



# Discrete lattice gauge theory via neural-network quantum states

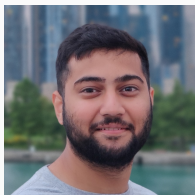
---

Anthony Ashmore

LPTHE, Sorbonne Université

Strings, Fields and Deep Learning – 16th January 2024

# Collaborators



Anuj Apte



Clay Córdova



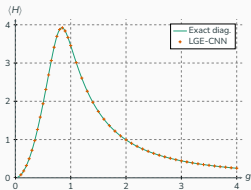
Tzu-Chen (Jimmy)  
Huang

# Overview

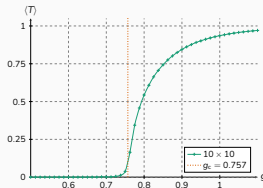
Lattice gauge theory provides a powerful tool for probing strongly coupled systems and critical phenomena

Monte Carlo and neural network methods provide a way to study these theories

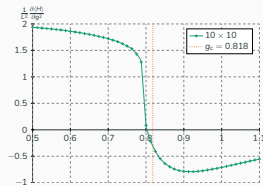
Gauge-invariant neural networks can model the wavefunction of these systems



Ground-state energy



Confinement transition



1st vs 2nd order

Neural-network quantum states

$\mathbb{Z}_N$  lattice gauge theory

Lattice gauge-equivariant CNNs

Application:  $\mathbb{Z}_2$  gauge theory

# Neural-network quantum states

---

# General problem

How do we find the wavefunction for interacting many-body quantum systems? For **gauge theories**, like the Standard Model?

- Wavefunction gives complete description of quantum state  $\Rightarrow$  can compute **expectation values**, etc.

Generic quantum state storage requires exponential amount of information

- Physical, low-energy states often require much less information to encode  $\Rightarrow$  amenable to classical computation

Focus on **ground-state wavefunctions**

# Two problems

Challenges: **storage** of wavefunction and **computing** expectations values  
e.g. system of  $N$  spins, exact wavefunction gives map

$$\Psi: \mathbb{Z}_2^N \rightarrow \mathbb{C},$$

with exponentially large Hilbert space:  $\dim = 2^N$

- Compress  $\Psi \Rightarrow$  **variational ansatz**

Exponentially large number of operations to compute expectation values

$$\langle H \rangle = \sum_{\sigma, \sigma'} \Psi(\sigma)^* \langle \sigma | H | \sigma' \rangle \Psi(\sigma')$$

- Reduce # operations by **Monte Carlo sampling**

# Neural network quantum states

Variational Monte Carlo has long history...

[Carleo, Troyer '16] introduced a new kind of variational ansatz: **neural network quantum state** (NNQS)

- Wavefunction  $\Psi$  represented by neural network with hidden layers
- Parameters (weights) of ansatz optimised by variational Monte Carlo + stochastic gradient descent to minimise  $\langle H \rangle$

State-of-the-art results for ground-state approximation and time evolution of quantum systems



# Some applications

## Ground-state approximation

- Ising model [Carleo, Troyer '17;...],
- $J_1$ - $J_2$  model [Nomura '20; Chen, Heyl '23;...]
- Continuous models [Pescia et al. '22; Lovato et al. '22; Zhao et al. '22,...]
- Transformers [Viteritti et al. '23]

Unitary evolution [Carleo et al. '17; Yaun et al. 19,...], excited states, state reconstruction, finite temperature, open system, and much more!

Combine **gauge-invariant** NNQS with variational Monte Carlo to model lattice gauge theories

Focus on **discrete**  $\mathbb{Z}_N$  gauge theories: proof-of-concept *and* (relatively) accurate predictions

# $\mathbb{Z}_N$ lattice gauge theory

---

# $\mathbb{Z}_N$ lattice gauge theory in 2 + 1d

Hamiltonian approach: space discretised, time continuous

- Spatial  $L \times L$  lattice with periodic boundary conditions
- Gauge field degrees of freedom live on links between lattice sites

Clock  $Q_\ell$  and shift  $P_\ell$  operators on link  $\ell$  satisfy a  $\mathbb{Z}_N$  algebra

$$\begin{aligned} P_\ell^\dagger P_\ell = Q_\ell^\dagger Q_\ell &= \mathbf{1}, & P_\ell^N = Q_\ell^N &= \mathbf{1}, \\ P_\ell^\dagger Q_\ell P_\ell &= e^{2\pi i/N} Q_\ell. \end{aligned}$$

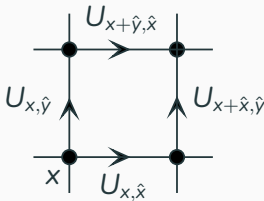
# $\mathbb{Z}_N$ lattice gauge theory in $2 + 1$ d

