Finite-size analysis of second-order phase transitions in classical spin models using machine learning

Presenter

Lev Shchur, Landau Institute

Coauthors

Diana Sukhoverhova, HSE University

Vladislav Chertenkov, HSE University

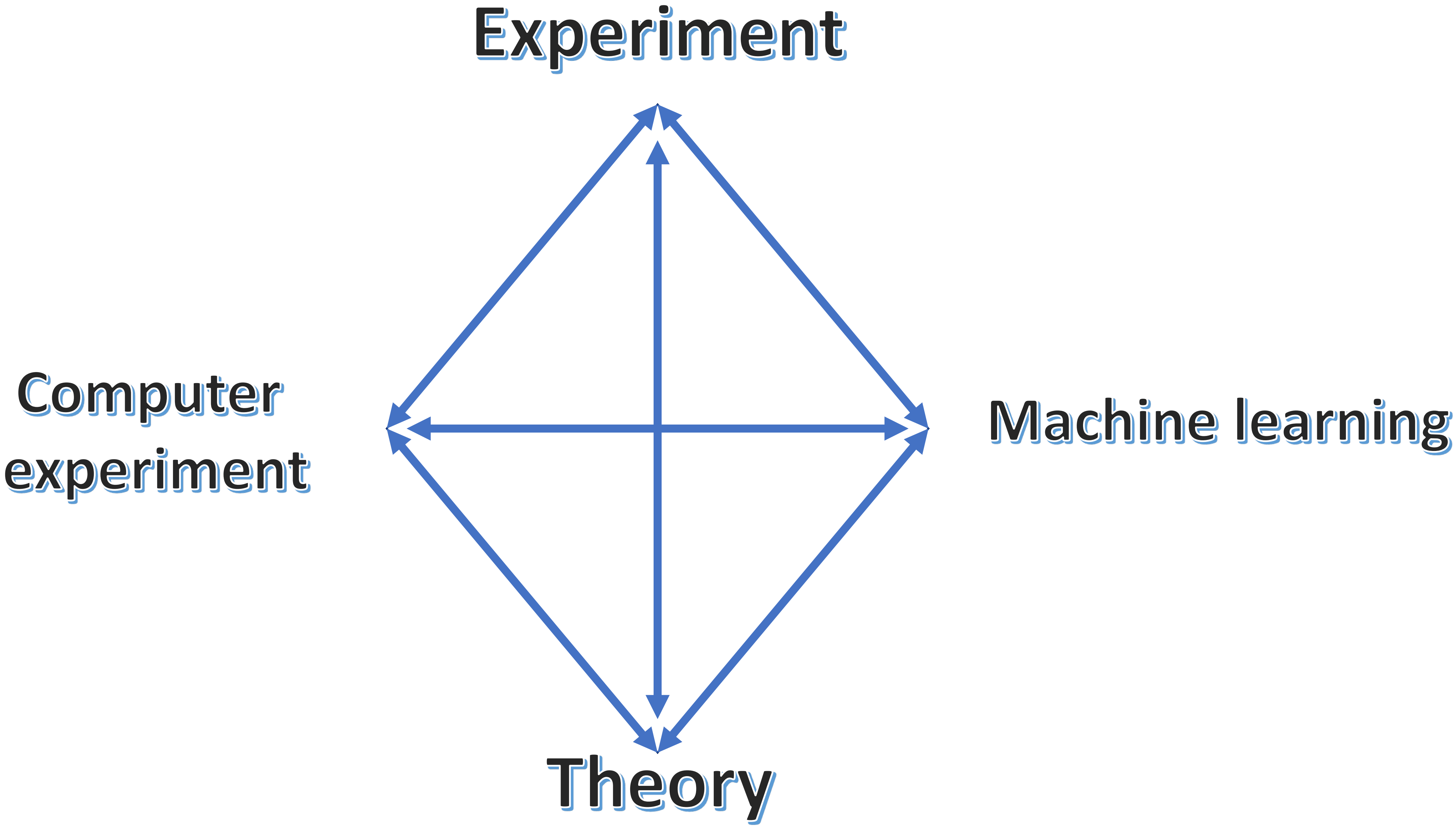
Evgenii Burovski, HSE University

Supported by Russian Science Foundation grant 22-21-00259

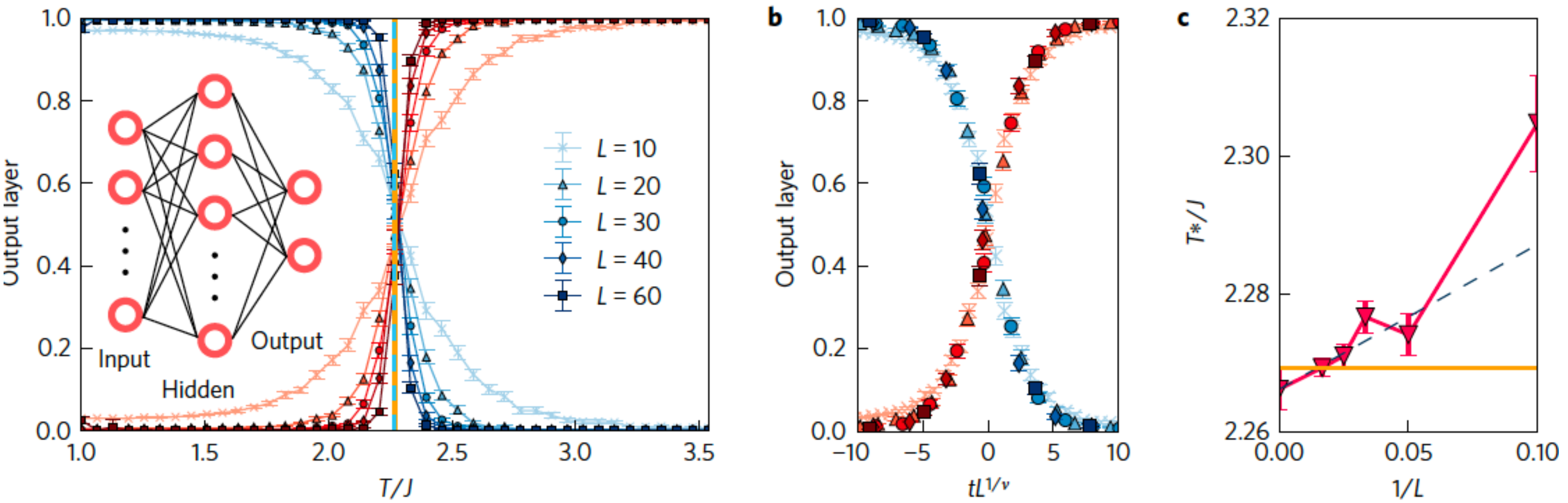
24th International NTZ-Workshop on New Developments in Computational Physics, Leipzig, December 20, 2023

Outline

1. Phase transition analysis with ***supervised*** machine learning :
    - Carrasquilla and Melko, «Machine learning phases of matter», Nature Physics 2017
    - Chertenkov, Burovski, and LS, «Finite-size analysis in neural network classification of critical phenomena», Physical Review E 108, L032102 (2023)
2. Anisotropy influence and transfer learning:
    - Training with isotropic model
    - Testing with anisotropic models
    - Prediction of critical temperature
    - Prediction of critical length exponent
3. Future work and conclusion


