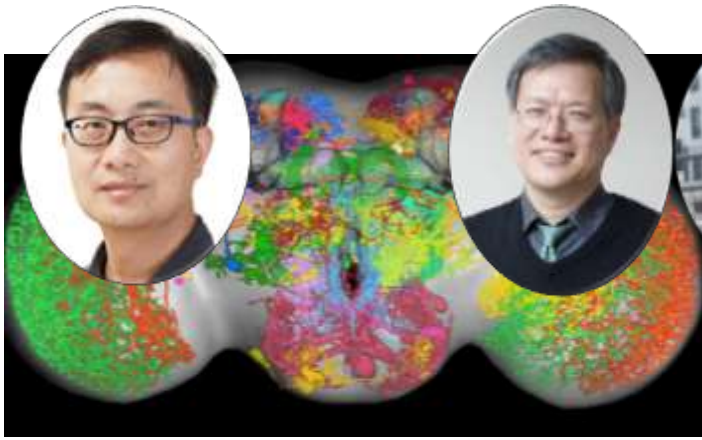
Exploring Quantum Many-Body Problems by Random Sampling Neural Networks and Self-Supervised Learning

Daw-Wei Wang
Physics Department, National Tsing Hua University
Physics Division, National Center for Theoretical Sciences

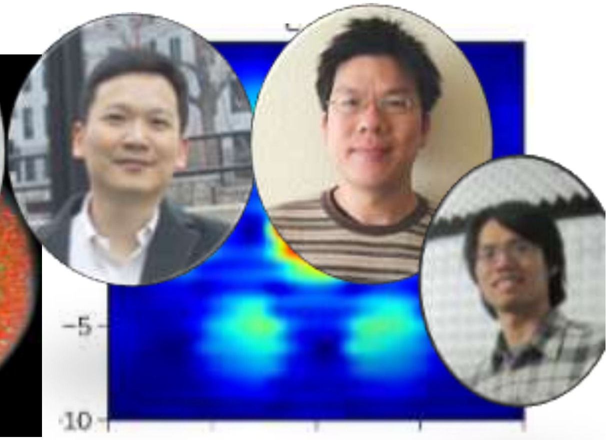
個人近年來與人工智慧相關的應用研究

(Artificial Intelligence for Fundamental Research Group)

天文：尋找初生恆星與重力波探測超新星(賴詩萍與潘國全)



腦科學：從果蠅腦神經影像解碼大腦訊息傳遞(羅中泉與江安世)



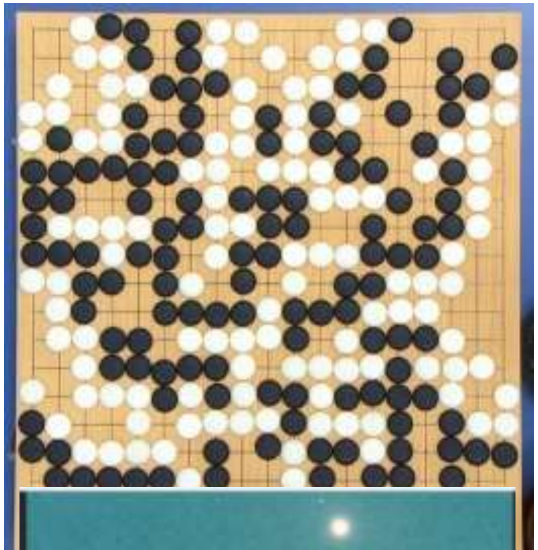
物理：多體系統與量子電腦(牟中瑜、陳柏中與林晏詳)

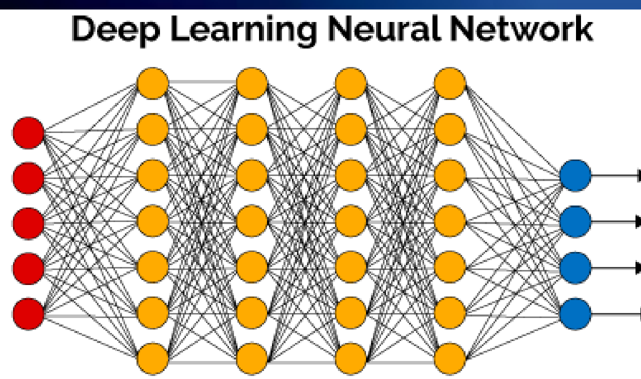
心理：學生自我傷害篩檢與預警系統(區國良、李昆樺、許育光、陳宜欣、劉奕汶等)

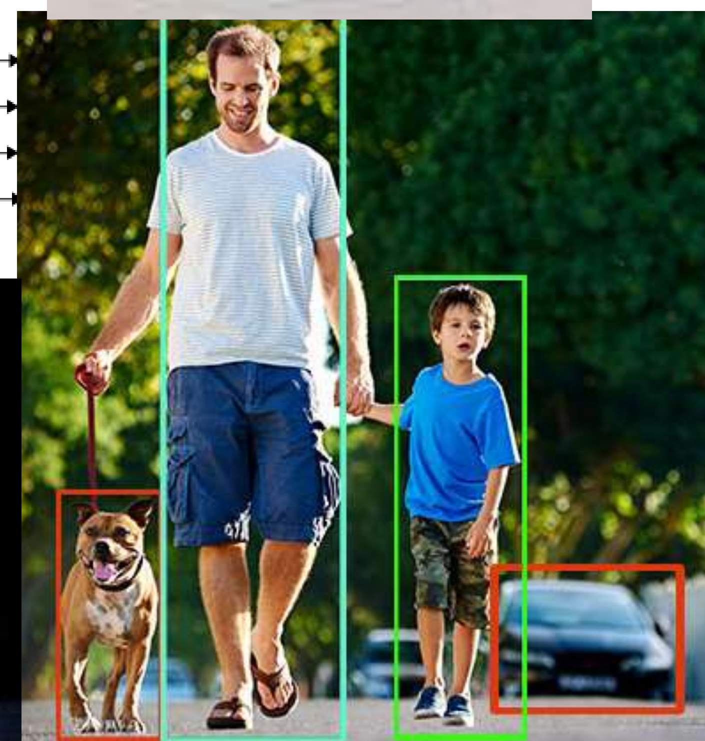
AI公共化：(林文源與其他)

司法：親權判決預測與司法院量刑資訊系統(林昀嫺、連孟琦、盧映潔、謝國欣)

AI is the most interdisciplinary research in human history







Outline :

I. Supervised, Unsupervised and Self-Supervised Learning

II. Identify Topological Phase Transitions from Experimental Data

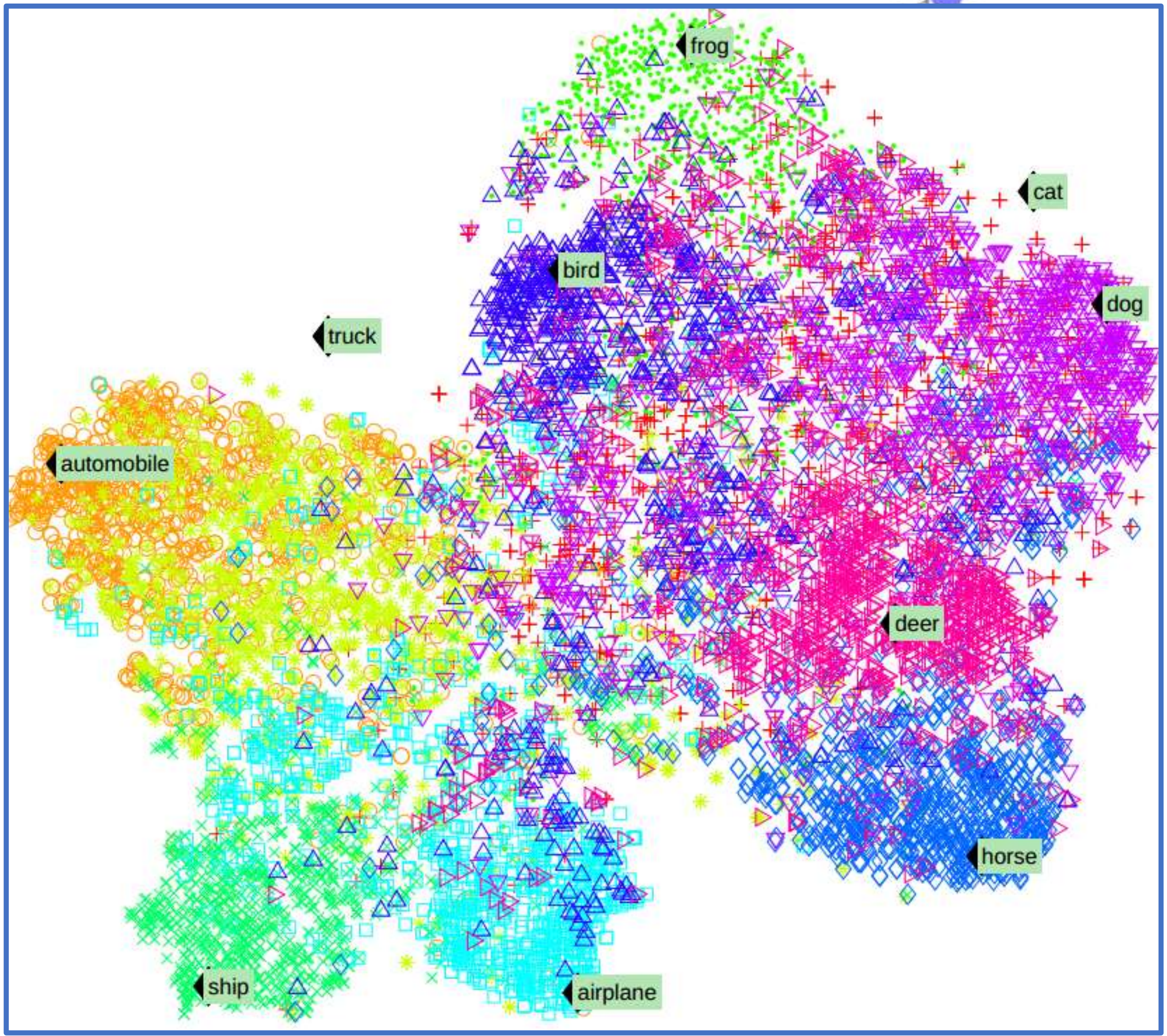
III. Random Sampling Neural Networks for Quantum Many-Body Problems

IV. Predict Long-Time Dynamics of Many-Body Systems

V. Summary and Outlook

I. Supervised, unsupervised and self-supervised learning

Standard supervised learning in computer vision

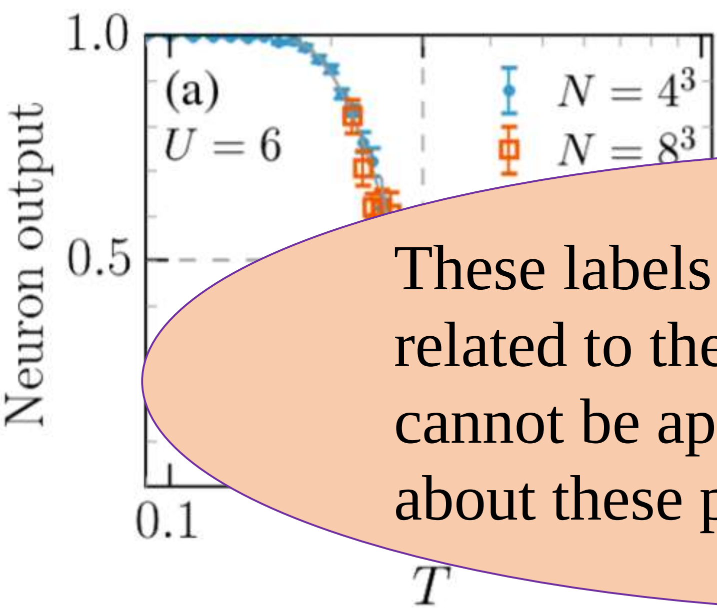


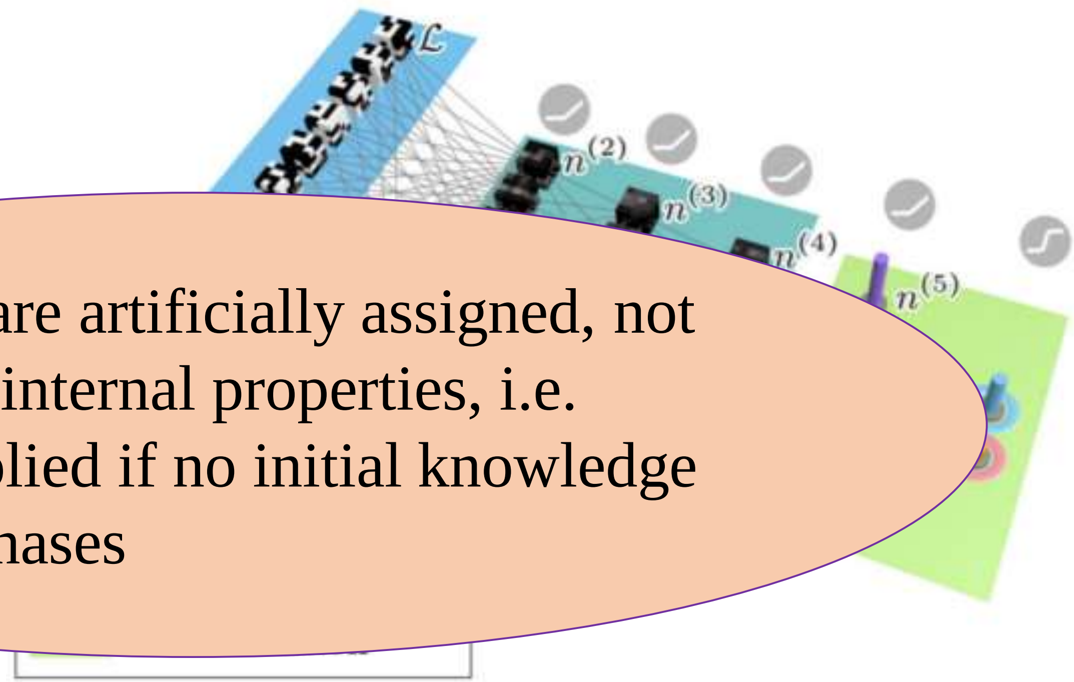
An example of supervised learning

PHYSICAL REVIEW X 7, 031038 (2017)

$$H = -t\sum_{\langle ij\rangle\sigma} c_{i\sigma}^{\dagger} c_{j\sigma} + U\sum_{i}\left(n_{i\uparrow} - \frac{1}{2}\right)\left(n_{i\downarrow} - \frac{1}{2}\right) - \mu\sum_{i\sigma} n_{i\sigma},$$

Input feature: QMC configurations in imaginary time
Output label: Anti-ferromagnetic (low T) or normal (high T)



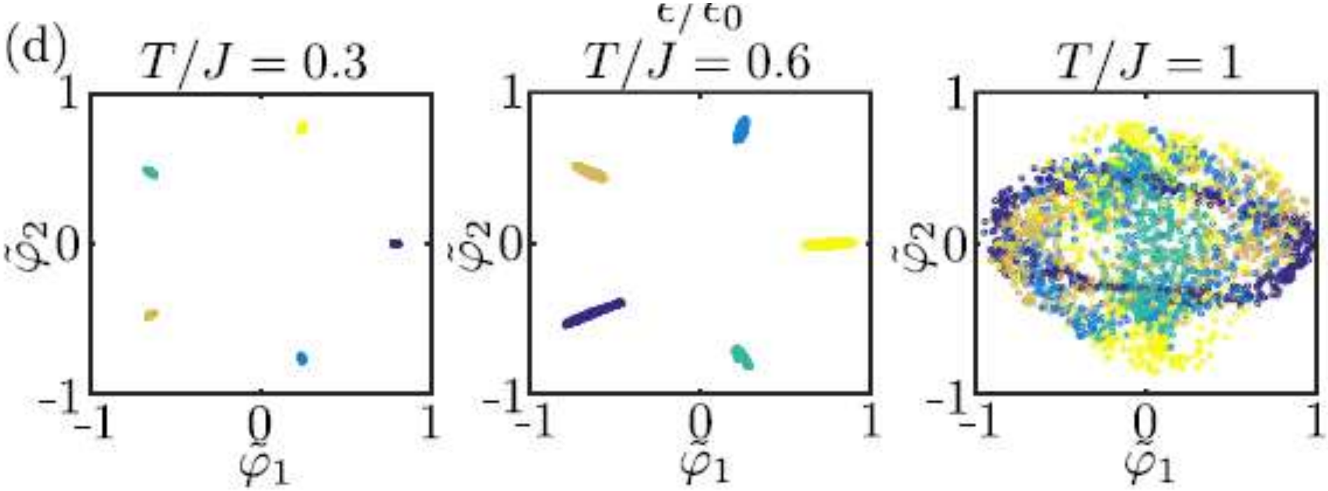


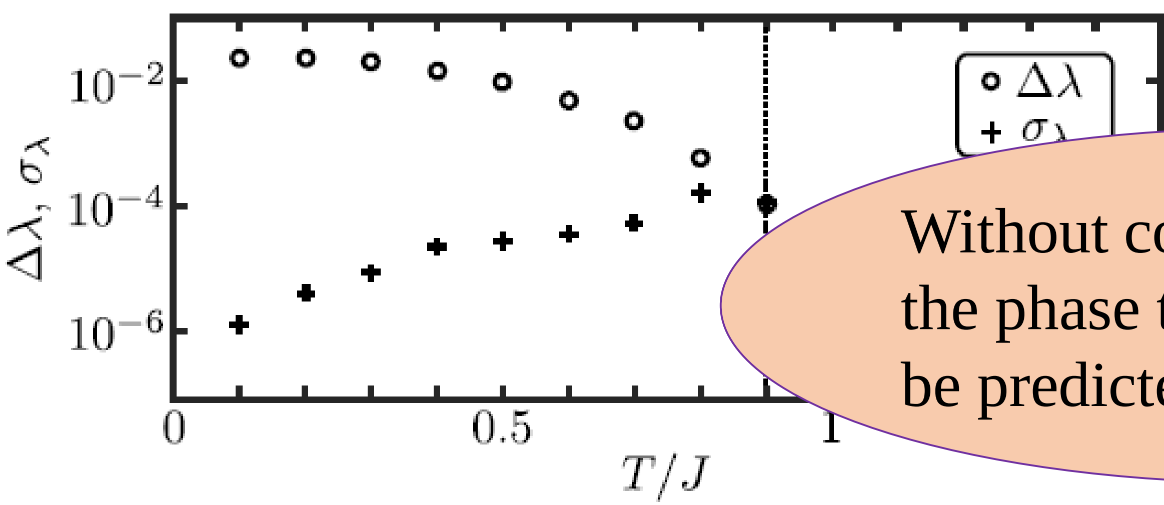
These labels are artificially assigned, not related to the internal properties, i.e. cannot be applied if no initial knowledge about these phases

