communications
physics

ARTICLE

https://doi.org/10.1038/s42005-024-01542-8 OPEN

Machine learning reveals features of spinon Fermi surface

Kevin Zhang[1], Shi Feng[2], Yuri D. Lensky[1,3], Nandini Trivedi[2] & Eun-Ah Kim[1,4,5,6]✉

With rapid progress in simulation of strongly interacting quantum Hamiltonians, the challenge in characterizing unknown phases becomes a bottleneck for scientific progress. We demonstrate that a Quantum-Classical hybrid approach (QuCl) of mining sampled projective snapshots with interpretable classical machine learning can unveil signatures of seemingly featureless quantum states. The Kitaev-Heisenberg model on a honeycomb lattice under external magnetic field presents an ideal system to test QuCl, where simulations have found an intermediate gapless phase (IGP) sandwiched between known phases, launching a debate over its elusive nature. We use the correlator convolutional neural network, trained on labeled projective snapshots, in conjunction with regularization path analysis to identify signatures of phases. We show that QuCl reproduces known features of established phases. Significantly, we also identify a signature of the IGP in the spin channel perpendicular to the field direction, which we interpret as a signature of Friedel oscillations of gapless spinons forming a Fermi surface. Our predictions can guide future experimental searches for spin liquids.

---

[1] Department of Physics, Cornell University, Ithaca, NY, USA. [2] Department of Physics, The Ohio State University, Columbus, OH, USA. [3] Google Research, Mountain View, CA, USA. [4] Radcliffe Institute for Advanced Studies, Cambridge, MA, USA. [5] Department of Physics, Harvard University, Cambridge, MA, USA. [6] Department of Physics, Ewha Womans University, Seoul, South Korea. ✉email: eun-ah.kim@cornell.edu

As our ability to simulate quantum systems increases, there is a corresponding need for determining how to characterize unknown phases realized in simulators. Going from measurements to the nature of the underlying state is a challenging inverse problem. Full quantum state tomography[1] of the density matrix is impractical. Although the classical shadow[2] scales better than full tomography, the approach does not prescribe to researchers the proper observables to evaluate. Viewing the inverse problem as a data problem invites adopting machine learning methods: a quantum-classical hybrid approach. Machine learning has been widely applied for characterizing quantum states[3]. Such methods have been most fruitful with symmetry-broken states, with a diverse set of approaches increasingly bringing more interpretability and reducing bias[4–6]. The characteristic features of ordered phases are ultimately local and classical, hence ML models tuned for image processing have readily learned such features. By contrast, past learning of quantum states defined without order parameters has relied on theoretically guided feature preparation[7,8]. However, such reliance on prior knowledge blocks the researchers' access to new insights into unknown states: the ultimate goal of simulating quantum states.

To push the limits of the nascent quantum-classical hybrid approach, we need a setting known to host a non-trivial quantum phase of unknown nature. Recent investigations into extended Kitaev models[9–12] have led to the observation of a mysterious intermediate gapless phase (IGP) sandwiched between the Kitaev spin liquid and the trivial polarized state under a non-perturbative [111] magnetic field[13–16], whose identification presents an interesting and important puzzle away from the perturbative limit. However, the nature of this field-induced IGP has raised debate in the community.

Several theories have shown evidence that supports a gapless quantum spin liquid phase with an emergent $U(1)$ spinon Fermi surface[15,17–19], while there are also mean field theories indicating that the low energy effective theory of the intermediate phase is gapped with a non-zero Chern number[20,21]. This tension between theories arises due to the challenge in determining the nature of the IGP that forms in a non-perturbative region under a magnetic field. Unlike the gapped topological phase adiabatically connected to the exactly solvable limit with known loop operators[7,22,23], the absence of measurable positive features for the possible candidate IGP states[18,19] also makes this problem a worthy challenge for machine learning.

We present a quantum-classical hybrid approach, QuCl, to reveal characteristic motifs associated with states without known signature features. We treat variational wavefunctions obtained from density matrix renormalization group (DMRG)[24–26] as output of a quantum simulator. Namely, we sample snapshots from the ground state and train an interpretable neural network architecture, i.e. the correlator convolutional neural network (CCNN)[4] (Fig. 1c). Based on the trained network, we use regularization path analysis[27] to determine the distinct correlation functions learned by the CCNN as characteristic features of the state captured by snapshots. We benchmark the performance of this hybrid approach on the known phases and confirm that the CCNN learned features are consistent with the known characteristic features. Importantly, we reveal the signature feature of the IGP to imply the existence of a spinon Fermi surface, as proposed in refs.[18,19].

Results and discussion

Model. The Kitaev-Heisenberg model under an external field is defined by

$$H=\sum_{\gamma=x,y,z}\sum_{\langle ij\rangle_\gamma}K_\gamma S_i^\gamma S_j^\gamma - J\sum_{\langle ij\rangle}\mathbf{S}_i\cdot\mathbf{S}_j-\sum_i\mathbf{h}\cdot\mathbf{S}_i \qquad (1)$$

where $\gamma=x,y,z$ enumerates the three colorings of bonds on the honeycomb lattice (Fig. 1a), and S^γ is the γ projection of spin-1/2 degrees of freedom on each site. We also add a uniform Zeeman field $\mathbf{h}$ along the [111] direction, i.e., the out-of-plane $\hat{e}_3$ direction in the lab frame (see also Supplementary Note III), as well as a ferromagnetic Heisenberg term of strength J. Here we investigate the antiferromagnetic Kitaev intereaction ($K>0$), for which a field in the [111] direction gives rise to an intermediate phase over a significant field regime. For the ferromagnetic Kitaev interaction, on the other hand, the intermediate phase is either absent or exists over a very small field regime[18]. Starting from the exactly solvable point at $K_x=K_y=K_z=1$, $h=J=0$, which is a nodal $\mathbb{Z}_2$ spin liquid[28], we consider three axes of the phase diagram that are controlled by the parameters h, J, and K_z (Fig. 1b).

