Supplementary Information to the paper "Many-body quantum sign structures as non-glassy Ising models"

Tom Westerhout, Mikhail I. Katsnelson, Andrey A. Bagrov

Supplementary Note 1: Reconstructing signs in unfrustrated models

In the main text, we have stated that, for ground states of unfrustrated quantum systems, the auxiliary classical Ising model is not frustrated either, and that allows one to unambiguously reconstruct the exact ground state sign structure without performing combinatorial optimization. In what follows, we provide a proof of this statement for $U(1)$-symmetric bi-local spin-1/2 Hamiltonians.

Let $\hat{H}$ be such a Hamiltonian. The most general form of a two-site interaction term is then

$$\hat{H}_{i,j} = \begin{pmatrix} a_{i,j} & 0 & 0 & 0 \\ 0 & b_{i,j} & 2J_{i,j} & 0 \\ 0 & 2J_{i,j} & c_{i,j} & 0 \\ 0 & 0 & 0 & d_{i,j} \end{pmatrix}, \tag{1}$$

where $\{a, b, c, d\}_{i,j}$ are generic constants that do not enter the couplings of the sign Ising model (and will thus be ignored in the rest of the section), and the only off-diagonal matrix elements $2J_{i,j}$ come from the $J_{i,j}(\sigma_i^x \otimes \sigma_j^x + \sigma_i^y \otimes \sigma_j^y)$ interaction terms that connect Hilbert space basis vectors with flipped spins:

$$|\ldots\uparrow_i\ldots\downarrow_j\ldots\rangle \longleftrightarrow |\ldots\downarrow_i\ldots\uparrow_j\ldots\rangle. \tag{2}$$

If, in the ground state expansion, a basis vector $|a\rangle = |v\rangle \otimes |\uparrow_i\downarrow_j\rangle$ has a sign $\mathcal{S}$ (with $|v\rangle$ defined over all sites except i and j), then $|b\rangle = |v\rangle \otimes |\downarrow_i\uparrow_j\rangle$ should have the sign equal to $-\mathrm{sign}(\mathcal{J}_{a,b}) \cdot \mathcal{S}$, where $\mathcal{J}_{a,b} = |\psi_a||\psi_b|\langle a|\hat{H}|b\rangle$. Then, for any closed loop on the Ising model graph, $|s_1\rangle \leftrightarrow |s_2\rangle \leftrightarrow \cdots \leftrightarrow |s_M\rangle \leftrightarrow |s_1\rangle$, we can define its parity as

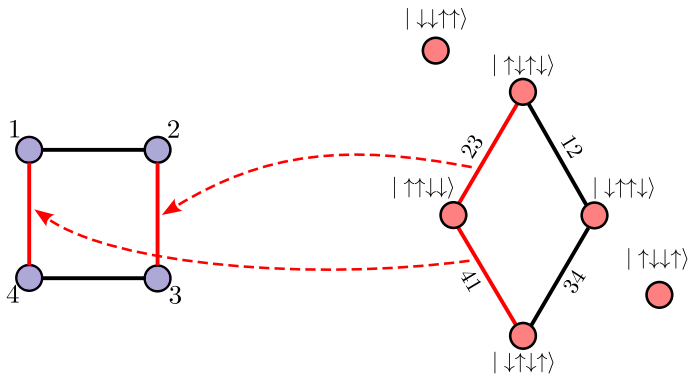
$$P_{\text{Ising}} = (-1)^M \prod_{n=1}^{M} \mathrm{sign}(\mathcal{J}_{s_n,s_{n+1}}), \quad \text{with } s_{M+1} = s_1 .$$

We say that the Ising model is not frustrated if all non-self-intersecting closed paths have parity $P = 1$ (which we call *even*). If this is the case, the sign of a single basis vector unambiguously defines the signs of all of its neighbors, and the full sign structure becomes apparent.