An example of unsupervised learning: 2D XY model

$$H(\mathbf{s}) = -\sum_{i \neq j} J_{ij}\, \mathbf{s}_i \cdot \mathbf{s}_j$$

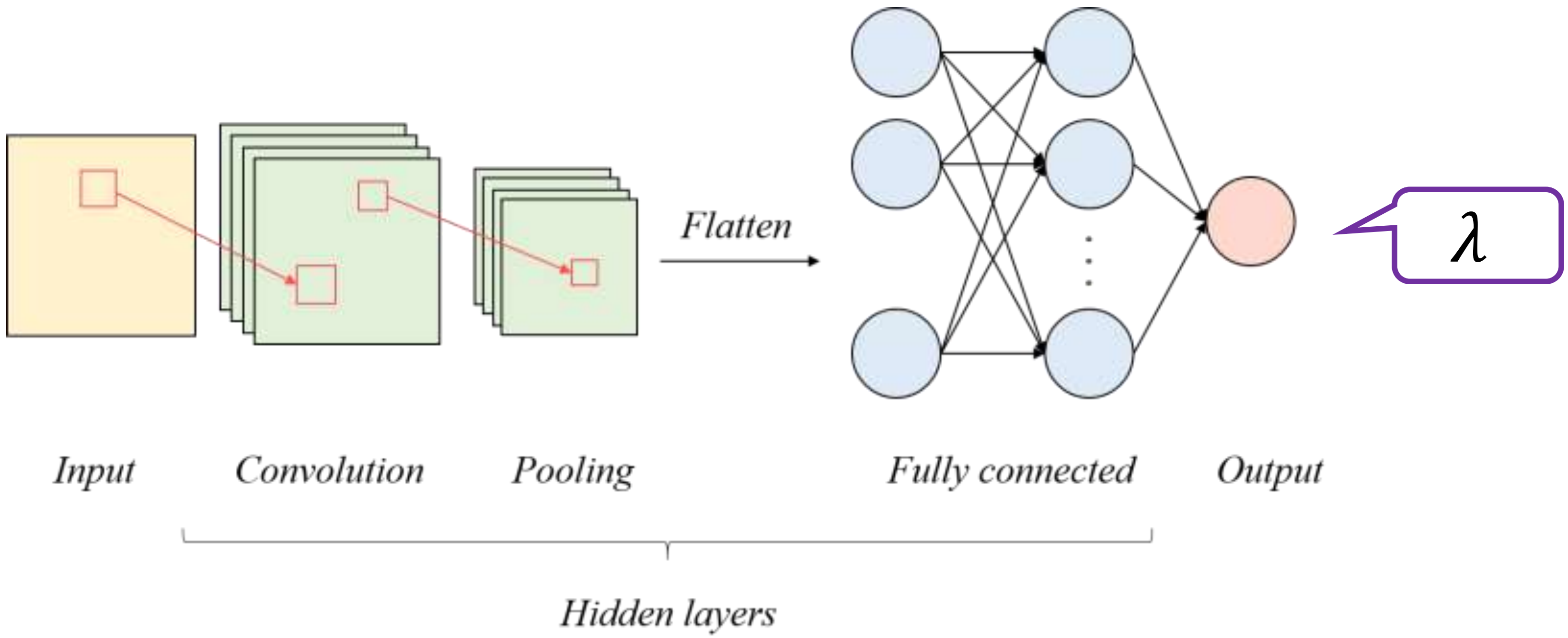
Nature Physics 15, 790–795 (2019)





Without correct labelling, the phase transition cannot be predicted accurately.

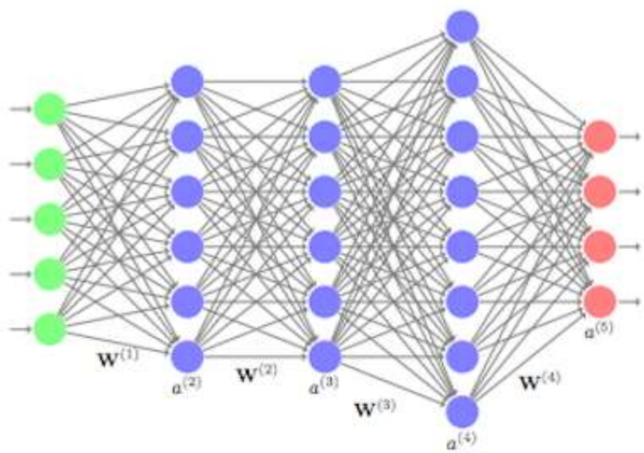
Self-Supervised Learning Method

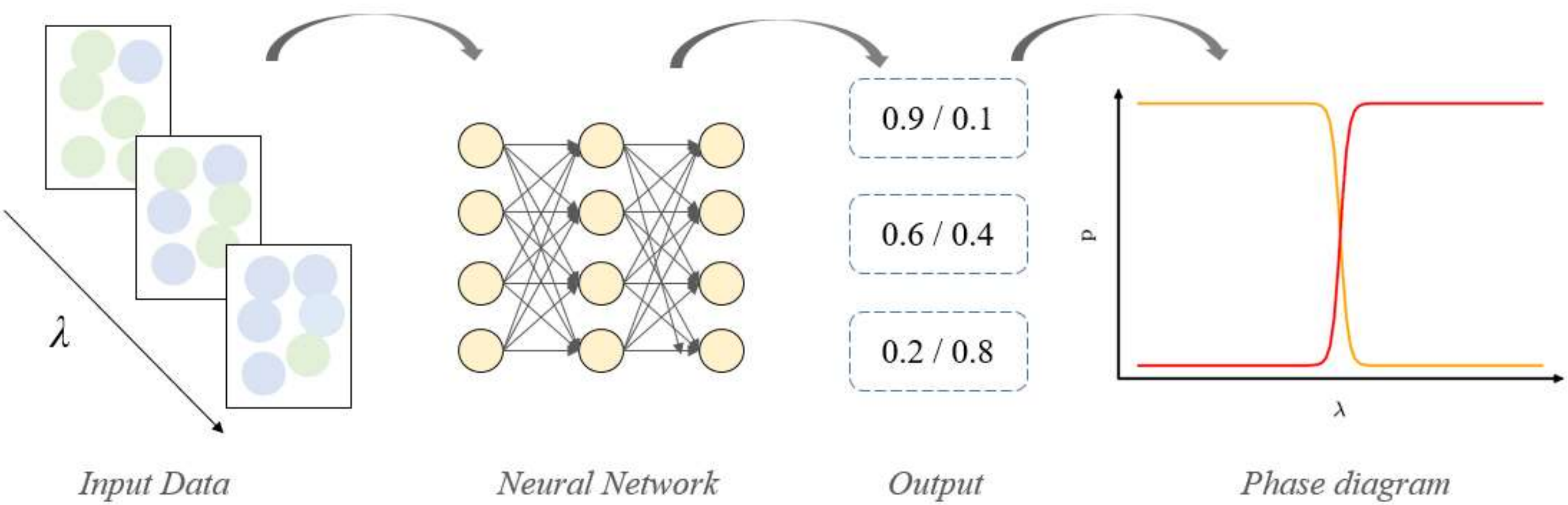


These labels are the physical quantities measured in the experiments or predicted by theoretical models, so that a physics-based functional relationship is simulated.

II. Identification of Topological Phase Transition by Self-Supervised Machine Learning Approach

(Ho and Wang, New J. Phys., 23, 083021 (2021))

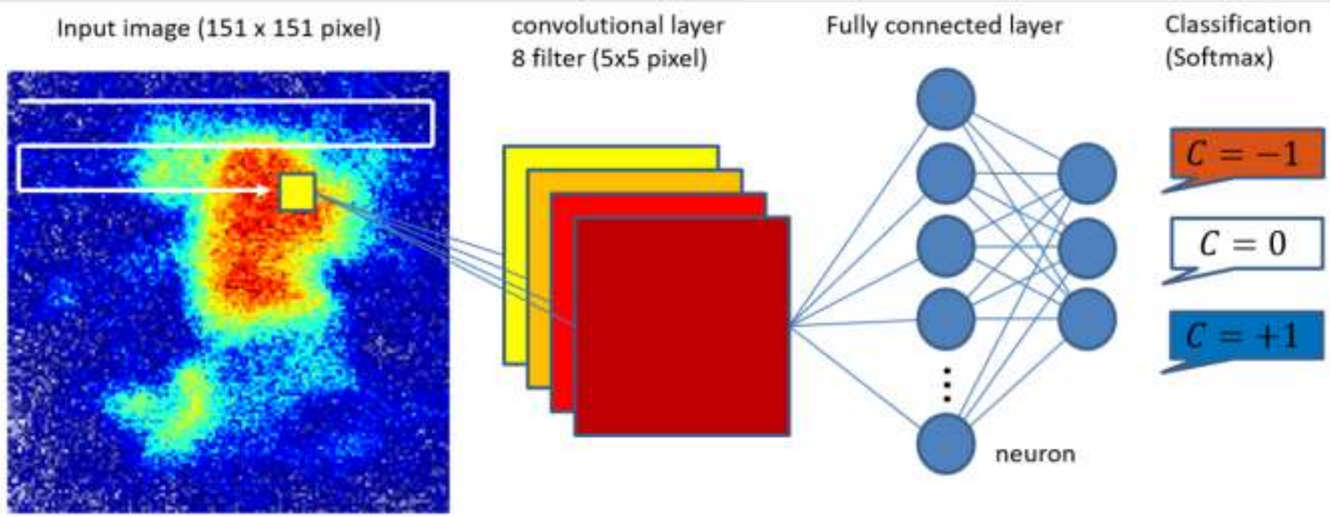


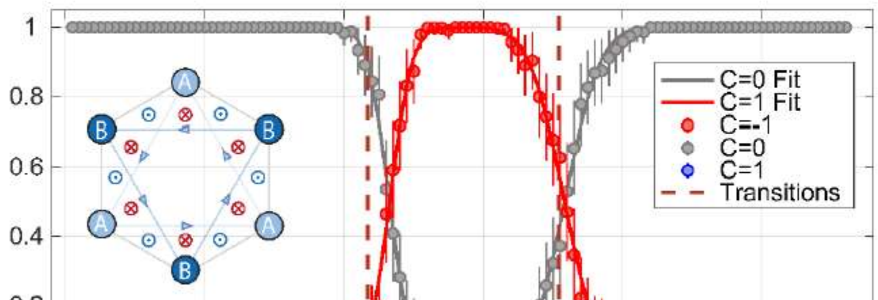


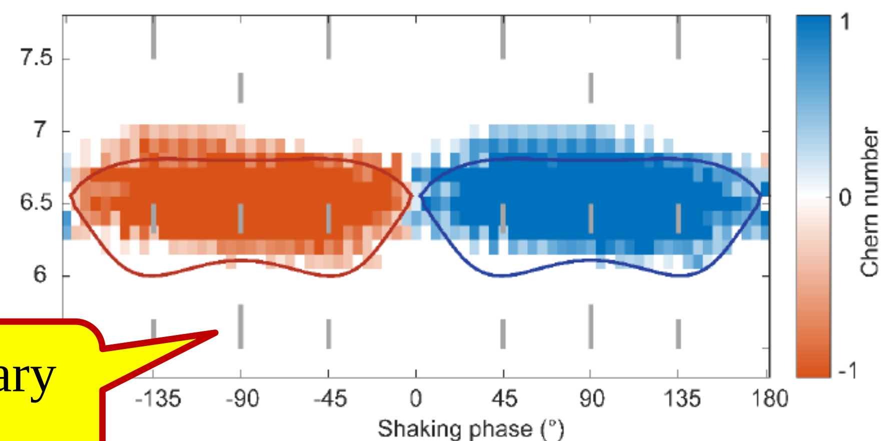
Existing work by supervised learning

Haldane model

$$\mathcal{H}_0 = \sum_i (-1)^i M c_i^\dagger c_i - \sum_{\langle i,j \rangle} t_1 c_i^\dagger c_j - \sum_{\langle\langle i,j \rangle\rangle} t_2 e^{i\phi_{ij}} c_i^\dagger c_j$$





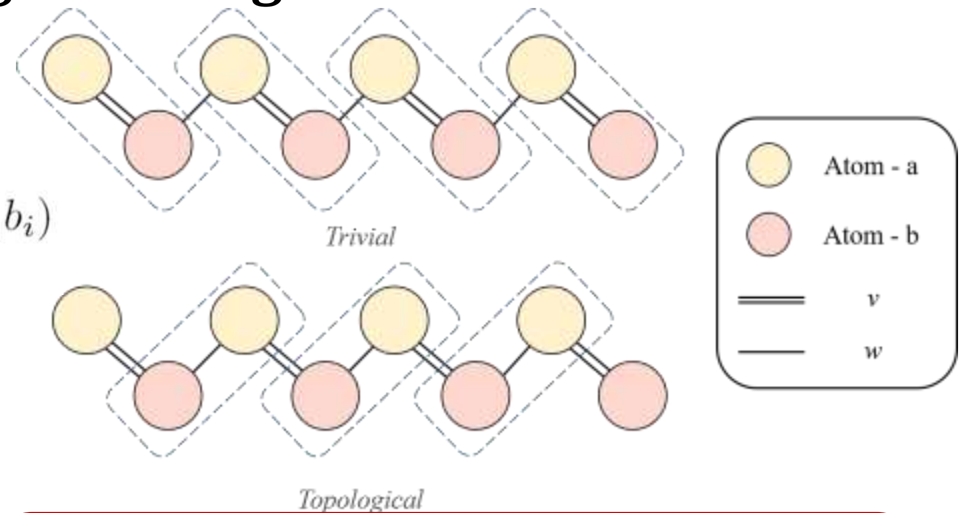


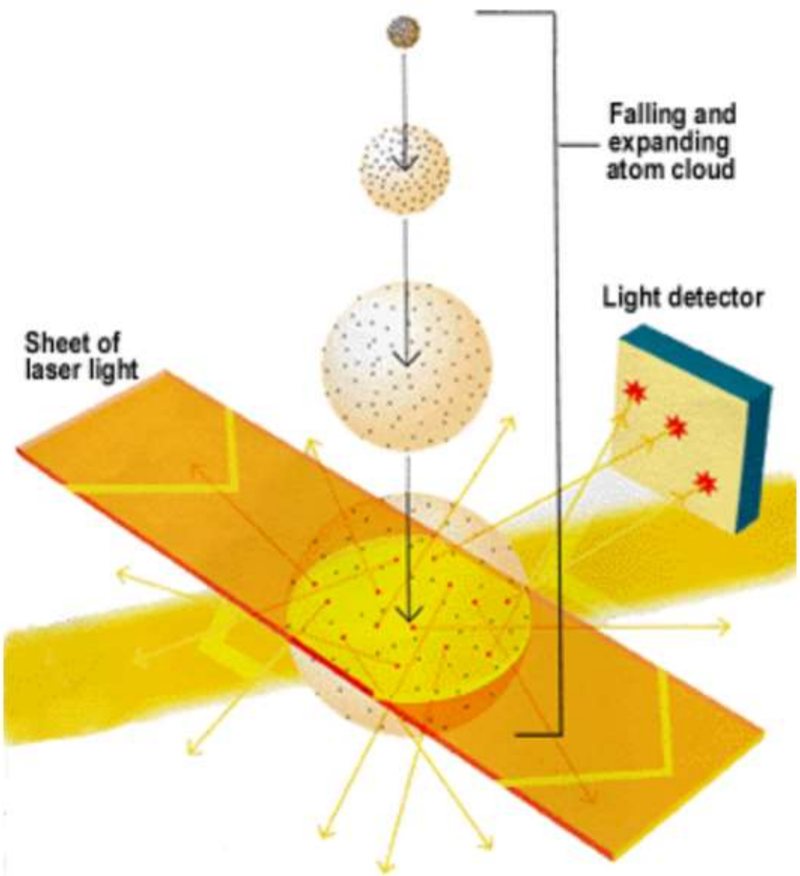
Q: If we do not know the phase boundary beforehand, can ML help to find it ?