For the J axis of the phase diagram in Fig. 1b, the system undergoes a sequence of transitions through magnetically ordered states[11]. For small values of J the system preserves time-reversal symmetry and the system remains a gapless $\mathbb{Z}_2$ spin liquid. As J is increased, the system acquires a zigzag magnetic order (also experimentally observed in α-RuCl$_3$[11,29,30]). At even larger values of J, the system eventually becomes a Heisenberg ferromagnet. On the other hand, a small magnetic field $h\|[111]$ breaks the time reversal symmetry of the Kitaev model and opens a gap in the spectrum of free majorana fermions, resulting in a CSL[28]. However, upon leaving the perturbative regime, numerical evidence through DMRG[19] and exact diagonalization[16,18] have shown that the system goes through an IGP before entering a partially polarized (PP) magnetic phase. Although the precise nature of the IGP is unknown, a $U(1)$ spinon Fermi surface has been proposed recently[18,19], which, as we are to show, is in agreement with our CCNN results. Finally, the Kitaev model has an exact solution along the K_z axis when $J=h=0$[28], where the system undergoes a transition to a gapped $\mathbb{Z}_2$ spin liquid upon increasing K_z. We use this axis for benchmarking the QuCl outcome to known exact results.

In order to generate a single snapshot from a wavefunction, we perform the following procedure sequentially on each site i of the lattice:

1. Find the reduced density matrix for site i, and exactly evaluate the expectation value of the spin operator projected along the chosen axis α_i.
2. Choose eigenvalue + or − with probabilities $P_+=\frac{1+\langle\sigma_i^\alpha\rangle}{2}$, $P_-=1-P_+$; record the eigenvalue and axis of projection.
3. Collapse the wavefunction onto the associated eigenstate of site i using the projector $|\pm\alpha\rangle_i\langle\pm\alpha|_i$.
4. Repeat 1-3 until every site is addressed.
5. Organize the snapshots into channels, one for each unique axis α_i; see Fig. 1c.

The choice of axis α_i is random for the J and K_z axes but tailored to the target phases for the h axis. The wavefunctions at phase space points of interest are obtained using DMRG on finite size systems composed of 6×5-unit-cell (60 sites). In Supplementary Note III we also show results from an extended 12×3-unit-cell (120 sites) cylinder geometry. In both cases we used a maximum of 1200 states, giving converged results with a truncation error $\sim10^{-7}$ or less in all phases. Within a phase, we generated 10,000 snapshots for each wavefunction in question. Each resulting snapshot forms a three-dimensional array of bit-strings, with two spatial dimensions and a "channel" dimension (see Fig. 1c). Such a collection of snapshots is a classical shadow of the quantum state[2]. Since our goal is to characterize a quantum state without prior knowledge of the best operator to measure, we treat the snapshot collection as data rather than using them to estimate an operator expectation value as in refs. [2,31].

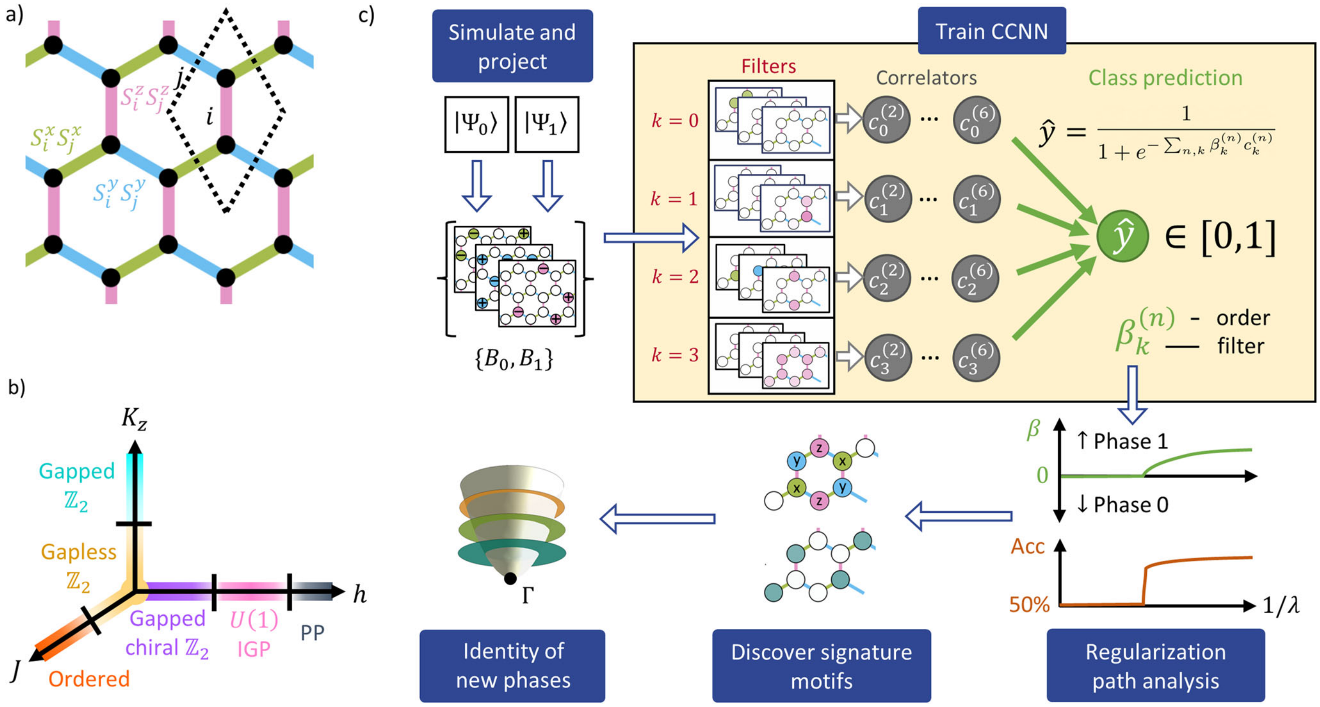
a)
$S_i^z S_j^z$
$S_i^x S_j^x$
$S_i^y S_j^y$
i
j
b)
K_z
Gapped $\mathbb{Z}_2$
Gapless $\mathbb{Z}_2$
Gapped chiral $\mathbb{Z}_2$
$U(1)$ IGP
PP
h
J
Ordered
c)
Simulate and project
$|\Psi_0\rangle$
$|\Psi_1\rangle$
$\{B_0, B_1\}$
Train CCNN
Filters
Correlators
Class prediction
$k=0$
$k=1$
$k=2$
$k=3$
$c_0^{(2)}$ ⋯ $c_0^{(6)}$
$c_1^{(2)}$ ⋯ $c_1^{(6)}$
$c_2^{(2)}$ ⋯ $c_2^{(6)}$
$c_3^{(2)}$ ⋯ $c_3^{(6)}$
$\hat{y} = \frac{1}{1+e^{-\sum_{n,k}\beta_k^{(n)} c_k^{(n)}}}$
$\hat{y} \in [0,1]$
$\beta_k^{(n)}$ - order — filter
β
0
↑ Phase 1
↓ Phase 0
Acc
50%
$1/\lambda$
Γ
Identity of new phases
Discover signature motifs
Regularization path analysis