States  $|q\rangle_\ell$  span Hilbert space on link  $\ell$  are eigenstates of clock operator:

$$Q_\ell |q\rangle_\ell = e^{2\pi i q_\ell / N} |q\rangle_\ell, \quad q_\ell \in \{0, \dots, N-1\}.$$

Shift operator is periodic lowering operator:  $P_\ell |q\rangle_\ell = |q-1\rangle_\ell$

A **configuration** of the  $2L^2$  links is given by a choice of phases  $\mathcal{U} = \{U_{x,\mu}\} \in \mathbb{C}^{2L^2}$ , where  $x$  denotes a lattice site and  $\mu \in \{\hat{x}, \hat{y}\}$  specifies the link



## $\mathbb{Z}_N$ lattice gauge theory in 2 + 1d

Under a gauge transformation  $\Omega$ , the link variables transform **non-locally**

$$U_{x,\mu} \mapsto T_\Omega U_{x,\mu} = \Omega_x U_{x,\mu} \Omega_{x+\mu}^\dagger.$$

Instead, the locally transforming data is encoded by (untraced) **Wilson loops**

$$U_{x,\mu\nu} \equiv U_{x,\mu} U_{x+\mu,\nu} U_{x+\nu,\mu}^\dagger U_{x,\nu}^\dagger.$$

Under gauge transformation

$$U_{x,\mu\nu} \mapsto T_\Omega U_{x,\mu\nu} = \Omega_x U_{x,\mu\nu} \Omega_x^\dagger.$$

# $\mathbb{Z}_N$ lattice gauge theory in 2 + 1d

Hamiltonian of pure  $\mathbb{Z}_N$  gauge theory

$$H = \frac{g^2}{2} \sum_{\ell} [1 - P_{\ell}] + \frac{1}{2g^2} \sum_{\ell_i \in \square} [1 - Q_{\ell_1}^{\dagger} Q_{\ell_2}^{\dagger} Q_{\ell_3} Q_{\ell_4}] + \text{h.c.}$$

- Coupling  $g$
- Electric term – sum over **links**  $\ell$
- Magnetic term – sum over  $1 \times 1$  **plaquettes**  $\square$

**Gauss' law** encoded by vertex operators  $\Theta_x$ :

$$\Theta_x = P_{x,\hat{x}} P_{x,\hat{y}} P_{x-\hat{x},\hat{x}}^{\dagger} P_{x-\hat{y},\hat{y}}^{\dagger}.$$

Local gauge invariance from  $[\Theta_x, H] = 0$

# $\mathbb{Z}_N$ lattice gauge theory in 2 + 1d

At fixed time, a **wavefunction**  $\Psi$  maps a configuration of phases  $\mathcal{U}$  to a complex number

$$\begin{aligned}\Psi: \mathbb{C}^{2L^2} &\rightarrow \mathbb{C}, \\ \mathcal{U} &\mapsto \Psi(\mathcal{U}).\end{aligned}$$

We seek the **ground-state wavefunction**  $\Psi_0$

- $H$  is time independent  $\Rightarrow \Psi_0$  is time independent
- Restrict to  $\Theta_x = 1$  sector  $\Rightarrow$  pure gauge theory
- Ground state is **gauge invariant**  $\Rightarrow \Psi_0(T_\Omega \mathcal{U}) = \Psi_0(\mathcal{U})$



# Lattice gauge-equivariant CNNs

---

Lattice gauge-equivariant convolutional neural networks (LGE-CNNs) introduced by [Favoni et al. '20]

- Built from gauge-equivariant convolutions and multiplications
- Can approximate arbitrary gauge invariant (or equivariant) functions of a lattice system

Originally used for supervised learning (traced Wilson loops and topological charge density)

Alternative gauge-equivariant networks constructed by [Luo et al. '22] for  $\mathbb{Z}_N$  and [Luo et al. '22] for  $U(1)$

# Network structure

At each layer, keep track of link-like variables  $\mathcal{U}$  and loop-like variables  $\mathcal{W}$

$$\mathcal{U} = \{U_{x,\mu}\}, \quad \mathcal{W} = \{W_{x,i}\},$$

where  $i$  is a **channel index** – can associate multiple  $\mathcal{W}$  elements to the same lattice site  $x$

- e.g.  **$1 \times 1$  plaquette** variables  $U_{x,\mu\nu} = U_{x,\mu} U_{x+\mu,\nu} U_{x+\nu,\mu}^\dagger U_{x,\nu}^\dagger$
- LGE-CNN constructs more general objects that transform like these

Each layer of an LGE-CNN acts on the pair  $(\mathcal{U}, \mathcal{W})$

- Input to network is the set of link variables  $\mathcal{U}$ , while  $\mathcal{W} = \emptyset$  initially