Experiment
Computer experiment
Machine learning
Theory

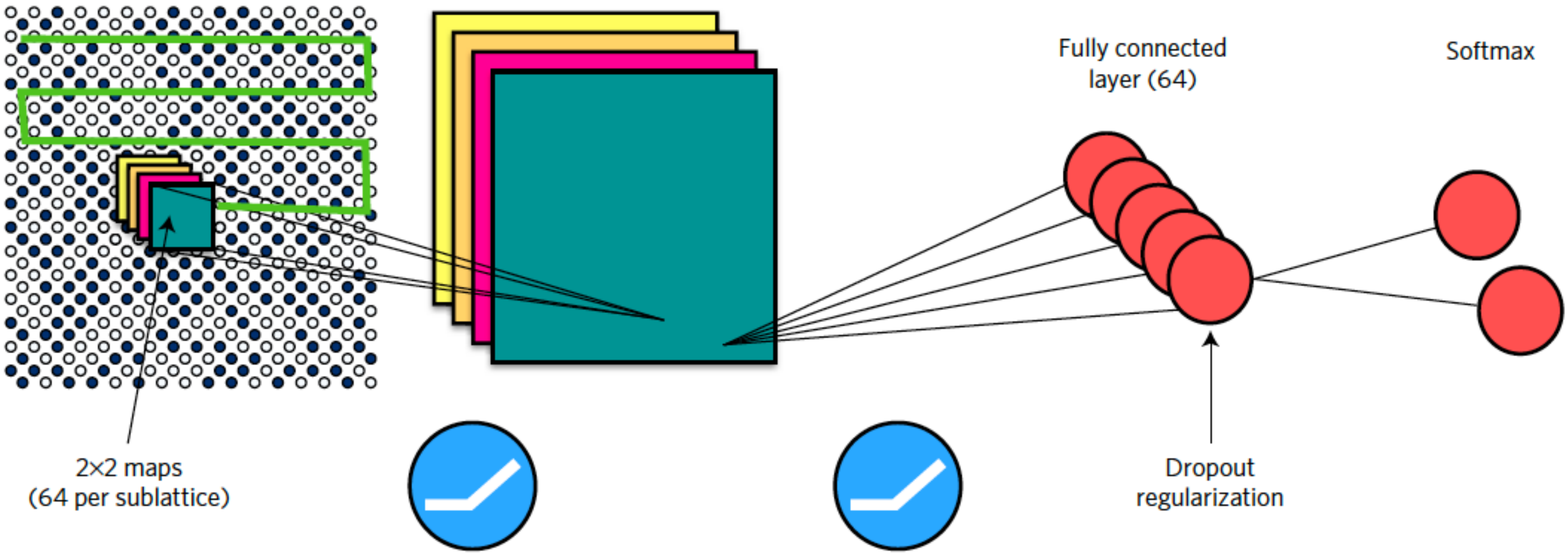
J. Carrasquilla and R.G. Melko, «Machine learning phases of matter», Nature Physics 13 (2017) 431



$\nu \simeq 1.0 \pm 0.2,$ $T_c/J \simeq 2.266 \pm 0.002$

- Data collapse to estimate correlation length exponent ν
- Linear fit to estimate T_c, (exact value 2.26918…)
- 2d Ising on triangular lattice $T_c/J = 3.65 \pm 0.01$ (exact value 3.64095…)
- FCNN failed for Ising lattice gauge theory with Hamiltonian $H = -J \sum_p \prod_{i \in p} \sigma_i^z$ (CNN used)

J. Carrasquilla and R.G. Melko, «Machine learning phases of matter», Nature Physics 13 (2017) 431



- Convolutional Neural Network (CNN)

Variation of the output function - V(T) or VOT

Averaging over the testing dataset, we define the average prediction, F^T,

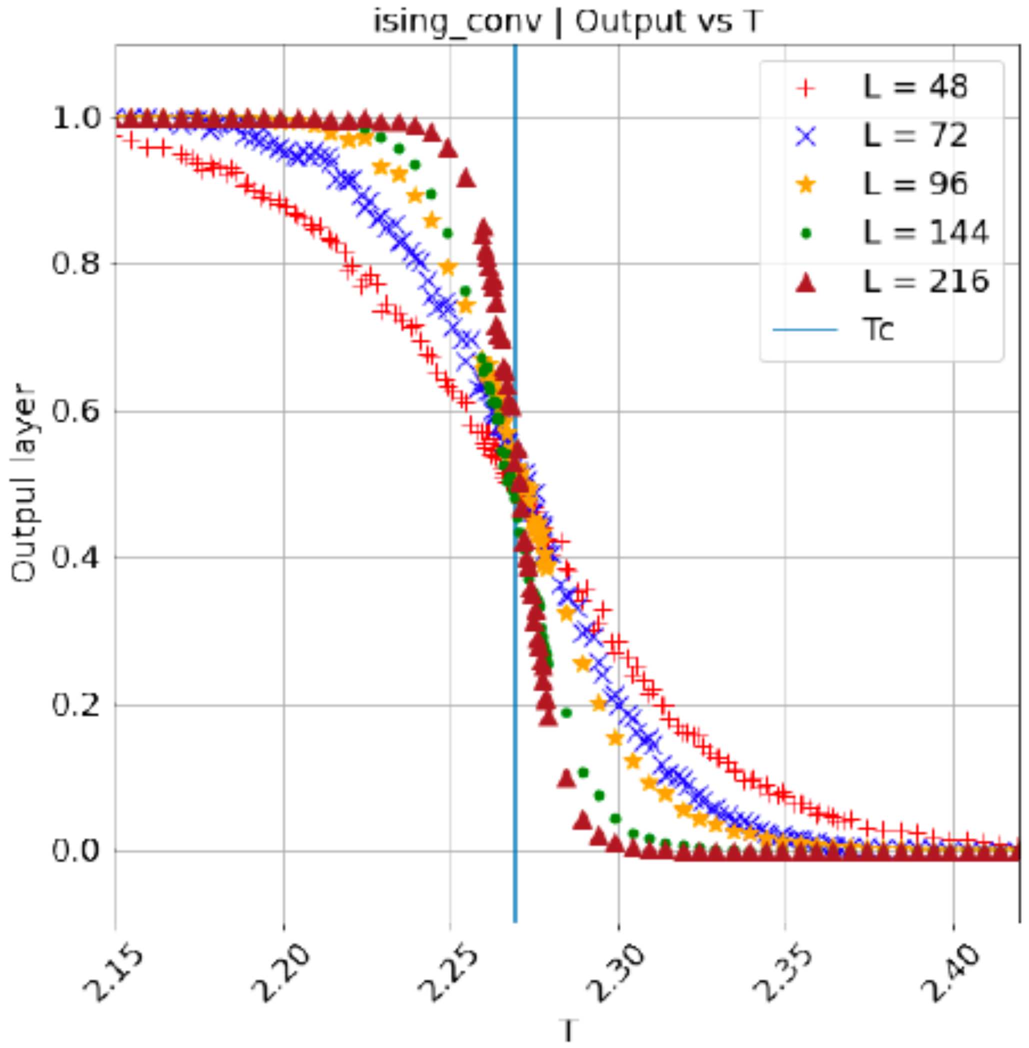
$$F^T = \frac{1}{N} \sum_{i=1}^{N} f_i^T \qquad (1)$$