Benno S. Rem et. al. Nature Physics (2019)

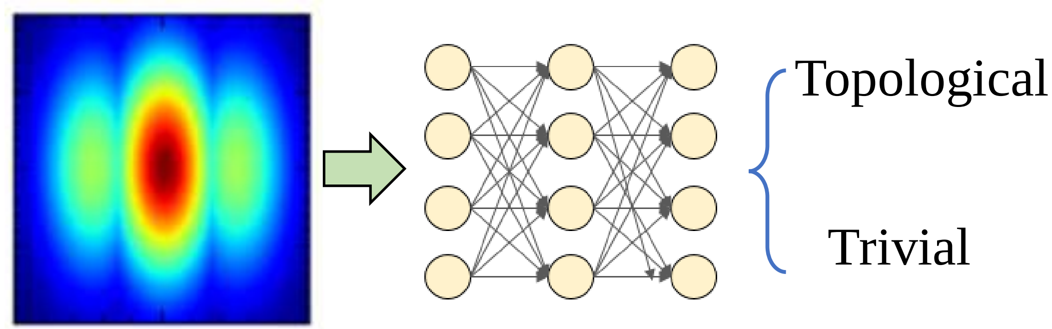
Example: 1D SSH model using ToF image

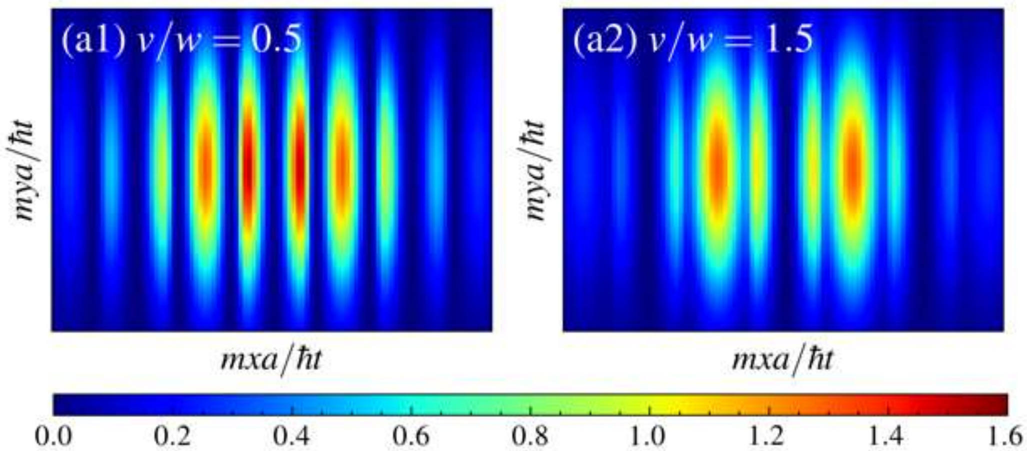
$$H = \sum_{i=1}^{L/2} v(a_i^\dagger b_i + b_i^\dagger a_i) + \sum_{i=1}^{L/2-1} w(b_i^\dagger a_{i+1} + a_{i+1}^\dagger b_i)$$

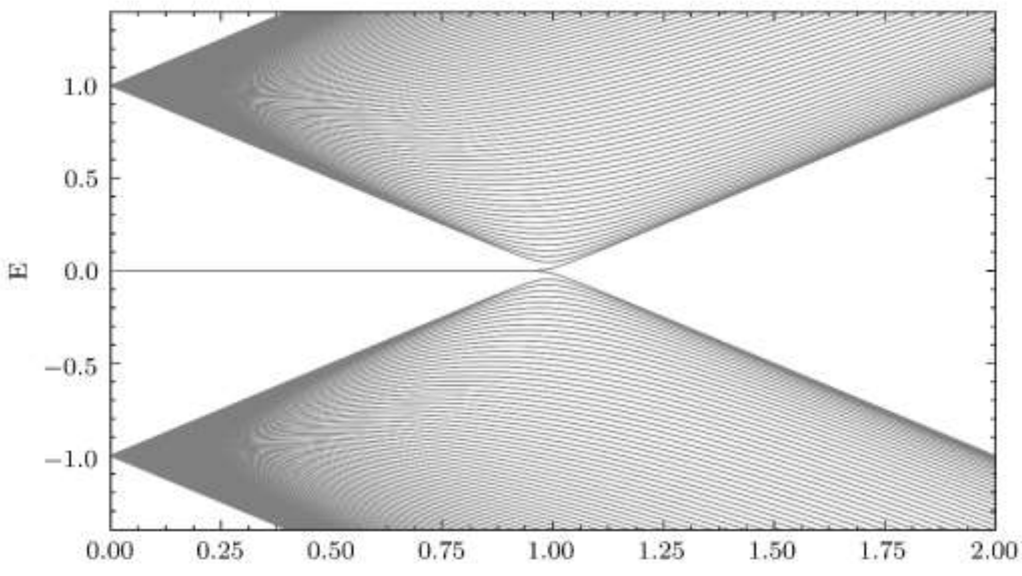


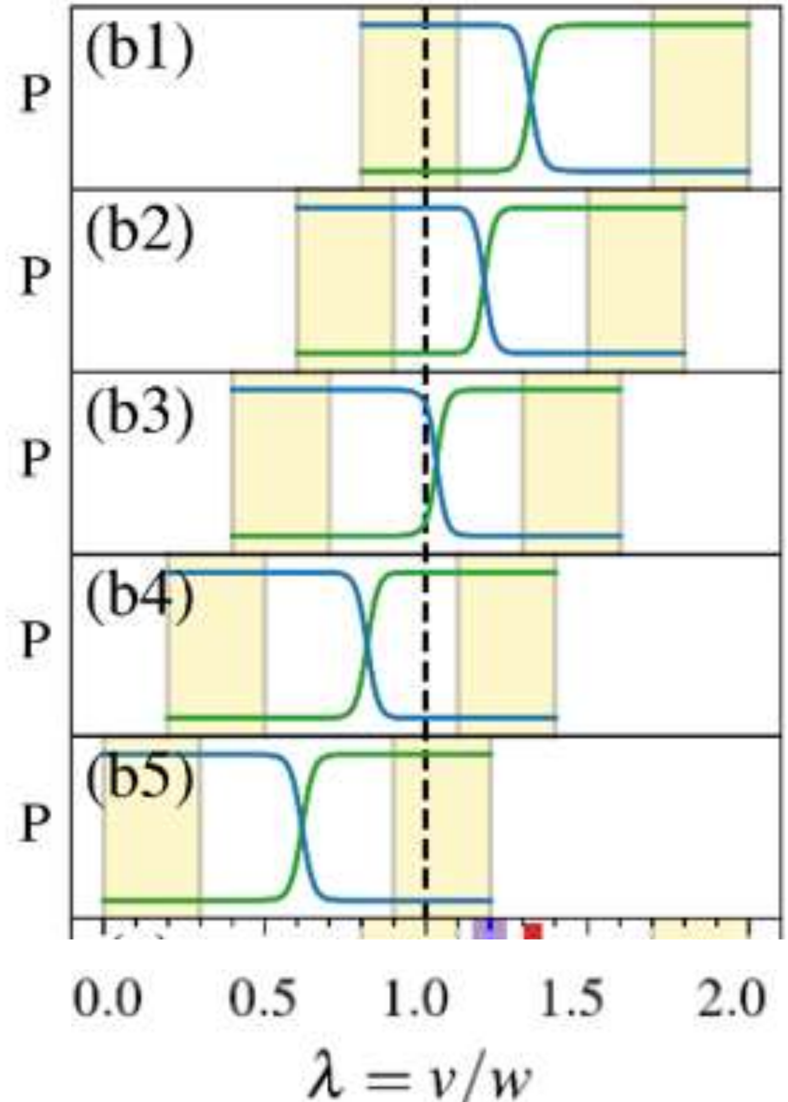


We first check if the predicted boundary could be affected by different training data, using ToF images.



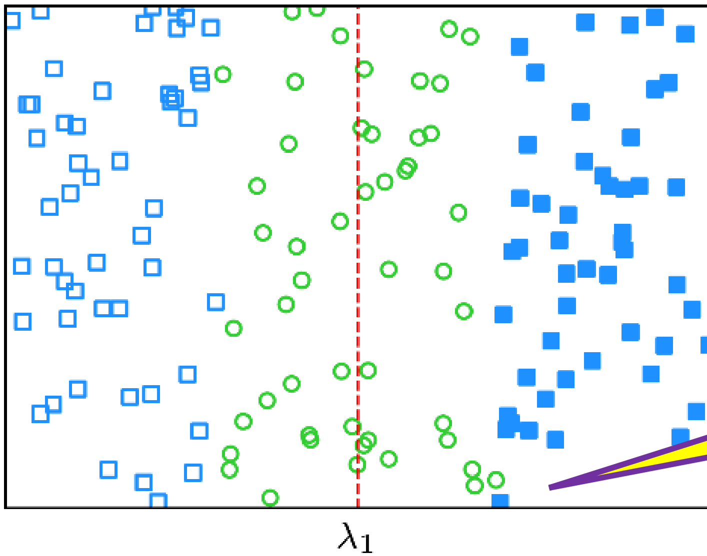




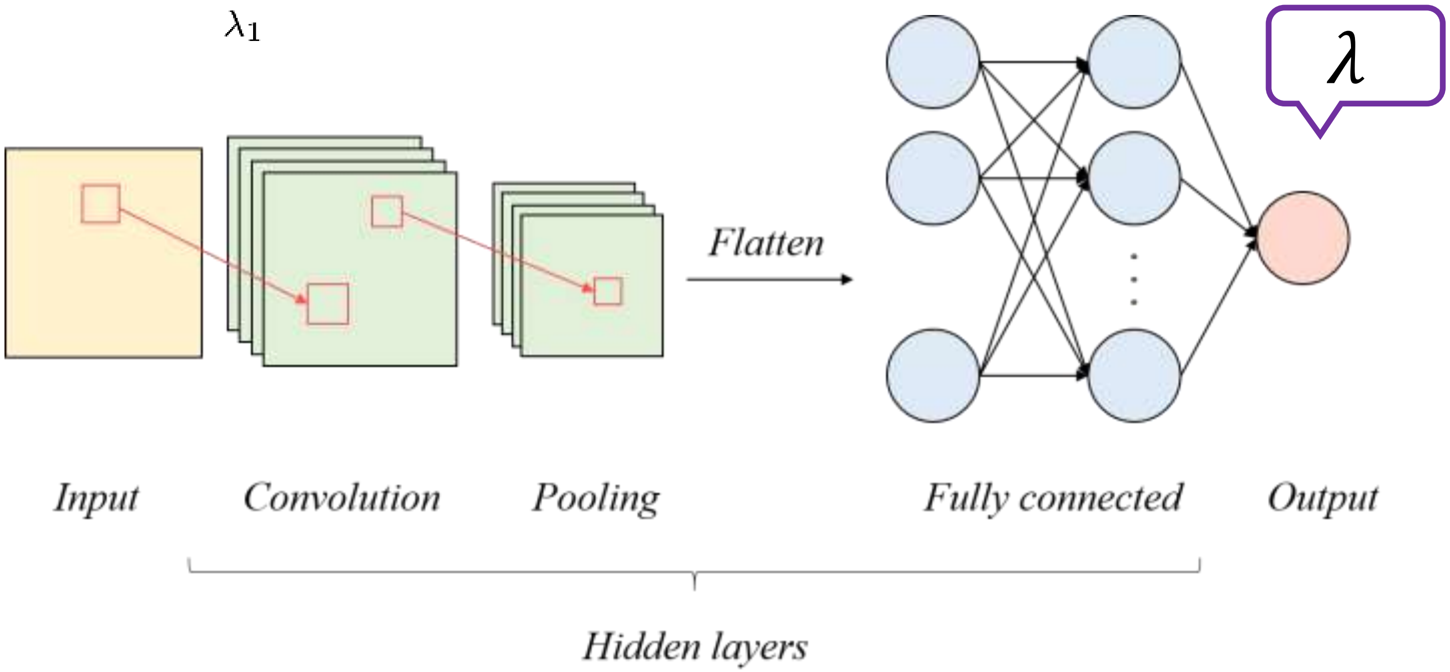


It is obvious that the “phase boundary” obtained by this approach is an artifact of machine learning, due to the data structure

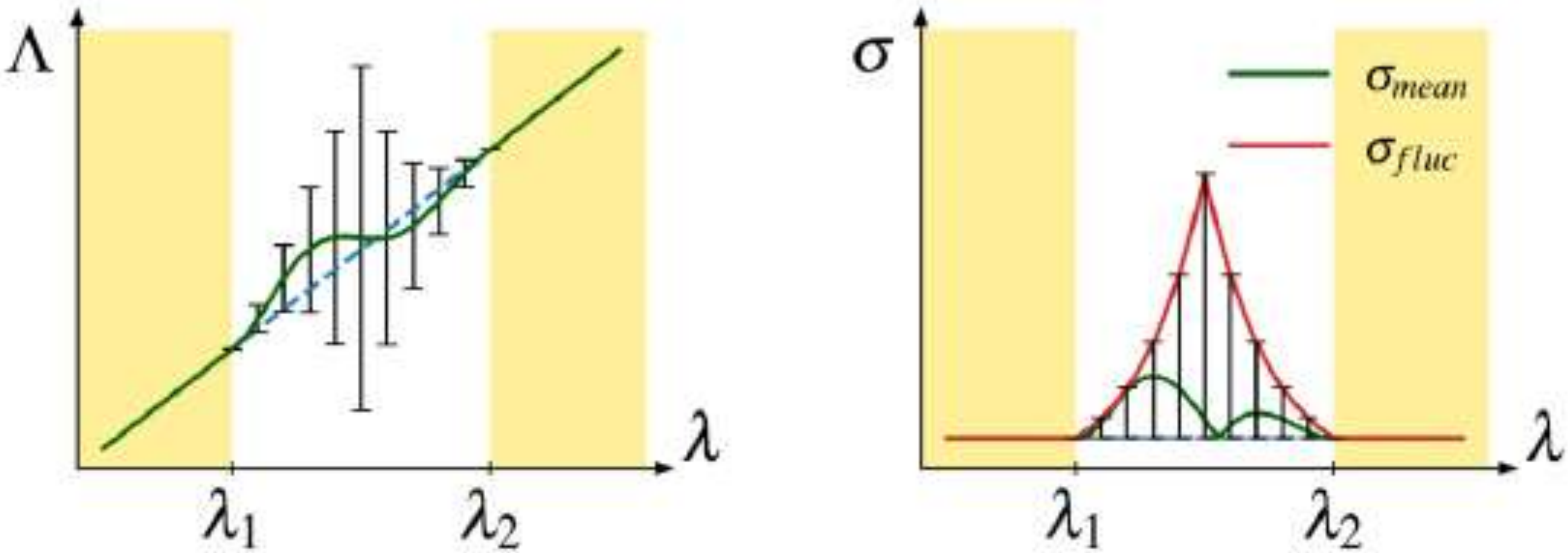
Our approach: Self-Supervised Learning Method



"Labels" (say phase I/II) are artificial. The model has to learn the true parameter to know where they are.



What to measure ? The deviation from the internal parameters

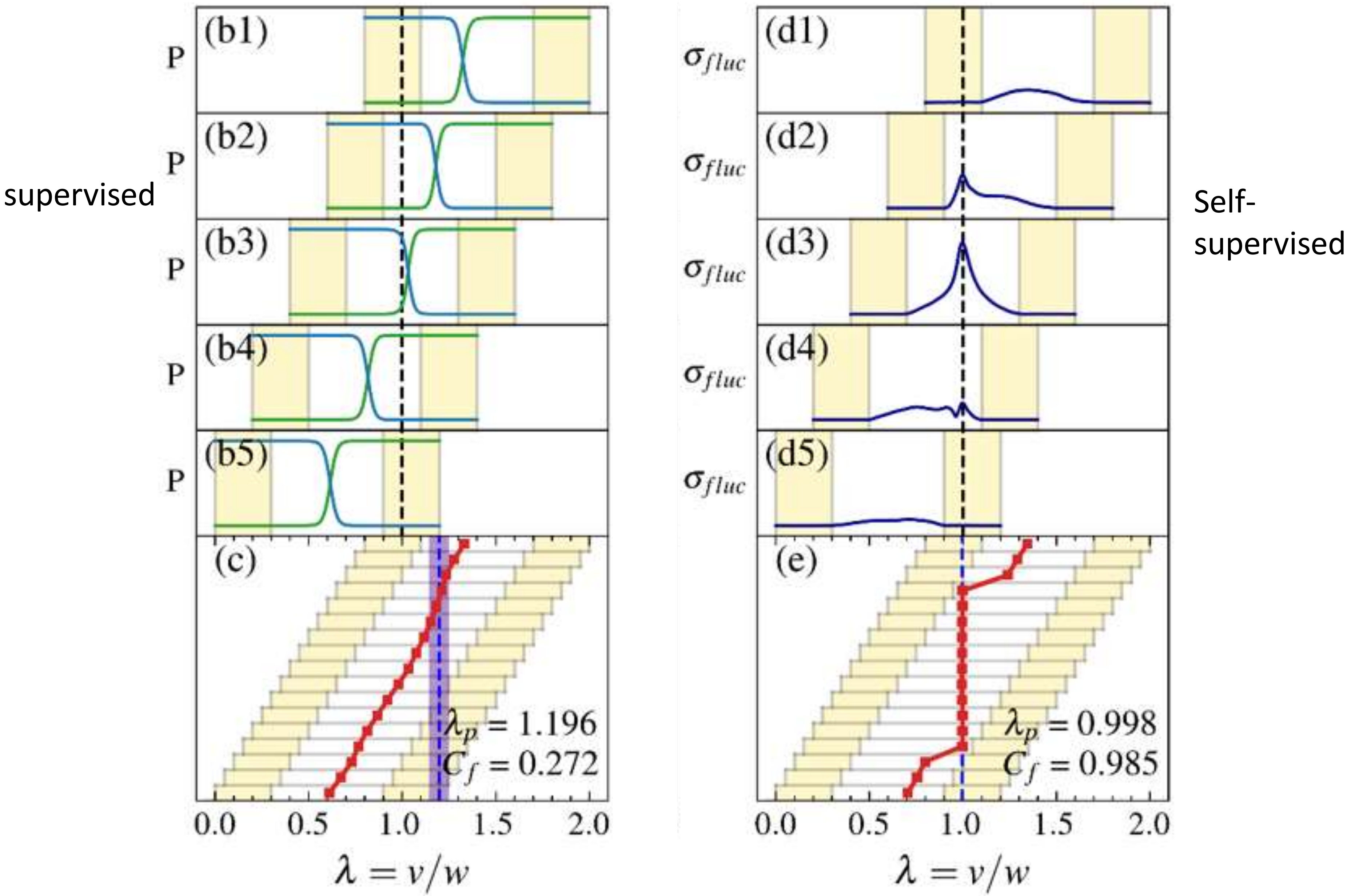


$$
\begin{aligned}
\sigma_{tot}^2(\lambda) &= \frac{1}{N_{tr}} \sum_{i=1}^{N_{tr}} |\Lambda_i(\lambda) - \lambda|^2 \\
&= |\bar{\Lambda}(\lambda) - \lambda|^2 + \frac{1}{N_{tr}} \sum_{i=1}^{N_{tr}} |\Lambda_i(\lambda) - \bar{\Lambda}(\lambda)|^2 \\
&\equiv \sigma_{mean}^2 + \sigma_{fluc}^2
\end{aligned}
$$

