Entanglement entropy with generative neural networks

Tomasz Stebel

Institute of Theoretical Physics,
Jagiellonian University, Kraków

NATIONAL SCIENCE CENTRE
POLAND

Project is supported by National Science Centre, grant no. 2021/43/D/ST2/03375.

with Piotr Białas, Piotr Korcyl and Dawid Zapolski

Based on:
CPC 281 (2022) 108502 and 2406.06193

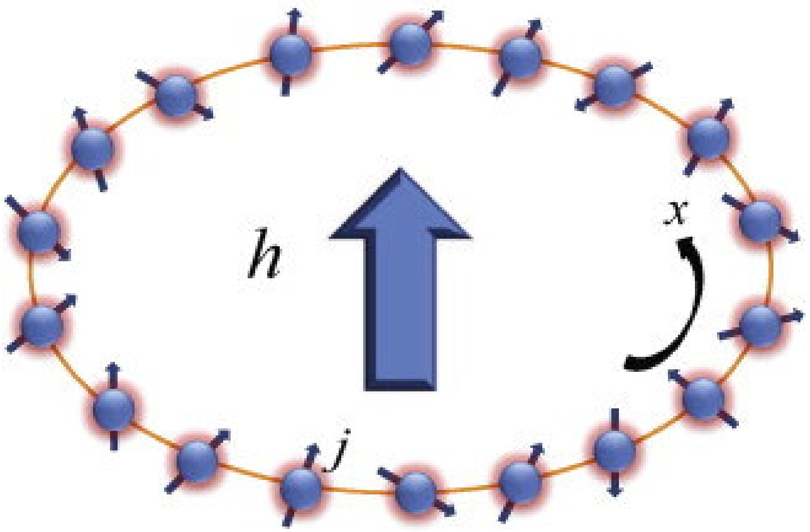
ML meets LFT,
26.07.2024

Plan

- 1d quantum Ising model
- Entanglement entropy
- Variatonal Autoregressive networks
- Hierarchical autoregressive networks (HAN)
- HAN for entanglement in Ising model
- Numerical results

1+1D quantum Ising model

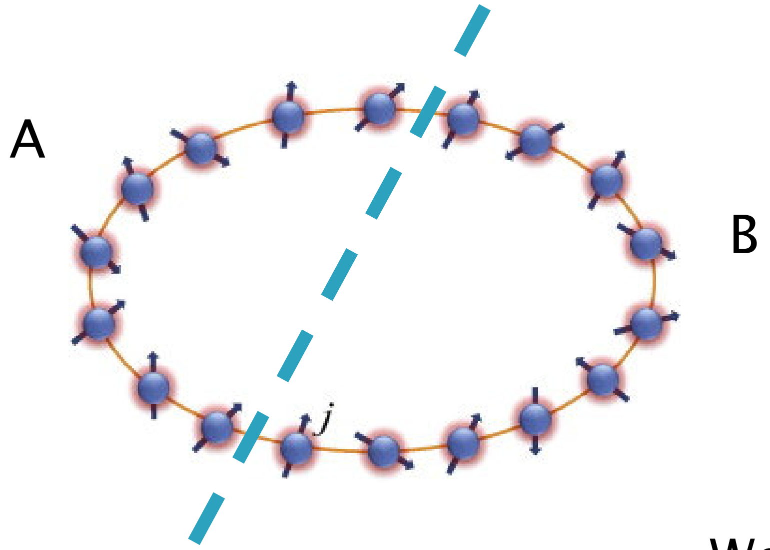
Spin chain with periodic boundary conditions (1+1D):



$$\hat{H} = -J \sum_{\langle i,j \rangle} \hat{\sigma}_i^z \hat{\sigma}_j^z - h \sum_i \hat{\sigma}_i^x,$$

where $\hat{\sigma}^x$, $\hat{\sigma}^z$ are Pauli matrices;
h–external magnetic field

Entanglement entropy



Density matrix of the system:

$$\rho_{ij} = \frac{\langle i | e^{-\beta H} | j \rangle}{\sum_i \langle i | e^{-\beta H} | i \rangle}$$

Normalization Z

We divide the system into 2 subsystems A and B.

Reduced density matrix of subsystem A: $\rho_A = \mathrm{Tr}_B\, \rho$

von Neumann entanglement entropy: $S(A) = -\,\mathrm{Tr}\, \rho_A \log \rho_A$

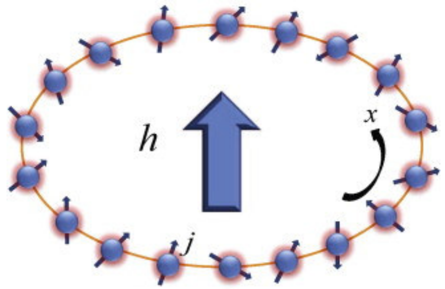
Rényi entropy of order n:

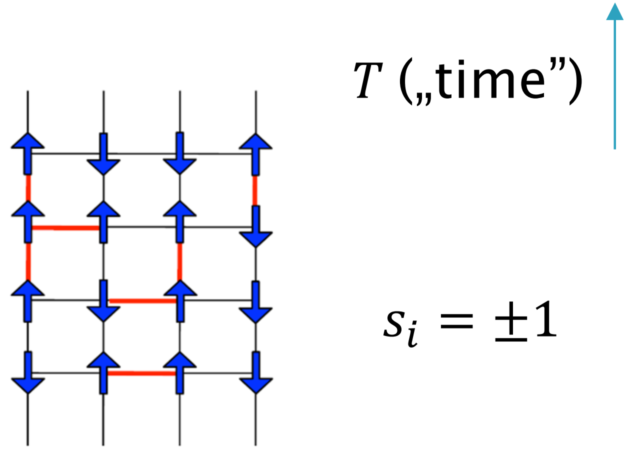
$$S_n(A) = \frac{1}{1-n} \log \mathrm{Tr}\, \rho_A^n$$

In this talk we focus on $n = 2$

Path integral formalism

To calculate $S_n(A) = \frac{1}{1-n} \log \operatorname{Tr} \rho_A^n$ one can use the path integral formalism:





1+1D quantum Ising model with transverse field

$$\hat{H} = -J \sum_{\langle i,j \rangle} \hat{\sigma}_i^z \hat{\sigma}_j^z - h \sum_i \hat{\sigma}_i^x,$$

2D classical Ising model (without external field)

$$E(\mathbf{s}) = -\beta \sum_{\langle i,j \rangle} s_i s_j$$

For simplicity we assume that couplings between spins are the same in space and time direction (specific choice of J, h and time discretization).