Fig. 1 The Kitaev-Heisenberg model and schematic of the Quantum-Classical (QuCl) approach. **a** Honeycomb lattice of Kitaev model with bond-dependent interactions indicated by the three different colored bonds. **b** Phase diagram of Kitaev-Heisenberg model in an external magnetic field (Eq (1)) long three axes of K_z, h, and J. **c** A schematic description of QuCl: (i) From a pair of variational wavefunctions $|\Psi_0\rangle$ and $|\Psi_1\rangle$, labeled projective measurements ("snapshots") B_0 and B_1 are generated. (ii) The collection of labeled snapshots are used to train the correlator convolutional neural network (CCNN). (iii) The CCNN is configured with four filters ($k = 0, \cdots, 3$) each with three channels, for the binary classification problem minimizing the distance between the prediction $\hat{y}$ and the label. (iv) Once the training is completed, we fix the filters and use regularization path analysis to reveal signature motifs of the two phases, 0 and 1, under consideration. The correlator weight β onsetting, upon reduction of the regularization strength λ, to a negative (positive) value signals a feature of the phase 0 (1).

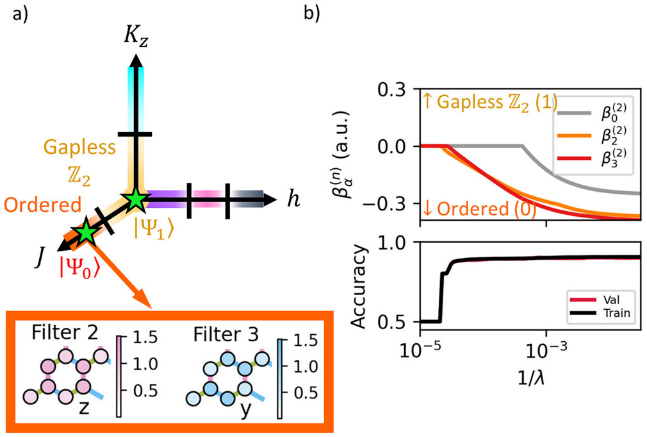

Fig. 2 Gapless $\mathbb{Z}_2$ v.s. Heisenberg ordered phase. **a** Ground state wavefunctions from the gapless $\mathbb{Z}_2$ phase and ferromagnetically ordered phase are obtained at the points on the J-axis marked by stars. The highlighted box shows the two most informative filters for the classification task. The pink and blue dots in filters correspond to projections in z and y basis. **b** The regularization path analysis results pointing to the filters in panel a as signature motifs of the ordered phase.

For each axis of the phase diagram, we set up a binary classification problem between a pair of phase space points, $|\Psi_0\rangle$ and $|\Psi_1\rangle$, each deep within a phase. The machine learning architecture of choice, CCNN, was introduced in Ref. [4] as an adaptation of a convolutional neural network where a controlled polynomial non-linearity splits into different orders of correlators for the neural network to use (Fig. 1c). Compared to the more standard CNN architecture, the CCNN has reduced expressibility due to using a low-order polynomial as the nonlinearity. However, at the expense of this reduction, we gain access to interpreting the network's learning that can be analytically connected to the traditional notion of correlation functions. Combined with regularization path analysis (RPA; see Methods and Supplementary Note I for details)[32], the CCNN can reveal spatial correlations or motifs that are characteristic of a given phase.

For a given channel α of filter k to be learned, $f_{k,\alpha}$, the CCNN samples correlators for each snapshot bit-string $B^\alpha(\mathbf{x})$ through an estimate for the n-th order spatially averaged correlator associated with filter k

$$c_k^{(n)} = \sum_{\mathbf{x}} \sum_{(\mathbf{a}_1,\alpha_1)\neq\ldots\neq(\mathbf{a}_n,\alpha_n)} \prod_{j=1}^{n} f_{k,\alpha_j}(\mathbf{a_j}) B^{\alpha_j}(\mathbf{x}+\mathbf{a_j}), \tag{2}$$

where the inner sum is over all n unique pairs of filter positions $\mathbf{a}$ and filter channels α. These correlator estimates are then coupled to coefficients $\beta_k^{(n)}$ of the linear layer (Fig. 1c; green arrows) according to $\hat{y} = \left[1 + \exp(-\sum_{n,k} \beta_k^{(n)} c_k^{(n)})\right]^{-1}$, where $0 \le \hat{y} \le 1$ is the CCNN output for the given input snapshot. We reserved 1000 samples from each wavefunction as a validation set, and used the remaining 9000 for training. The orders of correlators were restricted to be between 2 and 6, inclusive. We allowed the neural network to learn up to 4 different filters, corresponding to $0 \le k \le 3$. The training optimizes the model parameters, namely the filters and the weights, by comparing the output $\hat{y}$ to the training label (see Methods).

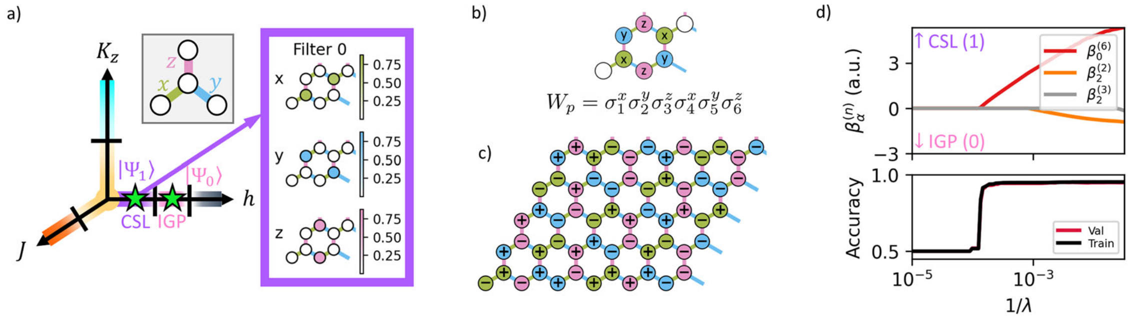

Fig. 3 The chiral spin liquid v.s. the intermediate gapless phase, benchmarking the indicator of the chiral spin liquid phase. a Ground state wavefunctions from the chiral spin liquid (CSL) phase and the intermediate gapless phase (IGP) are obtained at the points on the *h*-axis marked by stars. The highlighted box shows the most informative filter that signifies the CSL phase. Inset shows the bond anisotropy for the three colorings of bonds, e.g., *x* refers to a $S_i^x S_j^x$ coupling. **b** The plaquette operator W_p is a operator defined on the six sites around a hexagonal plaquette. **c** A sample snapshot from the CSL phase showing the measurement basis that makes the plaquette operator W_p accessible. **d** The regularization path analysis pointing to the six-point correlator of filter in panel a as indicator of the CSL.