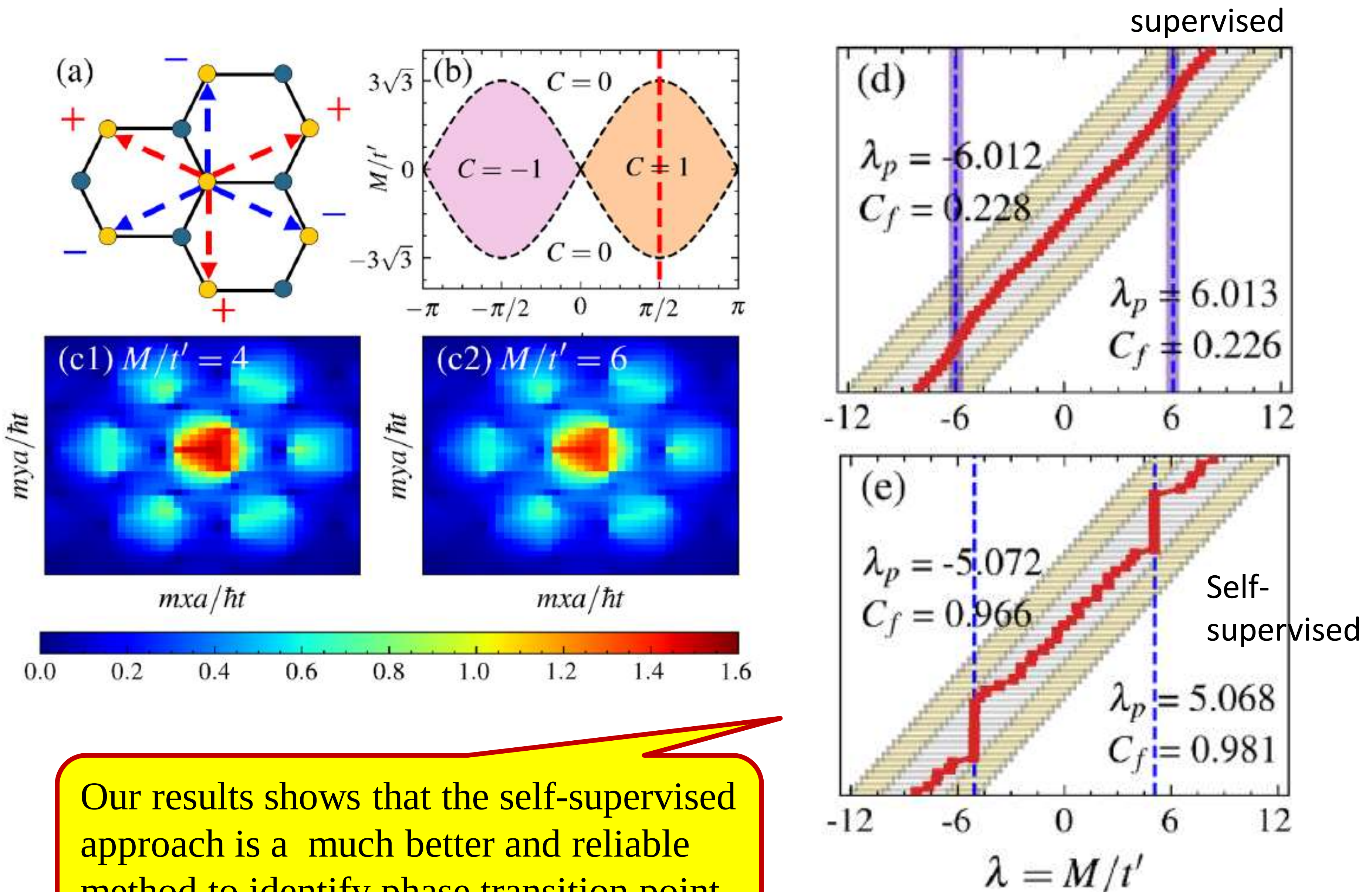
From model itself

From data structure

Results for 1D SSH model



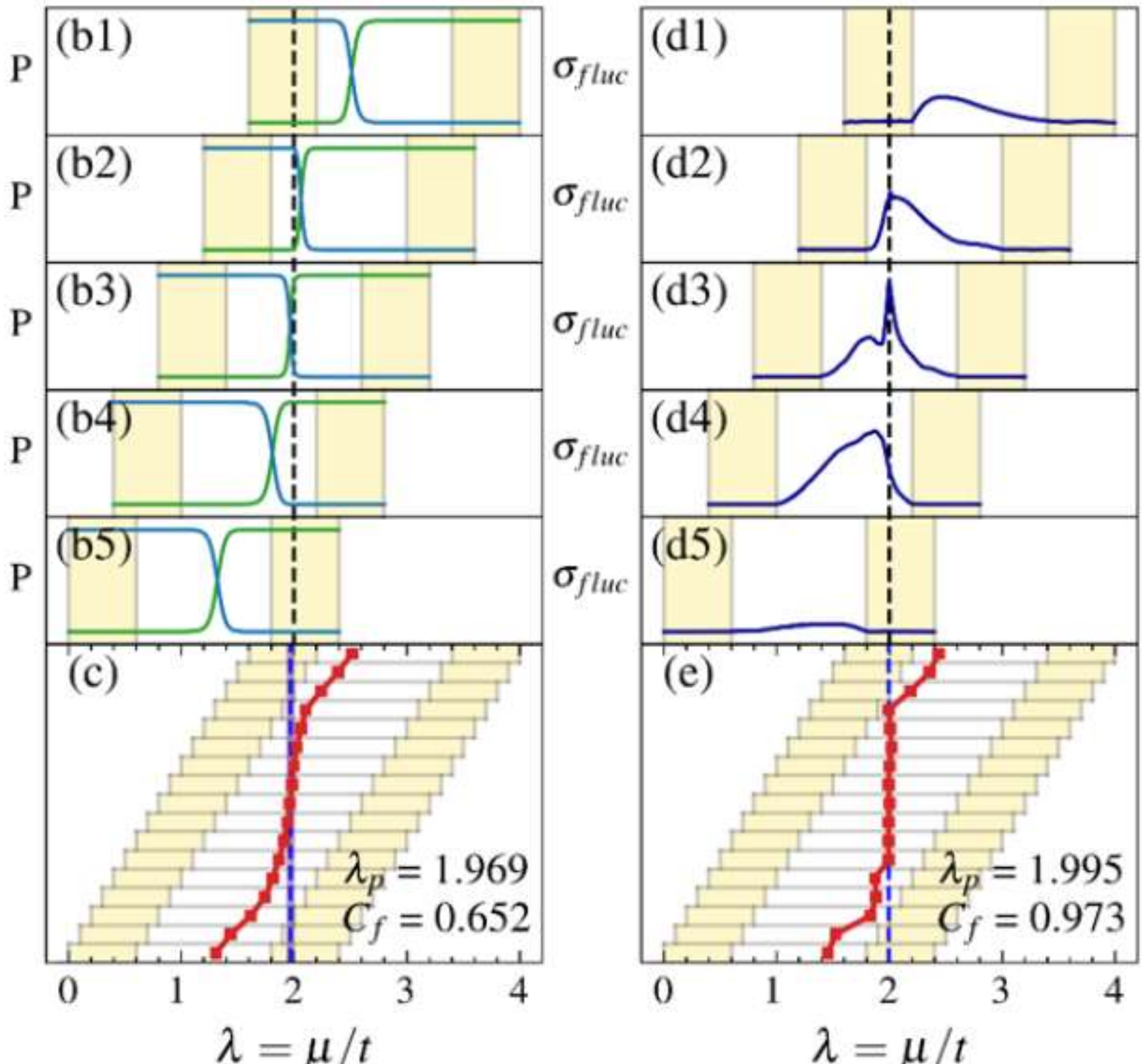
Results for 2D Haldane model



Our results shows that the self-supervised approach is a much better and reliable method to identify phase transition point.

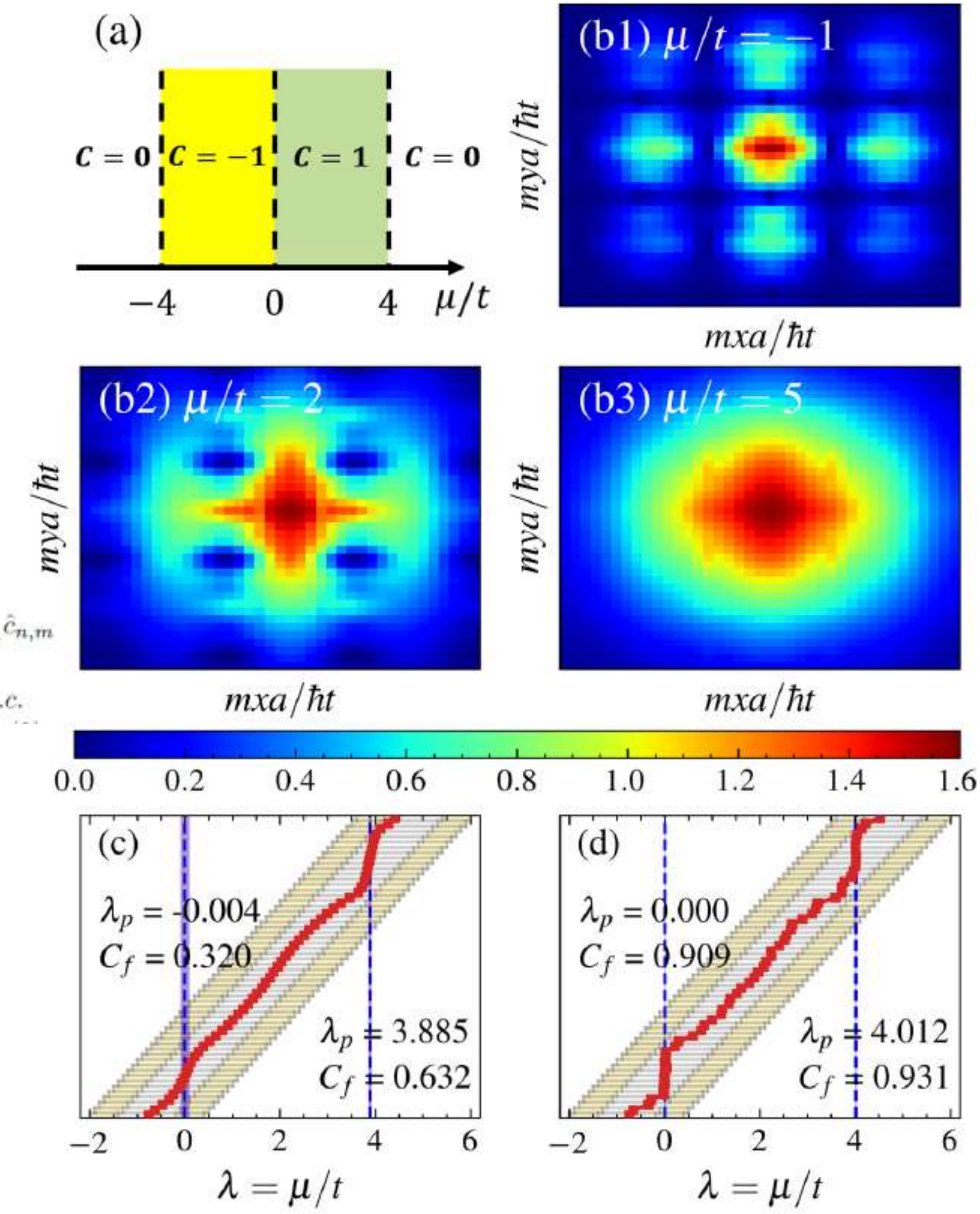
Results for 1D Kitaev model

$$\hat{H}_{Kitaev} = \sum_n (-\mu \hat{c}_n^\dagger c_n - t\hat{c}_{n+1}^\dagger c_n + \Delta \hat{c}_{n+1}^\dagger \hat{c}_n^\dagger) + h.c.$$

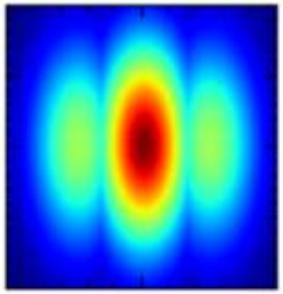


Results for 2D p-wave model

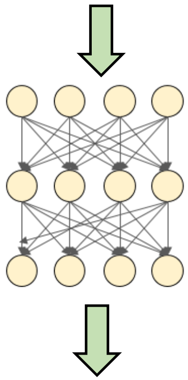
$$\hat{H}_{p-wave} = \sum_{n,m}(-\mu \hat{c}^\dagger_{n,m}\hat{c}_{n,m} - t\hat{c}^\dagger_{n+1,m}\hat{c}_{n,m} - t\hat{c}^\dagger_{n,m+1}\hat{c}_{n,m} + \Delta\hat{c}^\dagger_{n+1,m}\hat{c}^\dagger_{n,m} + i\Delta\hat{c}^\dagger_{n,m+1}\hat{c}^\dagger_{n,m}) + h.c.$$



State Function behind Self-Supervised Learning Method

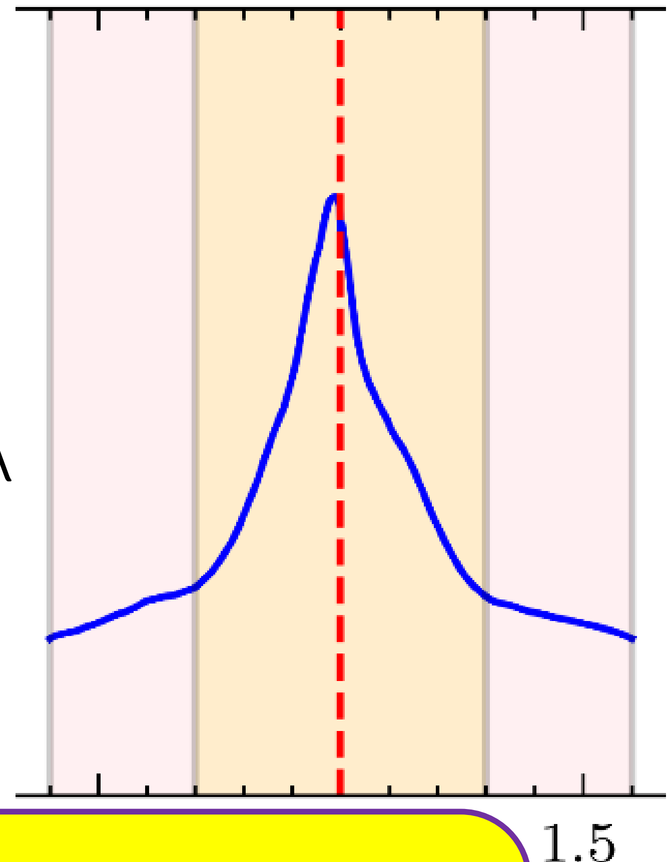


$\langle \hat{c}_{\mathbf{k}}^{\dagger} \hat{c}_{\mathbf{k}} \rangle$



λ

$$F(\langle \hat{c}_{\mathbf{k}}^{\dagger} \hat{c}_{\mathbf{k}} \rangle) = F(\hat{n}_{\mathbf{k}}) = \lambda$$



In thermodyn **veral state**
variables tha the system.
A state functi e type of
system, for ex).

Therefore, this self-supervised approach can be used to find new phase transition without knowing existing theoretical results !! → find new physics !

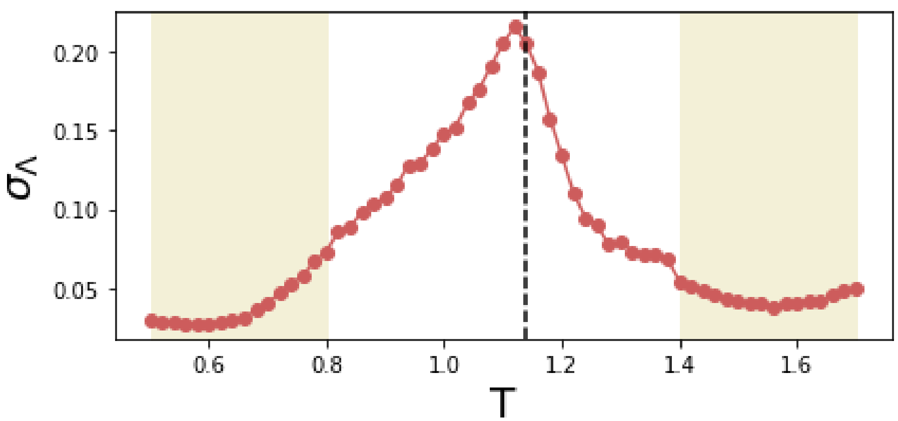
Further application to determine the types of phase transition

2D Potts model

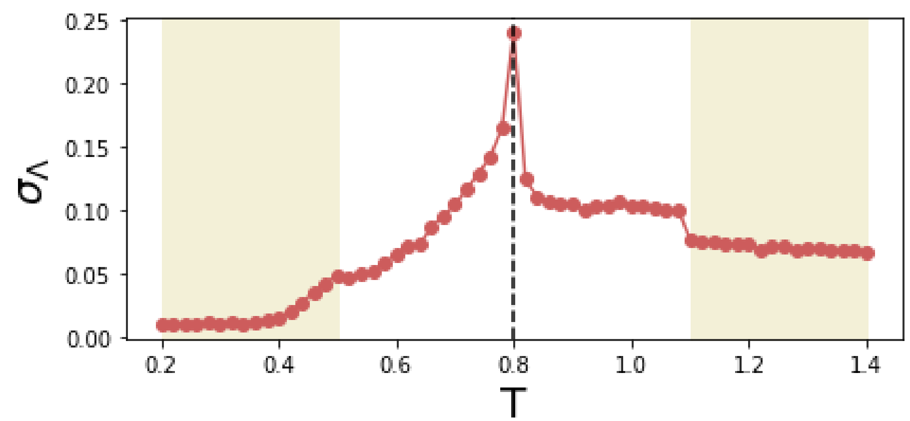
$$H = -J \sum_{<i,j>} \delta_{s_i, s_j}, \quad s_i \in \{0, 1, \ldots, q-1\} \qquad \beta J_c = \ln(1 + \sqrt{q})$$

- q = 2 : Ising model (BKT transition)
- q ≤ 4 : second order phase transition
- q > 4 : first order phase transition

q=2



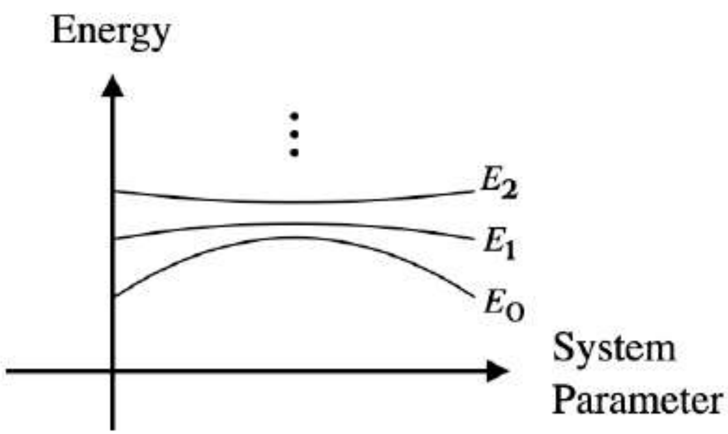
q=6

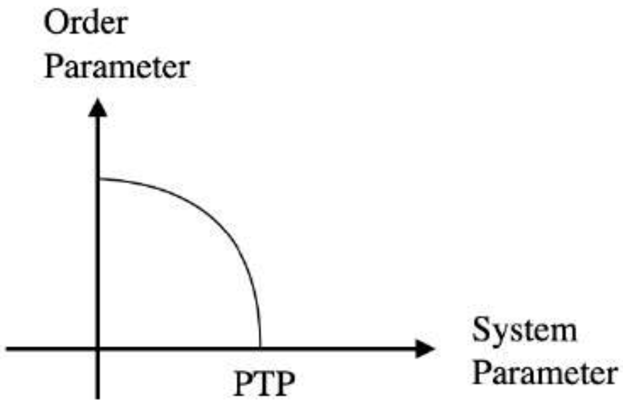


III. Random sampling neural network for quantum many-body problems

(Liu and Wang, Phys. Rev. B 103, 205107 (2021))

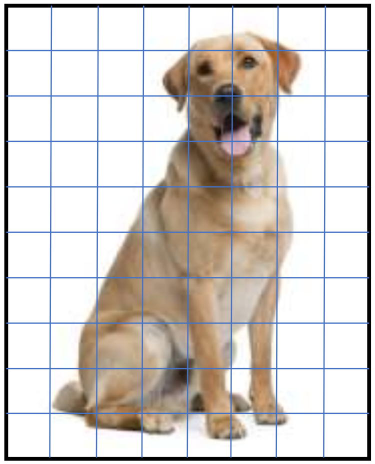
Long-standing Problems in Many-Body Physics: Finite-size Effect

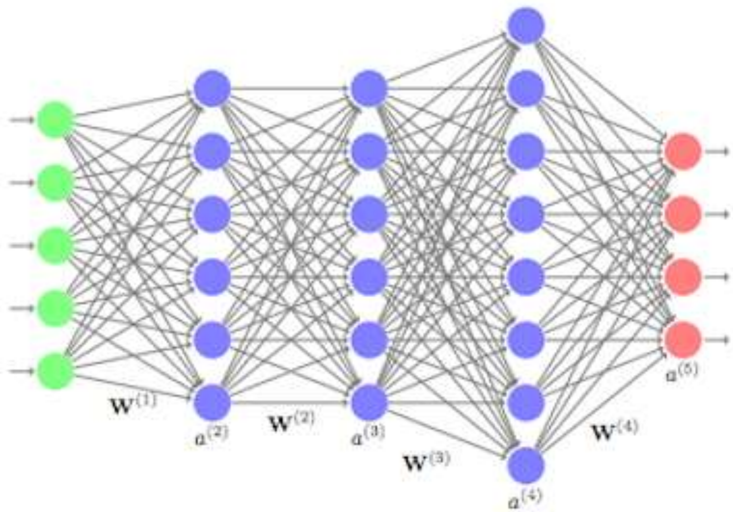




Very few interacting systems are exactly solvable. Numerical diagonalization is time-consuming for a large system size. For example, the Hilbert space of N spin ½ system is $2^N \times 2^N$, which is almost unsolvable even numerically for N >20. Numerical methods such as QMC, DMRG, TN etc. are usually limited by certain dimension or parameter regimes.