Note, that each closed path on the auxiliary Ising model graph corresponds to a closed path of the same length on the physical lattice (see Sup. Fig. 1). Each edge in the Ising model corresponds to an exchange of spins at some positions i and j, and it maps to the i, j-bond of the physical lattice (multiple Ising links can be mapped to the same physical link). By moving along a closed non-self-intersecting path, one flips every spin exactly two times, which implies that the corresponding bonds on the physical lattice form a closed path of the same length.

Parity of the loop on the physical lattice,

$$P_{\text{quantum}} = (-1)^M \prod_{n=1}^{M} \mathrm{sign}(J_{i_n,i_{n+1}}), \quad \text{with } i_{M+1} = i_1 ,$$



Supplementary Figure 1: **Mapping of a path on the auxiliary classical Ising model onto the physical quantum spin lattice.** On the left, we have a 4-site physical lattice, and on the right — the Hilbert space with a path on the auxiliary Ising model. Every edge on the right corresponds to a physical bond on the left connecting two sites. A continuous path on the right can result in a discontinuous one on the left as illustrated in red. Nevertheless, a closed loop on the Ising graph corresponds to a closed loop of the same length on the physical lattice.

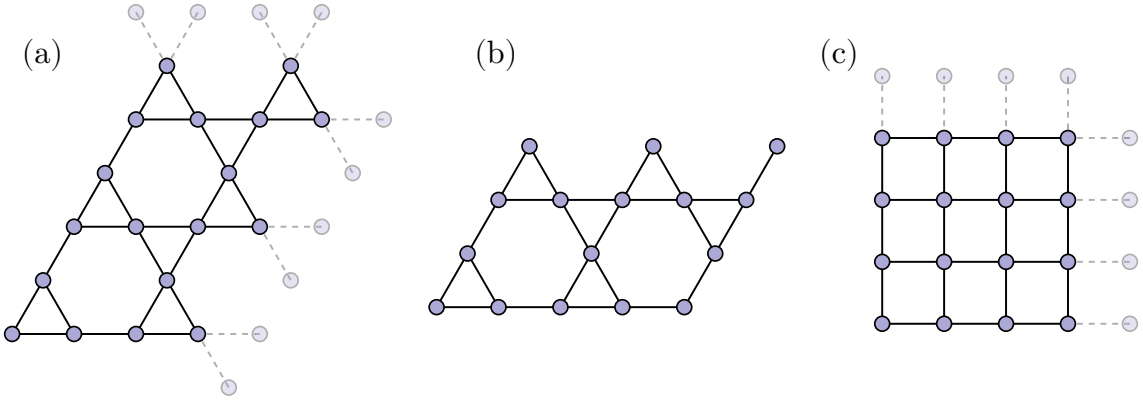
is then exactly the same as P_{Ising}, because $\text{sign}(J_{i,j}) = \text{sign}(\mathcal{J}_{a,b})$ for $|a\rangle$ and $|b\rangle$ that differ only by the exchange of spins at positions i and j.

Hence, if any closed path on the physical lattice of the quantum model has even parity, the resulting auxiliary sign Ising model lacks frustrations, and the pattern of signs can be straightforwardly reconstructed. For instance, for the Heisenberg antiferromagnet on a bipartite lattice, the auxiliary classical sign Ising model is also bipartite and antiferromagnetic, and the alternating pattern of signs is trivially reconstructed. However, this statement remains correct for any bi-local $U(1)$-invariant spin-1/2 quantum magnet where every closed loop on its lattice has even parity $P = 1$. For example, the Heisenberg magnet on the triangular lattice with ferromagnetic exchange along one axis and antiferromagnetic exchange along the other two axes.

Supplementary Note 2: Physical systems for numerical simulations

Numerical simulations were performed on small- to medium-sized clusters that were still amenable to exact diagonalization. In "Small quantum systems" subsection of the main text we considered the following models:

- Heisenberg model on 18-site Kagome lattice with periodic boundary conditions (see Sup. Fig. 2(a));
- Heisenberg model on 16-site Kagome lattice with open boundary conditions (see Sup. Fig. 2(b));
- J_1-J_2 model on 16-site square lattice with periodic boundary conditions (see Sup. Fig. 2(c)), J_2/J_1 ratio was set to 0.55;
- 16-site all-to-all connected model with random Heisenberg interactions $J_{ij} \sim \text{Normal}(\mu = 0, \sigma = 1)$.