$\beta = \beta_c = \frac{1}{2} \log(1 + \sqrt{2})$

Conformal symmetry → some results are available

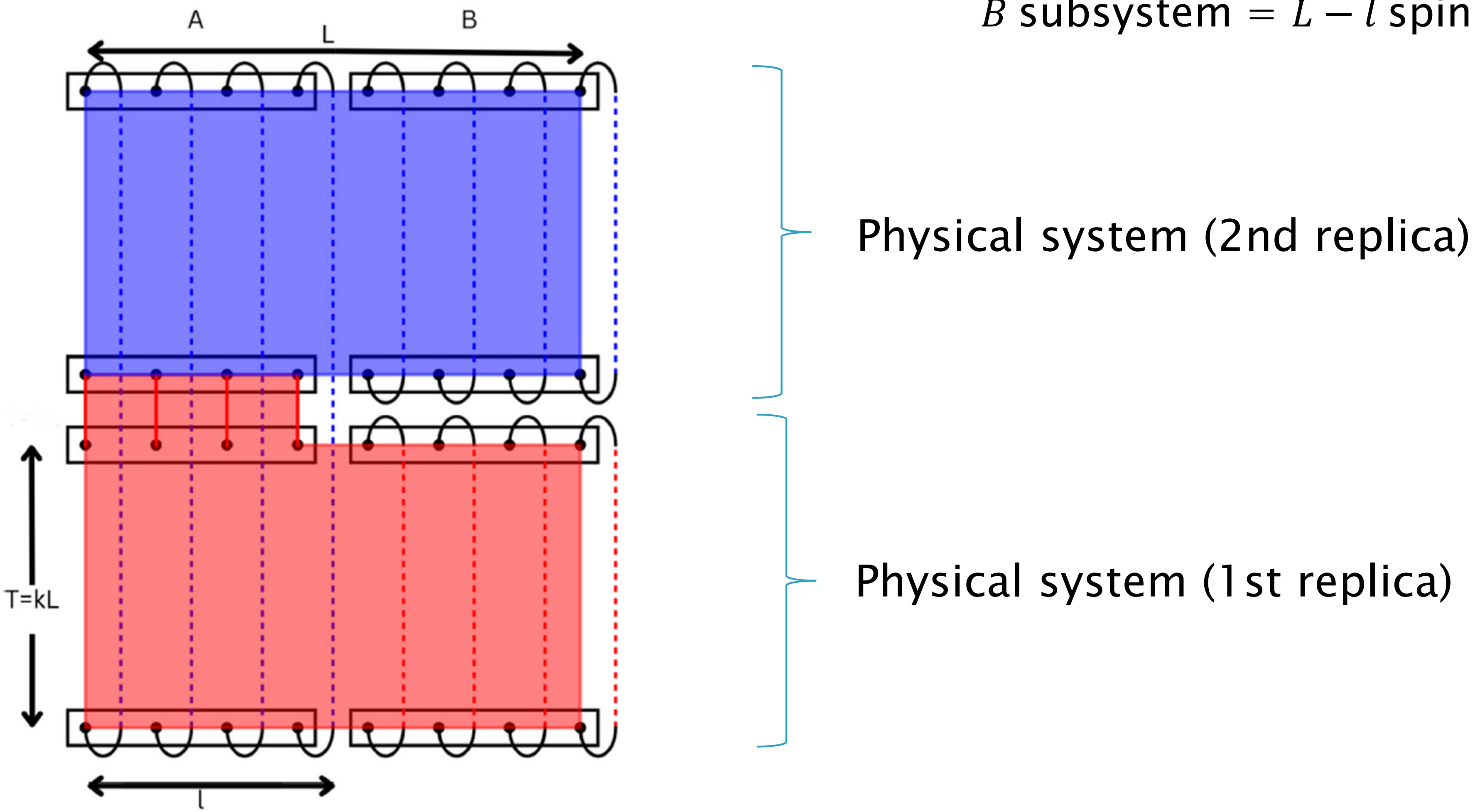
P. Calabrese and J. Cardy, Journal of Physics A: Mathematical and Theoretical, vol. 42, p. 504005, dec 2009.

Replica trick

To calculate Rényi entanglement entropy partition function is not enough.

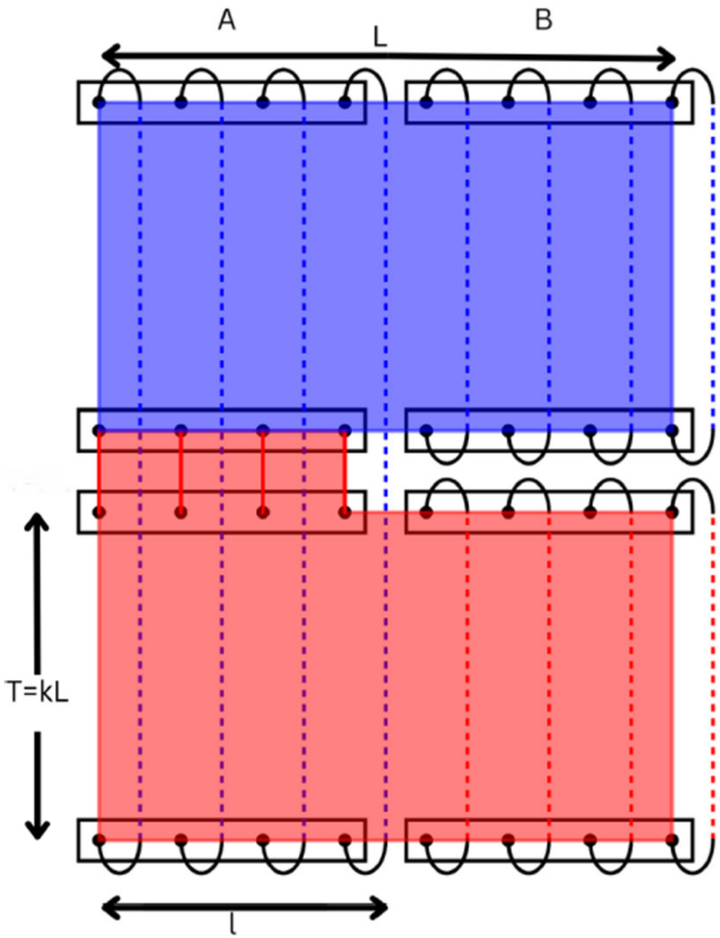
One can use replica trick:

A subsystem $= l$ spins,
B subsystem $= L - l$ spins



P. Calabrese and J. Cardy, Journal of Physics A: Mathematical and Theoretical, vol. 42, p. 504005, dec 2009.

Replica trick



Partition function of the n-replica system:

$$Z_n(A) = \sum_{\mathbf{s}_{(n)}} e^{-E_{(n)}\left(\mathbf{s}_{(n)}\right)}$$

Rényi entropy of order n:

$$S_n(A) = \frac{1}{1-n} \log \frac{Z_n(A)}{Z^n}$$

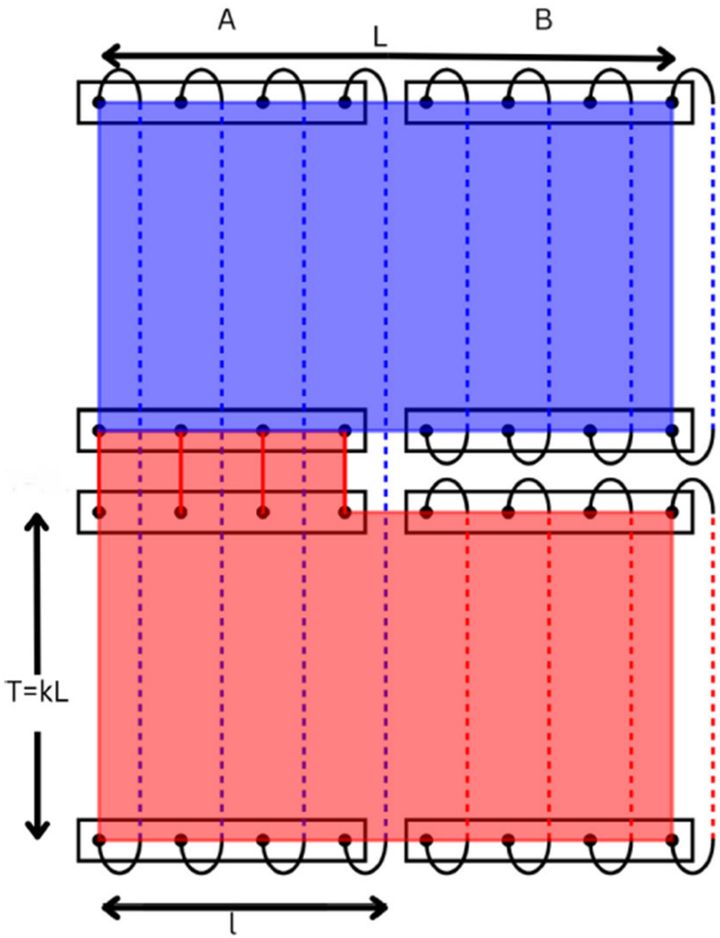
„Standard" partition function (1 replica)

Generative neural networks are capable to calculate partition functions...