But the Hamiltonian is Equivalent to a 2D Image !!





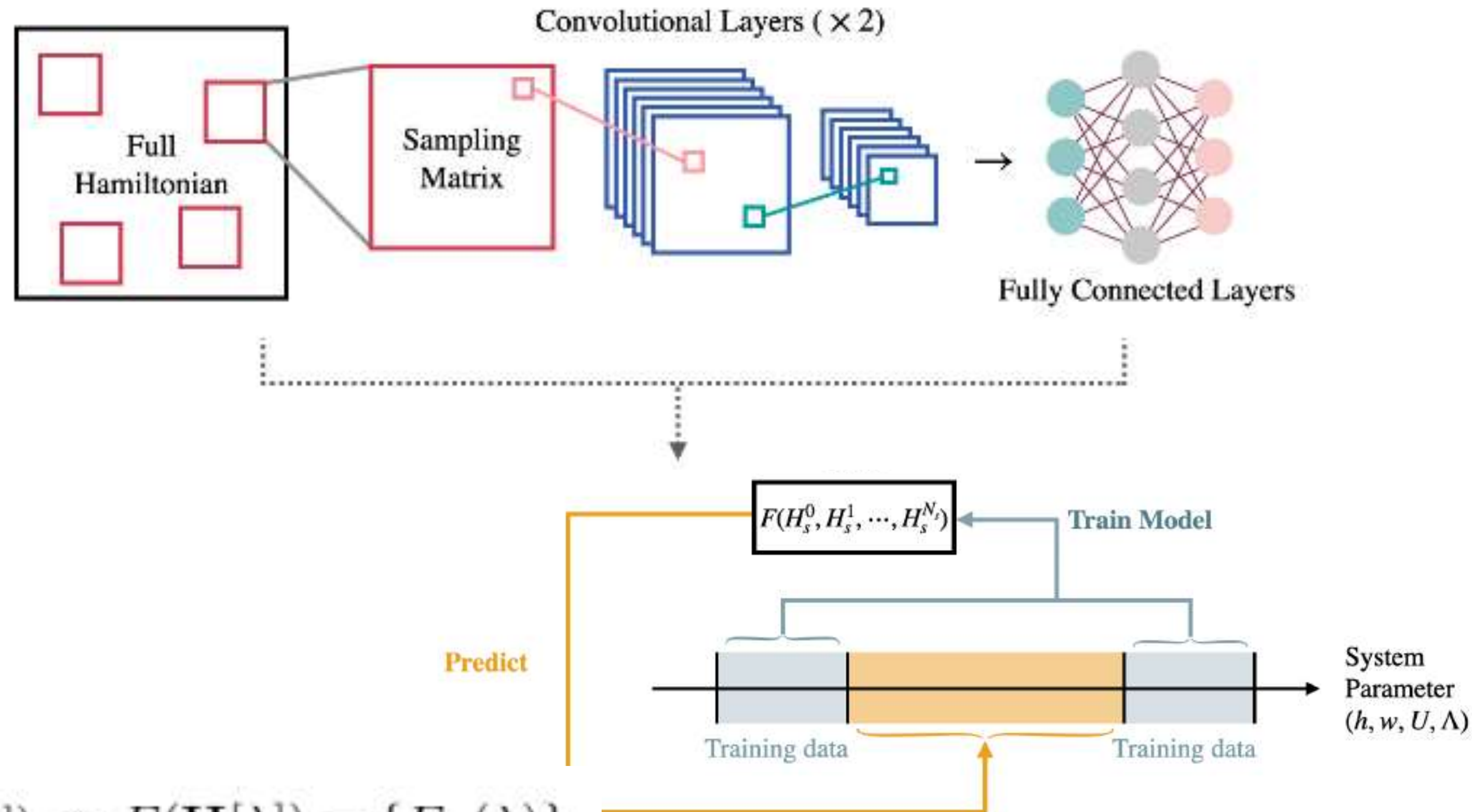
$$|\psi_j\rangle = [\uparrow, \uparrow, \downarrow, \uparrow, \uparrow, \downarrow, \cdots]$$

$$\hat{\mathcal{H}} = \begin{bmatrix} H_{00} & H_{01} & H_{02} & H_{03} & \cdots \\ H_{10} & H_{11} & H_{12} & H_{13} & \cdots \\ H_{20} & H_{21} & H_{22} & H_{23} & \cdots \\ H_{30} & H_{31} & H_{32} & H_{33} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

$$H_{i,j} = \left\langle \psi_i | \hat{\mathcal{H}} | \psi_j \right\rangle$$

For a given basis, the many-body Hamiltonian is nothing but like a large 2D image with two colors (real and imaginary part)…. and therefore maybe learned well by machine learning…

Our Approach: Random Sampling Neural Networks



$$F_{\mathrm{NN}}\left(\mathbf{H}[\lambda]\right) \Rightarrow F(\mathbf{H}[\lambda]) = \{E_n(\lambda)\}$$

$$\left|F_{\mathrm{RSNN}}\left(\{\mathbf{H}_S^{(m)}[\lambda], \mathbf{b}_S^{(m)}[\lambda]\}\right) - \{E_n(\lambda)\}\right| < \epsilon$$

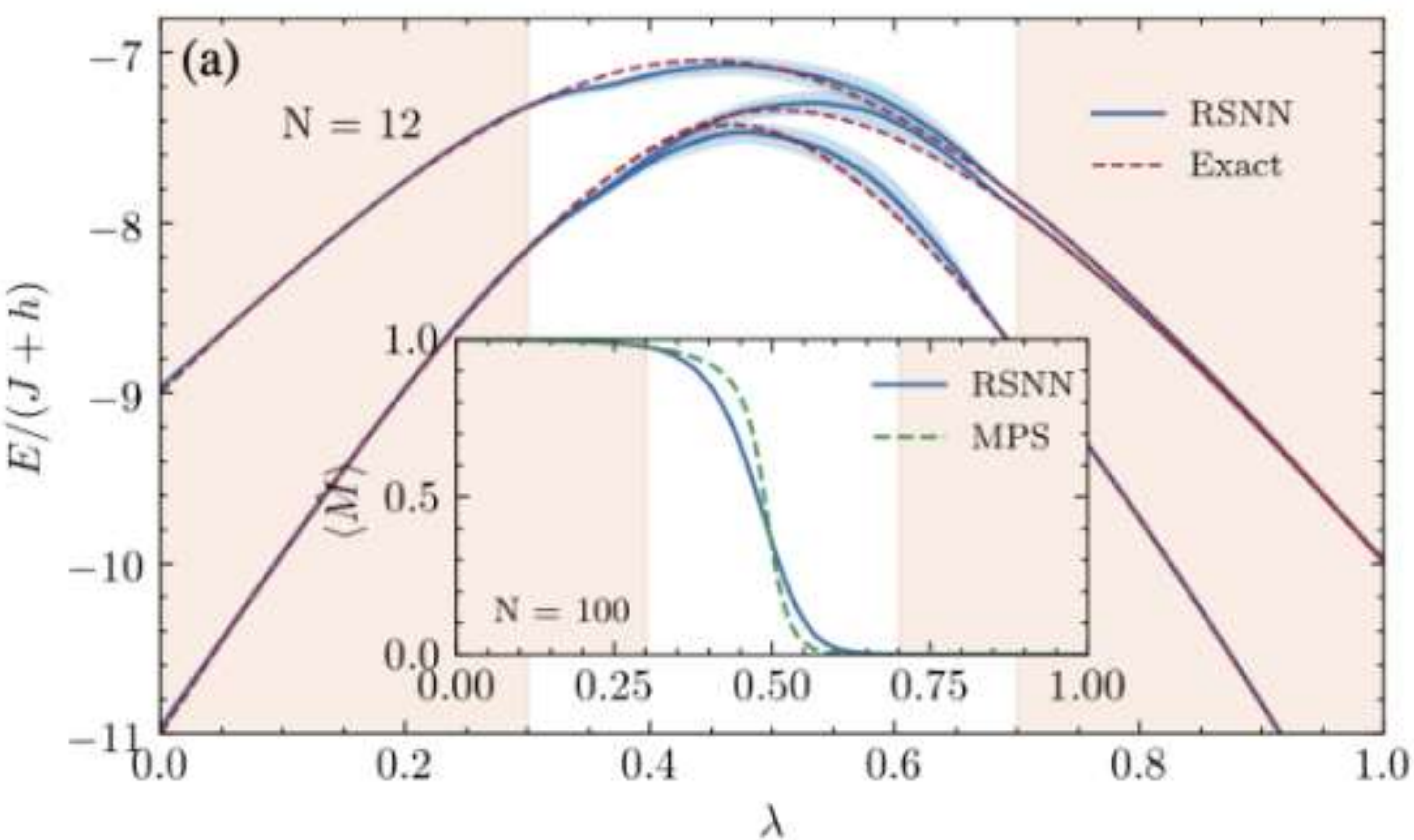
Example: 1D Ising Model with Transverse Field

$$H_{\mathrm{TFIM}} = -J\sum_{\langle i,j\rangle}^{N} \sigma_i^z \sigma_j^z - h\sum_{i}^{N} \sigma_i^x$$