Supplementary Figure 2: **Some of the studied Heisenberg models.** (a) 18-site Kagome lattice, (b) 16-site Kagome lattice, and (c) 16-site square lattice. Periodic boundaries in 2 directions are indicated with dashed lines.

For all of these models we restricted the Hilbert space to the sector of zero magnetization.

In Fig. 6 of the main text we considered a few larger models, namely:

- Heisenberg model on 32-site Pyrochlore cluster [1];
- Heisenberg model on 36-site Kagome lattice [2];
- 32-site all-to-all connected model with random Heisenberg interactions $J_{ij} \sim \text{Normal}(\mu = 0, \sigma = 1)$.

For both 16- and 32-site random Heisenberg models we chose the interactions J_{ij} to be distributed according to Normal (Gaussian) distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$. However, changing σ would only result in energy rescaling and would not affect our results.

Supplementary Note 3: Details of running the Simulated Annealing algorithm on small systems

In this section, we provide some more details about running SA algorithm [3] that was used to produce Fig. 2 in the main text. Specifically, we have implemented the algorithm [4] using tricks from [5] to improve its performance.

We determine the initial and final annealing temperatures automatically according to

$$\beta_{\min} = \log 2 / \Delta E_{\max} \,,$$
$$\beta_{\max} = \log 100 / \Delta E_{\min} \,.$$

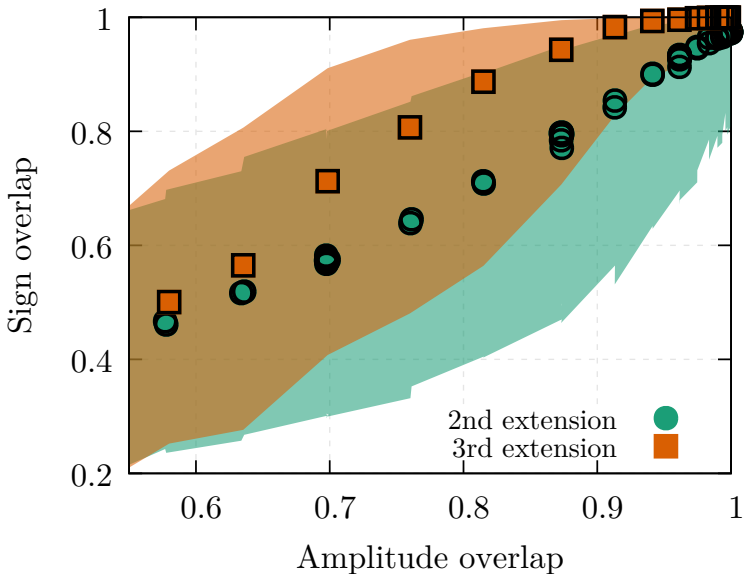
Here, $\Delta E_{\max}$ and $\Delta E_{\min}$ are maximal and minimal energy changes which can result from a single spin flip. We refer the reader to our implementation [4] of the algorithm that estimates $\Delta E_{\max}$ and $\Delta E_{\min}$. In all simulations, we use exponential annealing schedule because it produced better results than linear. The exponential schedule is defined by the following equation:

$$\beta(t) = \beta_{\min} \left(\frac{\beta_{\max}}{\beta_{\min}} \right)^{t/(T-1)} \,, \; t = 0 \ldots T-1$$

where T is the total number of sweeps.

For every model, we perform 10 runs with difference random number generator seeds (to roughly estimate the statistical errors). Within each run, for a given number of sweeps, we perform SA 1024 times to estimate the probability of obtaining 99.5% accuracy.