# Plaquette layer

First layer generates  $1 \times 1$  plaquettes from links

$$\text{Plaq: } U_{x,\mu} \mapsto U_{x,\mu\nu}$$

and stores them in  $\mathcal{W}$

- Keep those with positive orientation

# Equivariant convolution bilinear layer

Starting from  $\mathcal{W} = \{W_{i,x}\}$ , construct new objects by **parallel translating**

$$W'_{x,i} = U_{x,K\mu} W_{x+K\mu,i} U_{x,K\mu}^\dagger,$$

by  $K$  steps in  $\mu$  direction, and add these to  $\mathcal{W}'$

$$\mathcal{W}' = \{W_{i,x}, U_{x,\hat{x}} W_{x+\hat{x},i} U_{x,\hat{x}}^\dagger, \dots\}$$

where

- $\mu \in \{\hat{x}, \hat{y}\}$  runs over lattice directions
- $K \in \{0, \dots, \text{kernel size}\}$  determines the maximum distance to translate  $\mathcal{W}$  elements

# Equivariant convolution bilinear layer

Extend both  $\mathcal{W}$  and  $\mathcal{W}'$  by hermitian conjugate and identity elements, e.g.

$$\mathcal{W} \mapsto \{W_{i,x}, W_{i,x}^\dagger, \mathbf{1}_x\},$$

then multiply all local terms

$$(\mathcal{W}, \mathcal{W}') \mapsto W''_{x,i} = \sum_{jk} \alpha_{ijk} W_{x,j} W'_{x,k},$$

where

- $i \in \{1, \dots, N_{\text{out}}\}$  runs over number of **output channels**
- $\alpha_{ijk} \in \mathbb{C}$  are **trainable weights**

# Equivariant convolution bilinear layer

Combine these steps into a single layer

$$\mathbf{L-CB}: (\mathcal{U}, \mathcal{W}) \mapsto \{W''_{x,i}\}$$

where the layer is specified by a choice of **output channels** and **kernel size**  $(N_{\text{out}}, K)$

- Allows for a **bias** and acts as a **residual** layer

Stacking **L-CB** layers constructs arbitrary untraced Wilson loops [Favoni et al. '20]

- **Pla**q gives all  $1 \times 1$  Wilson loops
- Single **L-CB** output has linear combinations of  $1 \times 2$  and  $2 \times 1$  loops, plus original  $1 \times 1$  plaquettes and their squares

# Network structure

Output of network taken to be  $\log \Psi$ :

$$\log \Psi \equiv \text{Dense} \circ \text{Global Pooling} \circ \text{Trace} \circ \dots \circ \text{L-CB} \circ \text{Plaq}$$

- Complex weights: output can encode **non-trivial phase structure**
- **Translation invariant**

Implemented using **NetKet 3** [Vicentini et al. '21] and JAX/Flax

- Training minimises  $\langle H \rangle$  with respect to network weights via stochastic gradient descent
- Hilbert space too large to sum over exactly, compute expectation values etc. using sample of field configurations, chosen by **local Metropolis algorithm**



## Some practical comments

NetKet includes preconditioning of the gradient

- Use **stochastic reconfiguration** in all cases [Sorella et al. '07]

Use **transfer learning** when scanning over coupling

- Work from large to small coupling

Bilinear structure can easily lead to vanishing / exploding gradients – important to **initialise** network properly for **deep networks**

- No analytic results (e.g. Xavier initialisation) – find empirically, tune weights **layer-by-layer** to fix variance (c.f. LSUV [Mishkin, Matas '15])

Networks with 2 to 10 **L-CB** layers (500 to 15000 params), trained using single V100

## Application: $\mathbb{Z}_2$ gauge theory

---

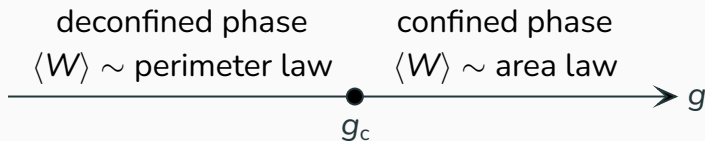
## $\mathbb{Z}_2$ gauge theory

For  $\mathbb{Z}_2$ , clock/shift are simply **Pauli operators**:  $Q \equiv Z$  and  $P \equiv X$