...by learning Boltzman probability distribution $p(s)$.

P. Calabrese and J. Cardy, Journal of Physics A: Mathematical and Theoretical, vol. 42, p. 504005, dec 2009.

Replica trick



Renyi entropy (RE): $$S_n(A) = \frac{1}{1-n} \log \frac{Z_n(A)}{Z^n}$$

When $T \to \infty$ (time direction) the RE will measure the entanglement of the ground state.

In the simulations we take:

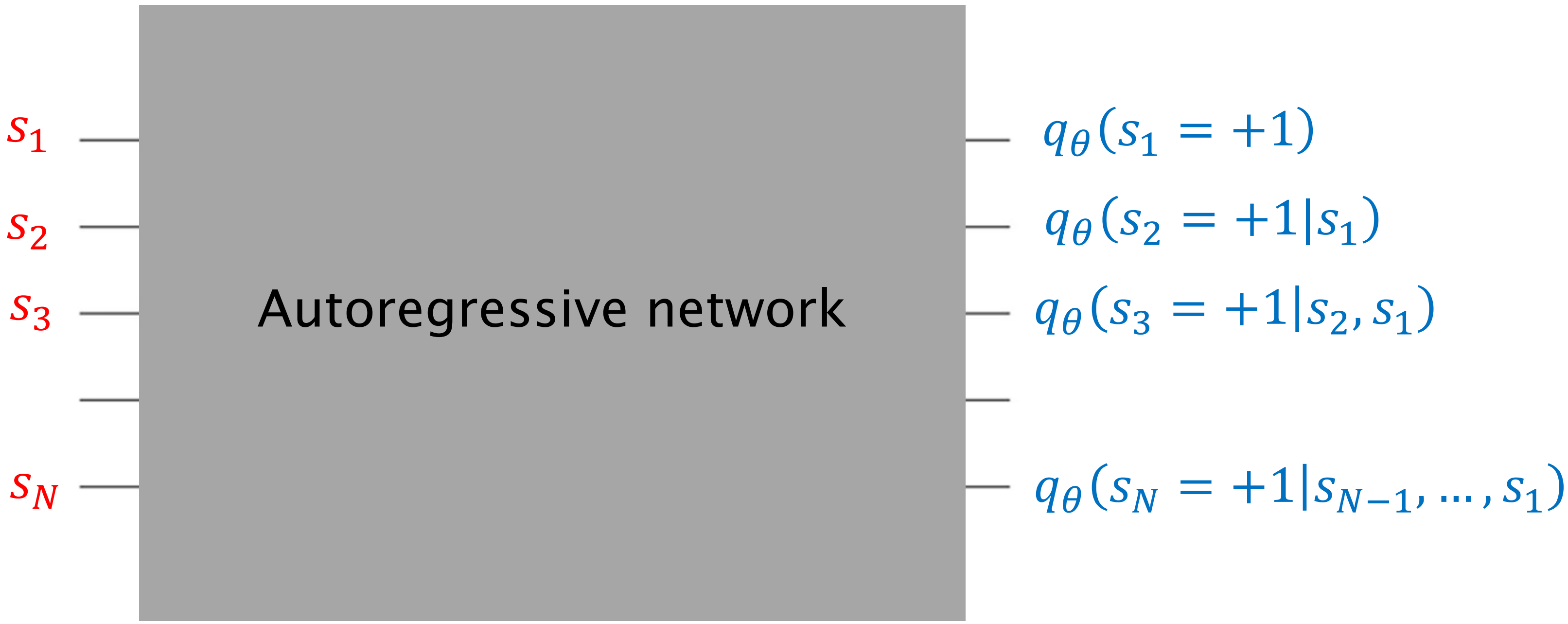
$$T = kL,$$

where $k \in \mathbb{Z}, k \gg 1$

D. Wu, L. Wang, and P. Zhang, Phys. Rev. Lett., vol. 122, p. 080602

Variatonal Autoregressive networks (VAN)

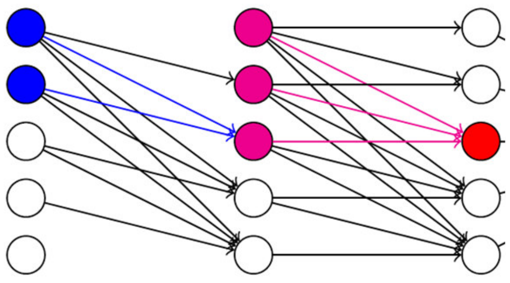
$$q_\theta(\boldsymbol{s}) = q_\theta(s_1)\, q_\theta(s_2|s_1)\, q_\theta(s_3|s_2, s_1) \ldots q_\theta(s_N|s_{N-1}, \ldots, s_1)$$



Input: spin configuration (value of each spin) $(\pm 1, \ldots \pm 1)$

Output: conditional probabilities

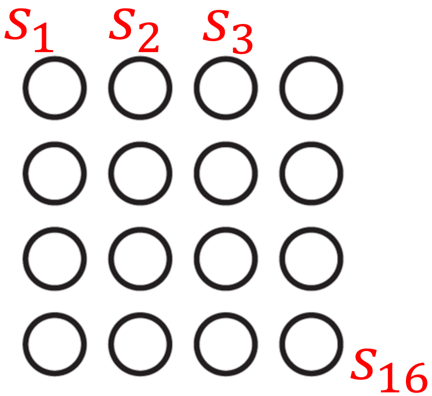
Autoregressive networks:



Half of the connections removed.

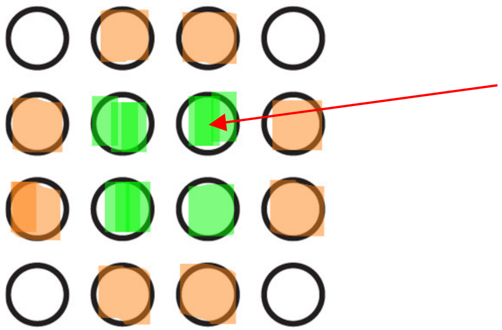
Białas, Korcyl, Stebel, Comput.Phys.Commun. 281 (2022) 108502

Hierarchical autoregressive networks



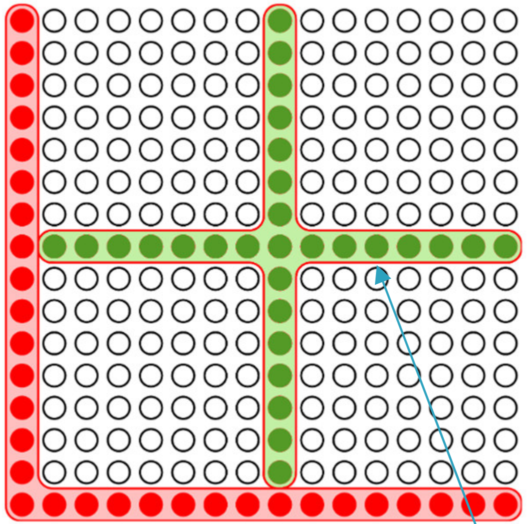
- It is there a better way to numerate the spins?