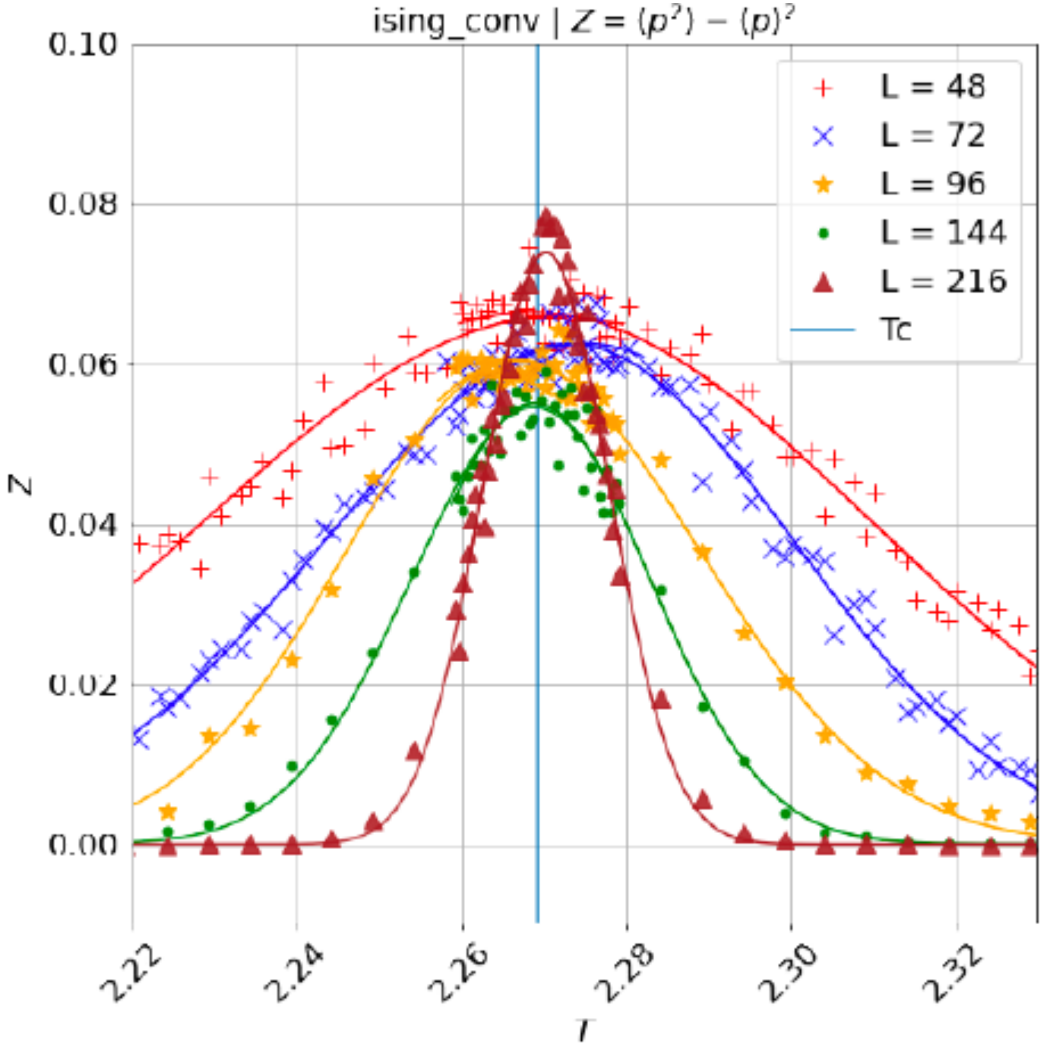
and its variance, V^T,

$$V^T = \frac{1}{N} \sum_{i=1}^{N} (f_i^T)^2 - \left(\frac{1}{N} \sum_{i=1}^{N} f_i^T \right)^2 . \qquad (2)$$

Chertenkov, Burovskii, and LS – PRE 108, L032102 (2023)

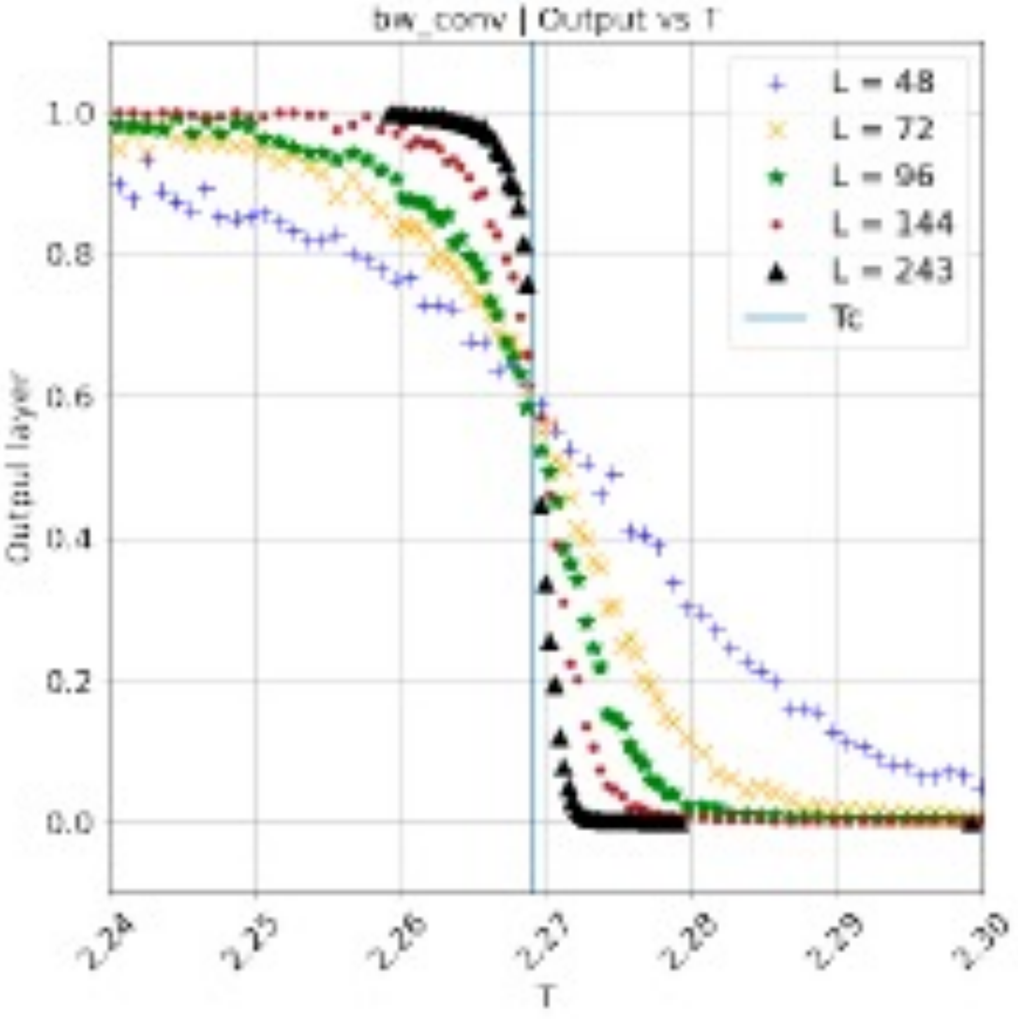
Output function and VOT – Ising model with CNN