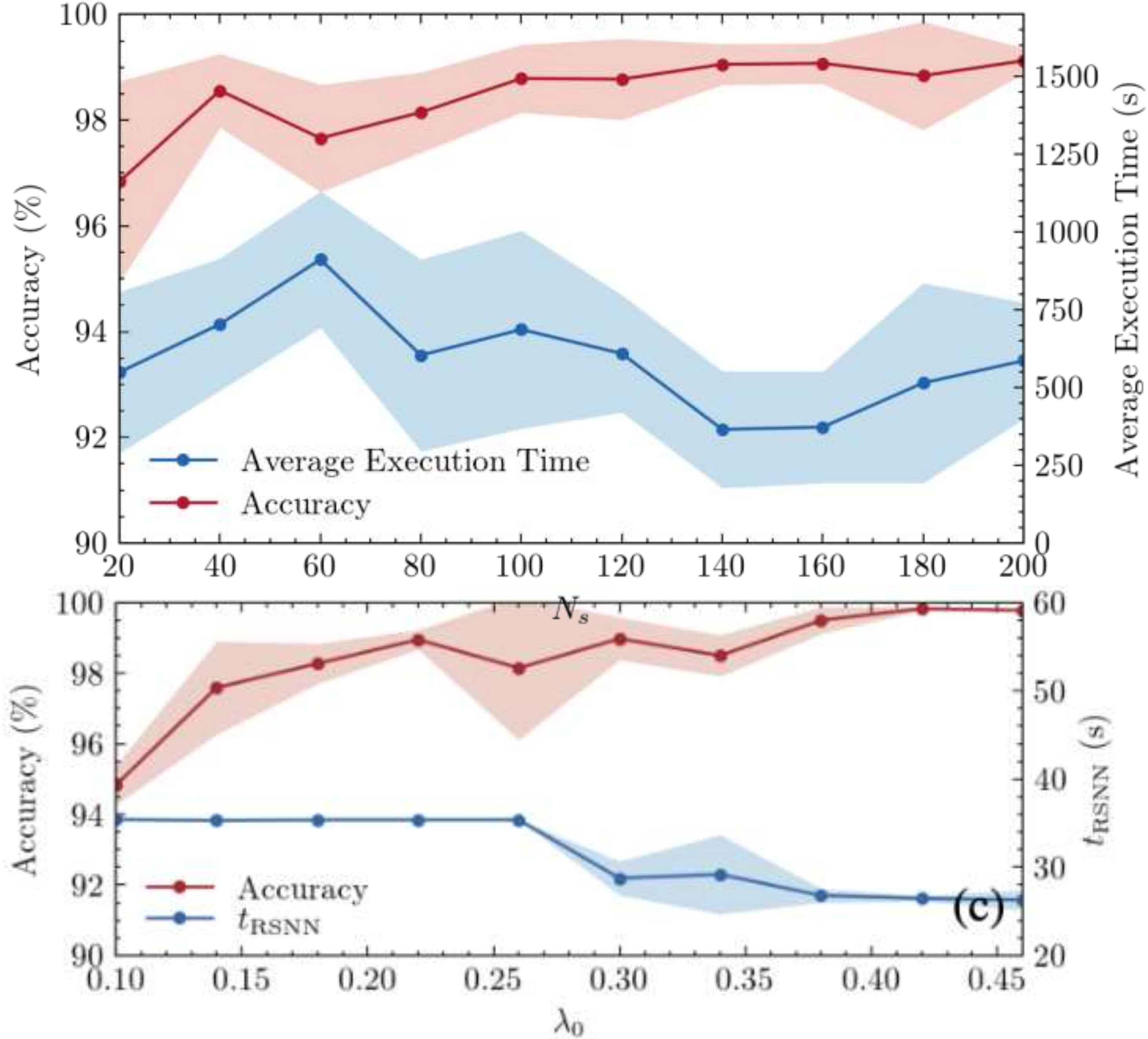
Order phase: $h < J$

Disorder phase: $h > J$

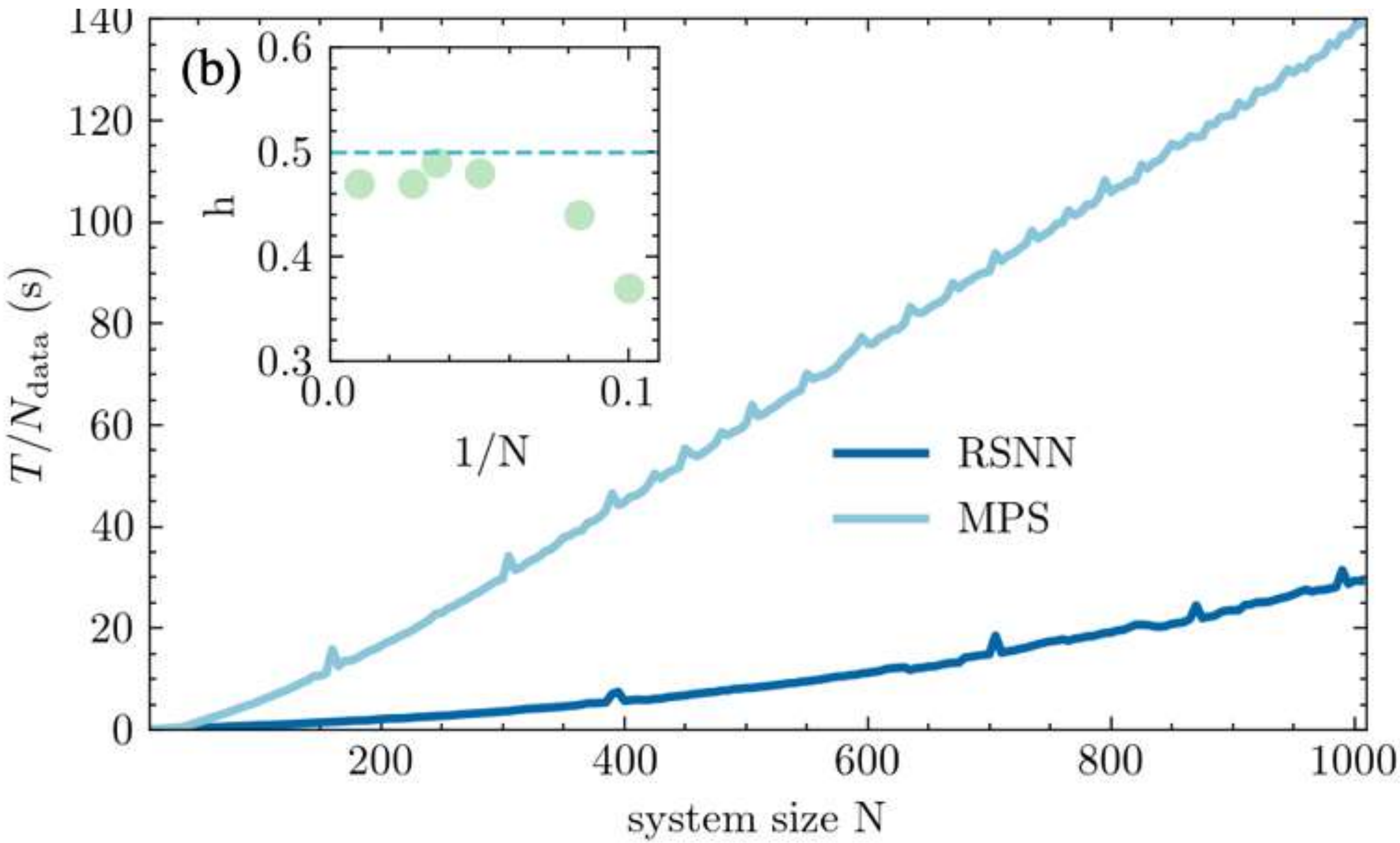
$$\lambda \equiv \frac{h}{h+J}$$



Prediction can be enhanced by using more training data



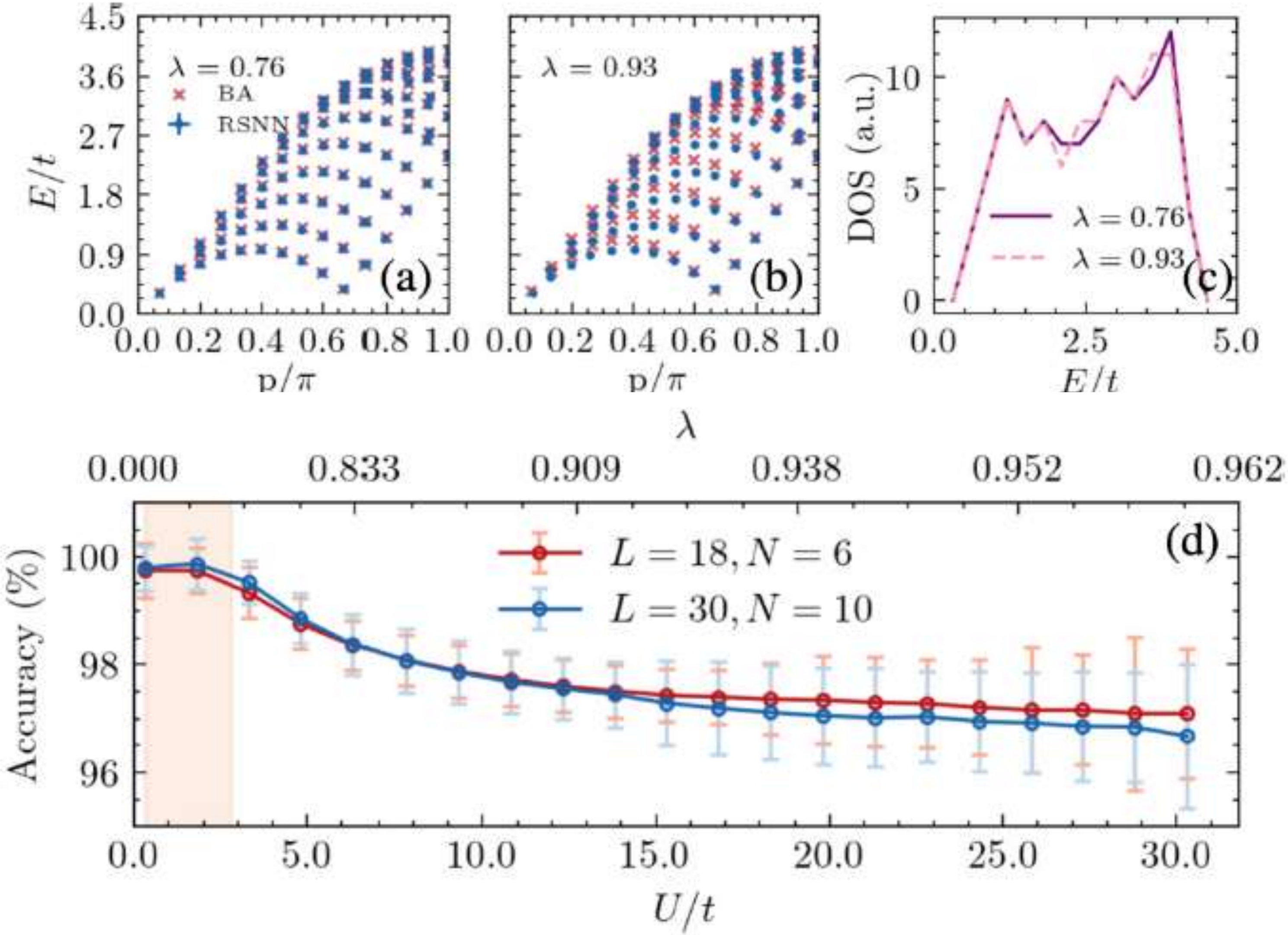
Much More Efficient Method for a Larger System Size



The major time of RSNN calculation is from data training, independent of the number of test data to generate….

Full Spectrum of 1D Fermi-Hubbard Model

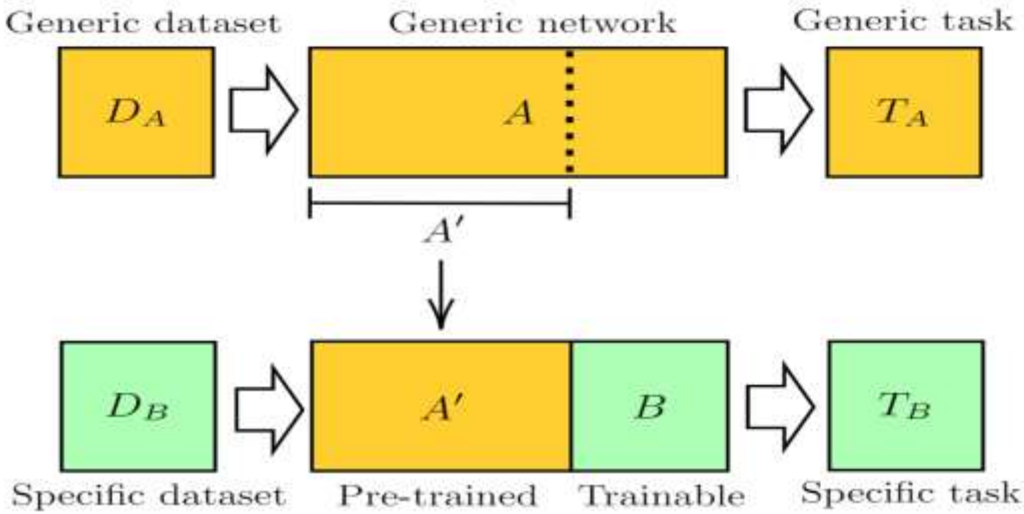
$$\hat{\mathcal{H}}_{\mathrm{FH}} = -t\sum_{i,s}\left(\hat{c}^{\dagger}_{i,s}\hat{c}_{i+1,s} + h.c.\right) + U\sum_{i}\hat{n}_{i,\uparrow}\hat{n}_{i,\downarrow},$$

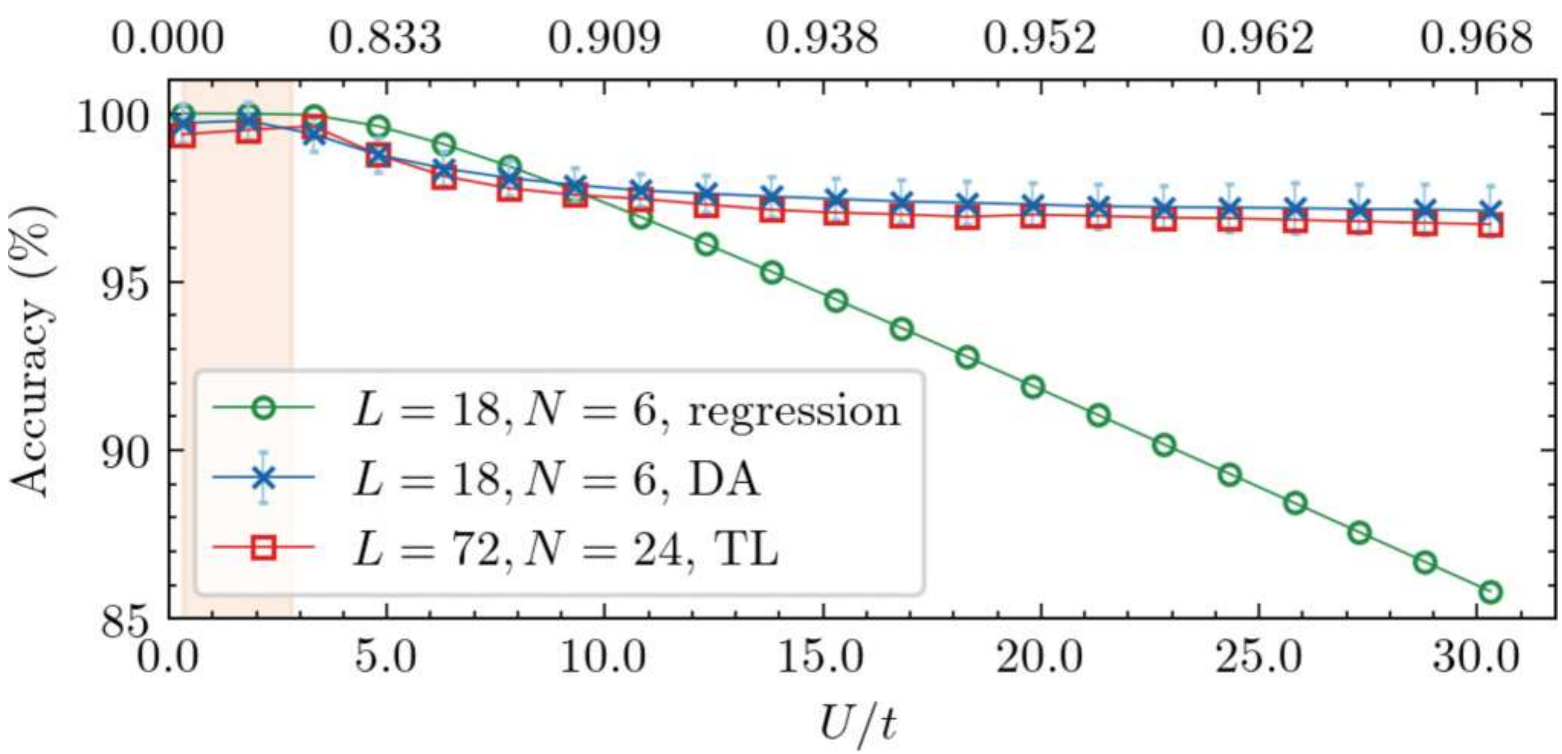


$$\lambda \equiv \frac{U}{U+t}$$

Transfer Learning for a larger system size

In order to calculate results of larger system size with a few known data only, we could pre-train the model and do the transfer learning…



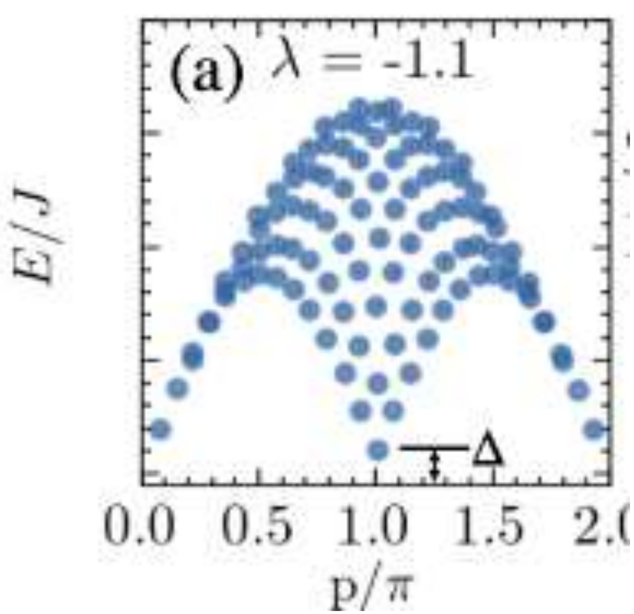


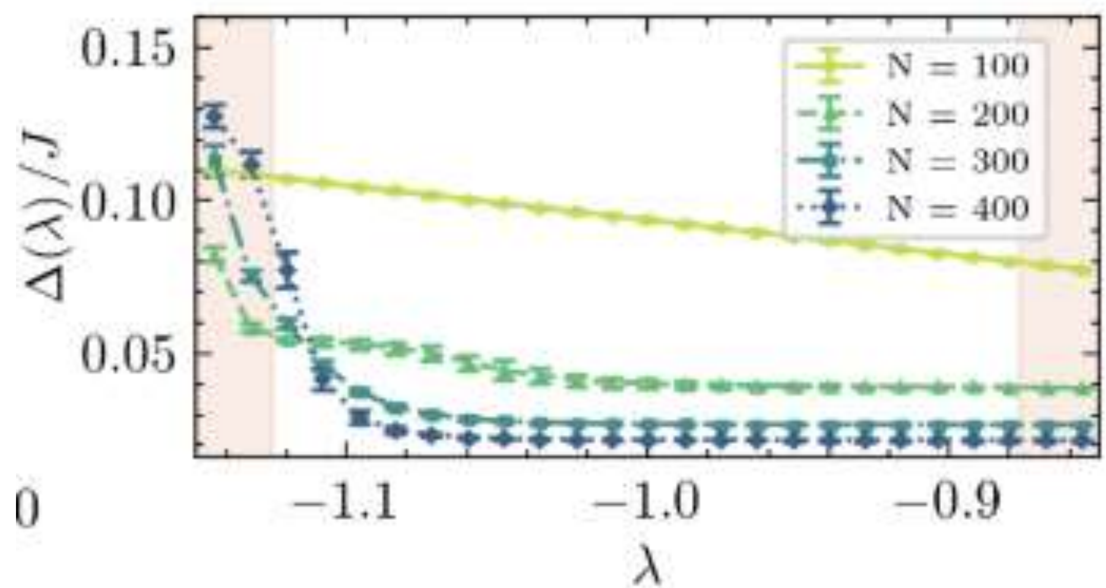
Finite Size Scaling for Excitation Gap of 1D XXZ Model

$$\hat{\mathcal{H}}_{XXZ} = -\frac{J}{2}\sum_{j=1}^{N}(\hat{\sigma}_j^x\hat{\sigma}_{j+1}^x + \hat{\sigma}_j^y\hat{\sigma}_{j+1}^y + \lambda\hat{\sigma}_j^z\hat{\sigma}_{j+1}^z),$$

$\Lambda < -1$	$-1 < \Lambda < 1$	$\Lambda > 1$
anti-ferromagnetism (gapped)	paramagnetism (gapless)	ferromagnetism (gapped)
↑↓↑↓↑↓↑	↑↓↑↑↓↑↑	↑↑↑↑↑↑↑

This model could be also mapped onto 1D hard-core boson *t-V* model (V is the nest nearest neighboring interaction) to describe Solid-Superfluid-Mott Insulator Transition.

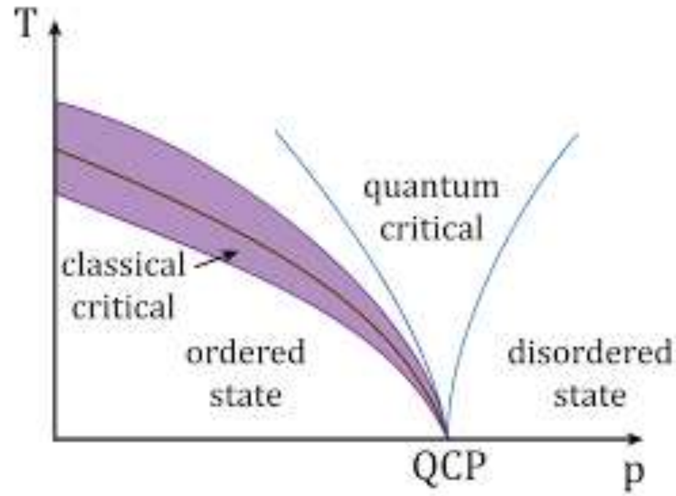


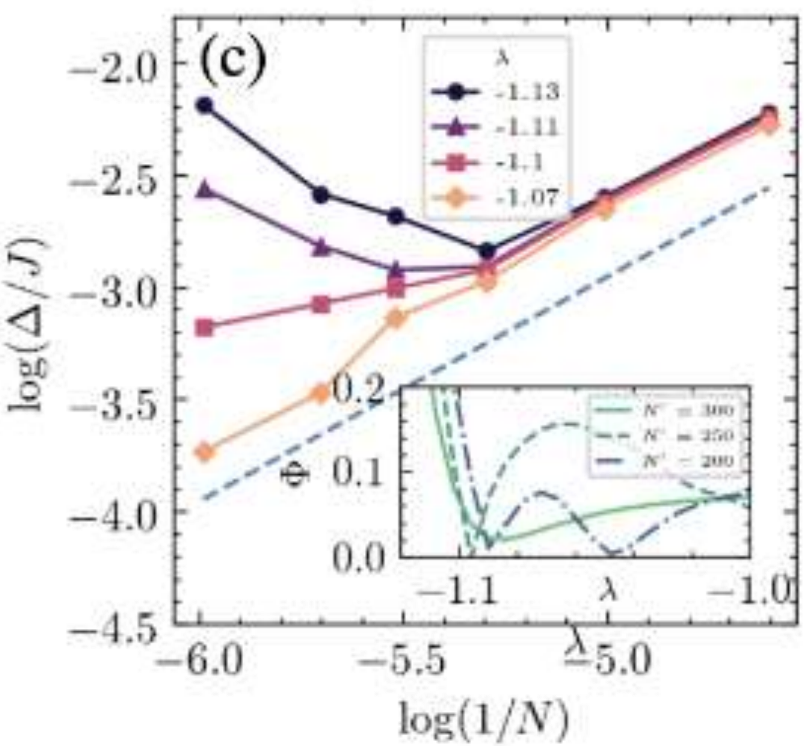


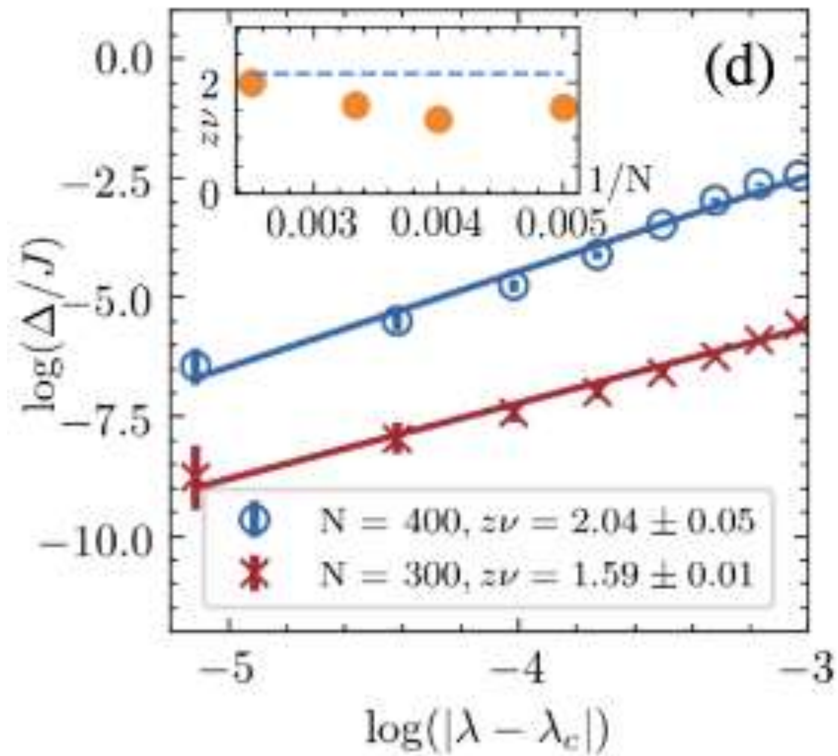
Quantum Critical Exponent

Critical exponent $z\nu$ is defined as:

$$\lim_{\Lambda \to \Lambda_c} \Delta(\Lambda) \propto |\Lambda - \Lambda_c|^{z\nu}$$