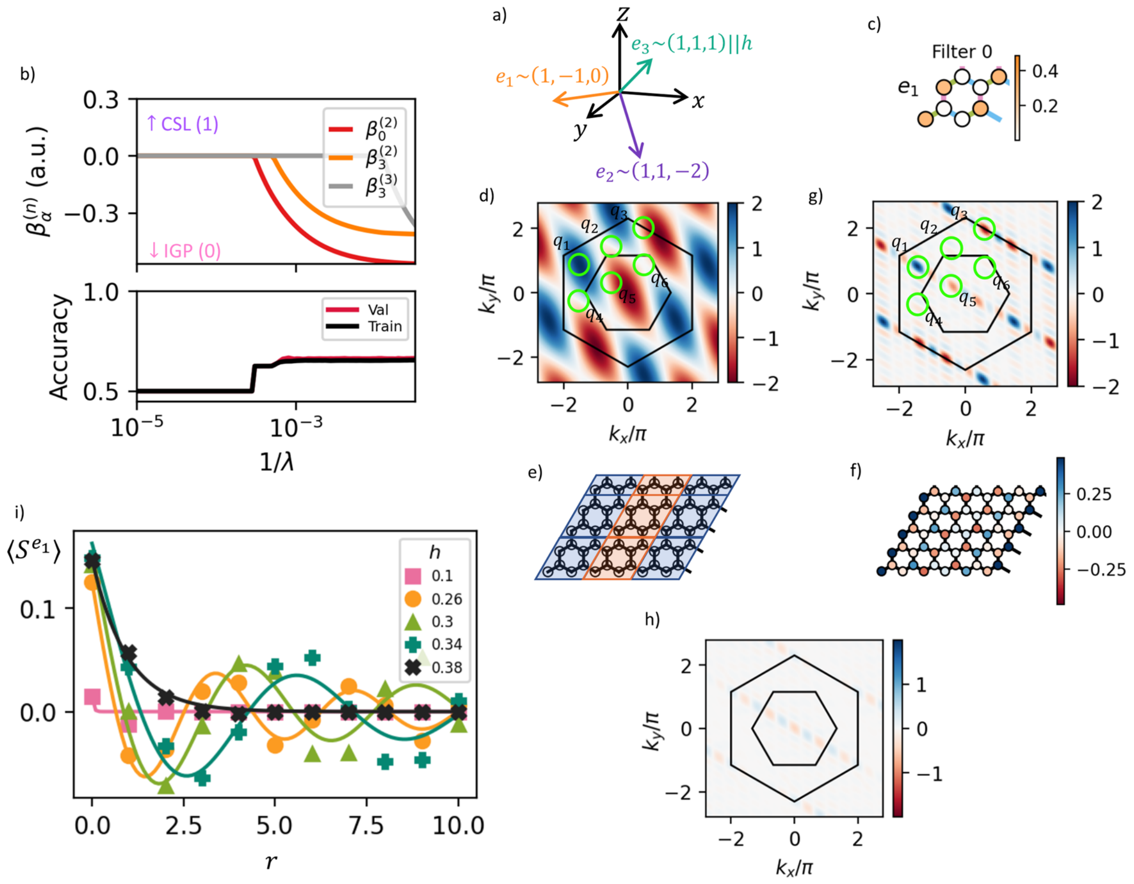

Fig. 4 The chiral spin liquid v.s. the intermediate gapless phase, discovering the indicator of the intermediate gapless phase. a The rotated basis vectors in relation to the cardinal axes; the external field is along e_3. The signature of the intermediate gapless phase (IGP) is targetted using e_1 basis snapshots from the same pair of wavefunctions as in Fig. 3a. **b** The regularization path analysis results pointing to the filter in **c** as an indicator of the IGP. c) Most significant filter learned by the correlator convolutional neural network to be associated with the IGP. **d** Fourier transform of filter in **c**, with black lines indicating first and extended Brillouin zones. Green circles mark six Bragg peaks associated with the filter tiling pattern in **e**. **e** Simplest possible tiling of filter shown in panel c, resulting in a superlattice of antiferromagnetic stripes. The Bragg peaks of this tiling pattern are marked by green circles in panels **d** and **g**. **f** On-site magnetization $\langle S_i^{e_1}\rangle$ of the wavefunction $|\Psi_0\rangle$ in the IGP. **g** Real part of Fourier transform of **f**, again with Bragg peaks of antiferromagnetic tiling marked (imaginary part is negligible). **h** Real part of Fourier transform of $\langle S_i^{e_1}\rangle$ of the wavefunction $|\Psi_1\rangle$ in the CSL phase showing no discernable features. **i** The perpendicular magnetization $\langle S^{e_1}(r)\rangle$ as a function of distance from boundary for various values of field strength *h*, showing decreasing modulation period with increasing field strength. Solid lines show fitted curve based on Equation (3).

Once the CCNN is successfully trained for a given phase, we uncover the characteristic motif that is most informative for the contrast using RPA[32]. For this, we fix the filters and relearn the weights of each learned correlation $\beta_k^{(n)}$ with regularization that penalizes the magnitude of the $\beta_k^{(n)}$'s with strength λ (see Methods). The $\beta_k^{(n)}$ that turns on at the lowest value of $1/\lambda$ points to the specific filter k and the correlation order (n) of that filter which is most informative for the contrast task. The sign of the onsetting $\beta_k^{(n)}$ reveals whether the associated correlation is a feature of phase 0 (− sign) or of phase 1 (+ sign); see Fig. 1(c).