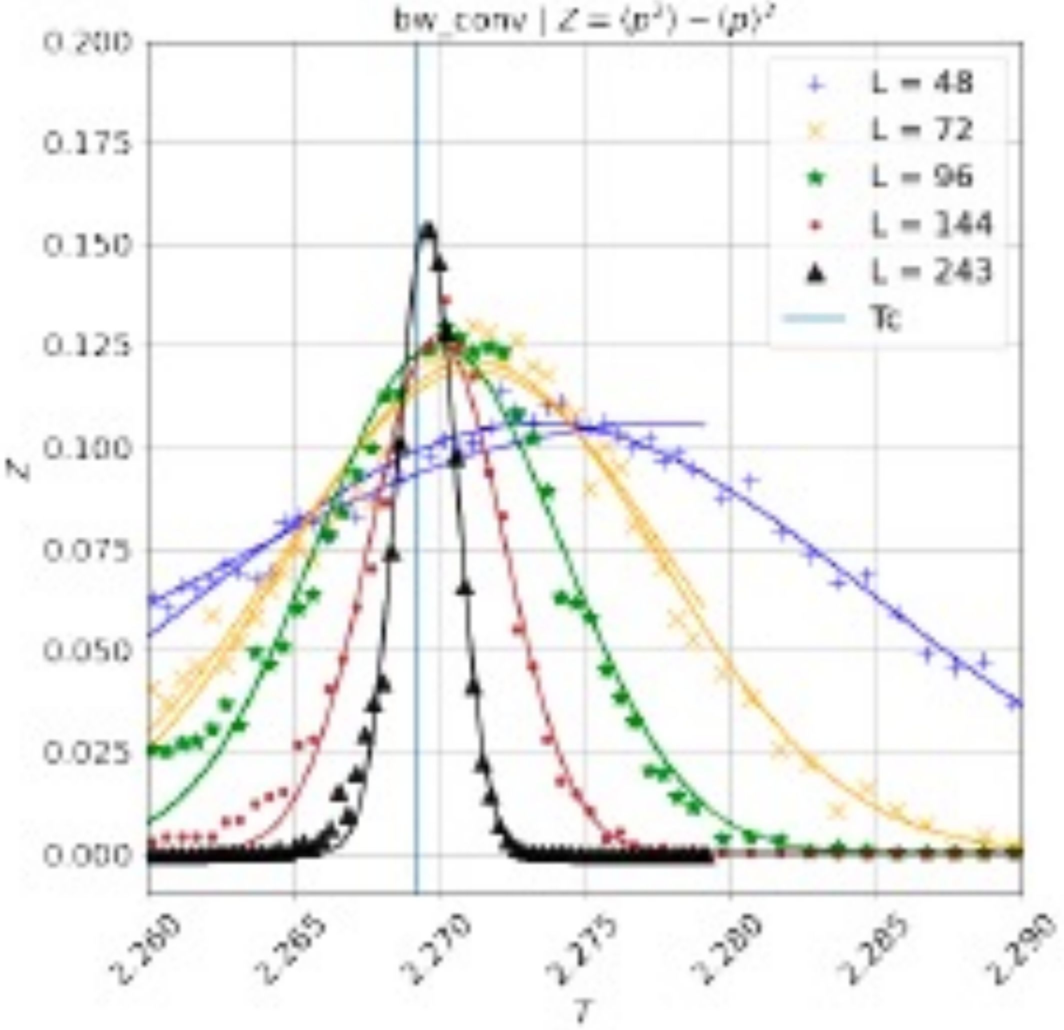




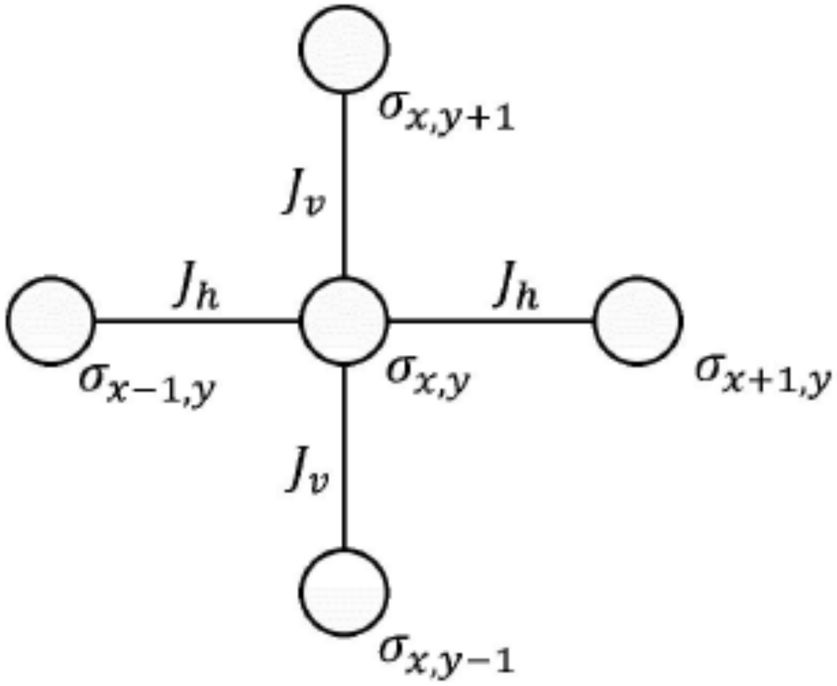
Chertenkov, Burovskii, and LS – PRE 108, L032102 (2023)

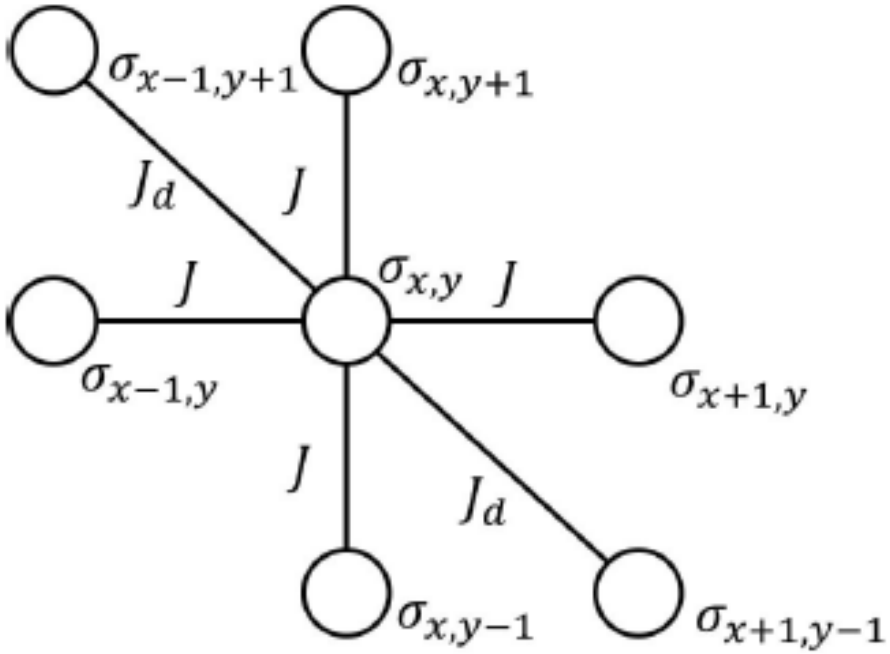
Baxter-Wu model with CNN (Intersection of outputs not at ½!))





Anisotropic Ising model on square lattice and anisotropic Ising model on triangular lattice





$$\mathcal{H} = -\sum_{(x,y)} \left(J_h \sigma_{x,y}\sigma_{x+1,y} + J_v \sigma_{x,y}\sigma_{x,y+1} \right)$$

$$\sinh\frac{2J_h}{k_B T_c}\,\sinh\frac{2J_v}{k_B T_c} = 1$$

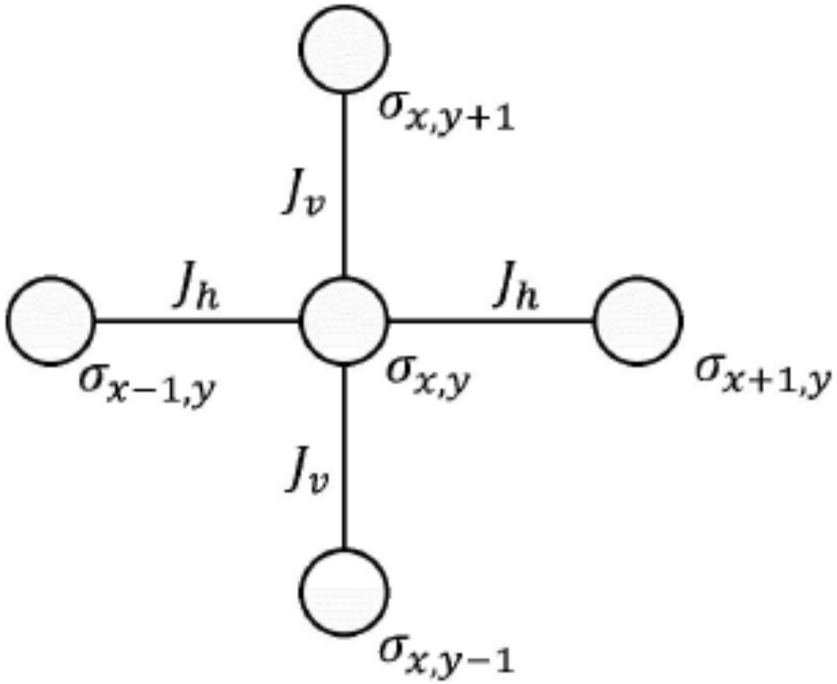
Onsager, 1941

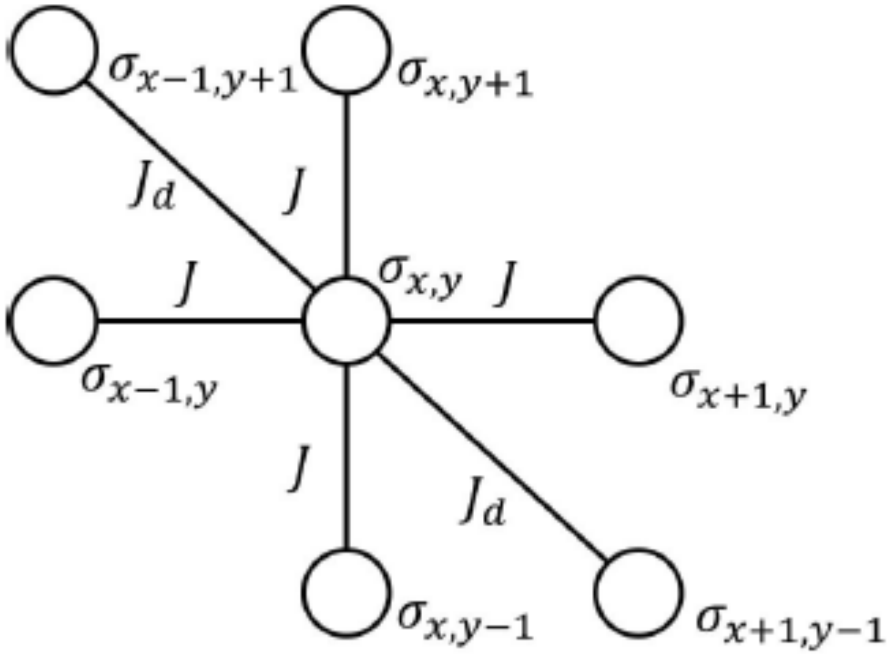
$$\mathcal{H} = -\sum_{(x,y)} \left(J\sigma_{x,y}(\sigma_{x+1,y} + \sigma_{x,y+1}) + J_d \sigma_{x,y}\sigma_{x+1,y+1} \right)$$

$$\left(\sinh\frac{2J}{k_B T_c}\right)^2 + 2\sinh\frac{2J}{k_B T_c}\,\sinh\frac{2J_d}{k_B T_c} = 1$$