Supplementary Note 4: Robustness of the algorithm to noise in amplitudes



Supplementary Figure 3: **Effect of the noise in the amplitudes on the quality of the reconstructed sign structure for random connected clusters $\mathcal{C}$.** The experiments were performed on the 36-site Kagome lattice. We show the sign overlap $O_{\mathcal{C}}^{\mathcal{S}}$ as a function of the amplitude overlap $O_{\mathcal{C}}^{\mathcal{A}}$. Green circles depict median values of the sign overlap computed using $\mathcal{C}^{(2)}$ extensions of the randomly sampled connected clusters, and orange squares — the median values computed using $\mathcal{C}^{(3)}$ extensions. Shaded regions show the interquartile range of $O_{\mathcal{C}}^{\mathcal{S}}$. For each point on the plot, 6000 clusters of sizes 50–1000 have been sampled to collect statistics.

To demonstrate the robustness of the algorithm with respect to noise in the amplitudes for larger systems, where greedy optimization cannot be performed on the full Hilbert space basis, we use the ground state of the 36-site Kagome antiferromagnet and corrupt its amplitudes in the same way as outlined in the main text. Then, we randomly sample clusters $\mathcal{C}$, construct their extensions $\mathcal{C}^{(2)}$ and $\mathcal{C}^{(3)}$, and use the greedy algorithm to retrieve sign structures on $\mathcal{C}$. From Sup. Fig. 3, one can see that, for larger systems, stability of the algorithm is less pronounced than for smaller systems. This is likely because of the interplay between errors resulting from the noise in the amplitudes and errors corresponding to uncertainty in the "external magnetic fields" h_i at the boundary of the cluster extension. However, if $O_{\mathcal{C}}^{\mathcal{A}} > 0.9$, optimization on $\mathcal{C}^{(2)}$ leads to good, and on $\mathcal{C}^{(3)}$ — to nearly perfect results. It should be noted that $O_{\mathcal{C}}^{\mathcal{A}} = 0.9$ corresponds to the noise level $\epsilon \approx 0.8$, and the maximal amplitude of the ground state, $\max_i |\psi_i^{exact}| = 0.017$, acquires corrupted values in the range $(0.007, 0.038)$.

Supplementary Note 5: Sign optimization within variational Monte Carlo loop

Here, we outline how the approach to reconstructing sign structures via combinatorial optimization of the auxiliary classical Ising model can be used in conjunction with variational Monte Carlo (vMC) algorithms to solve larger-scale problems. Consider a variational ansatz:

$$\psi_W = \sum_i \psi(\mathbf{W}, i)|i\rangle, \tag{3}$$

where $\mathbf{W}$ are variational parameters of the ansazt (e.g., neural network weights in case ψ_W is an NQS), $|i\rangle = |\dots\uparrow\dots\downarrow\dots\rangle$ are basis vectors of the Hilbert space, and the sum runs over the whole basis.

Traditionally, within the vMC framework, the optimal approximation to the quantum Hamiltonian $\hat{H}$ ground state is obtained through step by step minimization of the energy of the variational ansatz:

$$E_n = \frac{\langle\psi_{W_n}|\hat{H}|\psi_{W_n}\rangle}{\langle\psi_{W_n}|\psi_{W_n}\rangle}, \tag{4}$$

$$\mathbf{W}_{n+1} = \mathbf{W}_n - \gamma\frac{\partial E_n}{\partial \mathbf{W}}, \tag{5}$$

where n is the current optimization step, and γ is the learning rate.

The Hamiltonian expectation value is evaluated by means of Monte Carlo sampling as

$$E_W = \sum_i \frac{\langle i|\hat{H}|\psi_W\rangle}{\langle i|\psi_W\rangle}\left|\frac{\langle i|\psi_W\rangle}{\langle\psi_W|\psi_W\rangle}\right|^2 \simeq \frac{1}{K}\sum_{i=i_1}^{i_K}\frac{\langle i|\hat{H}|\psi_W\rangle}{\langle i|\psi_W\rangle} \equiv \frac{1}{K}\sum_{i=i_1}^{i_K} E_{loc}(i), \tag{6}$$

where K is the number of basis vectors sampled from probability distribution $p_W(i) = \left|\frac{\langle i|\psi_W\rangle}{\langle\psi_W|\psi_W\rangle}\right|^2$, which typically scales linearly with the size of the system, $K \sim 2000 \cdot N$.

