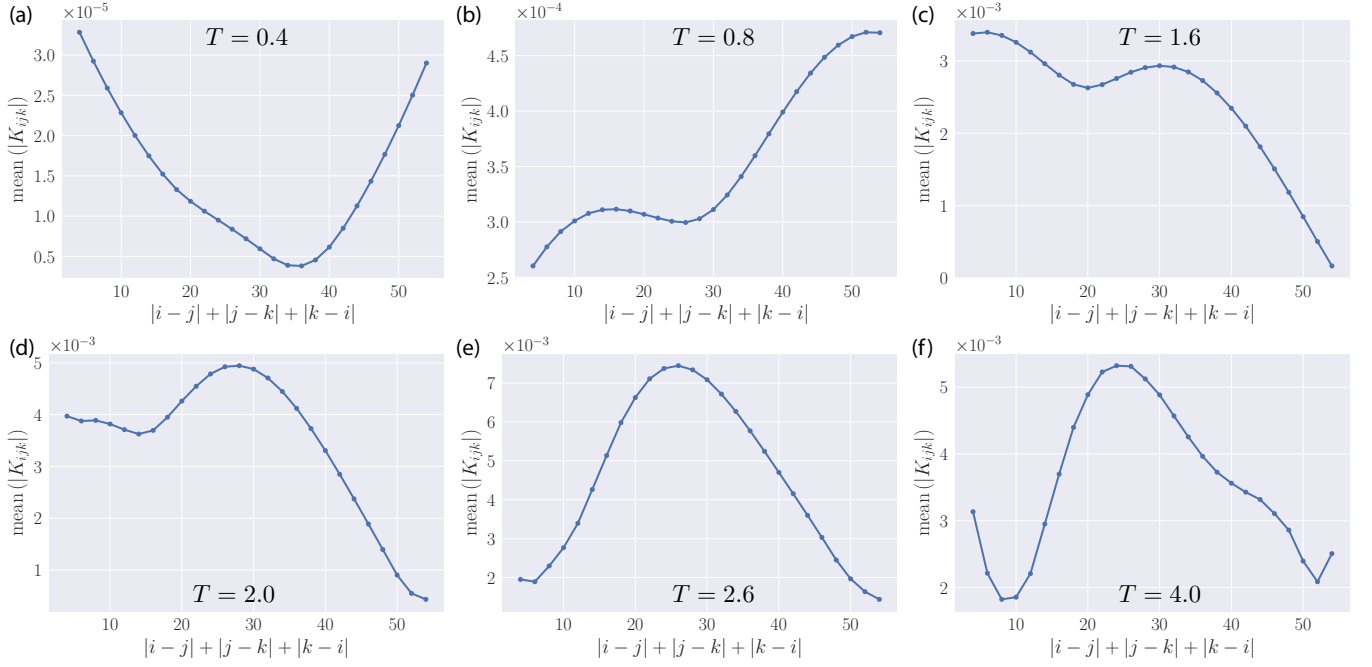


FIG. S13: Non-locality of the three-body interaction term in the exact effective Hamiltonian $H_{\text{eff}}(T)$. The parameters are $N_T = 28$, $L = 6$.

that all couplings of the effective spin model $\mathcal{H}_{\text{eff}}(T)$ can be uniquely determined from the following expressions:

$$\begin{aligned}
 I_0(T) &= \frac{1}{2^{N_T}} \sum_{s=1}^{2^{N_T}} C_s(T), \\
 G_j(T) &= \frac{1}{1!} \frac{1}{2^{N_T}} \sum_{s=1}^{2^{N_T}} C_s(T) h_{j;s}, \\
 J_{ij}(T) &= \frac{N_T}{2!} \frac{1}{2^{N_T}} \sum_{s=1}^{2^{N_T}} C_s(T) h_{i;s} h_{j;s}, \quad i \neq j \\
 K_{ijk}(T) &= \frac{N_T^2}{3!} \frac{1}{2^{N_T}} \sum_{s=1}^{2^{N_T}} C_s(T) h_{i;s} h_{j;s} h_{k;s}, \quad i \neq j, i \neq k, j \neq k \\
 &\dots
 \end{aligned} \tag{S4}$$