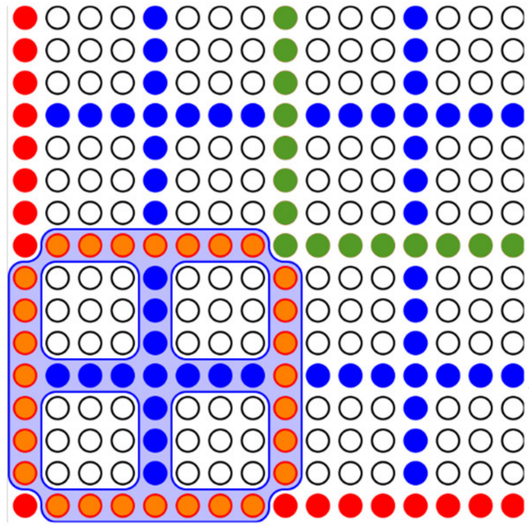
- We can use a property of Nearest Neighbour interactions:



Probability of green interior depends only on orange boundary (Hammersley-Clifford theorem)

Hierarchical autoregressive networks (HAN)





1st network

2nd network

3rd network
(called 4 times)

These networks have additional dependence on boundary surrounding them

Loss function and training

Training = adjust network weights θ such that $q_\theta(\boldsymbol{s})$ is as close to $p(\boldsymbol{s}) = Z^{-1}e^{-\beta E(\boldsymbol{s})}$ as possible.

Kullback–Leibler (KL) divergence

$$D_{\mathrm{KL}}(q_\theta \,\|\, p) = \sum_{\mathbf{s}} q_\theta(\mathbf{s}) \ln\left(\frac{q_\theta(\mathbf{s})}{p(\mathbf{s})}\right)$$

can measure a difference between two distributions.

$$D_{\mathrm{KL}}(q_\theta \,\|\, p) = \sum_{\mathbf{s}} q_\theta(\mathbf{s}) \ln\left(\frac{q_\theta(\mathbf{s})}{p(\mathbf{s})}\right) = \beta(F_q - F),$$

where

$$F_q = \frac{1}{\beta}\sum_{\mathbf{s}} q_\theta(\mathbf{s})\,[\beta E(\mathbf{s}) + \ln q_\theta(\mathbf{s})]$$

Variational free energy

Albergo, Kanwar, Shanahan, Phys.Rev.D 100 (2019) 3, 03451;
K. A. Nicoli et al., Phys. Rev. E, vol. 101, p. 023304,

Imperfection of training

- NN cannot learn $p(s)$ perfectly. We can however correct it. There are two ways to do this:

1) Neural Markov Chain Monte Carlo (NMCMC)

2) Neural Importance Sampling (NIS):

Reweighting observables

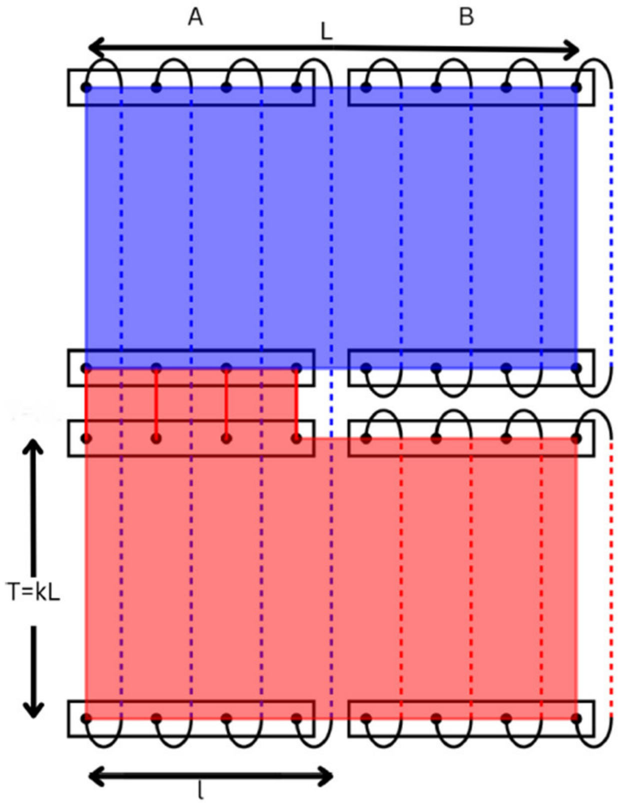
$$\langle \mathcal{O}(s) \rangle_p \approx \sum_i w_i \, \mathcal{O}(s_i)$$

$$\text{where } w_i = \frac{\hat{w}_i}{\sum_i \hat{w}_i} \text{ for } \hat{w}_i = \frac{e^{-\beta H(s_i)}}{q(s_i)}$$

Here I focus on 2) as it gives unbiased estimator of the partition function:

$$Z \approx \frac{1}{N} \sum_{i=1}^{N} \hat{w}(\mathbf{s}_i) \equiv \hat{Z}_N, \qquad \mathbf{s}_i \sim q_\theta$$

Coming back to entanglement…



Rényi entropy of order n:

$$S_n(A) = \frac{1}{1-n} \log \frac{Z_n(A)}{Z^n}$$

Partition function of the n-replica system:

$$Z_n(A) = \sum_{\mathbf{s}_{(n)}} e^{-E_{(n)}\left(\mathbf{s}_{(n)}\right)}$$

We use NIS to calculate $Z_2(A) \equiv Z_2(l)$

For this purpose, we need to train network to probability distribution given by „2-replica system energy" $E_2(s)$