More formally, the optimization algorithm can be outlined as follows:

$\mathbf{W}_0 \leftarrow$ random weights
$n \leftarrow 0$
while not converged **do**
  $\{|i_1\rangle, \dots |i_K\rangle\} \leftarrow p_{W_n}(i)$
  $E_n \leftarrow \frac{1}{K}\sum_{i=i_1}^{i_K}\frac{\langle\psi_{W_n}|\hat{H}|i\rangle}{\langle\psi_{W_n}|i\rangle}$
  $\mathbf{W}_{n+1} \leftarrow \mathbf{W}_n - \gamma\frac{\partial E_n}{\partial \mathbf{W}_n}$
  $n \leftarrow n+1$

Here, the variational ansatz ψ_W encodes the wave function coefficients — both their amplitudes and phases/signs: $\langle\psi_W|i\rangle = \psi(\mathbf{W}, i)$. To implement our idea of combinatorial optimization of sign structure, we shall instead use variational representation only for the amplitudes, which we denote as $A(\mathbf{W}, i)$, while storing signs $\mathcal{S}_i$ of relevant basis vectors explicitly: $\langle\psi_W|i\rangle = A(\mathbf{W}, i)\mathcal{S}_i$.

The Monte Carlo estimate of the energy of the combined representation becomes

$$E_W \simeq \frac{1}{K}\sum_{i=i_1}^{i_K} E_{loc}(i) = \frac{1}{K}\sum_{i=i_1}^{i_K}\frac{\sum_j A(\mathbf{W}, j)\mathcal{S}_j H_{ij}}{A(\mathbf{W}, i)\mathcal{S}_i}. \tag{7}$$

At every optimization step, one needs the signs of basis vectors $|i\rangle \in \{|i_1\rangle, \dots |i_K\rangle\}$, as well as of all the vectors $|j\rangle$ for which $\langle j|\hat{H}|i\rangle = H_{ij} \neq 0$ for some $|i\rangle$. Let us denote a vector $|i\rangle$ together with its Hamiltonian neighbours $\{|j\rangle : H_{ij} \neq 0\}$ as M_i, and the complete set of vectors for which the signs are to be reconstructed as

$$M = \bigcup_{i=i_1}^{i_K} M_i \tag{8}$$

The optimization algorithm then acquires the form:

1 $\mathbf{W}_0 \leftarrow$ random weights
2 $n \leftarrow 0$
3 **while** not converged **do**
4 $\quad \{|i_1\rangle, \dots |i_K\rangle\} \leftarrow p_{W_n}(i)$
5 $\quad M \longleftarrow \mathcal{C}^{(1)}(\{|i_1\rangle, \dots |i_K\rangle\})$
6 $\quad \{\mathcal{S}_i : i \in M\} \leftarrow \text{Greedy}(\mathcal{C}^{(2)}(M))$
7 $\quad E_{W_n} \leftarrow \frac{1}{K} \sum_{i=i_1}^{i_K} \frac{\sum_j A(\mathbf{W}_n, j)\mathcal{S}_j H_{ij}}{A(\mathbf{W}_n, i)\mathcal{S}_i}$
$\quad$ `// In the following,` $\{\mathcal{S}_i : i \in M\}$ `are kept fixed during differentiation`
8 $\quad \mathbf{W}_{n+1} \leftarrow \mathbf{W}_n - \gamma \frac{\partial}{\partial \mathbf{W}_n} E_n$
9 $\quad n \leftarrow n + 1$