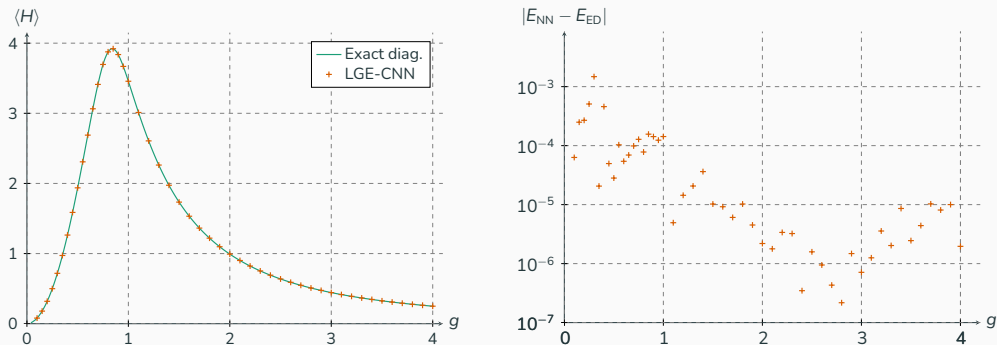
By duality, equivalent to classical Ising model in 3d

- In universality class of Ising CFT3 [Wilson, Fisher '72]

Phase transition for spontaneous breaking of  $\mathbb{Z}_2^{(1)}$  emergent **one-form symmetry**

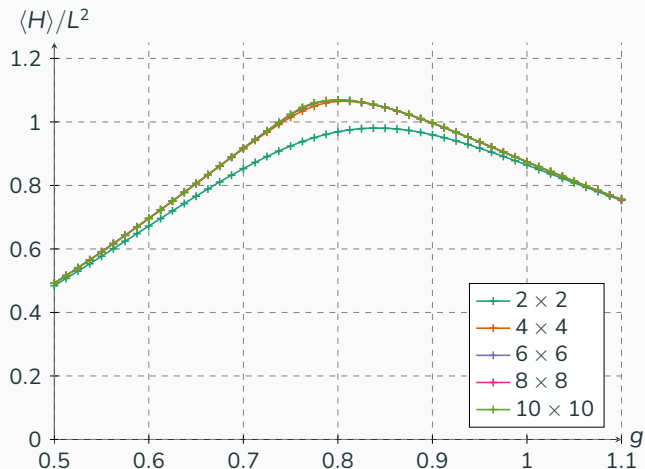


# Ground-state energy



**Figure 1:** Energy  $\langle H \rangle$  for on a  $2 \times 2$  lattice as a function of coupling  $g$ . Networks trained for up to 500 iterations with 4096 MC samples per step. Training stopped once the variance,  $\text{Var } H = \langle H^2 \rangle - \langle H \rangle^2$ , stabilised to a value of 0.0001 or less.

# Ground-state energy



**Figure 2:** Expectation value  $\langle H \rangle$  for  $\mathbb{Z}_2$  gauge theory for varying lattice sizes as a function of coupling  $g$ , focused on the region around the critical coupling.

# Locating the phase transition

Most accurate identification comes from **dual spin system** [Blöte, Deng '02]

- $\mathbb{Z}_2$  gauge theory  $\leftrightarrow$  TFIM  $\leftrightarrow$  limit of classical Ising model in 3d

$$g_c = 0.757051$$

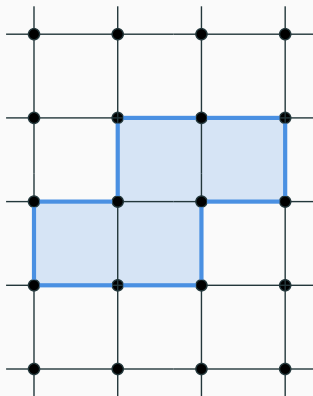
Previous work on **gauge-equivariant NNs** [Luo et al. '20]

- String tension on  $12 \times 12$  lattice

$$g_c \approx 0.7$$

We search for the phase transition – diagnosed by **order** and **disorder** parameters

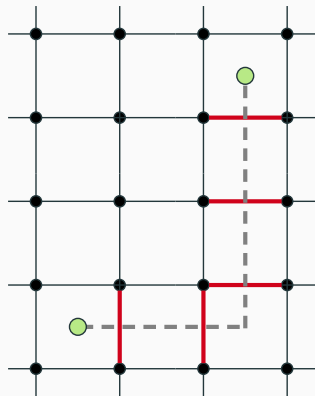
# Order parameter



**Wilson loop** of  $Z$  operators,  $W_\Gamma = \prod_\Gamma Z$ , along closed path  $\Gamma$  on lattice [Wegner '71; Kogut '79]

- In deconfined phase,  $\langle W \rangle$  decays with **perimeter law**
- In confined phase,  $\langle W \rangle$  decays with **area law**

# Disorder parameter



't Hooft string of  $X$  operators,  $T = \prod_{\tilde{\Gamma}} X$ , along open path  $\tilde{\Gamma}$  between two points on *dual* lattice

- $\langle T \rangle$  independent of path due to Gauss' law
- In deconfined phase, creates a pair of quasi-particles (magnetic monopoles) –  $\langle T \rangle$  decays exponentially with distance
- In confined phase,  $\langle T \rangle$  independent of distance – “monopoles condensed”