Houtappel, 1950

Protocol of transfer learning





Training on samples with $J_h=J_v=J$
Testing with varying J_v with $J_h=J$
- Ortogonal anisotropy

Training on samples with $J_d=0$
Testing with varying J_d
- Diagonal anisotropy

Supervised ML : 1000 samples at each temperature for training and 500 samples for testing

Finite-size analysis of F(T) and V(T) – critical temperature and correlation length exponent ν

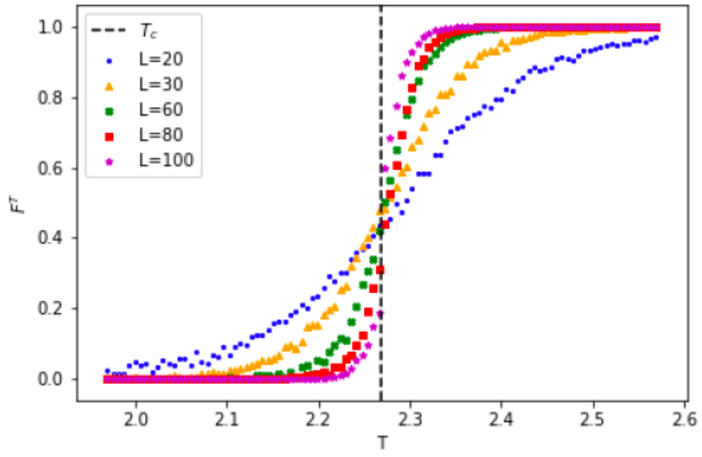
1. After testing: plot V(T;L) and approximate with Gaussian
2. Estimate maximum position $T^*(L)$
3. Estimate critical temperature $T^* = T^*(L) + \frac{A}{L^b}$

Ferdinand and Fisher, 1969
b=1/ν

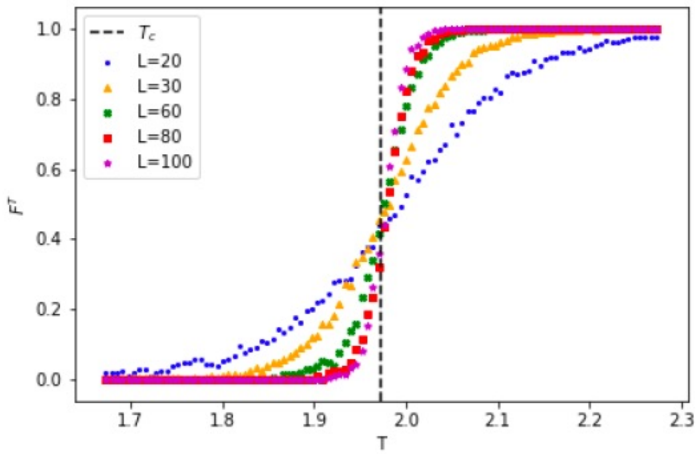
$$F(T;L) = \frac{1}{N_t}\sum_{i=1}^{N_t} f_i(T;L) \qquad V(T;L) = \sqrt{\frac{1}{N_t}\sum_{i=1}^{N_t}\left(f_i(T;L)\right)^2 - \left(\frac{1}{N_t}\sum_{i=1}^{N_t} f_i(T;L)\right)^2}$$

Ising model with ortogonal anisotropy – ouput function

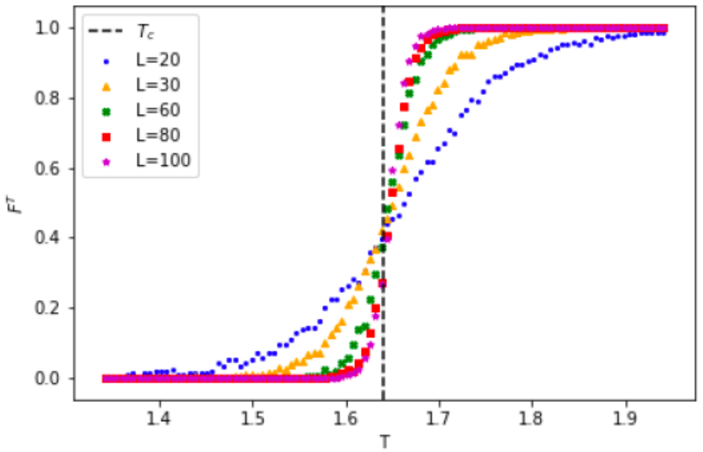
$$\sinh\frac{2J_h}{k_BT_c}\ \sinh\frac{2J_v}{k_BT_c}=1$$



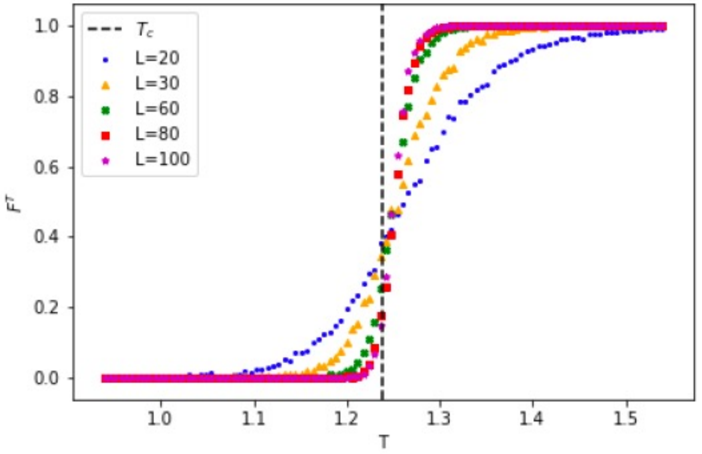
1



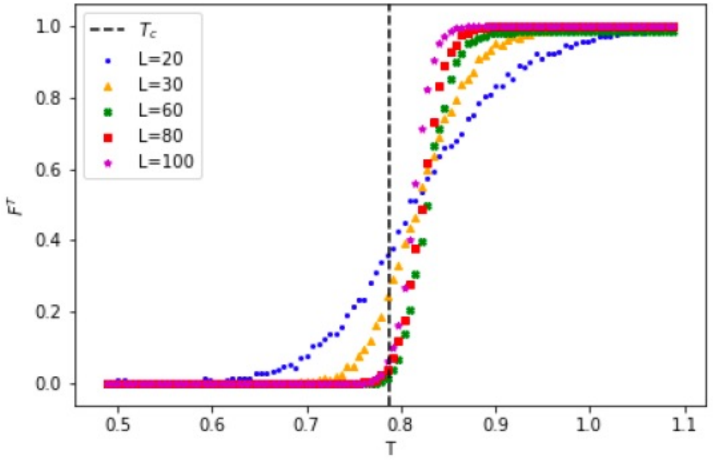
0.5



0.25



0.125



0.0625

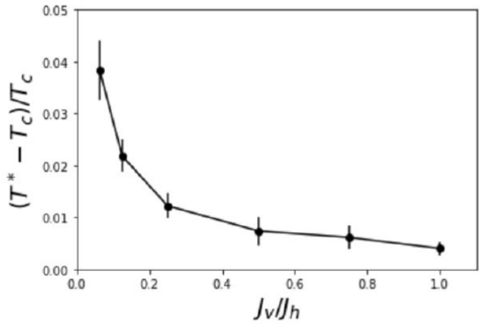


Fig. 2. Deviation of the critical temperature prediction T^* from T_c as function of the coupling ratio J_v/J_h. T^* estimated using method of output functions intersection [1].