Line 6 of this loop requires elaboration. In the main text of the paper, we have shown how greedy optimization can be used to find signs of connected (and not too large) clusters in the Hilbert space basis. Here, set M is different from those clusters in two ways. First, with high probability this set itself, as well as its extension $\mathcal{C}^{(2)}(M)$ (or a higher order one), have many disconnected components. Secondly, it can be pretty large. For example, if the underlying quantum Heisenberg Hamiltonian is defined on a regular lattice with coordination number z, every basis vector $|i\rangle$ has up to zN neighbours, and the size of M is

$$\#M < zNK \sim 8 \cdot 10^7 \tag{9}$$

for $K = 2000 \cdot N$, $z = 4$, $N = 100$. Its second extension would have

$$\#\mathcal{C}^{(2)}(M) < \#M \cdot z^2 N^2 < 1.28 \cdot 10^{13}. \tag{10}$$

In practice, this upper bound is never reached because many vectors share common neighbours (and the actual size of M is about two orders of magnitude smaller), but anyway the resulting set is way too large to run combinatorial optimization on it.

To prove applicability of the combinatorial approach to retrieving the signs, we shall now show two things: 1) optimization of signs can be performed on each disconnected component of $\mathcal{C}^{(2)}(M)$ independently; 2) with probability approaching 1, *all* extensions $\mathcal{C}^{(2)}(M_i)$ are disconnected from each other, and the optimization algorithm can be executed for each of K small independent Ising subsystems.

The first point can be proven by simply noticing that, to correctly compute $E_{loc}(i)$, one needs to know only the relative signs of $|i\rangle$ and $|j\rangle$, and a flip of both $\mathcal{S}_i$ and $\{\mathcal{S}_j\}$ does not change the value of $E_{loc}(i)$:

$$E_{loc}(i) = \frac{\sum_j A(\mathbf{W}, j)\mathcal{S}_j H_{ij}}{A(\mathbf{W}, i)\mathcal{S}_i} = \frac{\sum_j A(\mathbf{W}, j)(-\mathcal{S}_j) H_{ij}}{A(\mathbf{W}, i)(-\mathcal{S}_i)}. \tag{11}$$

This implies that, if two sampled basis vectors $|i_1\rangle$ and $|i_2\rangle$ have non-intersecting extensions of their neighbourhoods M_{i_1} and M_{i_2}: $\mathcal{C}^{(2)}(M_{i_1}) \cap \mathcal{C}^{(2)}(M_{i_2}) = \emptyset$, their signs structures can be optimized independently. Though as a result of independent optimization, they may acquire incorrect relative sign, each of the two sign structures is nevertheless a good approximation to the true one (which is also defined up to an overall flip) and gives the correct result for $E_{loc}(i_1)$ or $E_{loc}(i_2)$ correspondingly.

However, if two or more extensions intersect, their joint sign structure must be optimized as a whole. Hence, the combinatorial optimization of signs would be possible only if large clusters unifying many extensions $\mathcal{C}^{(2)}(M_i)$ do not generally form.

This can be proven by constructing the lower bound on the probability P that the following statement holds:

$$\forall\ i \neq j \in \{i_1, \dots i_K\} :\ \ \mathcal{C}^{(2)}(M_i) \cap \mathcal{C}^{(2)}(M_j) = \emptyset. \tag{12}$$

Let us sample basis vectors $|i_1\rangle, |i_2\rangle, |i_3\rangle, \dots$ from the uniform distribution. Once the first vector $|i_1\rangle$ is sampled (then $M_{i_1} = \mathcal{C}^{(1)}(|i_1\rangle)$), the second sampled vector must not belong to $\mathcal{C}^{(6)}(|i_1\rangle)$ to ensure that $\mathcal{C}^{(2)}(M_{i_1}) \cap \mathcal{C}^{(2)}(M_{i_2}) = \emptyset$. The size of $\mathcal{C}^{(6)}(|i_1\rangle)$ is bounded from above by z^6N^6 (much smaller in practice), so the probability of this event is

$$p(\mathcal{C}^{(2)}(M_{i_1}) \cap \mathcal{C}^{(2)}(M_{i_2}) = \emptyset) \geq 1 - \frac{z^6N^6}{\mathcal{C}^N_{N/2}}. \tag{13}$$