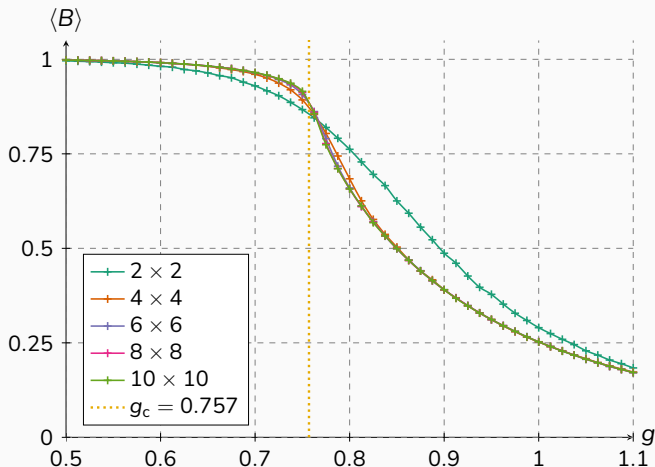
# Identifying the critical coupling

Ordered phase	Disordered phase
$g < g_c$	$g > g_c$
$\langle W \rangle \sim$ perimeter law	$\langle W \rangle \sim$ area law
$\langle T \rangle \sim$ exp. decay with distance	$\langle T \rangle \sim$ constant
electric flux lines condensed	mag. monopoles condensed
$\mathbb{Z}_2^{(1)}$ broken	$\mathbb{Z}_2^{(1)}$ preserved
“deconfined”	“confined”

[Rayhaun, Williamson '23]

To identify the critical coupling, we look for signs of the confinement transition

# Order parameter



**Figure 3:** Lattice average of the 1-plaquette Wilson loop (magnetic flux energy,  $B = \sum_{\square} W_{\square}$ ). The dotted line indicates the expected critical coupling at  $g_c = 0.757$ .

# Confinement

Classic signal of confinement: Wilson loops have **area-law decay**

Generically

$$\langle W_\Gamma \rangle \sim \exp(-\kappa P_\Gamma - \sigma A_\Gamma),$$

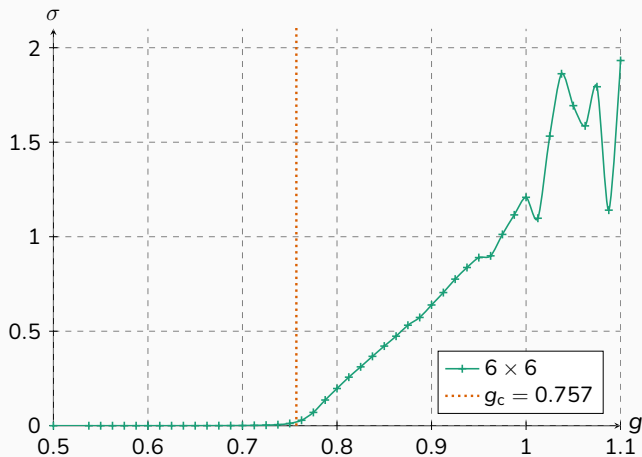
where  $\sigma$  is **string tension** – zero in deconfined phase and positive in confined phase

- Not a “clean” diagnostic – there is always a perimeter-law contribution

Can estimate  $\sigma$  via **Creutz ratio** [Creutz ‘80]

$$\sigma \approx -\log \frac{\langle W_{l \times l} \rangle \langle W_{(l-1) \times (l-1)} \rangle}{\langle W_{(l-1) \times l} \rangle \langle W_{l \times (l-1)} \rangle}$$

# String tension



**Figure 4:** String tension estimated from Creutz ratio for  $l = 2$  on a  $6 \times 6$  lattice. For  $g > g_c$ , Wilson loops decay rapidly with size. Estimate of  $\sigma$  challenging with Monte Carlo sampling – ratio of very small numbers.

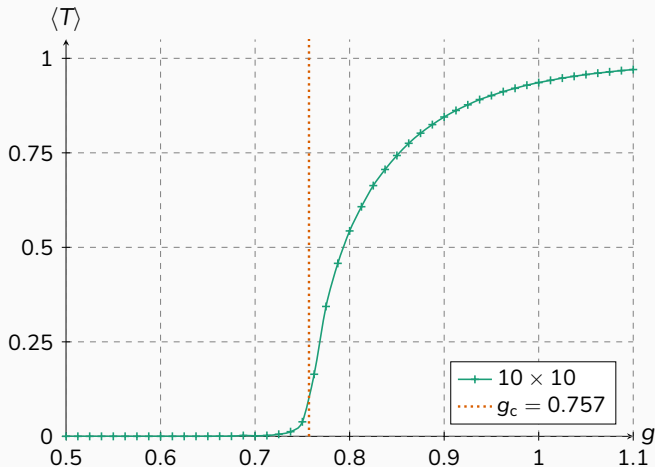
# Identifying the critical coupling

Ordered phase	Disordered phase
$g < g_c$	$g > g_c$
$\langle W \rangle \sim$ perimeter law	$\langle W \rangle \sim$ area law
$\langle T \rangle \sim$ exp. decay with distance	$\langle T \rangle \sim$ constant
electric flux lines condensed	mag. monopoles condensed
$\mathbb{Z}_2^{(1)}$ broken	$\mathbb{Z}_2^{(1)}$ preserved
“deconfined”	“confined”