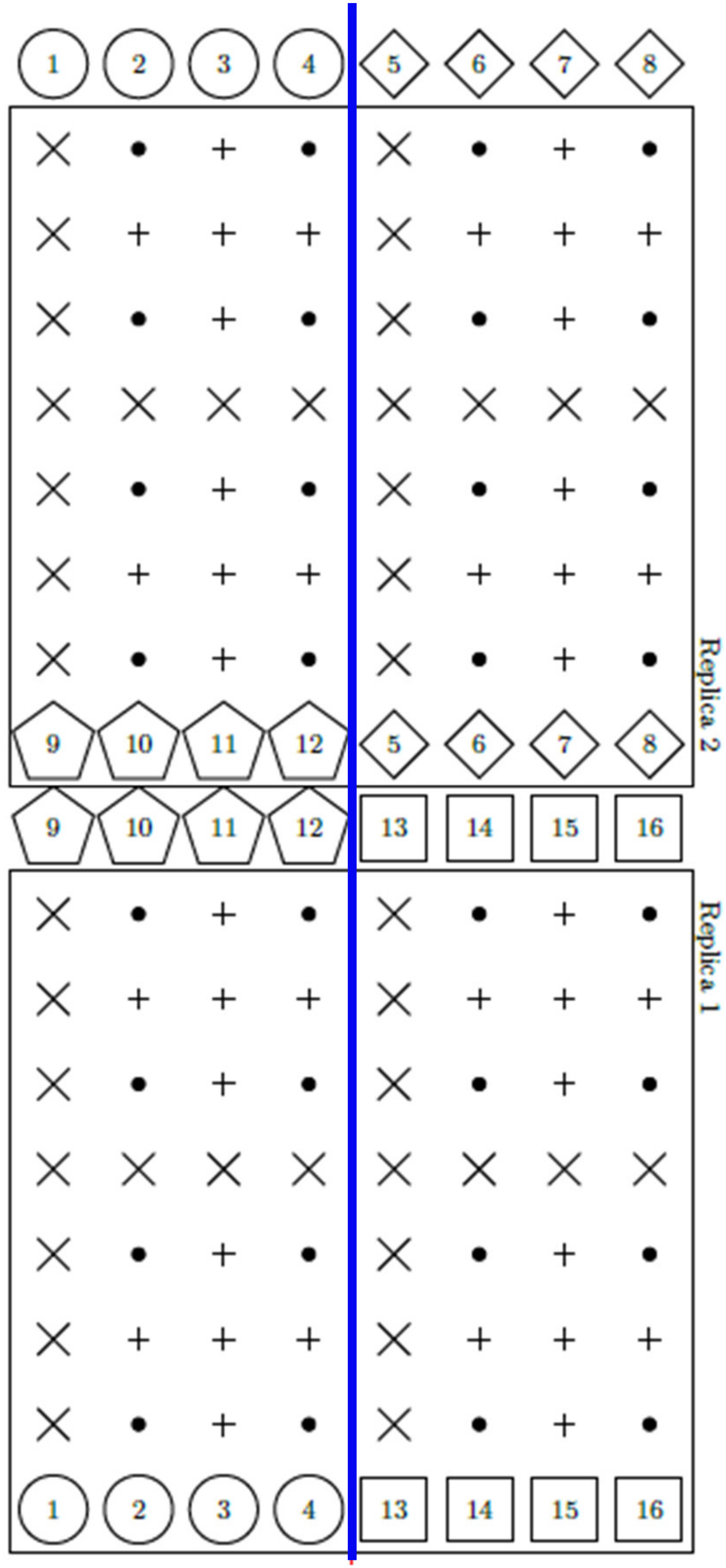
$$p_2\left(s_{2-replica}\right) = e^{-\beta\, E_2(s_{2-replica})} / \; Z_2$$

Generating 2-replica system

Numbers/marks denotes spins

Repeated number means that the spin is a copy

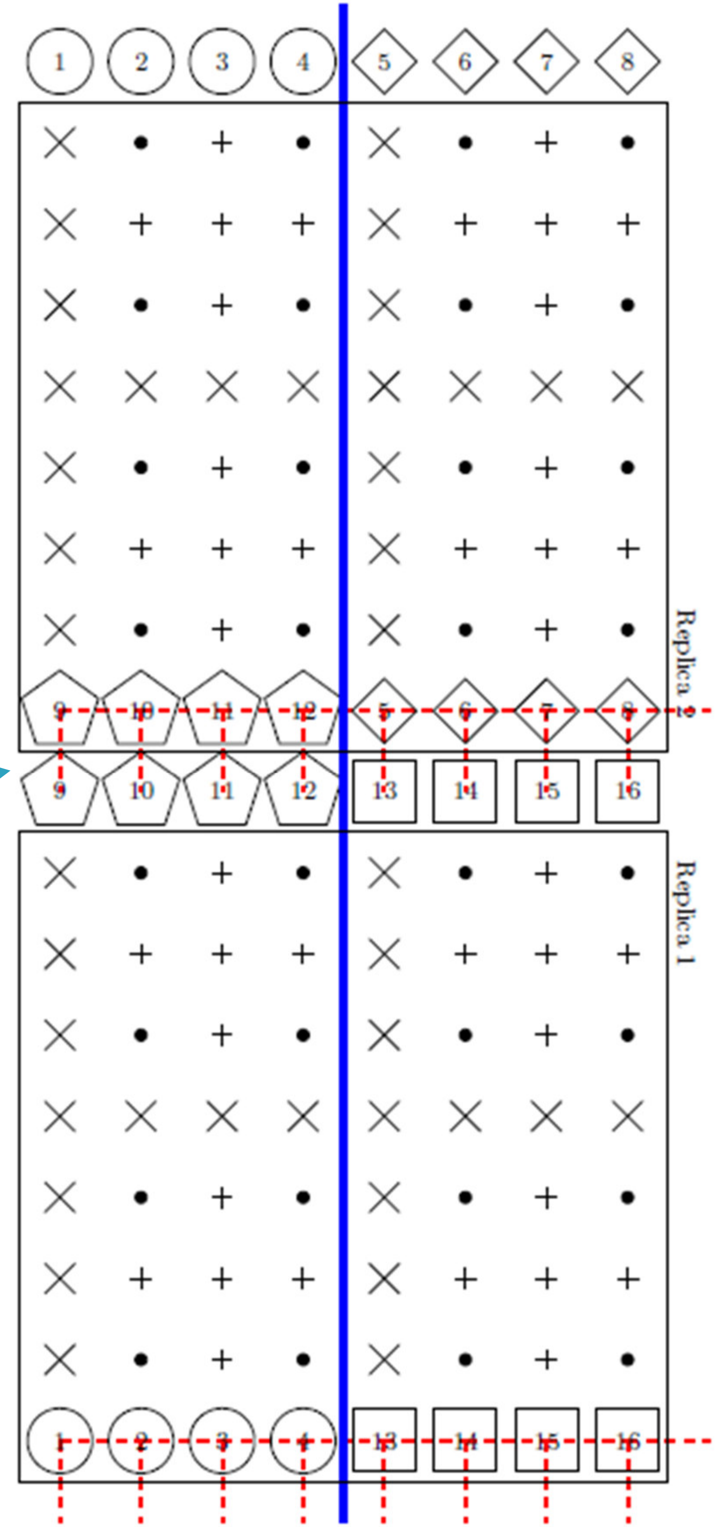
Marks denotes the hierarchy level.



Energy of 2-replica system

To calculate energy of this system E_2 we need to avoid double counting.

The red dashed lines denote interactions which are removed comparing to standard periodic boundary conditions.



Entropic C-function

- We consider derivative of $S_2(l)$ w.r.t. system size (with proper normalization) :

$$C_n(l) = \left[\frac{L}{\pi}\sin\left(\frac{\pi l}{L}\right)\right]^{D-1}\frac{1}{|\partial A|}\frac{1}{1-n}\times$$
$$\times \lim_{\epsilon\to 0}\frac{1}{\epsilon}\log\frac{Z_n(l)}{Z_n(l+\epsilon)}.$$

Known as entropic C-function.

It is UV finite in Quantum Field Theory (contrary to entanglement entropy).

Practical reason: there are more results for C_2 than for S_2.