and analogously for the higher-order terms. To derive these expressions, we note that in the set of all protocols $h_j/4 = \pm 1$ an equal number of times. For this reason, $\sum_s h_{j;s} = 0$ at any fixed time step (time-lattice site) j . Combining this observation with the fact $h_j^2/16 = 1$ yields the expressions above.

For the sake of simplicity and tractability, we can truncate the effective spin model, keeping all possible n -body interactions ($n = 1, 2, 3$), and neglect any higher-order ones. This leads to an approximate classical spin model which

we denote by $\mathcal{H}_{\text{approx}}^{(n)}(T)$. For instance:

$$\begin{aligned}\mathcal{H}_{\text{approx}}^{(1)}(T) &= C_0(T) + \sum_j G_j(T) h_j, \\ \mathcal{H}_{\text{approx}}^{(2)}(T) &= C_0(T) + \frac{1}{N_T} \sum_{i \neq j} J_{ij}(T) h_i h_j, \\ \mathcal{H}_{\text{approx}}^{(1+2)}(T) &= C_0(T) + \sum_j G_j(T) h_j + \frac{1}{N_T} \sum_{i \neq j} J_{ij}(T) h_i h_j, \\ \mathcal{H}_{\text{approx}}^{(1+2+3)}(T) &= C_0(T) + \sum_j G_j(T) h_j + \frac{1}{N_T} \sum_{i \neq j} J_{ij}(T) h_i h_j + \frac{1}{N_T^2} \sum_{i \neq j \neq k} K_{ijk}(T) h_i h_j h_k,\end{aligned}\tag{S5}$$

and so forth. We emphasize that G_j , J_{ij} and K_{ijk} are the exact coupling strengths which depend parametrically on the protocol duration T but are independent of the truncation order n , see Eq. (??). To quantify each of these approximate Hamiltonians, we define the mean error

$$\mathcal{E}(T) = \frac{1}{2^{N_T}} \sum_{s=1}^{2^{N_T}} |C_s(T) - C_{s,\text{approx}}^{(n)}(T)|,\tag{S6}$$

where C_s denotes the exact and $C_{s,\text{approx}}^{(n)}$ – the approximate log-fidelity to order n , respectively.

In the following we always restrict the total number of bangs to $N_T = 28$, and vary the total protocol duration T . Figure ?? shows the exact effective on-site field strength $G_j(T)$ for six protocol durations T . It follows that the optimal protocol (i.e. the lowest energy configuration) of $\mathcal{H}_{\text{approx}}^{(1)}$ for $T < T_c$ is a single step at time $T/2$. Interestingly, this is precisely the form of the optimal protocol in the overconstrained phase [?]. This is backed up by Fig. ??a (blue line), which shows that the mean error generated by $\mathcal{H}_{\text{approx}}^{(1)}$ is indeed smallest in the overconstrained control phase. In the glassy phase for $T > T_c$, however, the form of the G_j field changes gradually, and the higher-order terms become more important.

Figures ?? and ?? show the exact effective spin-spin interaction $J_{ij}(T)$. First, we notice that it is not sparse, but features finite all-to-all couplings on all the bonds. Moreover, $J_{ij}(T)$ keeps the same sign over large portions of neighbouring spin bonds (i, j) for all T we consider. Therefore, we can anticipate that it would hardly be possible for the optimal protocol to satisfy all two-body couplings simultaneously. To quantify the bond satisfiability of the optimal protocol, let us define the frustration parameter inspired by the k -SAT problem and spin-glass physics:

$$\Phi(T) = \frac{\min_s [C_s^{(n)}(T)] + \sum_{i_1, \dots, i_n} |\tilde{J}_{i_1 \dots i_n}(T)|}{\sum_{i_1, \dots, i_n} |\tilde{J}_{i_1 \dots i_n}(T)|},\tag{S7}$$

where $\tilde{J}_{i_1 \dots i_n}$ denotes all couplings up to and including order n , and $|\cdot|$ denotes the absolute value. The tilde means that the couplings $\tilde{J}_{i_1 \dots i_n}$ are first normalized such that $h_j \in \{\pm 1\}$. With this definition, the frustration parameter $\Phi(T)$ is normalized between zero and unity with zero signaling no frustration and unity – maximum possible frustration, respectively. To gain intuition for this quantity, notice that the energy function \mathcal{H}_{eff} is a sum of terms, each with its own coupling $\tilde{J}_{i_1 \dots i_n}$. In the absence of frustration, we can find a global minimum of the energy by finding the minimum for each term in \mathcal{H}_{eff} , and $\sum_{i_1, \dots, i_n} |\tilde{J}_{i_1 \dots i_n}(T)|$ will be equal to the negative minimum energy. Since the first term in the numerator of $\Phi(T)$ is just the energy of the lowest spin configuration, in the absence of frustration the two terms in the numerator will cancel and $\Phi(T) = 0$. In general, we can measure frustration by asking how different minimizing the individual terms of \mathcal{H}_{eff} is, from the true minimum of the sum of terms. This is what the quantity $\Phi(T)$ measures. Figure ??b demonstrates the high degree of frustration in the effective model.

Including the all-to-all three-body interaction term $K_{ijk}(T)$ improves the error \mathcal{E} in the log-fidelity spectrum only marginally, cf. Fig. ??a. This means that even more complicated higher-order multi-body terms are needed in order to fully capture the underlying physics of the effective model. To show the non-locality of the mean three-body interactions $\text{mean}(|K_{ijk}|)$, we adopt the following measure: (i) we fix the perimeter of the triangle spanned by three interacting spins $|i - j| + |j - k| + |k - i|$, (ii) we categorize all three-spin interactions according to this perimeter, and (iii) we compute the mean of their absolute value. The result is shown in Fig. ??. We attribute this nonlocal behaviour to the original quantum state preparation problem being causal: in other words, the value of the protocol at a later time depends on all possible values taken at previous times (up to symmetries). The complexity of the effective classical spin model is also sustained by the frustration parameter, which remains high upon adding the higher-order terms, see Fig. ??b.