[Rayhaun, Williamson '23]

Cleaner to look at decay of 't Hooft string

# Disorder parameter



**Figure 5:** Lattice average of the 't Hooft string operator near the critical point for  $L = 10$ . Distance  $L/2 = 5$  between ends of 't Hooft string.

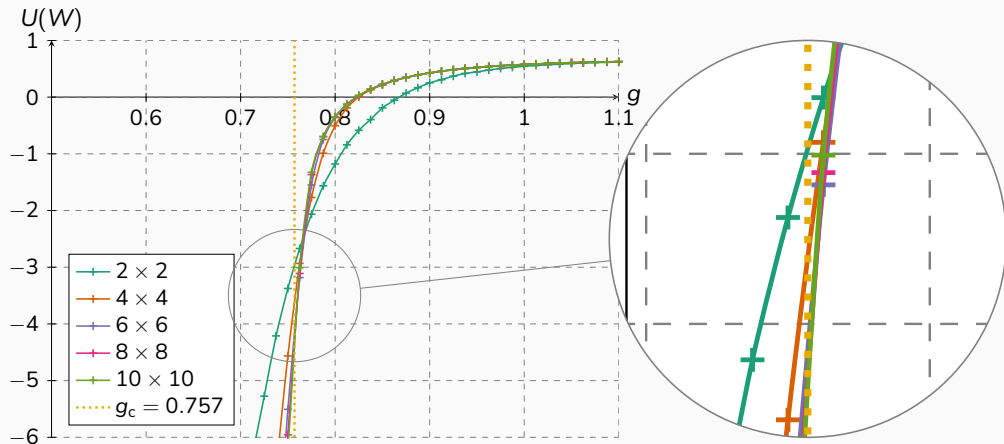
## A more accurate estimate

For a continuous phase transition, can identify critical point using the **Binder cumulant** of the order parameter [Binder '81]

$$U(W) = 1 - \frac{\langle W^4 \rangle_c}{3 \langle W^2 \rangle_c^2}$$

- At leading order in  $L^{-1}$ , **finite-scaling theory** predicts that the Binder cumulant has a universal value for all lattice sizes at the critical point
- Implies that the curves for different  $L$  must **cross at the critical coupling**
- Next-order terms in  $L^{-1}$  lead to a **small offset**

# Binder cumulant



**Figure 6:** Binder cumulant  $U(W)$  of 1-plaquette Wilson loops. The dotted yellow line indicates is the critical coupling at  $g_c = 0.757$ .



# Extrapolation

To extrapolate to  $L \rightarrow \infty$  limit, use **BST extrapolation** [Bulirsch, Stoer '64; Henkel, Patkos '87; Henkel, Schütz '88]

- Look at pairs  $(L, g_{\text{crossing}})$ :

$(2, 0.767512)$   $(4, 0.768072)$   $(6, 0.760667)$   $(8, 0.758856)$

- Construct a sequence which converges to  $L \rightarrow \infty$  limit more rapidly

Extrapolation predicts

$$g_c = 0.756(2)$$

- Compare with 0.757051 from dual spin model

# Curve collapse

Near to the critical point, **leading-order** finite-scaling theory predicts observables are functions of

$$\tilde{g} = L^{1/\nu}(g - g_c)/g_c$$

e.g. 't Hooft string  $\langle T \rangle$  should scale as

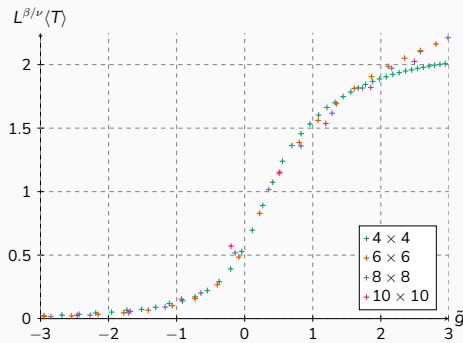
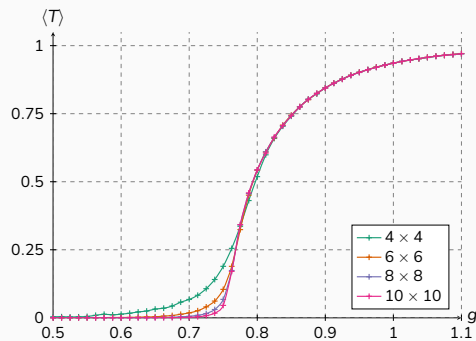
$$\langle T \rangle = L^{-\beta/\nu} t(\tilde{g})$$