After discretization:

$$C_n(l) \approx \frac{L}{2\pi}\sin\left(\frac{\pi l}{L}\right)\frac{1}{1-n}\log\frac{Z_n(l-\frac{1}{2})}{Z_n(l+\frac{1}{2})},$$

Numerical results

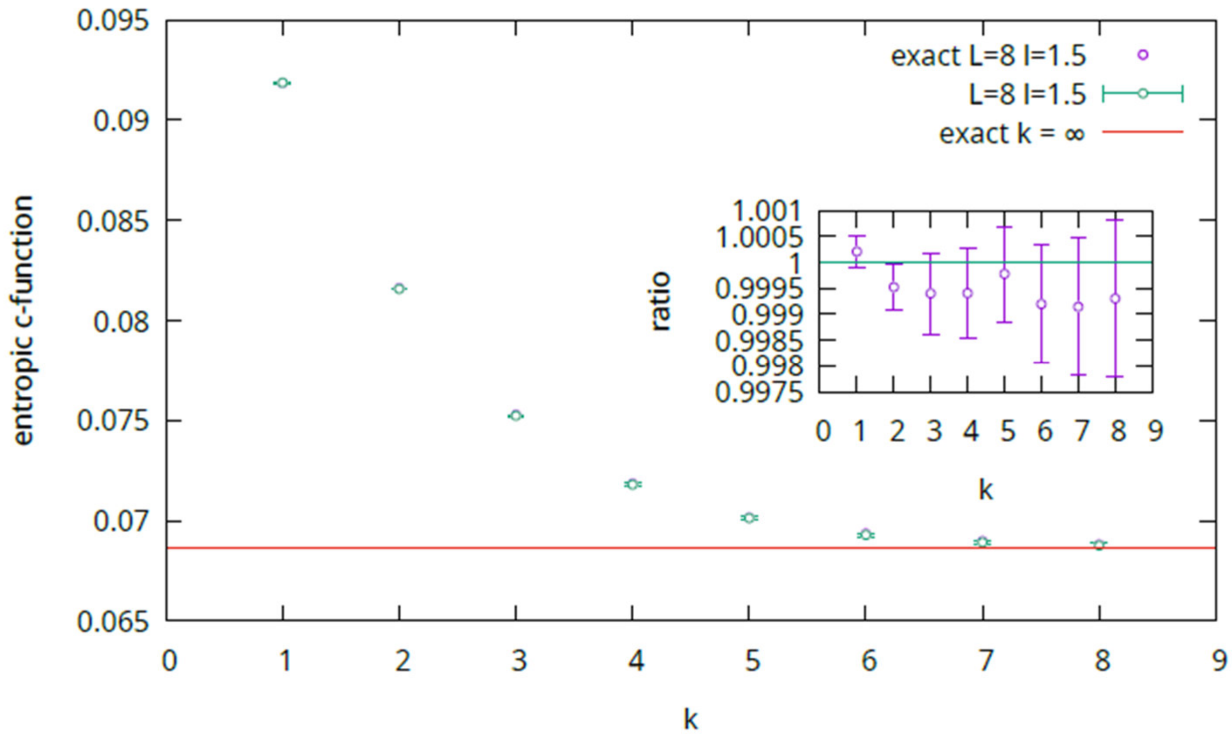
Dependence on T

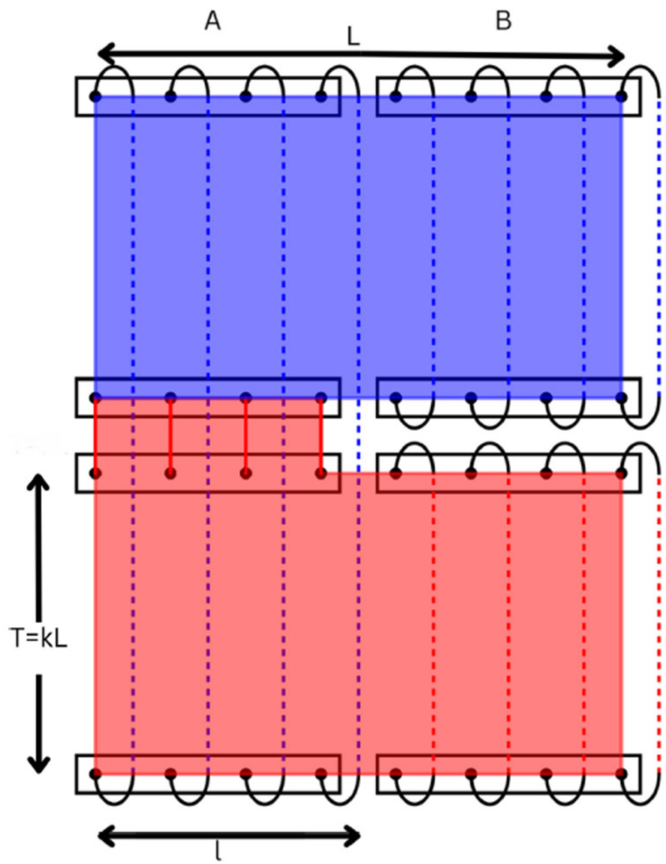
$$T = kL,$$

where $k \in \mathbb{Z}, k \gg 1$

We start with small system of 8 spins: L=8.

One can get exact results for such small size using transfer matrix method.

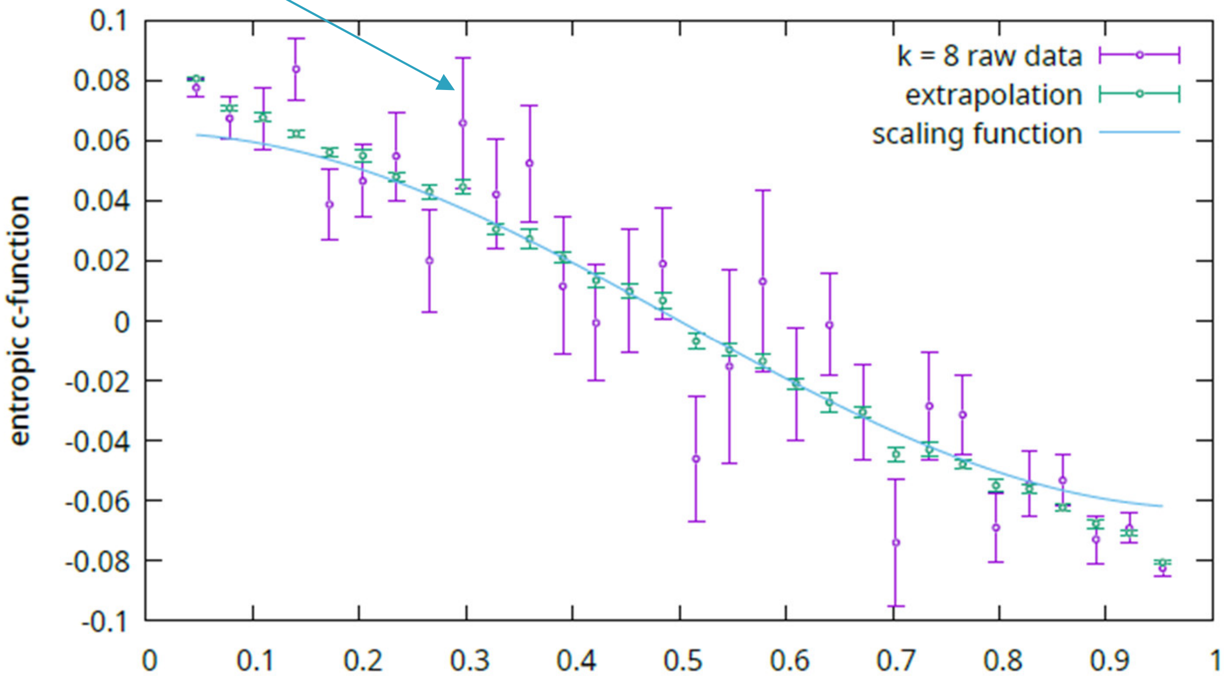




L=32 system (32 quantum spins)

For k=8 and L=32 our 2-repica system has $2^{14} = 16384$ classical spins.