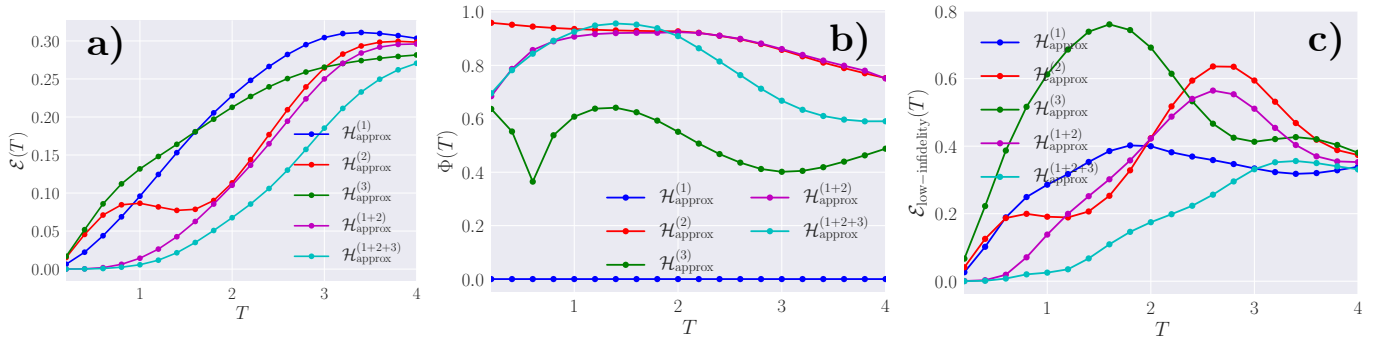


FIG. S14: a) Average difference between the spectrum of the truncated effective models and the exact fidelity spectrum as a function of the protocol duration T . b) Frustration order parameter for the truncated effective models as a function of the protocol duration T . c) Same as a), but with the effective model computed only on the 200000 lowest log-fidelity protocols. The legends show which terms are kept in the effective Hamiltonian (see text). The parameters used are $N_T = 28$, $L = 6$.

Finally, in Fig. ??c, we computed the effective couplings by summing only over the lowest 200000 log-fidelity protocols. As expected, the computed error increases w.r.t. to Fig. ??a.

B. Better-Fidelity (Low log-Fidelity) Coupling Strengths

Many properties of the phases of classical spin systems are usually determined by their low-energy states. Therefore, one might wonder if, despite the results presented above, a *local* effective spin model $\mathcal{H}_{\text{ML}}^{(n)}(T)$ still exists, which captures only the physics of the better-fidelity states [corresponding to the low negative log-fidelity part of the spectrum of the exact $\mathcal{H}_{\text{eff}}(T)$]. In other words, at least in principle, there exists the possibility that the completely non-local character of the exact effective couplings discussed above originates from bad-fidelity states, and we want to rule that out. This check is important, since the number of the bad-fidelity states is exponentially large in N_T , compared to that of the good-fidelity states, as a consequence of any two randomly chosen quantum states being with high probability orthogonal in the high-dimensional Hilbert space of the quantum many-body system.

To address this concern, we order the protocols [classical spin states] according to their fidelities, and impose a cut-off, keeping only the better fidelities [in practice, we keep the best 2×10^5 out of a total of 2^{28} protocols for $N_T = 28$ bangs]. In the following, we shall refer to this set as the low log-fidelity manifold (in analogy with the low-energy manifold of classical spin models). Based on this data, we employ ideas from Machine Learning to learn only those properties of the coupling of the exact effective spin-energy model $\mathcal{H}_{\text{eff}}(T)$ [? ? ?], which influence the low log-fidelity manifold. The learning problem being linear in the coupling strengths, we can employ Ridge and Lasso regression to fit an effective spin-energy model to the log-fidelity data. While Ridge regression assumes that the resulting learned coupling strengths are all-to-all, Lasso is particularly suited for finding sparse couplings. The ML model $\mathcal{H}_{\text{ML}}^{(n)}(T)$ differs from the effective one in the previous section, in that it should approximately reproduce only the better fidelities, while all information about the bad protocols is discarded.

Starting with the set of better fidelities and the corresponding states, we divide it into a training and a test data set in proportion 5:3. We train our ML models $\mathcal{H}_{\text{ML}}^{(n)}(T)$ using an L^2 -cost function, trying out different regularisation strength hyperparameters, and select the one which results in the best performance [?]. The model is trained by only using the training data, while we measure its performance on both the training and test data sets. We denote the exact log-fidelities by C_s , and the predicted ones – by $C_{s,\text{ML}}^{(n)}$, where s here runs up to the data set size N . The measure of performance is the quantity R^2 , defined as

$$R^2(T) = 1 - \frac{\sum_{s=1}^N \left| C_s(T) - C_{s,\text{ML}}^{(n)}(T) \right|^2}{\sum_{s=1}^n \left| C_s(T) - N^{-1} \sum_{s=1}^N C_{s,\text{ML}}^{(n)}(T) \right|^2}. \quad (\text{S8})$$