Ising model with ortogonal anisotropy – ouput function P(T) analysis a la Carrasquilla and Melko

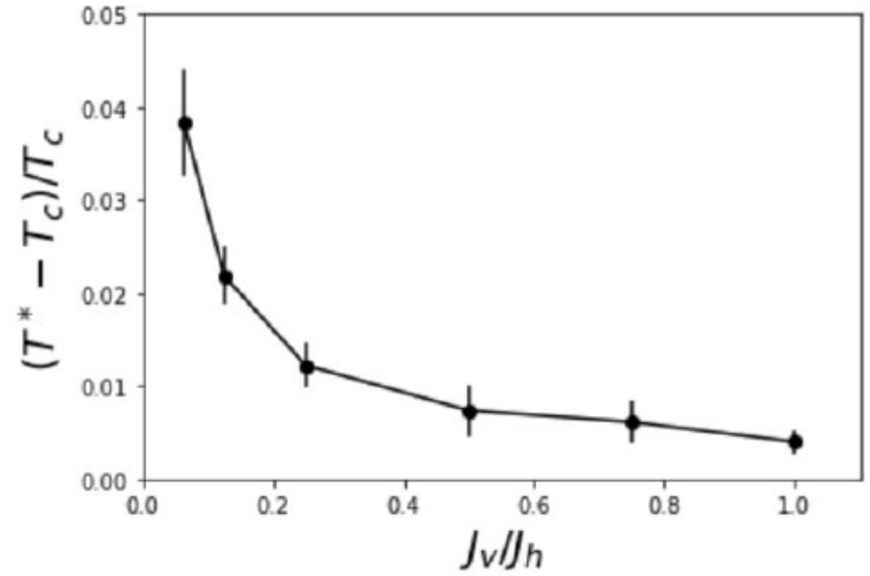
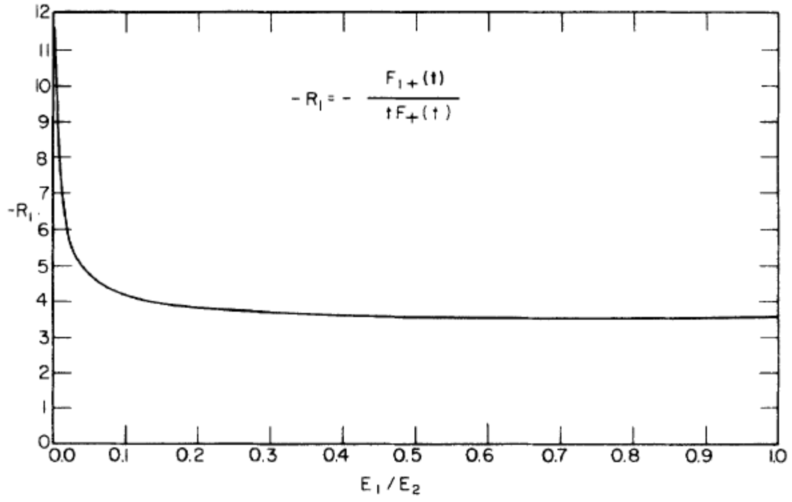


Fig. 2. Deviation of the critical temperature prediction T^* from T_c as function of the coupling ratio J_v/J_h. T^* estimated using method of output functions intersection [1].



Wu, MCoy, Tracy and Barouch, 1976

spin-spin correlation function was analytically calculated for the Ising model on the square lattice in the thermodynamic limit

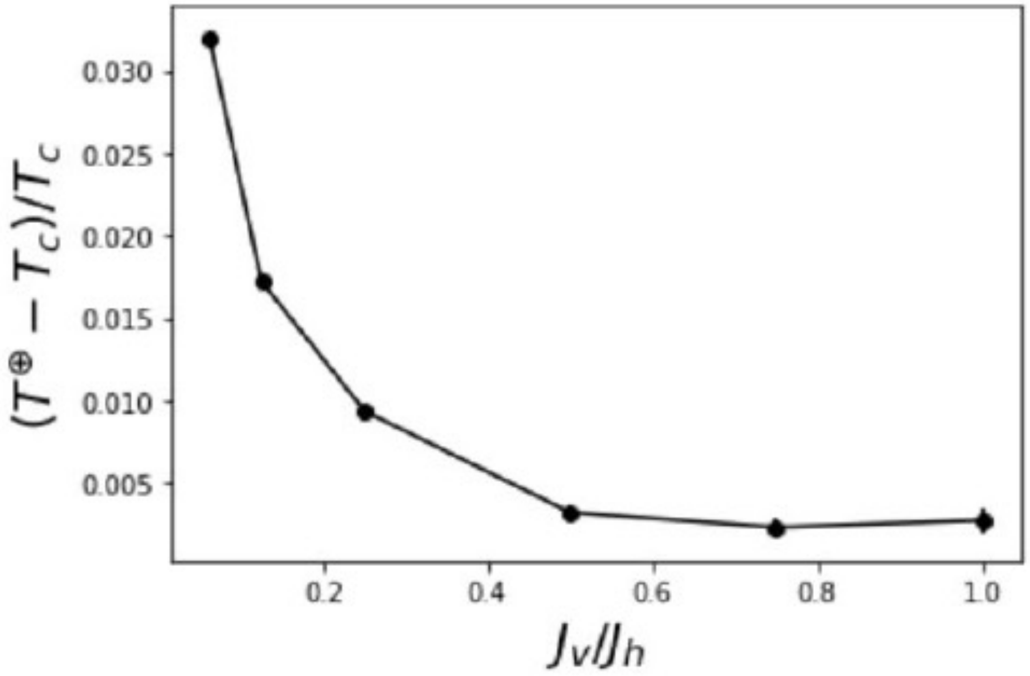
$$C\left(\sigma(0)\sigma(R)\right) = \frac{F_\pm}{R^{1/4}} + O(1/R^{5/4}), \qquad (11)$$

where R is the distance between any two spins $\sigma(0)$ and $\sigma(R)$, and F_+ and F_- are the amplitudes in the FM and PM phases, respectively. The figure 1 of the paper [26] shows dependence of the ratio of the amplitudes F_+/F_- which is similar to those deviation shown in the Fig. 2.

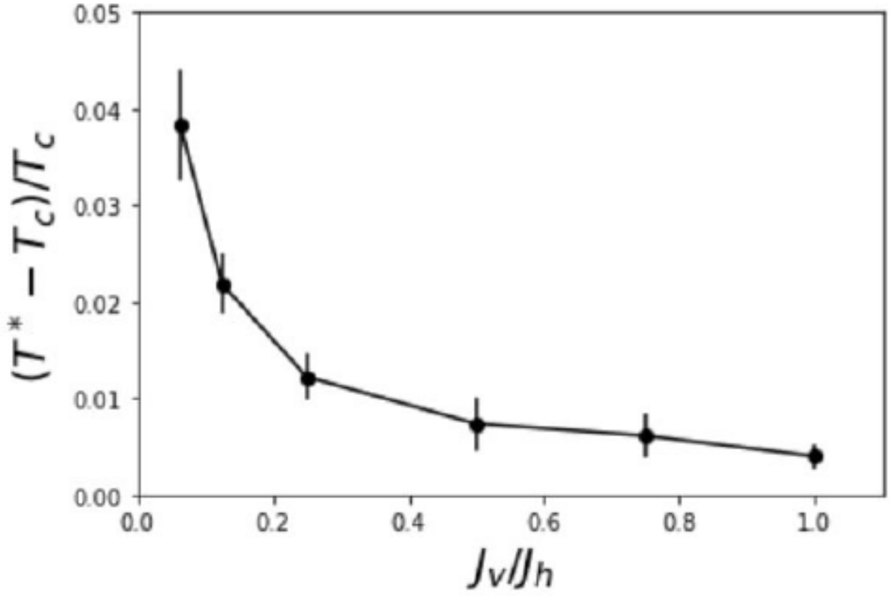
Ising model with ortogonal anisotropy – variation of ouput function V(T) analysis