The training of such systems is challenging and the resulting statistical errors for C_2 are large:



$x = l/L$

Fitting k –dependence and extrapolation $k \to \infty$

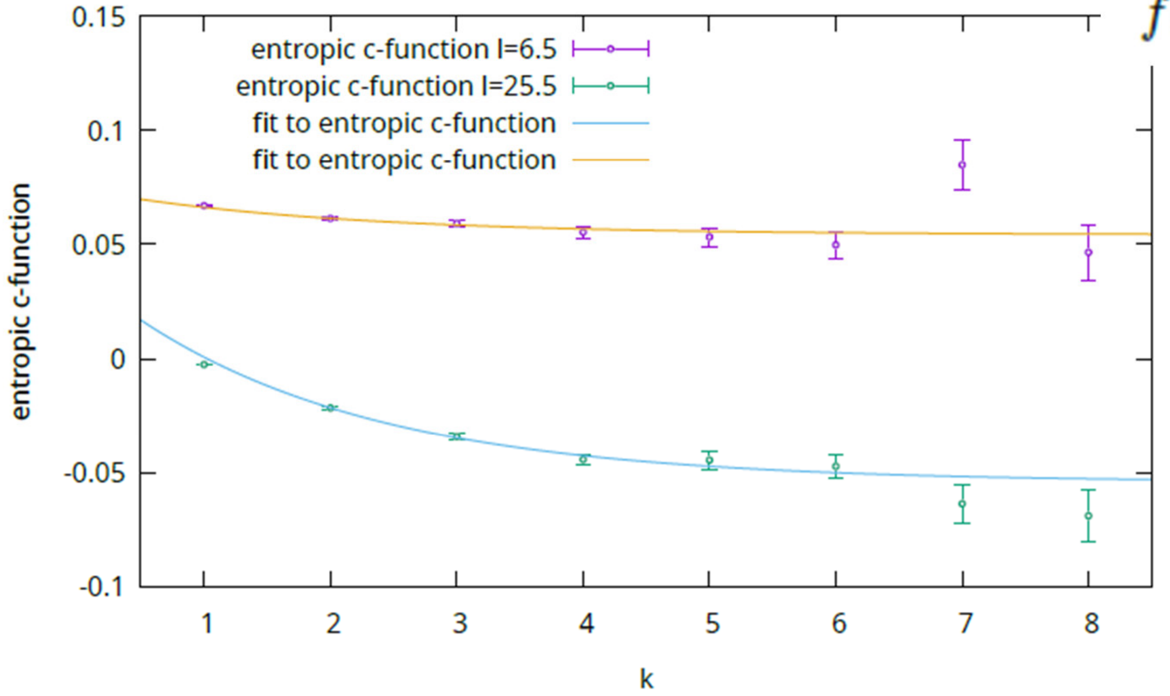
We use properties for $k\to\infty$:

1) $C_2(x) = -C_2(1-x)$

2) subleading corrections for x and $1-x$ have the same dependence on k.

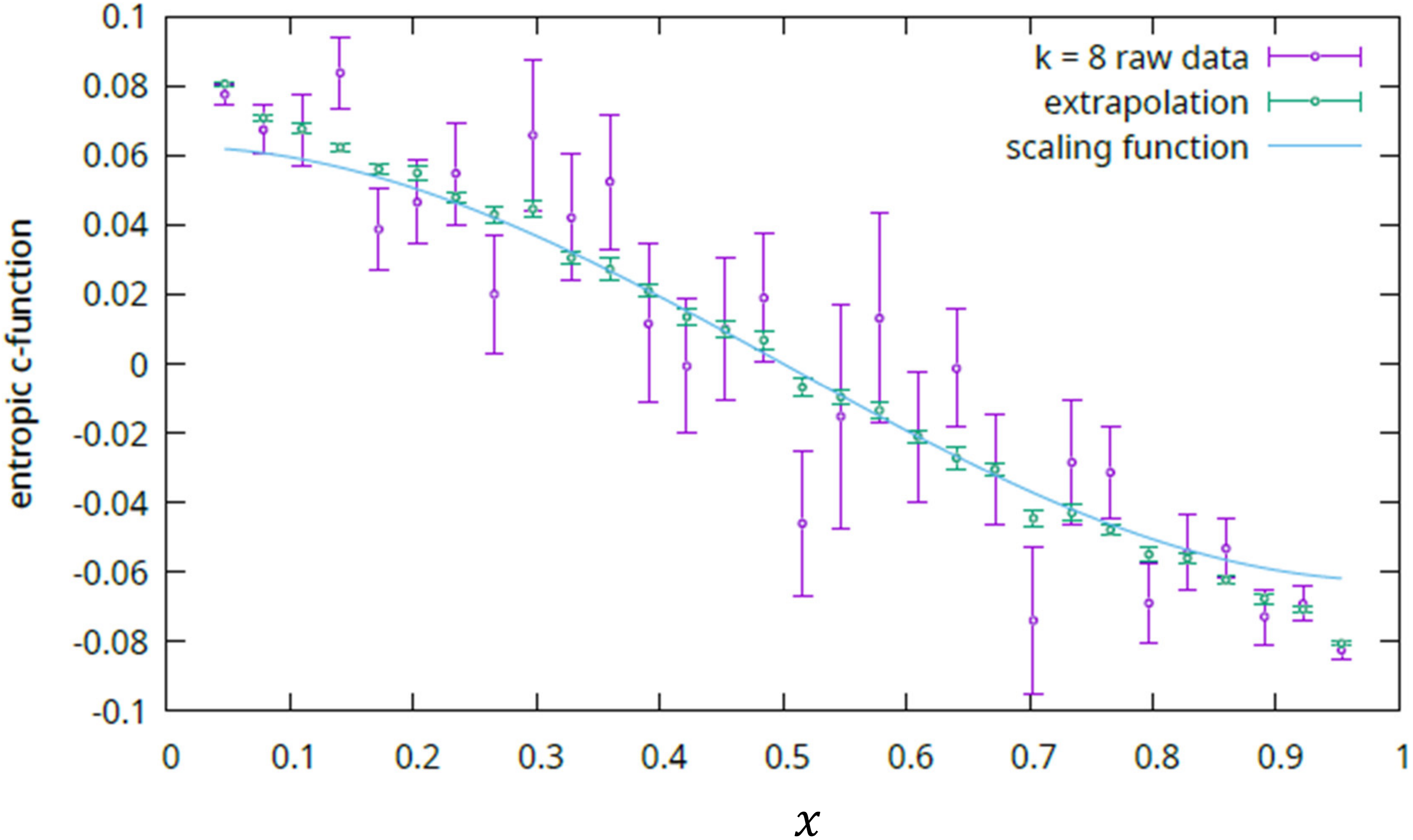
Model for fit:

$$f(k) = a + be^{-mk} \qquad \text{for } x = x_0,$$
$$f(k) = -a + \hat{b}e^{-mk} \qquad \text{for } x = 1 - x_0.$$