This quantity is unity if the predicted data matches the true data, while its deviation from unity quantifies how good the ML model is. We evaluate R^2 for the Ridge and Lasso regression on both the training and the test data, see Fig. ??. In general, one can expect two types of behavior: (i) if the training (solid line) and the test (dashed line) values of R^2 are not on top of each other, it means the model failed to learn the correct properties of the data which determine the physics, most likely due to overfitting. (ii) sometimes the train and test curves are indeed very close,

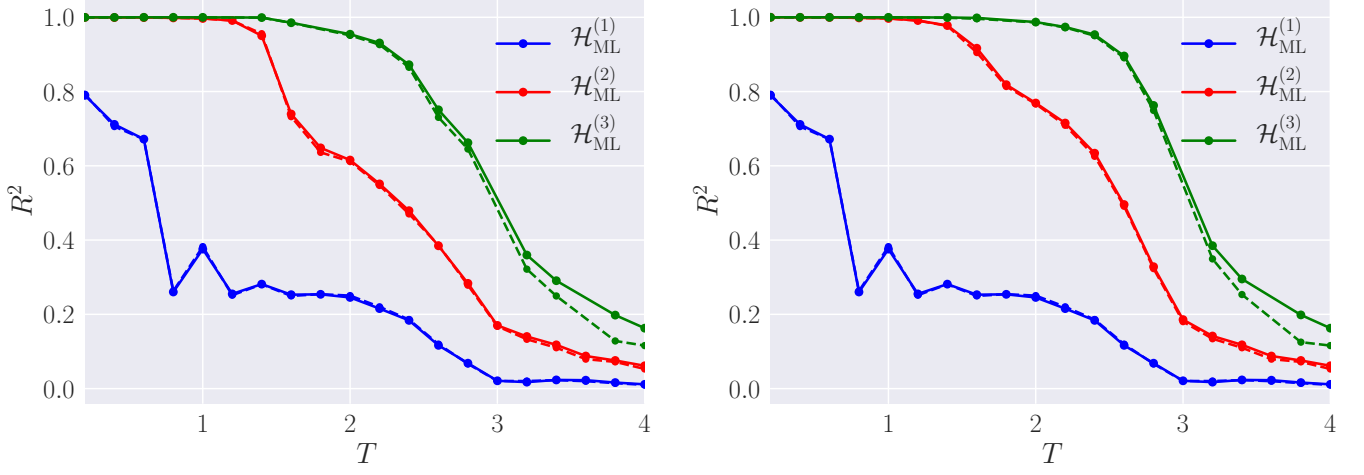


FIG. S15: R^2 of model performance vs. the protocol duration T for the three model classes considered for Lasso (left) and Ridge regression (right). The solid and dashed lines correspond to the training and test data, respectively. The similarity between the test and training R^2 -curves shows the low degree of overfitting. The parameters are $N_T = 28$, $L = 6$.

but the value of R^2 deviates from unity. This means that the true model which generated the data points lies outside the model class we assumed to look for a solution in. In other words, the model learned everything there is to learn about the data *within the model class under consideration*. In our problem, since the true model (??) is linear in the coupling strengths, this suggests that we have to include higher-order multi-body interaction terms in $\mathcal{H}_{\text{ML}}^{(n)}(T)$, i.e. consider $n > 3$.