- $\nu$  and  $\beta$  are correlation length and magnetisation **critical exponents**

**Curve collapse:** plot of  $L^{\beta/\nu} \langle T \rangle$  vs  $\tilde{g}$  *independent* of  $L$  near to  $\tilde{g} = 0$

- Extract  $g_c$ ,  $\beta$  and  $\nu$  by minimising distance between curves

# Collapse of disorder parameter



**Figure 7:** Lattice average of the 't Hooft string for varying lattice size. Distance  $L/2$  between ends of 't Hooft string.

# Critical exponents and the Ising CFT

Ising: [Blöte, Deng '02]

$$g_c = 0.757051 \quad \nu = 0.629971 \quad \beta = 0.326419$$

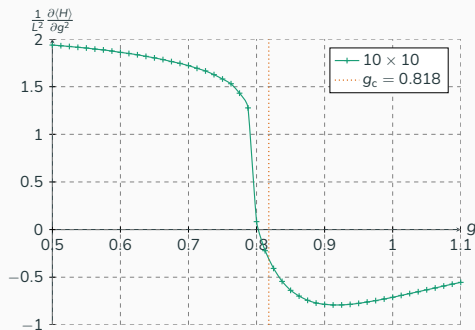
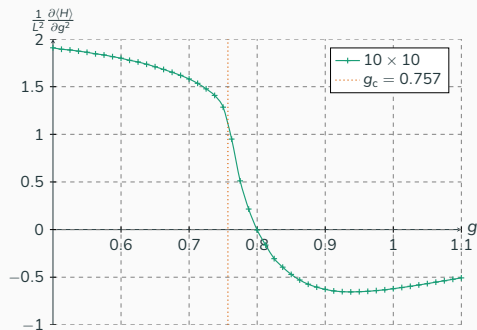
Curve collapse:

$$g_c \approx 0.766 \quad \nu \approx 0.610 \quad \beta \approx 0.320$$

Only includes **leading-order**  $L^{-1}$  behaviour (worse than BST extrapolation)

- Larger lattice sizes should give more accurate results as next-order corrections in  $L^{-1}$  suppressed (Ising data for  $L = 2, \dots, 48$ )

## $\mathbb{Z}_2$ vs $\mathbb{Z}_3$ : a first-order phase transition?



**Figure 8:** Derivative of  $\langle H \rangle / L^2$  with respect to  $g^2$  for  $\mathbb{Z}_2$  (left) and  $\mathbb{Z}_3$  (right).

First-order phase transition  $\Rightarrow$  derivative of energy is **discontinuous**

- **Finite correlation length**, but can be much larger than achievable system size [Binder '81]

# Summary

Neural-network quantum states are a powerful tool for studying interacting quantum systems

Gauge-equivariant CNNs extend this tool to lattice gauge theories

Provide accurate and flexible approximation of ground state for  $\mathbb{Z}_N$  theories

Can be used to probe phase transitions, critical exponents, and more

# Future directions

## Deeper networks

- NetKet and JAX scale well with number of GPUs
- Phase vs amplitude learning?

Symmetries: time reversal, reflection, etc.

Extension to  $U(1)$  and  $SU(N)$  (in progress)

- LGE-CNN structure remains the same
- How does complexity scale for non-abelian groups?

Include matter fields at lattice sites

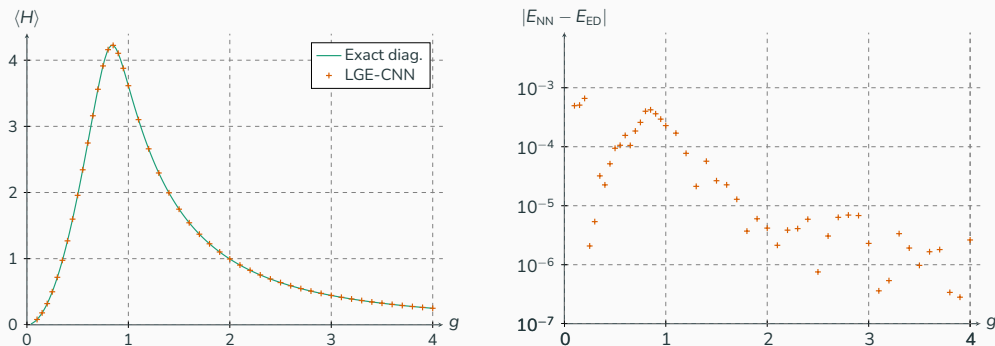
- Lattice QED/QCD?