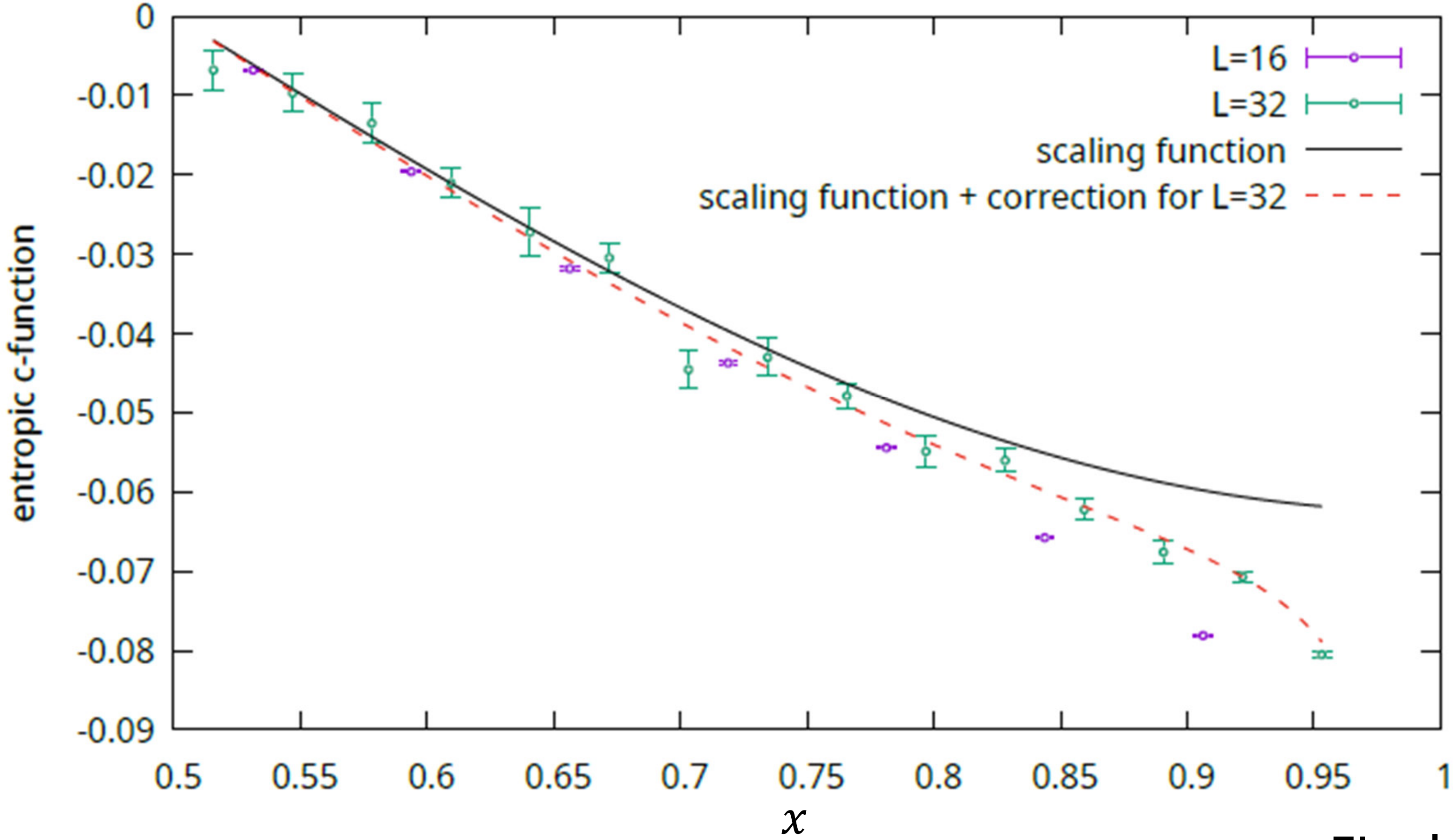


Simultaneous fit for two sets of data (4 parameters and 16 data points)

L=32 system: k=8 vs extrapolation



Comparison with theoretical calculations (and the other method)

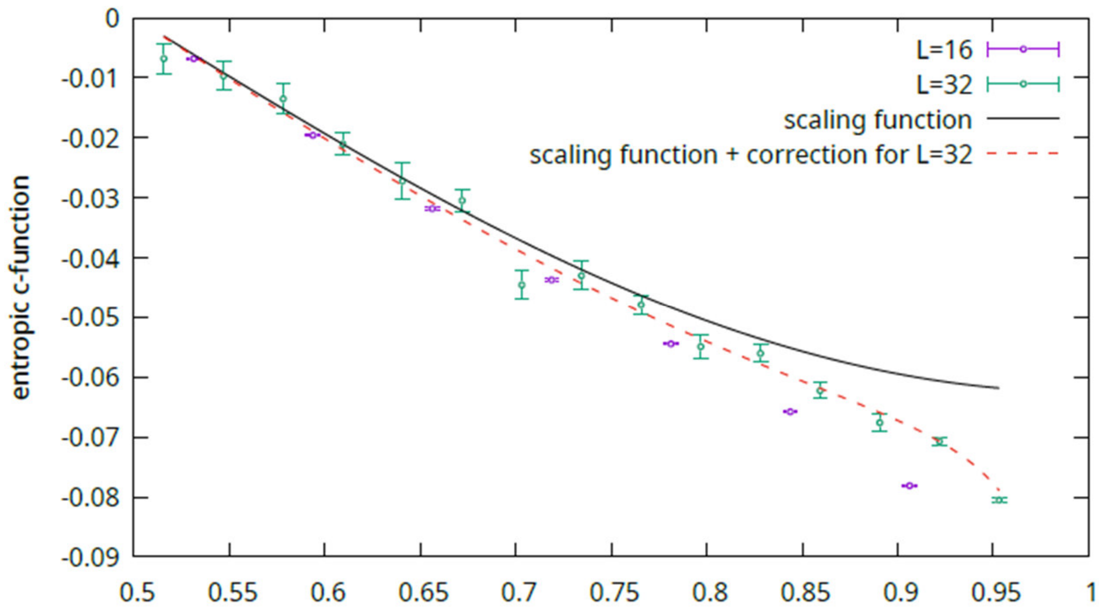


Conformal field theory prediction:

$$C_2(x) = \frac{1}{16}\cos(\pi x) + \frac{\kappa}{2L}\cot(\pi x)$$

Final size correction

Comparison with theoretical calculations (and the other method)



Conformal field theory prediction:

$$C_2(x) = \frac{1}{16}\cos(\pi x) + \frac{\kappa}{2L}\cot(\pi x)$$

Parameter κ was obtained in:

A. Bulgarelli and M. Panero, JHEP06, p. 030, 2023.

by fitting above formula to C_2 obtained using Jarzynski theorem.

See talk by Elia

Our result agrees with Bulgarelli&Panero.

Summary

- Autoregressive networks can be used to calculate partition functions.
- Using replica trick one can express the Renyi entanglement entropy in terms of partition function of system with specific boundary conditions.
- We calculated $n = 2$ entanglement entropy for 1d quantum Ising model, with 32 spins.

- Can this method be competitive to other methods (tensor networks, Jarzynski equation, etc...)?

- Better NN architectures? (see talk by Ankur)

- Can calculate entanglement in Lattice Field theories with this method?

Thank you