Gapless $\mathbb{Z}_2$ versus Heisenberg phases. As a benchmark, we first focus on the phases along the J axis (Fig. 2a). At intermediate J, the system has zigzag order, while it is a Heisenberg ferromagnet for large J. We trained the CCNN to distinguish wavefunctions from the two points marked by stars in Fig. 2a, at $J/K = 0$ and $K/J = 0$, corresponding to the Kitaev spin liquid and Heisenberg ferromagnetic states, respectively. The snapshots were generated by choosing a random axis from x, y, or z for each site. The RPA shown in Fig. 2b reveals that the most informative correlation functions are the two-point functions of filters 2 and 3 presented in Fig. 2a. The negative sign of the onsetting β's means these features are positive indicators of the ordered phase (see Supplementary Note I). Given that the correlation length vanishes at the exactly solvable point at the origin (phase 0), the network's choice to focus on features of phase 1 is sensible. Moreover, the learned motif of phase 1 is clearly consistent with a ferromagnetic correlation. Hence this benchmarking confirms that the CCNN's learning is consistent with our theoretical understanding when both phases 0 and 1 are known.

Chiral spin liquid. Next, we contrast the CSL phase (phase 1) and the IGP (phase 0) along the h axis (Fig. 3a). Neither of these phases is characterized by a local order parameter. However, the chiral phase is known to be a $\mathbb{Z}_2$ quantum spin liquid characterized by non-local Wilson loop expectation values[28]. To confirm that such non-local information can be learned with our architecture, we first use snapshots with a fixed basis shown in Fig. 3c so that the architecture can access the necessary information. The RPA with the positive onset of $\beta_0^{(6)}$ (Fig. 3d) implies that a sixth-order correlator of the filters shown in Fig. 3a is learned to be the key indicator of phase (1), the CSL phase. Remarkably, the relevant correlator $\langle\sigma_1^z\sigma_2^x\sigma_3^y\sigma_4^z\sigma_5^x\sigma_6^y\rangle$ is exactly the expectation value of the Wilson loop associated with the plaquette p consisting of the six sites $\langle W_p\rangle$, shown in Fig. 3b. Theoretically, $\langle W_p\rangle \approx 1$ implies the state is well-described by the $\mathbb{Z}_2$ gauge theory of the zero-field gapless phase[28]. The fact that none other than $\langle W_p\rangle$ was learned to contrast the CSL phase from the intermediate gapless phase reveals that the latter is a distinct state. However, discovering the indicator of the intermediate phase requires a different approach, as we discuss below.

Intermediate gapless phase. We next discuss how we discover the physically meaningful features of the IGP (phase 0). Previous work has focused on mapping out the low energy excitations $S(\mathbf{q}, \omega \approx 0)$ in momentum space. Also, in real space, the spin-spin correlations averaged over all directions show power-law decay, indicating gapless spin excitations for intermediate fields. However, it has not been clear how to translate these correlations to positive signatures of a particular state that can be experimentally detected. While the QuCl approach has the potential to reveal such signatures, we have to first overcome a ubiquitous challenge accompanying using ML for scientific discoveries: the need to guide the machine away from trivial features. While the unbiased pursuit of representative feature in data is the benefit of using ML, a non-trivial cost is that the neural network's learning can be dominated by features that are trivial from the physicist's perspective. The neural network's propensity to make decisions based on what appears most visible to the network means it is essential that we guide the CCNN away from the trivial yet dominant difference between phase 0 and phase 1: the field-driven magnetization along the e_3 axis (see Supplementary Note II). This basic requirement for extracting meaningful information using ML led us to supply CCNN with snapshots in the basis orthogonal to the field direction, such as e_1 basis (see Fig. 4a). This decision to guide the CCNN away from trivial features led to a sought-after discovery.

The RPA shown in Fig. 4b unambiguously points to two-point correlators of the filter shown in Fig. 4c as a signature feature of the IGP. As is clear from its Fourier transform shown in Fig. 4d, the filter implies the emergence of a length scale in the e_1 component of the magnetization. Given that the e_1 direction is perpendicular to the direction of h-field, the repeating arrangement of the motif the filter is detecting must be antiferromagnetic. One such ansatz we conjecture shown in Fig. 4e will single out specific momenta points marked in Fig. 4d from the Fourier intensity of the filter (see Supplementary Note IV for more details). To confirm this conjecture, we explicitly measure the per-site e_1-magnetization, $\langle S^{e_1}(r)\rangle$, of the two states. The measurement outcome (Fig. 4f) and its Fourier transform (Fig. 4g) confirms indeed the IGP state has a modulating e_1-magnetization that we inferred from the CCNN learned filter and the ansatz tiling the filter. Furthermore, the contrast between Fourier transforms from the IGP Fig. 4g and from the CSL Fig. 4h establishes that the pattern and the associated length scale are unique features of the IGP. Remarkably, we find such modulation to be consistent with a conjecture[15–19,22,33] that the IGP is a $U(1)$ spin liquid with a spinon Fermi surface. Note here that in a translationally invariant system the corresponding quantity is the two-point spin-spin correlation function $\langle S^{e_1}(0)S^{e_1}(r)\rangle$ and its Fourier transform.

As detailed in Supplementary Note III, $\langle S^{e_1}(r)\rangle$ can be mapped to fermionic spinon density in the Kitaev model. If spinons are gapless and deconfined to form a spinon Fermi surface, the Friedel oscillation of the spinon density due to the open boundary[34–37] will be reflected in the modulation of $\langle S^{e_1}(r)\rangle$:

$$\langle S^{e_1}(r)\rangle \sim \langle n_1(r)\rangle \sim \frac{k_F}{\pi}\left[1 - \frac{\sin(2k_F r + \theta)}{2k_F r + \theta}\right] + C \qquad (3)$$

where C and θ are constants, and r the distance measured from the boundary. We confirm the spinon Friedel oscillation origin of the observed modulations by fitting the $\langle S^{e_1}(r)\rangle$ measured at different field strengths h to Eq. (3). The resulting excellent fit in Fig. 4i shows that the modulation period increases with the increase in the perpendicular field h. This is consistent with a mean field picture in which the magnetic field plays the role of the chemical potential; the spinon bands successively get depleted upon increasing field until the system enters a trivial phase through a Liftshitz transition[38]. Evaluation of S^{e_1} on a longer 20 unit cell system of a 3-leg ladder shows a modulation pattern that agrees with the 6-leg ladder results in Fig. 4i (see Supplementary Note III).