CNN and Ising model

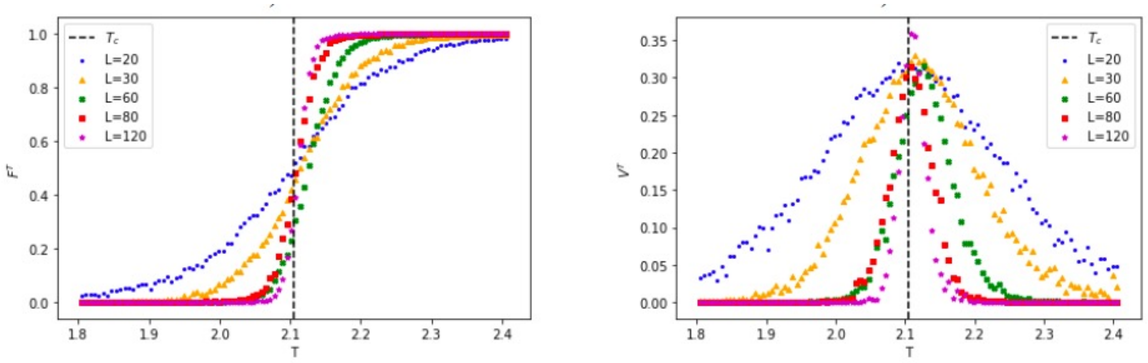
V(T) analysis



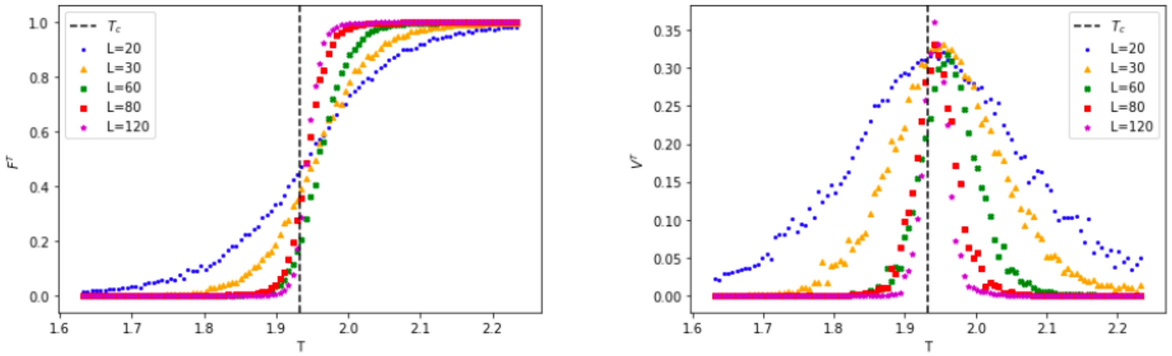
P(T) analysis



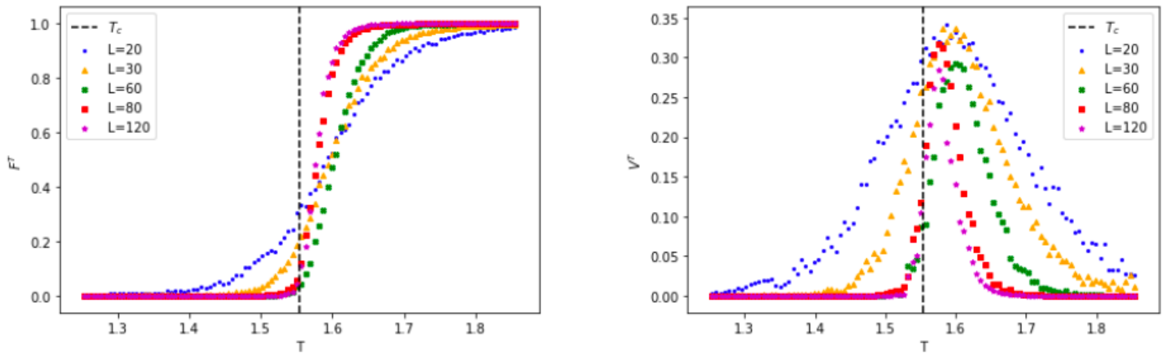
Ising model with diagonal anisotropy – ouput function P(T) and variation V(T)