We apply Machine Learning with three different model types: (i) a non-interacting spin model $\mathcal{H}_{\text{ML}}^{(1)}(T)$ where we only learn the on-site local field values $G_j(T)$, (ii) a more sophisticated interacting model $\mathcal{H}_{\text{ML}}^{(1+2)}(T)$ where we learn both the on-site $G_j(T)$ and the two-body interaction $J_{ij}(T)$, and (iii) a more general spin energy function $\mathcal{H}_{\text{ML}}^{(1+2+3)}(T)$ containing all possible one, two and three-body interactions $G_j(T)$, $J_{ij}(T)$ and $K_{ijk}(T)$. Figure ?? shows the $R^2(T)$ as a function of the protocol duration T in the three cases for the optimal regularization strength for Ridge and Lasso regression. First, note that all training and test curves are on top of each other, which means that our model learned the underlying correlations reliably. Second, notice that Ridge regression always outperforms the Lasso regression. Recalling that the Lasso regularization tries to enforce sparse couplings, this results backs up our conclusion that the effective classical spin model is non-local even when it comes to the low log-fidelity (i.e. better-fidelity) states. Last, observe how, at a fix protocol duration T in the glassy phase, enlarging the model type leads to a better performance, yet there is always a protocol duration at which the R^2 deviates significantly from unity. If we extrapolate this behavior, deeper in the glassy phase all multi-body interactions will most likely be required to keep the ML model performance close to unity. This means that, even for the set of better fidelities, it is insufficient to consider only local one and two-body terms in $\mathcal{H}_{\text{ML}}^{(n)}(T)$ for T in the glass phase. We thus conclude that the non-locality of the effective classical spin model is a property featured by the entire log-fidelity spectrum, and is not inflicted solely by the majority of bad protocols.

Figures ??, ?? and ?? display the better-fidelity coupling strengths obtained using ML with Ridge Regression for six protocol durations T . One can compare these to the exact couplings from Figs. ??, ?? and ??. Unlike the exact coupling strengths, they depend on the model class.