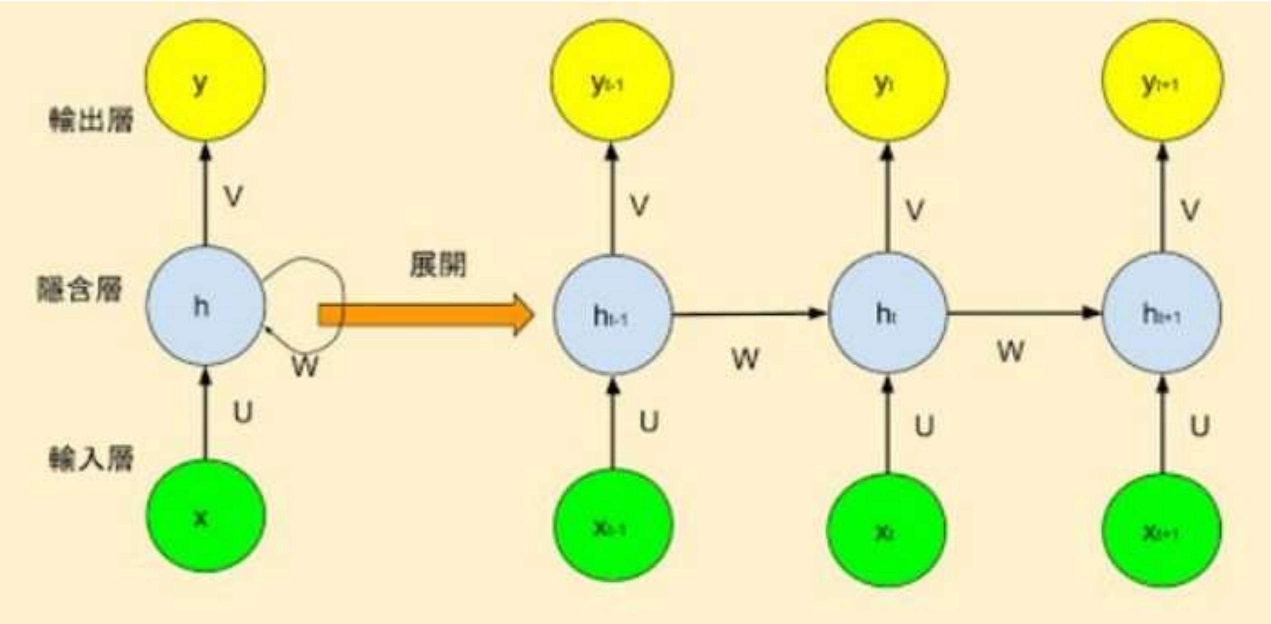
IV. Long Time Dynamics of Many-Boy Systems

(with Guang-Ting Chou, in preparation)

Quantum Dynamics includes High Energy Properties

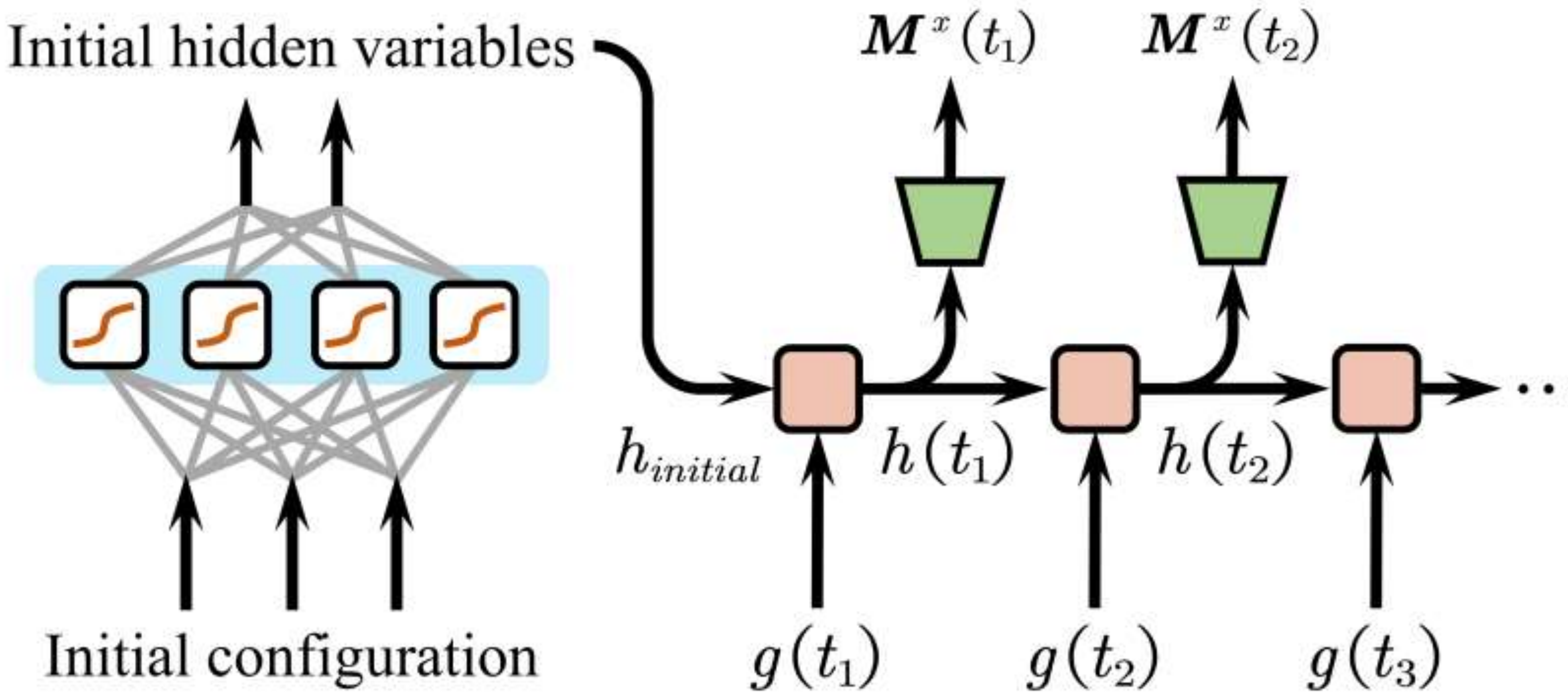
Different from static properties of a many-body system, where we usually want to know ground state or low energy excitation properties only. **Dynamical properties require information from higher energy wavefunctions**, which makes it more difficult.

Our approach: Treat time sequent dynamics as a “sentence” in Natural Language



Implement the Initial Condition

However, for natural language problem, the initial condition is not important (randomly chosen), while it is very crucial for physical problems. Therefore, we have to implement the initial condition (wavefunction or configuration or initial values) through an additional neural network.



Quantum Inspired Recurrent Neural Networks

- Initial Condition Encoder and Loss function

$$|\psi(0)>$$

$$h_{init} = E(C_I)$$

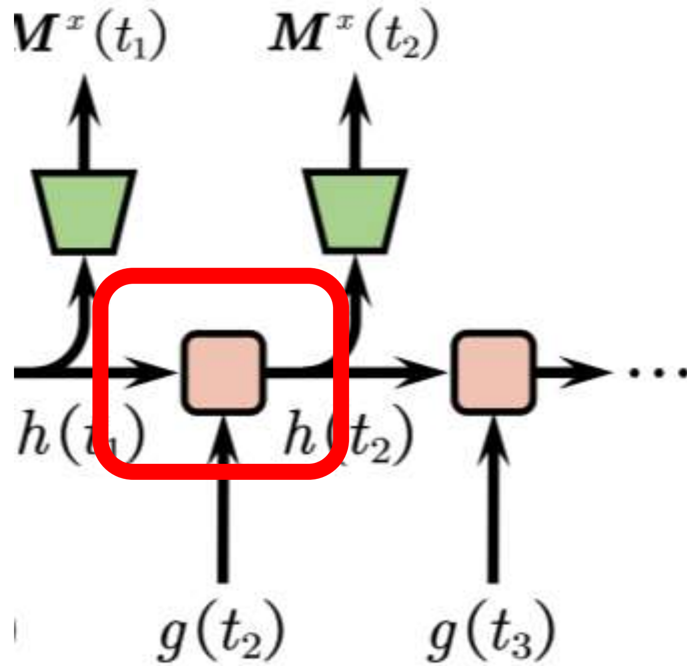
$$L_i = \sum_{t=0}^{T} (O_t{}' - O_t)^2$$

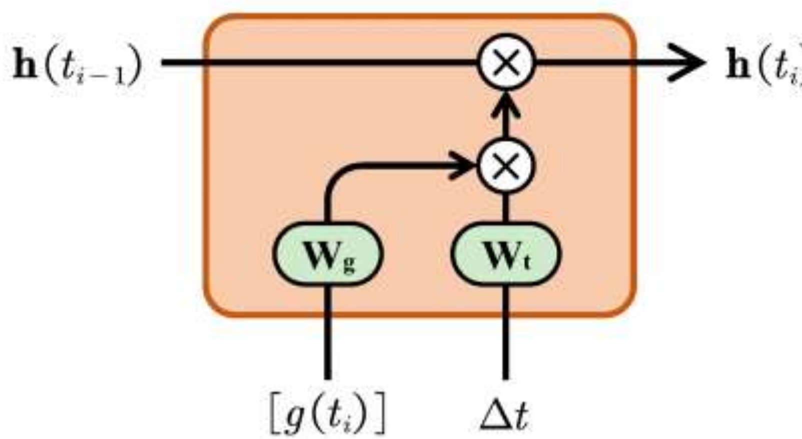
- Time Evolution Operator

$$|\psi(t)> = \sum_n e^{-iE_n t/\hbar}|n><n|\psi_0>$$

$$\phi = f_g(W_g \cdot g(t_i)) f_t(W_t \cdot \Delta t)$$

$$h(t_i) = exp(i\phi) h(t_{i-1})$$

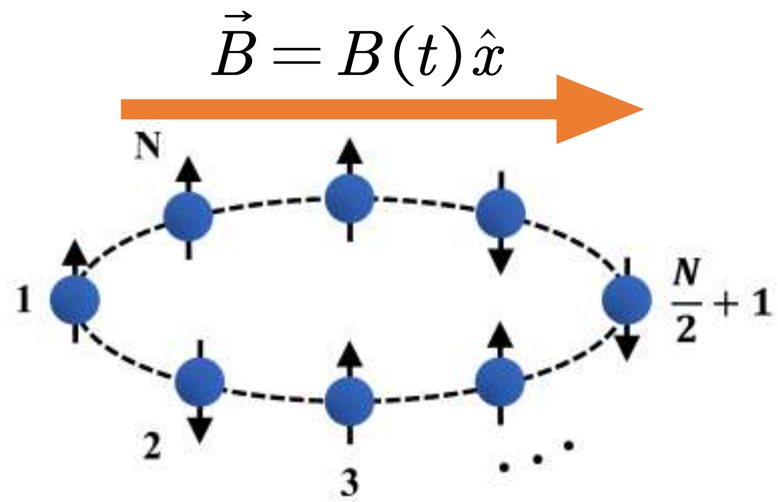




Example: 1D Ising Model with Time-Dependent Field

- The Hamiltonian

$$H = -J\sum_{n=1}^{N}(\sigma_n^z\sigma_{n+1}^z + g(t)\sigma_n^x)$$



- The Jordan–Wigner transformation

$$\sigma_n^x = 1 - 2c_n^\dagger c_n$$

$$\sigma_n^z = -\Pi_{m<n}(1-2c_m^\dagger c_m)(c_n + c_n^\dagger)$$

$$H = -JgN + 2J\sum_k (g-\cos k)c_k^\dagger c_k - \mathrm{i}J\sum_k \sin k(c_{-k}^\dagger c_k^\dagger + c_{-k}c_k)$$

$$c_k(t) = u_k(t)\eta_k + v_k(t)\eta_{-k}^\dagger$$

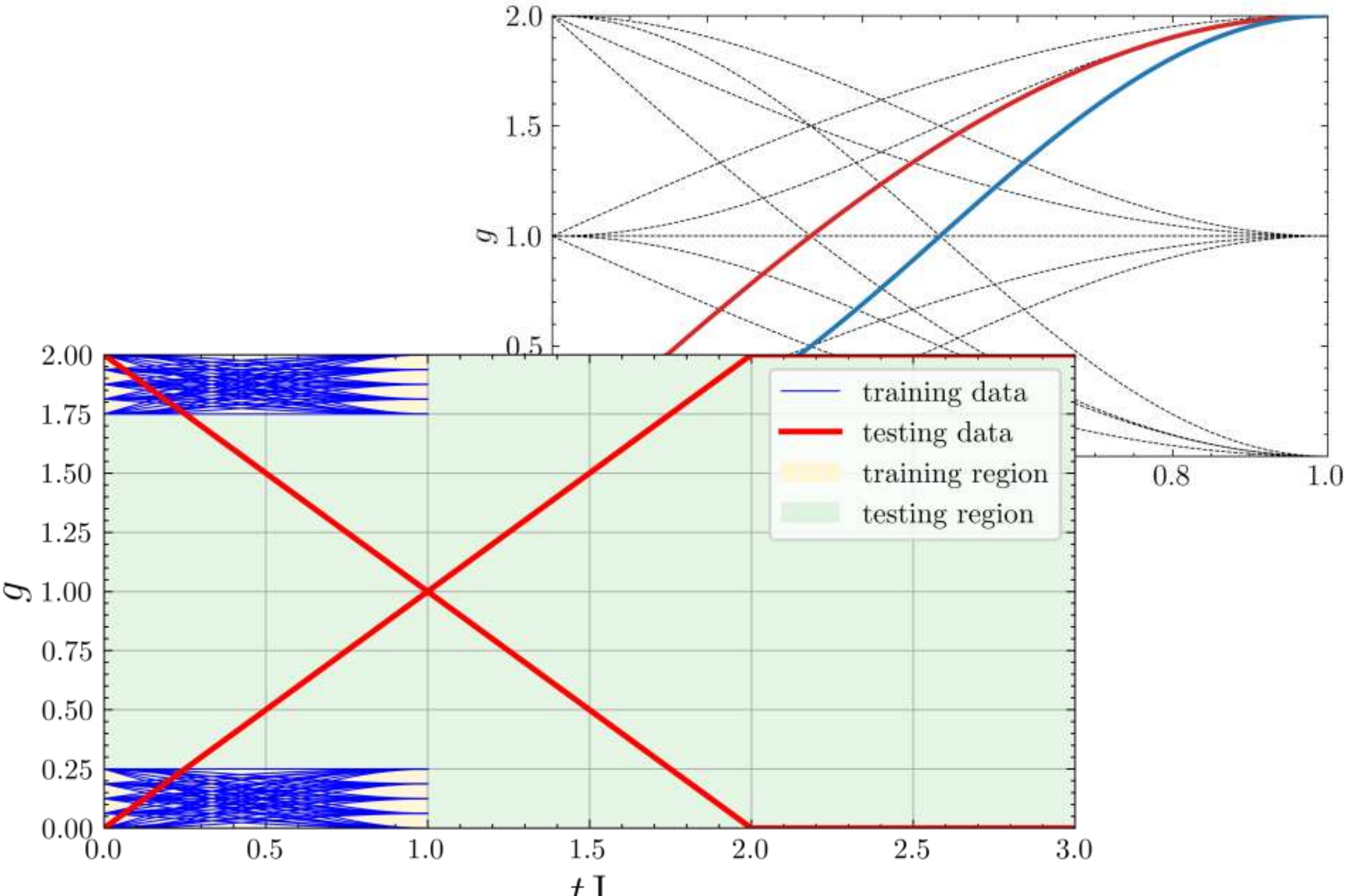
$$A_k(t) = 2J[g(t) - \cos k]$$

$$B_k = 2J\sin k$$

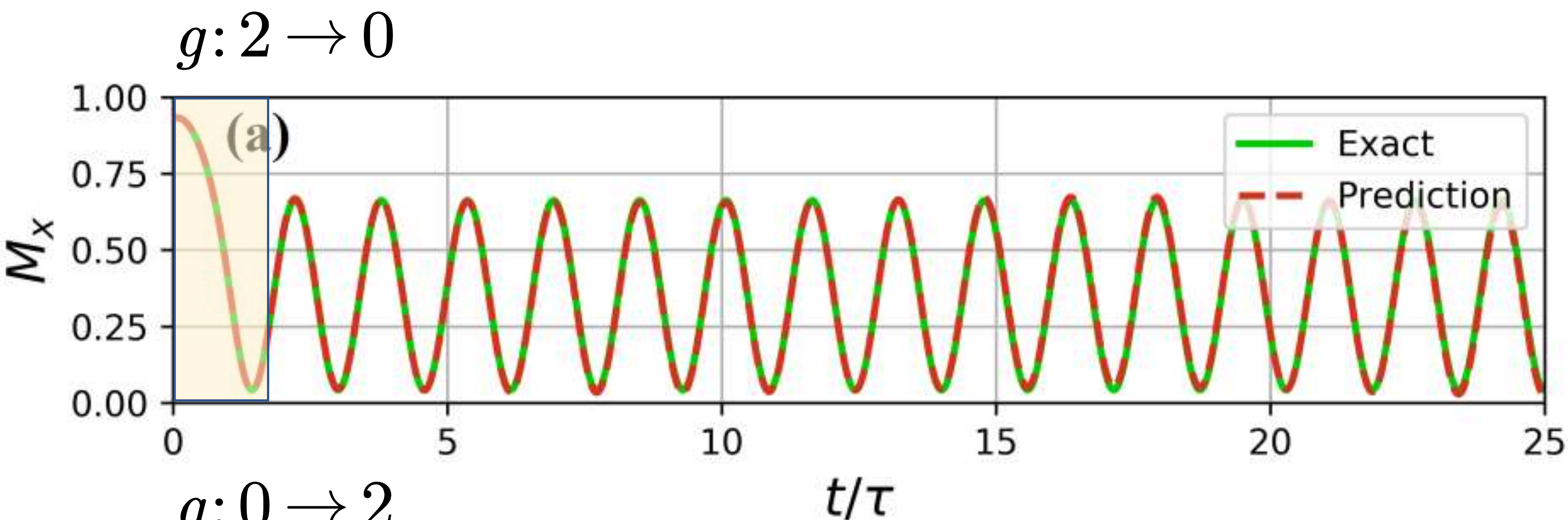
$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} u_k(t) \\ v_{-k}^*(t)\end{pmatrix} = \begin{pmatrix} A_k(t) & B_k \\ B_k & -A_k(t)\end{pmatrix}\begin{pmatrix} u_k(t) \\ v_{-k}^*(t)\end{pmatrix}$$