When the third vector is sampled, it must not belong to $\mathcal{C}^{(6)}(|i_1\rangle) \bigcup \mathcal{C}^{(6)}(|i_2\rangle)$, which has a maximal size of $2z^6N^6$, etc. The lower bound on the probability that all sampled vectors produce disconnected clusters in the Hilbert space basis, and their sign structures can be optimized independently, is then

$$P > \prod_{n=1}^{K}(1 - \frac{nz^6N^6}{\mathcal{C}^N_{N/2}}) > (1 - \frac{Kz^6N^6}{\mathcal{C}^N_{N/2}})^K. \tag{14}$$

E.g., for $z = 4$, $N = 100$, $K = 2 \cdot 10^5$, it gives

$$P > (1 - 8.12 \cdot 10^{-9})^{200000} \simeq 0.9983, \tag{15}$$

and it should be kept in mind that formation of small connected components unifying several M_i sets does not cause problems for the greedy algorithm.

From here, the $\mathcal{O}(N^4 \log N)$ complexity scaling can deduced. Sizes of individual connected $\mathcal{C}^{(2)}(M_i)$ clusters scale as $\mathcal{O}(N^3)$ (because $\#M_i = zN$), with the number of clusters K being linear in N. Total runtime of the greedy algorithm then amounts to $K\mathcal{O}(N^3 \log N^3) \sim \mathcal{O}(N^4 \log N)$.

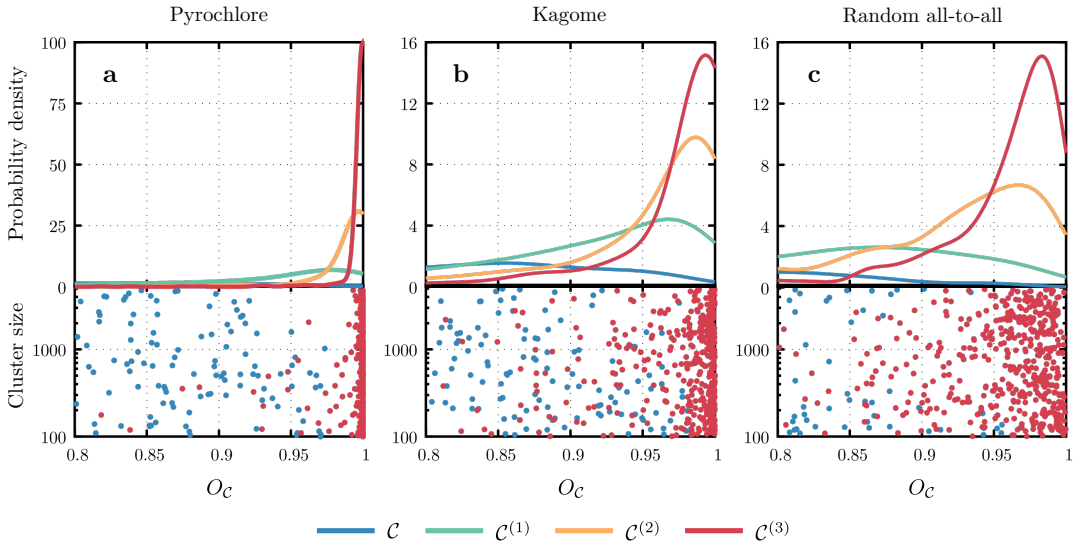
In practice, the sampling occurs from the probability distribution set by the wavefunction amplitudes, $p(i) \sim |A(\mathbf{W}, i)|^2$, which is not uniform, but if this leads to formation of unwanted large connected clusters, those can be broken up by flattening the probability distribution through importance sampling.

Supplementary Note 6: Reconstructing sign structure on random clusters with Simulated Annealing

In the main text, we relied on the greedy algorithm to reconstruct signs on random small connected clusters sampled from the Hilbert space basis of several middle-size quantum

models – the Heisenberg antiferromagnet on 32-site pyrochlore lattice, 36-site Kagome lattice, and 32-site fully connected graph with random couplings. In Sup. Fig. 4, similar results obtained with SA algorithm are shown.