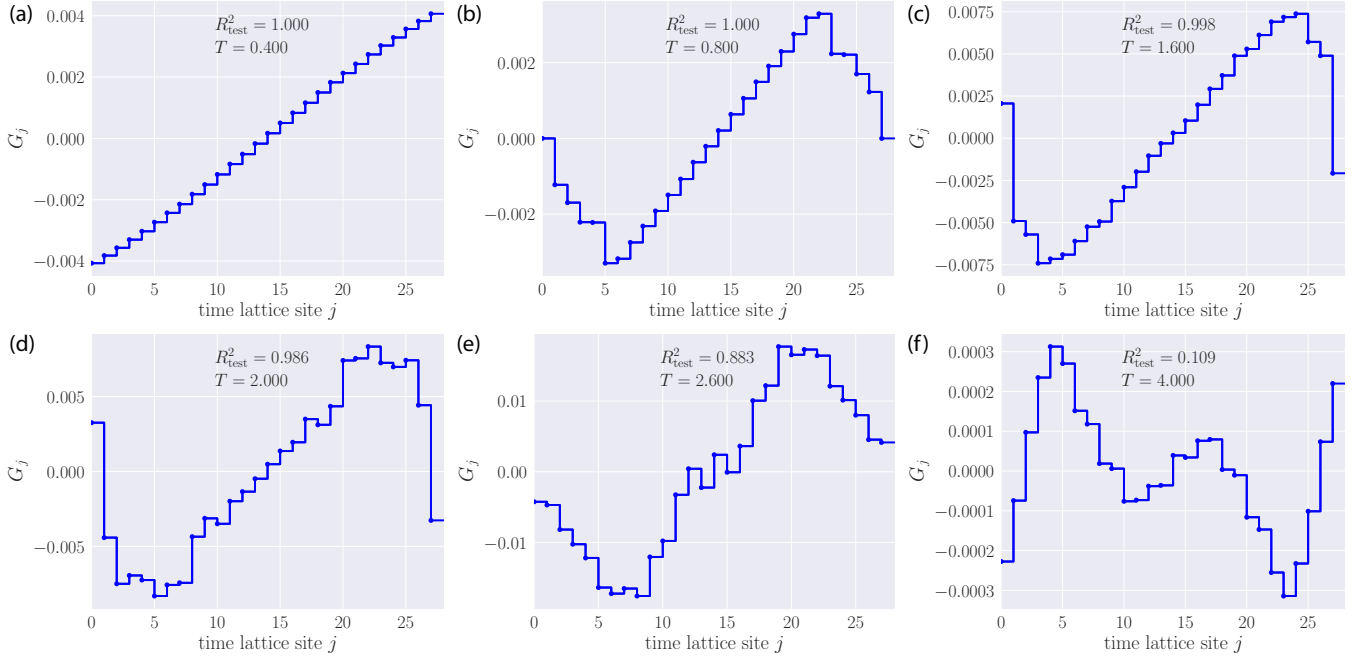


FIG. S16: Spatial dependence of the single-spin (‘onsite magnetic field’) term in the ML Hamiltonian $H_{\text{ML}}^{(1+2+3)}(T)$. The time lattice sites correspond to the bangs of the bang-bang protocols used to prepare the state. The parameters are $N_T = 28$, $L = 6$. The inverse Ridge regression regularisation strength is $\lambda = 10^{-9}$.

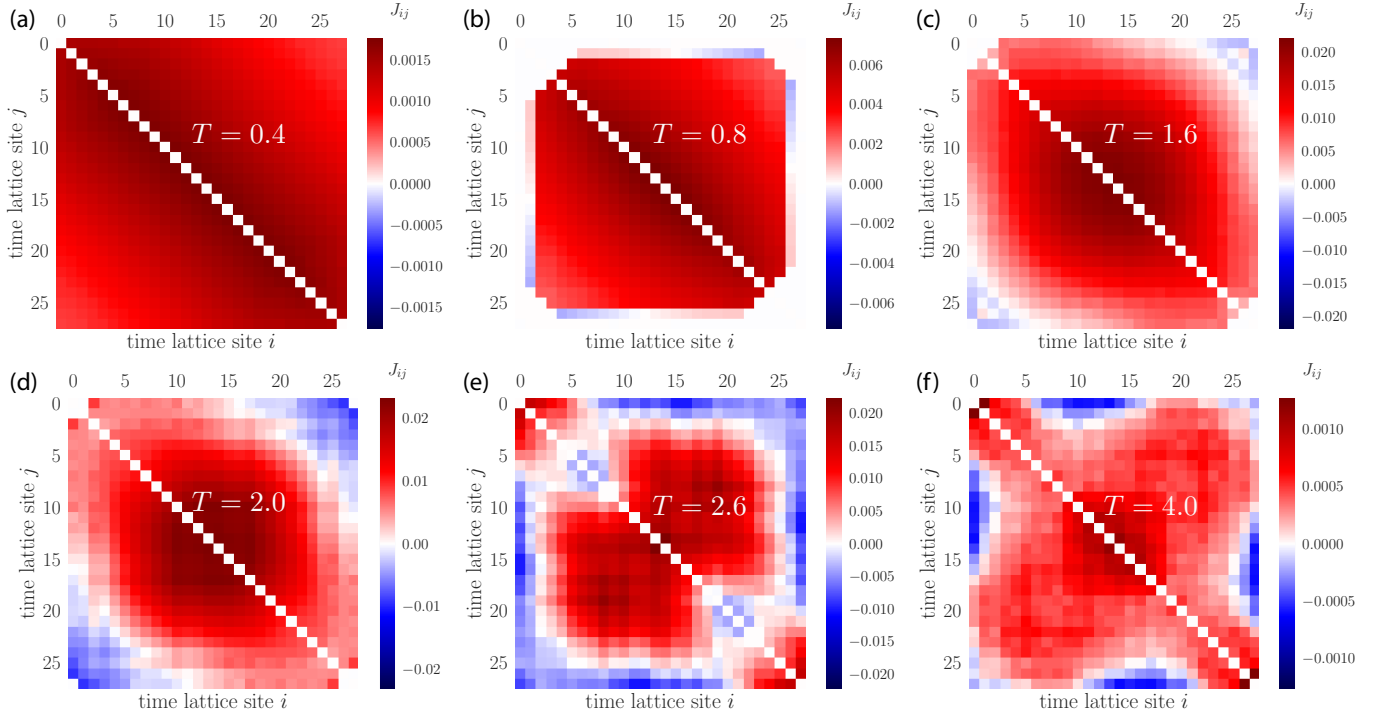


FIG. S17: Spatial dependence of the two-body interaction term in the ML Hamiltonian $H_{\text{ML}}^{(1+2+3)}(T)$. The time lattice sites correspond to the bangs of the bang-bang protocols used to prepare the state. The parameters are $N_T = 28$, $L = 6$. The inverse Ridge regression regularisation strength is $\lambda = 10^{-9}$.