Gapless $\mathbb{Z}_2$ vs gapped $\mathbb{Z}_2$. Finally, we contrast the gapless and gapped $\mathbb{Z}_2$ phases along the K_z axis as a sanity check in distinguishing two spin liquid phases. As one tunes K_z, the model Eq. (1) is known to go through a phase transition between a gapless $\mathbb{Z}_2$ and a gapped $\mathbb{Z}_2$ spin liquid phases[28]. However, since both phases have only short-range correlations in the ground state the

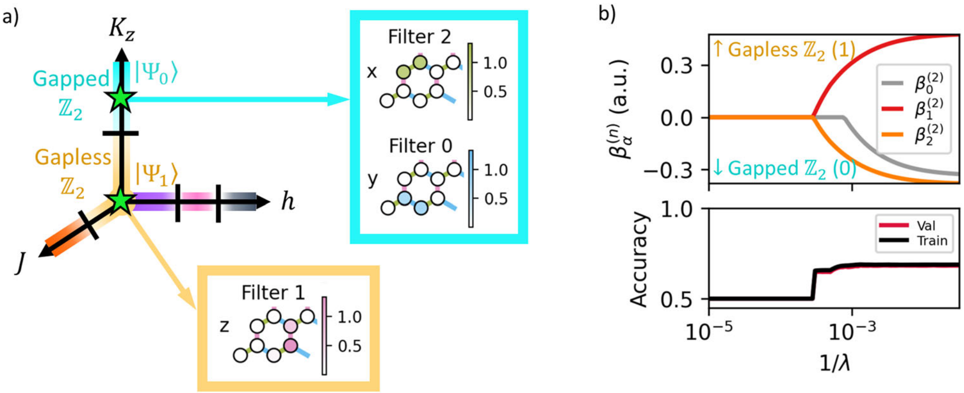

Fig. 5 Gapless v.s. gapped $\mathbb{Z}_2$ spin liquid, benchmarking distinguishing two spin liquids. **a** Ground state wavefunctions from the gapless $\mathbb{Z}_2$ phase and gapped $\mathbb{Z}_2$ phase are obtained at the points on the K_z-axis marked by stars. The highlighted boxes shows the most informative filters that signifies the each phases. **b** The regularization path analysis results associate two-point correlators of filters 0 and 2 to the gapped phase and that of filter 1 to the gapless phase.

distinction cannot be learned from the correlation lengths, unlike usual transitions between a gapless phase and a gapped phase. Hence it is a non-trivial benchmarking test for QuCl-based state characterization. Contrasting the two points marked by stars in Fig. 5a, again using the random basis snapshots, we find signature motifs consistent with exact solutions. Specifically, the RPA (Fig. 5b) shows that nearest-neighbor correlation functions of x and y axes are a feature of the gapped $\mathbb{Z}_2$ phase while the z axis nearest-neighbor correlation function is the feature of the gapped $\mathbb{Z}_2$ phase. These results are consistent with the exact solution of the zero-field Kitaev model[39,40].

Conclusion. The significance of our findings is threefold. Firstly, we gained insight into the intermediate field spin liquid phase in the Kitaev-Heisenberg model. Confronted by two complementary predictions: a gapless spin liquid based on exact diagonalization and DMRG versus a gapped spin liquid in the same region from mean field theory, an identification of a positive signature for either possibilities was critical. The need for guiding the CCNN away from a trivially changing feature led to the discovery that it is critical to focus on snapshots taken along a direction e_1 perpendicular to the magnetic-field e_3 axis. Remarkably, the network then learned a geometric pattern characteristic of Friedel oscillations of spinons in the IGP. This observation strongly supports earlier theoretical proposals of a spinon Fermi surface in the IGP, thus advancing our understanding of this phase.

Secondly, our discovery translates to a prediction for experiments by providing a direct evidence of spinon FS in the modulated magnetization and the spin–spin correlations perpendicular to the field direction along e_1. Such a feature in the computational data has been previously missed since the focus has been on isotropic spin correlation $\langle \mathbf{S}_i \cdot \mathbf{S}_j \rangle$ which is dominated by the e_3 component. Our results can guide future experimental searches for spin liquids with spinon Fermi surfaces.

Finally, on a broader level, we have demonstrated that hidden features of a quantum many-body state can be discovered using QuCl: a data-centric approach to snapshots of the quantum states, employing an interpretable classical machine learning approach. Conventionally, quantum states have been studied through explicit and costly evaluation of correlation functions. However, when the descriptive correlation function is unknown in a new phase, the conventional approach gets lost in the overwhelming space of expensive calculations. Although our method does not explicitly evaluate the correlation functions that it extracts, snapshots that can be readily treated with QuCl will enable computationally efficient identification of new phases associated with a quantum state, including topological states or states with hidden orders. Finally, our method is also broadly applicable to searches for physical indicators of states prepared on quantum simulators which are naturally accessed through projective measurements.

Methods

In this section, we describe the architecture of the neural network and the training procedure. The CCNN, as first proposed in ref. [4], consists of two layers: the correlator convolutional layer and the fully connected linear layer. We fed as input to the CCNN three-spin-channel (two-spin-channel for rotated basis measurements) snapshot data. Since the CCNN was originally applied to square lattice data at its conception, we reinterpreted our hexagonal lattice geometry as a rectangular grid with a 1×2 unit cell forming its two-site basis. We modified the convolutional layer to consist of 4 different learnable filters of dimension 2×2 unit cells, for a total receptive field of 8 sites each. To accommodate the 1×2 unit cell, we also introduced a horizontal stride of 2 in the convolution operation between filters and snapshots.