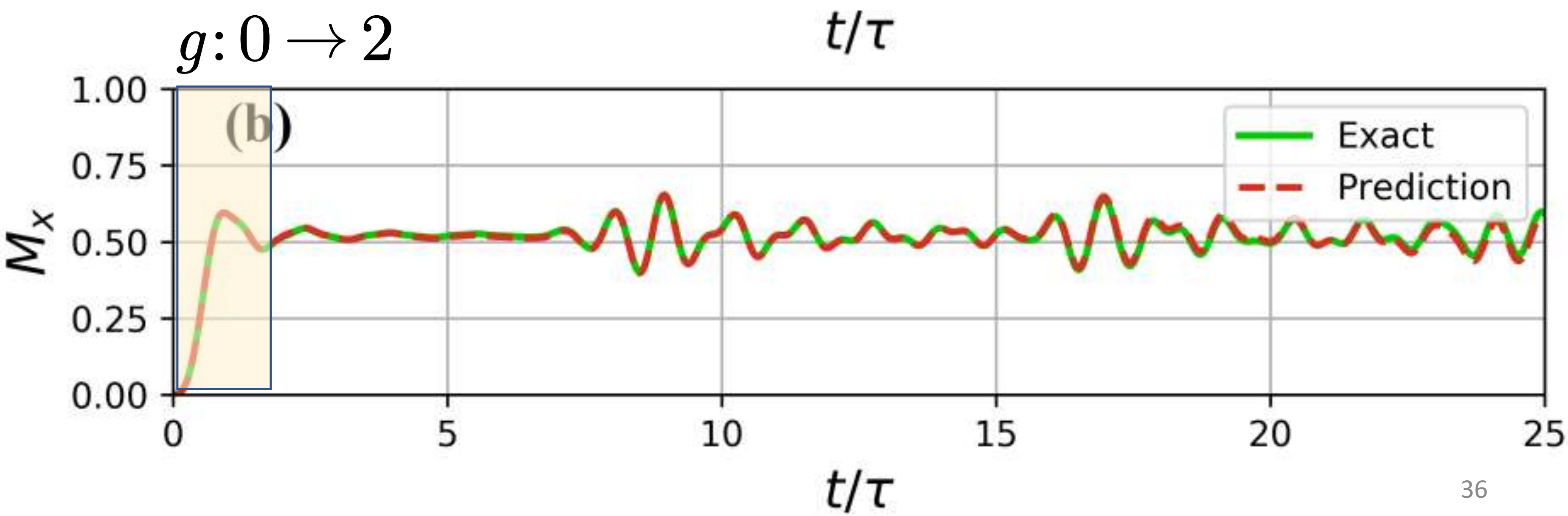
B. Damski, Phys. Rev. Lett. **95**, 035701 (2005).

Training Method in the Short Time and Test Regime

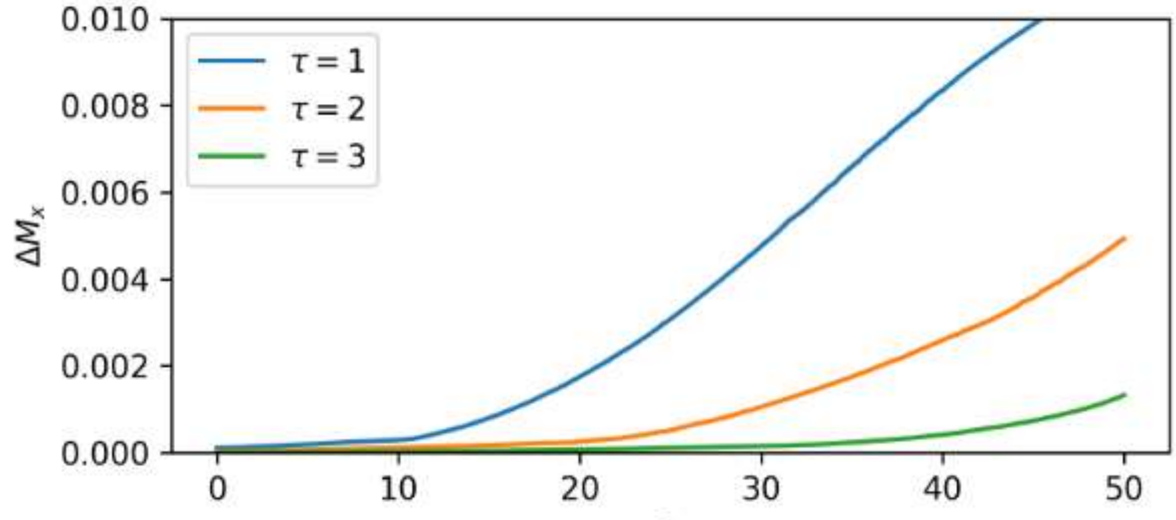


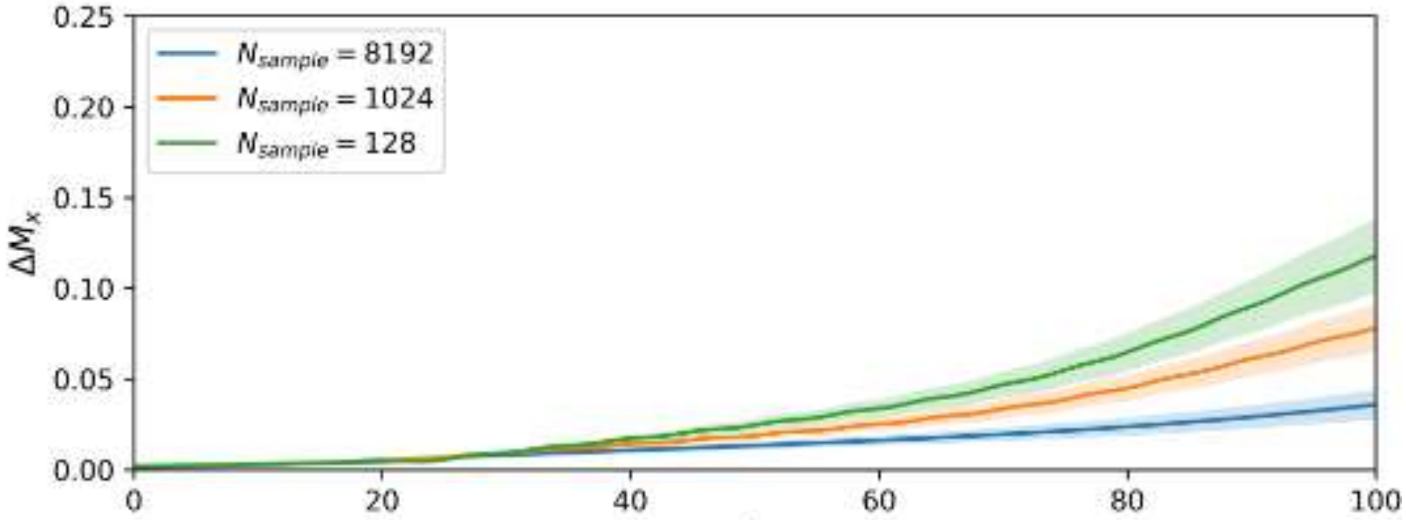
Results: Transverse Magnetization



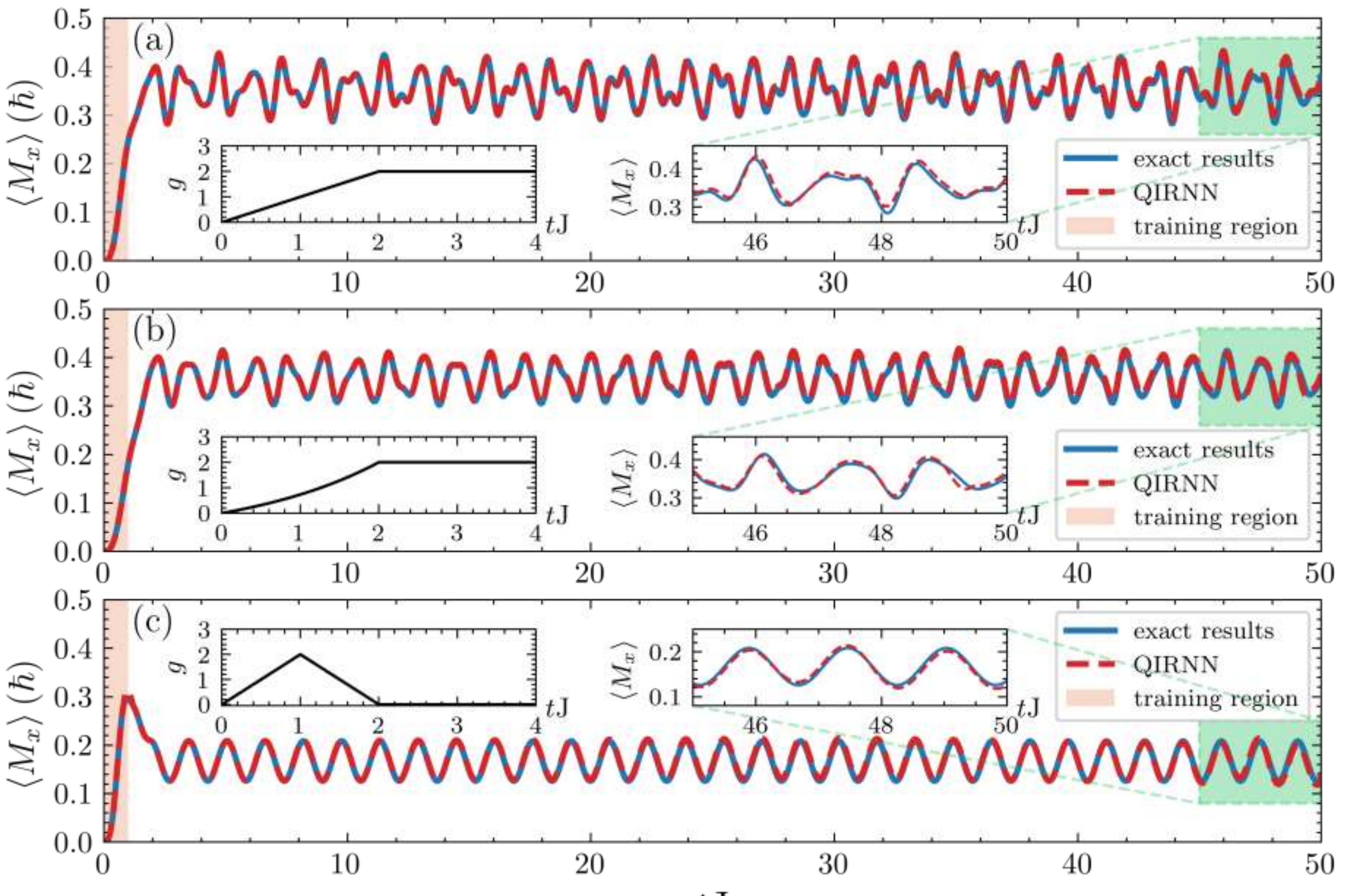


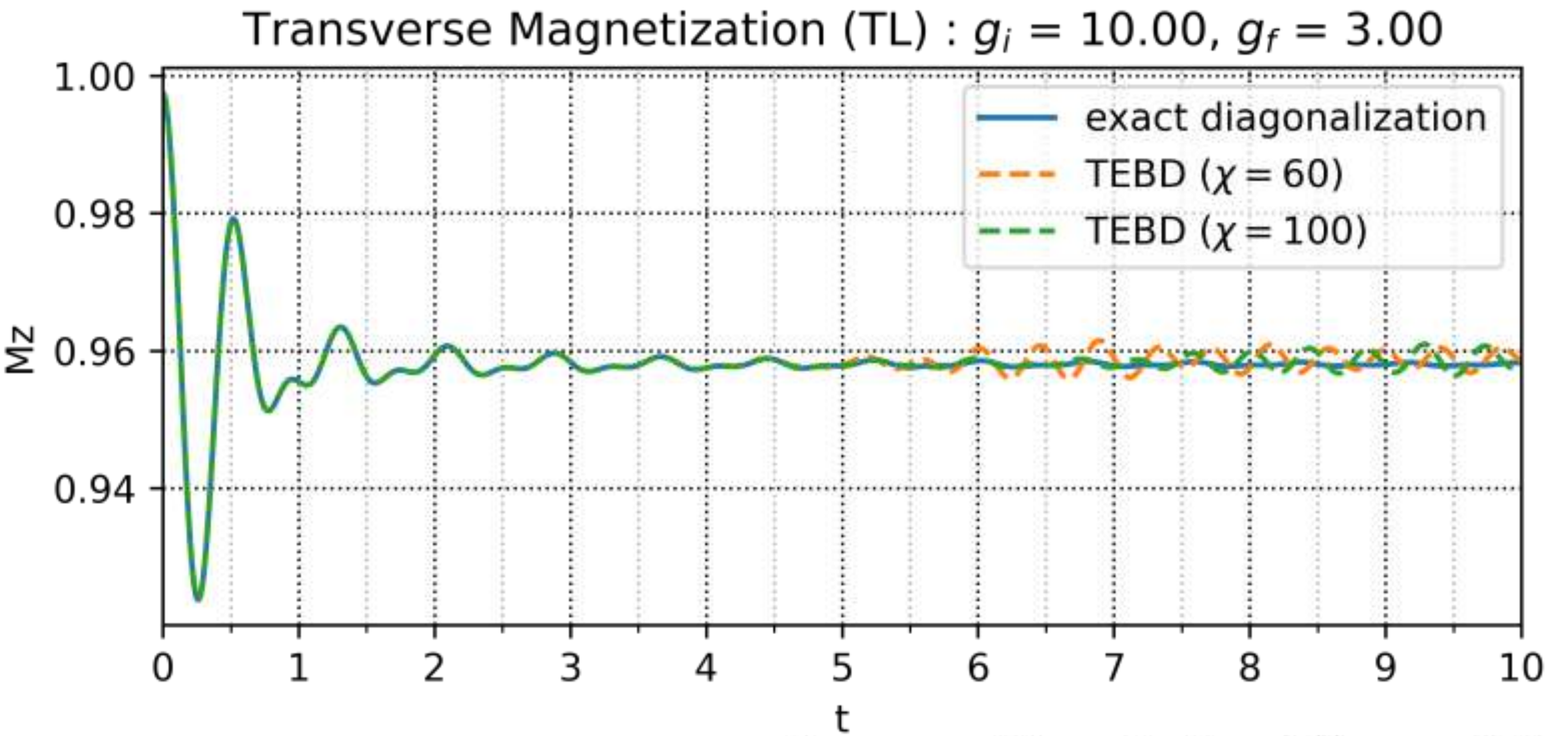
Deviation due to Different Training Parameters

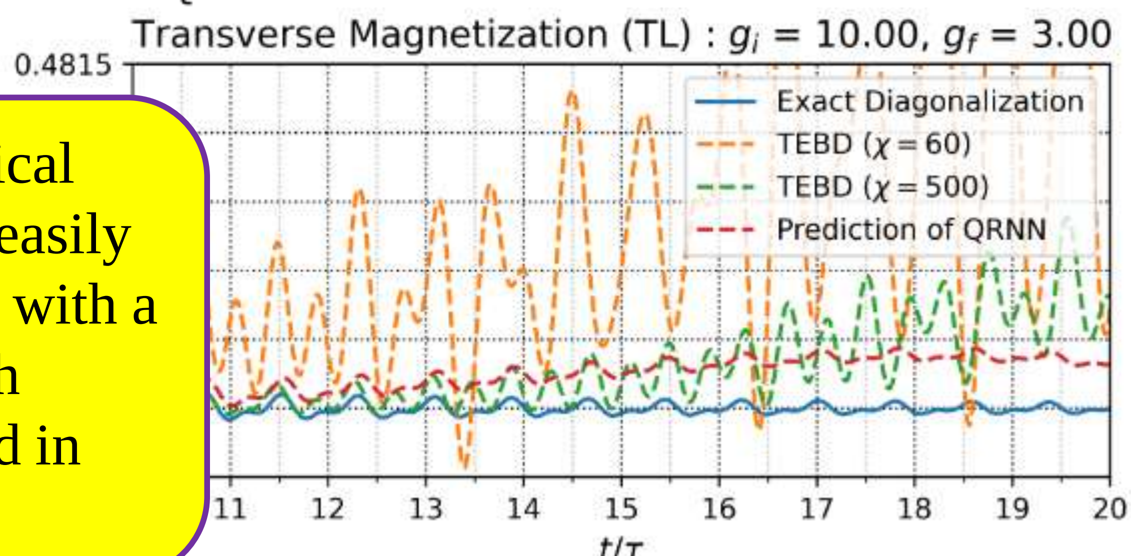




Different Quench Methods







Compare to existing numerical method, our QIRNN could easily predict long-term dynamics with a much shorter time and much better accuracy, even trained in the shorter time regime.

V. Summary and Outlook

We have shown that it is possible to use Machine Learning method to explore **new knowledge for fundamental research**, after properly combined with the unstructured experimental data and partially known theory.