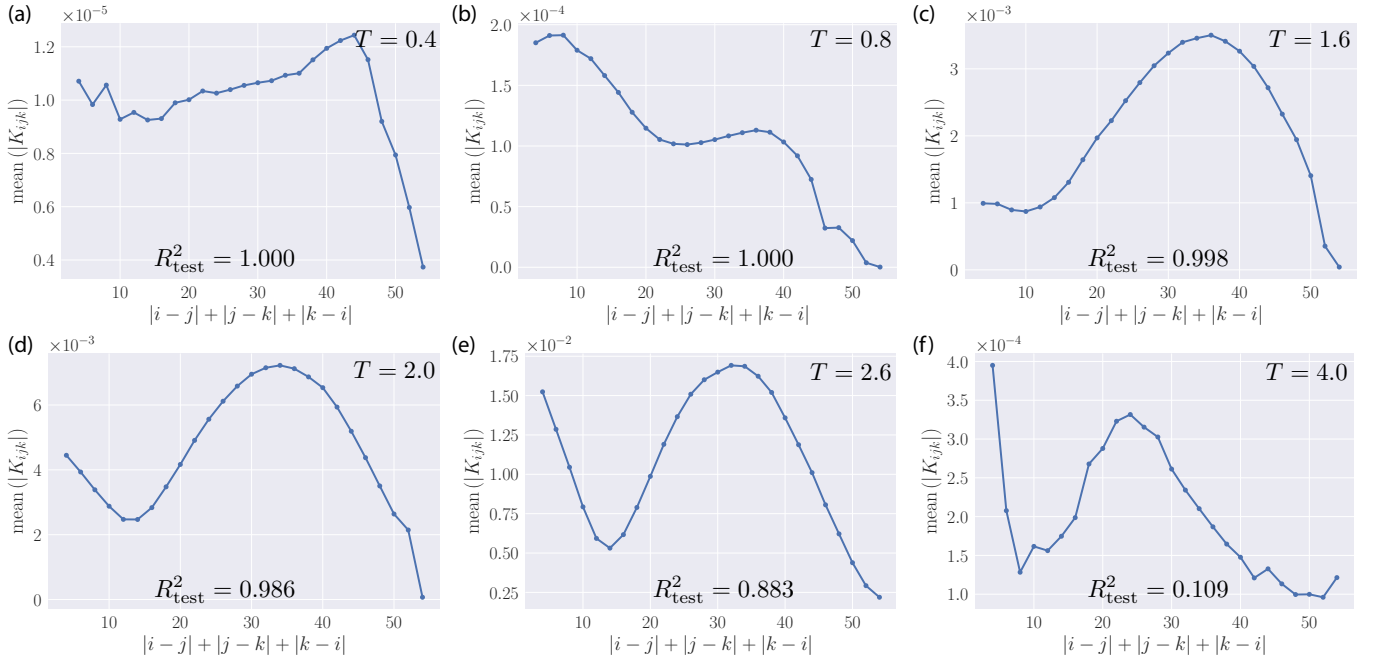


FIG. S18: Non-locality of the tree-body term in the ML Hamiltonian $H_{\text{ML}}^{(1+2+3)}(T)$. The time lattice sites correspond to the bangs of the bang-bang protocols used to prepare the state. The parameters are $N_T = 28$, $L = 6$. The inverse Ridge regression regularisation strength is $\lambda = 10^{-9}$.