We considered 500 random clusters. Their sizes were distributed logarithmically between 100 and 5000. For each cluster and each of its Hamiltonian extensions, we ran simulated annealing 64 times using 5000 sweeps. The probability density in the top panel was obtained from the raw data using the Gaussian kernel density estimation algorithm.



Supplementary Figure 4: **Quality of the sign structures reconstructed using SA on random small connected clusters.** The results are shown for **a** the 32-site pyrochlore lattice, **b** 36-site Kagome lattice, and **c** 32-site all-to-all connected model with random couplings. The top panel shows the distribution (or the probability density function) of clusters as a function of the overlap $O_{\mathcal{C}}$. We run SA on the original cluster $\mathcal{C}$ as well as on its Hamiltonian extensions $\mathcal{C}^{(1)}$, $\mathcal{C}^{(2)}$, and $\mathcal{C}^{(3)}$. The bottom panel shows the raw data from which the probability densities were estimated. Only $\mathcal{C}$ and $\mathcal{C}^{(3)}$ are shown to reduce the visual noise. There is no correlation between the overlap $O_{\mathcal{C}}$ and the cluster size.

Supplementary Note 7: Extension to not time-reversal-symmetric Hamiltonians

If the quantum lattice model is not time-reversal symmetric, its Hamiltonian is inherently complex-valued. In that case, the phase structure cannot be reduced to binary signs, and the suggested mapping onto the classical Ising model is impossible. Instead, one can reconstruct the phase structure by optimizing the energy of a continuous rotor model, also known as the XY model. Indeed, the quantum state is

$$|\psi\rangle = \sum_i |\psi_i| e^{i\phi_i} |i\rangle, \tag{16}$$

and the corresponding energy is

$$E = \langle\psi|\hat{H}|\psi\rangle = \sum_{i,j}\langle i|\hat{H}|j\rangle|\psi_i||\psi_j|e^{i(\phi_i-\phi_j)} = \tag{17}$$

$$\sum_{i,j}\frac{1}{2}\left(\langle i|\hat{H}|j\rangle e^{i(\phi_i-\phi_j)} + \langle j|\hat{H}|i\rangle e^{i(\phi_j-\phi_i)}\right)\cdot|\psi_i||\psi_j| =$$

$$\sum_{i,j}\mathrm{Re}\left[\langle i|\hat{H}|j\rangle e^{i(\phi_i-\phi_j)}\right]\cdot|\psi_i||\psi_j| =$$

$$\sum_{i,j}\mathcal{J}_{i,j}\cos\left(\Delta_{i,j}+\phi_i-\phi_j\right),$$

where we have expressed the Hamiltonian matrix elements as $\langle i|H|j\rangle = |H_{i,j}|\cdot e^{i\Delta_{i,j}}$ and have defined $\mathcal{J}_{i,j} = |H_{i,j}|\cdot|\psi_i||\psi_j|$.

Supplementary references

[1] N. Astrakhantsev, T. Westerhout, A. Tiwari, K. Choo, A. Chen, M. H. Fischer, G. Carleo and T. Neupert, Broken-symmetry ground states of the heisenberg model on the pyrochlore lattice, Physical Review X **11**(4), 041021 (2021).

[2] P. W. Leung and V. Elser, Numerical studies of a 36-site kagome antiferromagnet, Physical Review B **47**(9), 5459 (1993).

[3] S. Kirkpatrick, C. D. Gelatt Jr and M. P. Vecchi, Optimization by simulated annealing, science **220**(4598), 671 (1983).

[4] https://github.com/twesterhout/ising-glass-annealer (2021).

[5] S. V. Isakov, I. N. Zintchenko, T. F. Rønnow and M. Troyer, Optimised simulated annealing for ising spin glasses, Computer Physics Communications **192**, 265 (2015).