The filter weights are learnable nonnegative numbers indicated by $f_{\alpha,k}(a)$, where $0\le k\le 3$ indexes the filter, $1\le \alpha\le 3$ indexes the channel of the weight, and a is a spatial coordinate. The weights are convolved with the input snapshots using the recursive algorithm described in ref. [4] to produce per-snapshot correlators as

$$C_k^{(n)}(x) = \sum_{(\mathbf{a}_1,\alpha_1)\neq\ldots\neq(\mathbf{a}_n,\alpha_n)} \prod_{j=1}^{n} f_{k,\alpha_j}(\mathbf{a_j}) B^{\alpha_j}(\mathbf{x}+\mathbf{a_j}), \tag{4}$$

where $C_k^{(n)}(x)$ is the position-dependent n-th order correlator of filter k, and $B^{\alpha_j}(\mathbf{x}+\mathbf{a_j})$ indicates the snapshot value at location $\mathbf{x}+\mathbf{a_j}$ in channel α_j. The correlator estimates are then defined as the spatially-averaged correlators, $c_k^{(n)} = \sum_x C_k^{(n)}(x)$, which are coupled to coefficients $\beta_k^{(n)}$ of the linear layer, and summed to produce the logistic regression classification output

$$\hat{y} = \frac{1}{1+\exp(-\sum_{n,k}\beta_k^{(n)} c_k^{(n)})} \tag{5}$$

so that they are constrained to the range $0\le\hat{y}\le 1$. For a visual overview of the architecture, see Fig. 6.

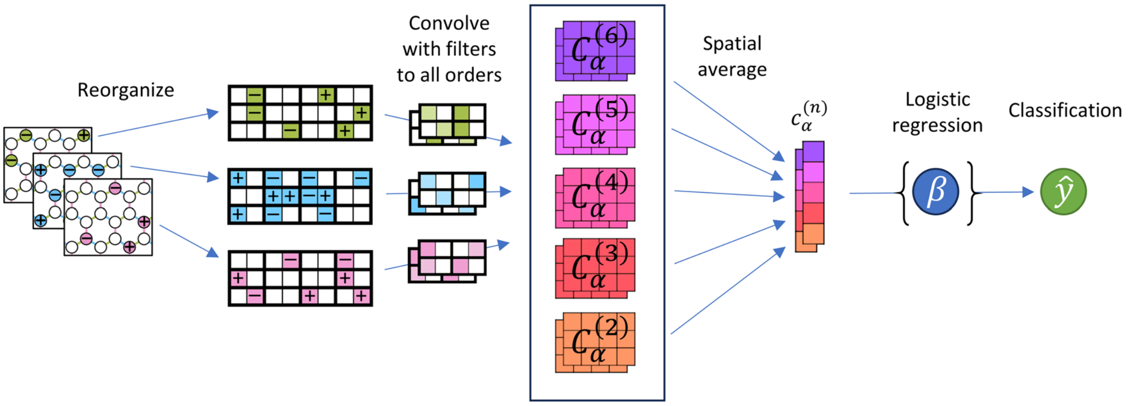

Fig. 6 Visualization of correlator convolutional neural network architecture. Input snapshots $B^\alpha(x)$ (with α enumerating the spin channels) are shaped into rectangular data arrays with a 1 × 2 unit cell. The inputs are convolved with filters to obtain orders $n = 2 \ldots 6$ of correlations $C_k^{(n)}(x)$ (Equation (4)), then spatially averaged to obtain correlator estimates $c_k^{(n)}$. Correlator estimates are then used in a logistic regression to predict the class of the input snapshot.

During training, the weights of the network are updated with stochastic gradient descent to optimize the loss function

$$\begin{aligned} L(y,\hat{y}) = {} & -y\log\hat{y} - (1-y)\log(1-\hat{y}) \\ & + \gamma_1 \sum_{\alpha,k,a} |f_{\alpha,k}(a)| + \gamma_2 \sum_{\alpha,k,a} f_{\alpha,k}(a)^2 \end{aligned} \tag{6}$$

where $y \in \{0, 1\}$ is the ground truth label of the snapshot, and γ_1 and γ_2 are L1 and L2 regularization strengths, respectively. We took $\gamma_1 = 0.005$ and $\gamma_2 = 0.002$. The training was performed for 20 epochs consisting of 9000 snapshots each with a learning rate of 0.006, using Adam stochastic gradient descent. For the regularization path analysis, the weights f are kept fixed, and the model is retrained in the same way with loss function

$$L(y,\hat{y}) = -y\log\hat{y} - (1-y)\log(1-\hat{y}) + \gamma\sum_{k,n} |\beta_k^{(n)}|, \tag{7}$$

where γ is the regularization strength to be swept over.

Data availability

Data is available upon request to the authors.

Code availability

Code is available upon request to the authors.

Received: 20 July 2023; Accepted: 24 January 2024;
Published online: 14 February 2024