**Extra**

---

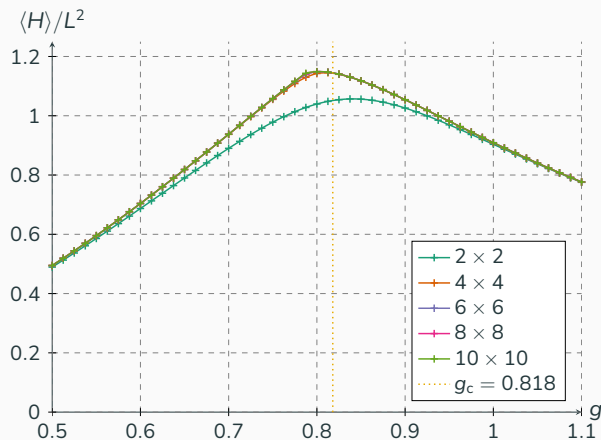


# Ground-state energy



**Figure 9:** Expectation value  $\langle H \rangle$  for  $\mathbb{Z}_3$  gauge theory on a  $2 \times 2$  lattice as a function of coupling  $g$ . The solid teal line is the exact diagonalization result, computed using NetKet. The inset shows the difference between the ground state energies calculated using the neural network and exact diagonalisation.

# Ground-state energy



**Figure 10:** Expectation value  $\langle H \rangle$  for  $\mathbb{Z}_3$  gauge theory for varying lattice sizes as a function of coupling  $g$ , focused on the region around the critical coupling.

# Variational Monte Carlo with NetKet

Given a state  $|\psi\rangle$ , one can compute  $\langle\hat{O}\rangle$  via

$$\langle\hat{O}\rangle = \langle\psi|\hat{O}|\psi\rangle = \sum_{\sigma} |\psi(\sigma)|^2 \sum_{\eta} \frac{\langle\sigma|\hat{O}|\eta\rangle\langle\eta|\psi\rangle}{\langle\sigma|\psi\rangle} = \mathbb{E}_{\sigma\sim|\psi(\sigma)|^2} [E_{\text{loc}}(\sigma)],$$

where  $\sigma$  and  $\eta$  are gauge field configurations,  $\mathbb{E}_{\sigma\sim|\psi(\sigma)|^2}$  means take the expectation value over  $\sigma$ , with  $\sigma$  distributed according to the PDF  $|\psi(\sigma)|^2$ , and we have defined the “local energy”

$$E_{\text{loc}}(\sigma) = \sum_{\eta} \frac{\langle\sigma|\hat{O}|\eta\rangle\langle\eta|\psi\rangle}{\langle\sigma|\psi\rangle}.$$

# Variational Monte Carlo with NetKet

In order to evaluate this expectation value via variational Monte Carlo, one needs to:

1. Restrict to a finite sum over  $\sigma$  by sampling the PDF,  $|\psi(\sigma)|^2$ . This is provided by the Monte Carlo variational state interface of NetKet.
2. Compute the “connected elements” for the samples  $\{\sigma\}$ , i.e. those for which  $\langle \sigma | \hat{O} | \eta \rangle \neq 0$ . For each  $\sigma$ , one can restrict to a (usually much) smaller set of configurations, labelled by  $\{\eta\}$ . Then compute the matrix elements themselves.
3. Calculate the local energies  $E_{\text{loc}}(\sigma)$  given the matrix elements, the sets  $\{\sigma\}$  and  $\{\eta\}$ , and the state  $|\psi\rangle$ . Again, this implementation is provided by NetKet.
4. The statistical average of the local energies weighted by  $|\psi(\sigma)|^2$ .

# Stochastic reconfiguration

Stochastic gradient descent with a non-trivial weight-space metric,

c.f. **imaginary time evolution** [Sorella et al. '07]

$$(1 - \epsilon H)|\Psi\rangle \approx e^{-\epsilon H}|\Psi\rangle = e^{-\epsilon H} \sum_{i=0} c_i |i\rangle = e^{-\epsilon E_0} c_0 |0\rangle + e^{-\epsilon E_1} c_1 |1\rangle + \dots$$

- Iterating this will project a trial wavefunction onto the ground state

Construct a weight update rule so that the updated state is  $(1 - \epsilon H)|\Psi\rangle$

$$\theta_\alpha \mapsto \theta_\alpha - \eta (S^{-1})_{\alpha\beta} R_\beta$$

where

$$O_\alpha |\psi\rangle = \partial_\alpha |\psi\rangle, \quad R_\alpha = \langle O_\alpha^\dagger H \rangle - \langle O_\alpha^\dagger \rangle \langle H \rangle, \quad S_{\alpha\beta} = \langle O_\alpha^\dagger O_\beta \rangle - \langle O_\alpha^\dagger \rangle \langle O_\beta \rangle$$