References

1. James, D. F. V., Kwiat, P. G., Munro, W. J. & White, A. G. Measurement of qubits. *Phys. Rev. A* **64**, 052312 (2001).
2. Huang, H.-Y., Kueng, R. & Preskill, J. Predicting many properties of a quantum system from very few measurements. *Nat. Phys.* **16**, 1050–1057 (2020).
3. Carrasquilla, J. Machine learning for quantum matter. *Adv. Phys.: X* **5**, 1797528 (2020).
4. Miles, C. et al. Correlator convolutional neural networks as an interpretable architecture for image-like quantum matter data. *Nat. Commun.* **12**, 3905 (2021).
5. Arnold, J. & Schäfer, F. Replacing neural networks by optimal analytical predictors for the detection of phase transitions. *Phys. Rev. X* **12**, 031044 (2022).
6. Miles, C. et al. Machine learning discovery of new phases in programmable quantum simulator snapshots. *Phys. Rev. Res.* **5**, 013026 (2023).
7. Zhang, Y., Melko, R. G. & Kim, E.-A. Machine learning $\mathbb{Z}_2$ quantum spin liquids with quasiparticle statistics. *Phys. Rev. B* **96**, 245119 (2017).
8. Huang, H.-Y., Kueng, R., Torlai, G., Albert, V. V. & Preskill, J. Provably efficient machine learning for quantum many-body problems. *Science* **377**, eabk3333 (2022).
9. Hermanns, M., Kimchi, I. & Knolle, J. Physics of the Kitaev model: fractionalization, dynamic correlations, and material connections. *Annu. Rev. Condens. Matter Phys.* **9**, 17–33 (2018).
10. Knolle, J. & Moessner, R. A field guide to spin liquids. *Annu. Rev. Condens. Matter Phys.* **10**, 451–472 (2019).
11. Gohlke, M., Verresen, R., Moessner, R. & Pollmann, F. Dynamics of the Kitaev-Heisenberg model. *Phys. Rev. Lett.* **119**, 157203 (2017).
12. Trebst, S. & Hickey, C. Kitaev materials. *Phys. Rep.* **950**, 1–37 (2022).
13. Zhu, Z., Kimchi, I., Sheng, D. N. & Fu, L. Robust non-Abelian spin liquid and a possible intermediate phase in the antiferromagnetic Kitaev model with magnetic field. *Phys. Rev. B* **97**, 241110 (2018).
14. Gohlke, M., Moessner, R. & Pollmann, F. Dynamical and topological properties of the Kitaev model in a [111] magnetic field. *Phys. Rev. B* **98**, 014418 (2018).
15. Jiang, Y.-F., Devereaux, T. P. & Jiang, H.-C. Field-induced quantum spin liquid in the Kitaev-Heisenberg model and its relation to α-RuCl$_3$. *Phys. Rev. B* **100**, 165123 (2019).
16. Ronquillo, D. C., Vengal, A. & Trivedi, N. Signatures of magnetic-field-driven quantum phase transitions in the entanglement entropy and spin dynamics of the Kitaev honeycomb model. *Phys. Rev. B* **99**, 140413 (2019).
17. Jiang, H.-C., Wang, C.-Y., Huang, B. & Lu, Y.-M. Field induced quantum spin liquid with spinon Fermi surfaces in the Kitaev model (2018).
18. Hickey, C. & Trebst, S. Emergence of a field-driven U(1) spin liquid in the Kitaev honeycomb model. *Nat. Commun.* **10**, 530 (2019).
19. Patel, N. D. & Trivedi, N. Magnetic field-induced intermediate quantum spin liquid with a spinon Fermi surface. *Proc. Natl Acad. Sci. USA* **116**, 12199–12203 (2019).
20. Jiang, M.-H. et al. Tuning topological orders by a conical magnetic field in the Kitaev model. *Phys. Rev. Lett.* **125**, 177203 (2020).
21. Zhang, S.-S., Halász, G. B. & Batista, C. D. Theory of the Kitaev model in a [111] magnetic field. *Nat. Commun.* **13**, 399 (2022).
22. Feng, S., Agarwala, A., Bhattacharjee, S. & Trivedi, N. Anyon dynamics in field-driven phases of the anisotropic kitaev model. *Phys. Rev. B* **108**, 035149 (2023).
23. Liu, K., Sadoune, N., Rao, N., Greitemann, J. & Pollet, L. Revealing the phase diagram of Kitaev materials by machine learning: cooperation and competition between spin liquids. *Phys. Rev. Res.* **3**, 023016 (2021).
24. White, S. R. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.* **69**, 2863–2866 (1992).
25. White, S. R. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B* **48**, 10345–10356 (1993).
26. Fishman, M., White, S. & Stoudenmire, E. The ITensor software library for tensor network calculations. *SciPost Physics Codebases* 4 https://doi.org/10.21468/SciPostPhysCodeb.4 (2022).
27. Efron, B., Hastie, T., Johnstone, I. & Tibshirani, R. Least angle regression. *Ann. Stat.* **32**, 407–499 (2004).
28. Kitaev, A. Anyons in an exactly solved model and beyond. *Ann. Phys.* **321**, 2–111 (2006).
29. Jackeli, G. & Khaliullin, G. Mott insulators in the strong spin-orbit coupling limit: from Heisenberg to a quantum compass and Kitaev Models. *Phys. Rev. Lett.* **102**, 017205 (2009).

30. Kasahara, Y. et al. Majorana quantization and half-integer thermal quantum Hall effect in a Kitaev spin liquid. *Nature* **559**, 227–231 (2018).
31. Ferris, A. J. & Vidal, G. Perfect sampling with unitary tensor networks. *Phys. Rev. B* **85**, 165146 (2012).
32. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B (Methodol.)* **58**, 267–288 (1996). 2346178.
33. Pradhan, S., Patel, N. D. & Trivedi, N. Two-magnon bound states in the Kitaev model in a [111] field. *Phys. Rev. B* **101**, 180401 (2020).
34. White, S. R., Affleck, I. & Scalapino, D. J. Friedel oscillations and charge density waves in chains and ladders. *Phys. Rev. B* **65**, 165122 (2002).
35. Mross, D. F. & Senthil, T. Charge Friedel oscillations in a Mott insulator. *Phys. Rev. B* **84**, 041102 (2011).
36. He, W.-Y., Xu, X. Y., Chen, G., Law, K. T. & Lee, P. A. Spinon Fermi surface in a cluster mott insulator model on a triangular lattice and possible application to 1T- TaS2. *Phys. Rev. Lett.* **121**, 046401 (2018).
37. Ruan, W. et al. Evidence for quantum spin liquid behaviour in single-layer 1T-TaSe2 from scanning tunnelling microscopy. *Nat. Phys.* **17**, 1154–1161 (2021).
38. Feng, S., Alvarez, G. & Trivedi, N. Gapless to gapless phase transitions in quantum spin chains. *Phys. Rev. B* **105**, 014435 (2022).
39. Baskaran, G., Mandal, S. & Shankar, R. Exact results for spin dynamics and fractionalization in the Kitaev model. *Phys. Rev. Lett.* **98**, 247201 (2007).
40. Feng, S., He, Y. & Trivedi, N. Detection of long-range entanglement in gapped quantum spin liquids by local measurements. *Phys. Rev. A* **106**, 042417 (2022).

Acknowledgements

We thank Leon Balents, Natasha Perkins, John Tranquada, and Simon Trebst for helpful discussions. K.Z. acknowledges support by the NSF under EAGER OSP-136036 and the Natural Sciences and Engineering Research Council of Canada (NSERC) under PGS-D-557580-2021. Y.L. and E.A.K. acknowledge support by the Gordon and Betty Moore Foundation's EPiQS Initiative, Grant GBMF10436, and a New Frontier Grant from Cornell University's College of Arts and Sciences. E.A.K. acknowledges support by the NSF under OAC-2118310, EAGER OSP-136036, the Ewha Frontier 10-10 Research Grant, and the Simons Fellowship in Theoretical Physics award 920665. S.F. acknowledges support from NSF Materials Research Science and Engineering Center (MRSEC) Grant No. DMR-2011876, and N.T. from NSF-DMR 2138905.

Author contributions

E-AK and NT conceived the idea and supervised the project. KZ, YL, and E-AK formed the machine learning and data processing strategies. SF performed the DMRG optimization and data processing. KZ performed the machine learning analysis. All authors contributed to interpreting the results and writing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42005-024-01542-8.

Correspondence and requests for materials should be addressed to Eun-Ah Kim.

Peer review information *Communications Physics* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2024

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com