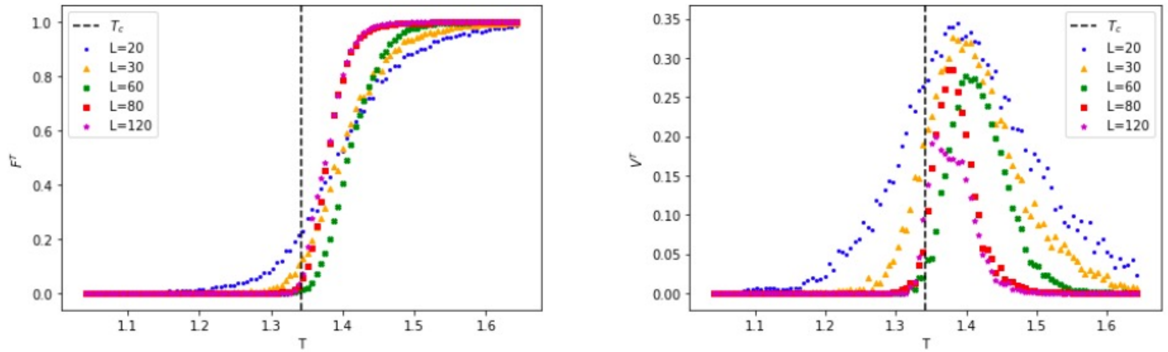
- 0.1



- 0.2

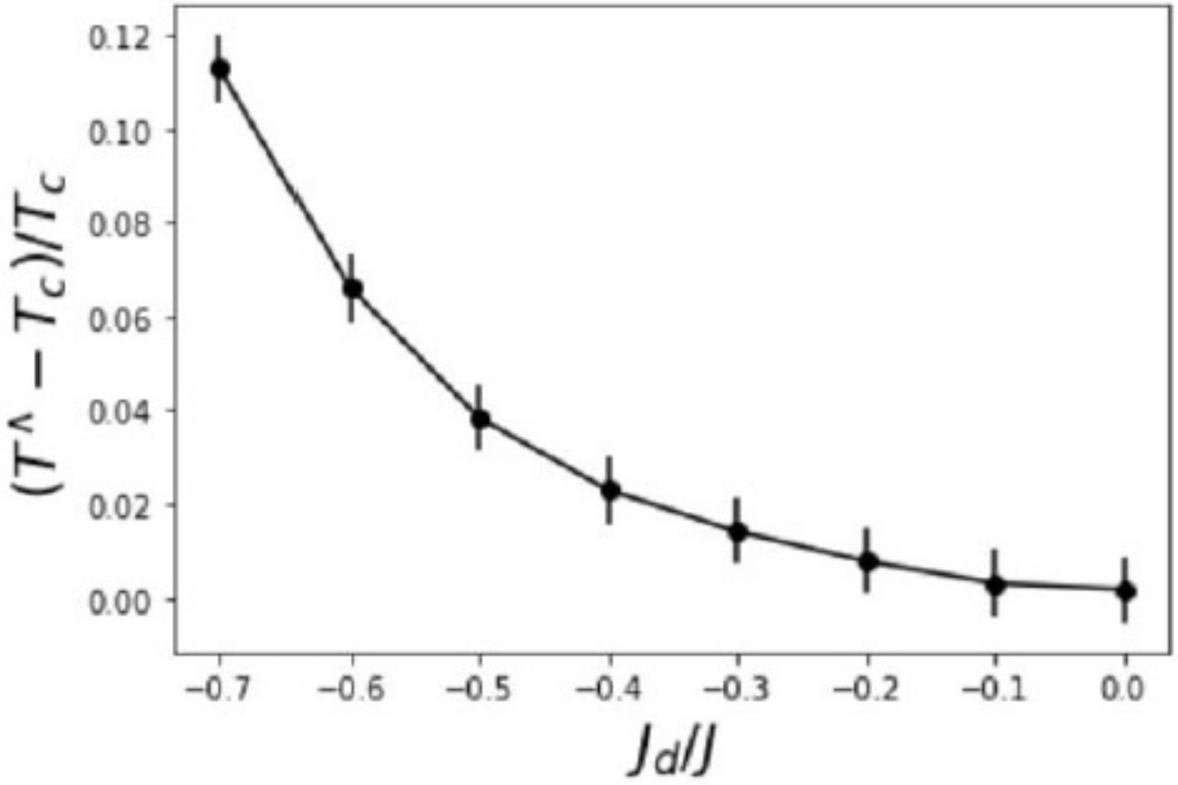


- 0.4



- 0.5

Ising model with diagonal anisotropy – variation of ouput function V(T) analysis

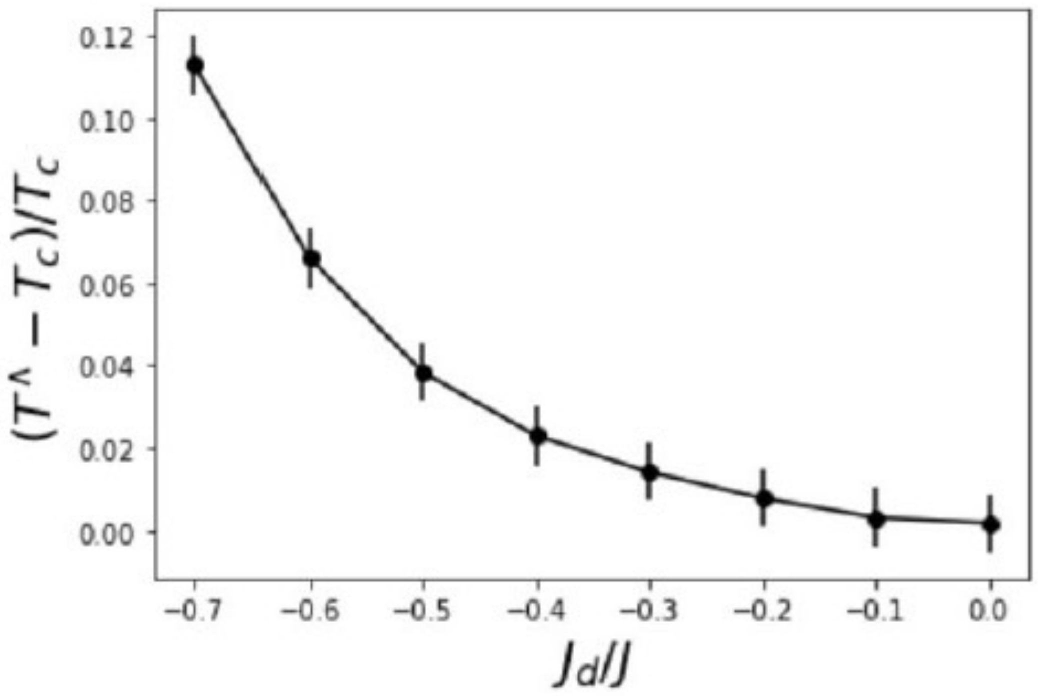


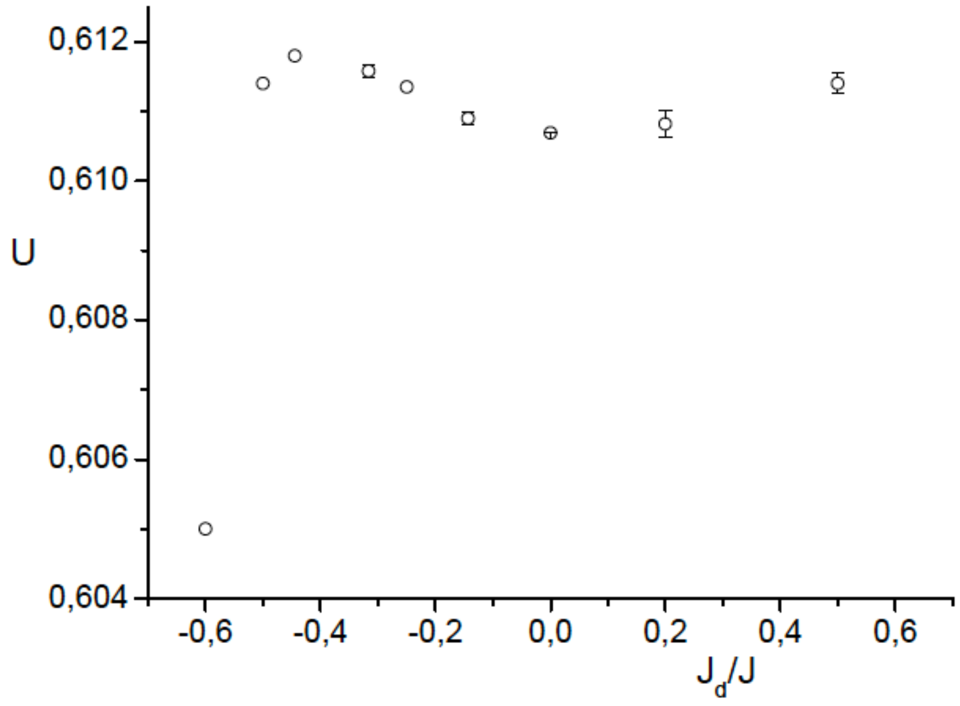
$1/\nu$

J_v/J_h	$b^{\oplus}$
1.0	1.12(3)
0.75	1.09(3)
0.5	1.07(4)
0.25	1.06(6)
0.125	0.98(14)
0.0625	1.02(9)

Sukhoverhova, Burovskii, and LS – Springer Proceedings in Physics, accepted

Ising model with diagonal anisotropy – variation of ouput function V(T) analysis





Binder cumulant, Selke and LS, 2009

Conclusion

1. The main result of the research is that the widely used technique for extraction of the critical temperature directly from the dependence of the output function is not universal. The value of the output function at the critical temperature depends on the anisotropy of the model under study, the architecture of the deep network, and some parameters of the deep network application.
2. A more reliable approach is to analyze the variation of the output function.
3. The negative result is that NN predicts the critical temperature of an anisotropic models with a visible displacement.
4. The positive result of the research is that neural networks trained on an isotropic model predict well the class of universality of anisotropic models.
5. Transfer-learning is possible within the same universality class.
6. Number of iterations (epochs) of education process matters.

Supported by Russian Science Foundation grant 22-21-00259

Chertenkov and LS – unpublished

Transfer-learning between universality classes

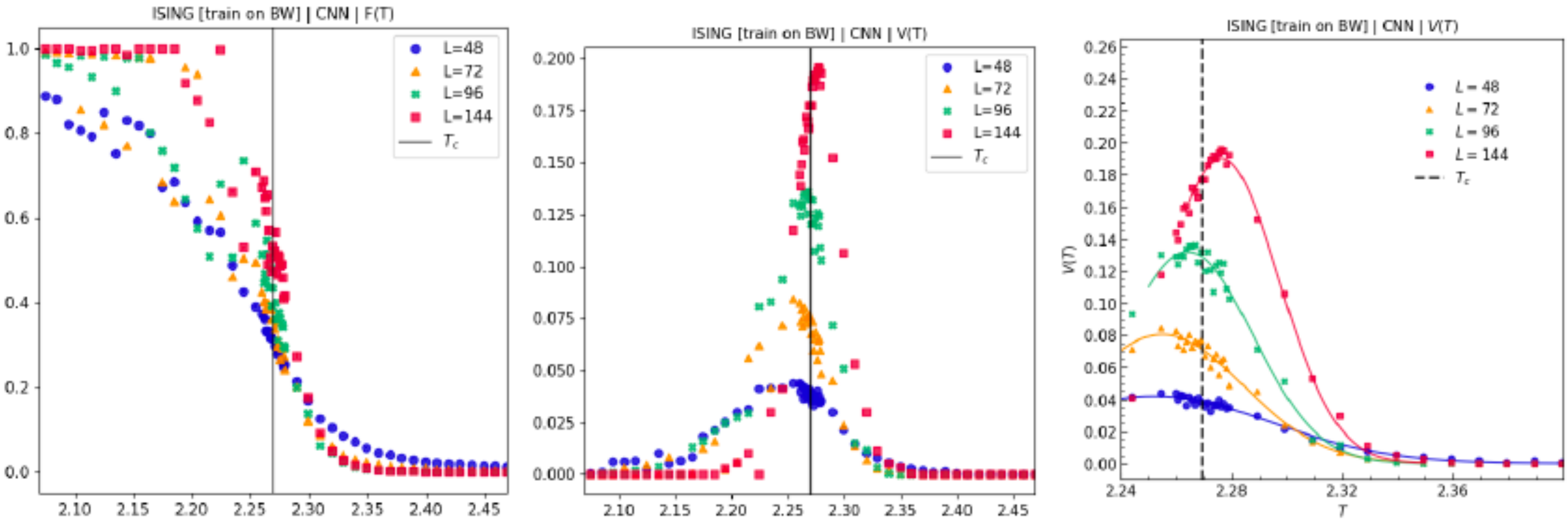


Рис. 5: Функции FT, VT (полная и правая часть) для модели Изинга и архитектуры CNN, обученной на конфигурациях модели Бакстера-Ву.

Dependence on the number of epochs

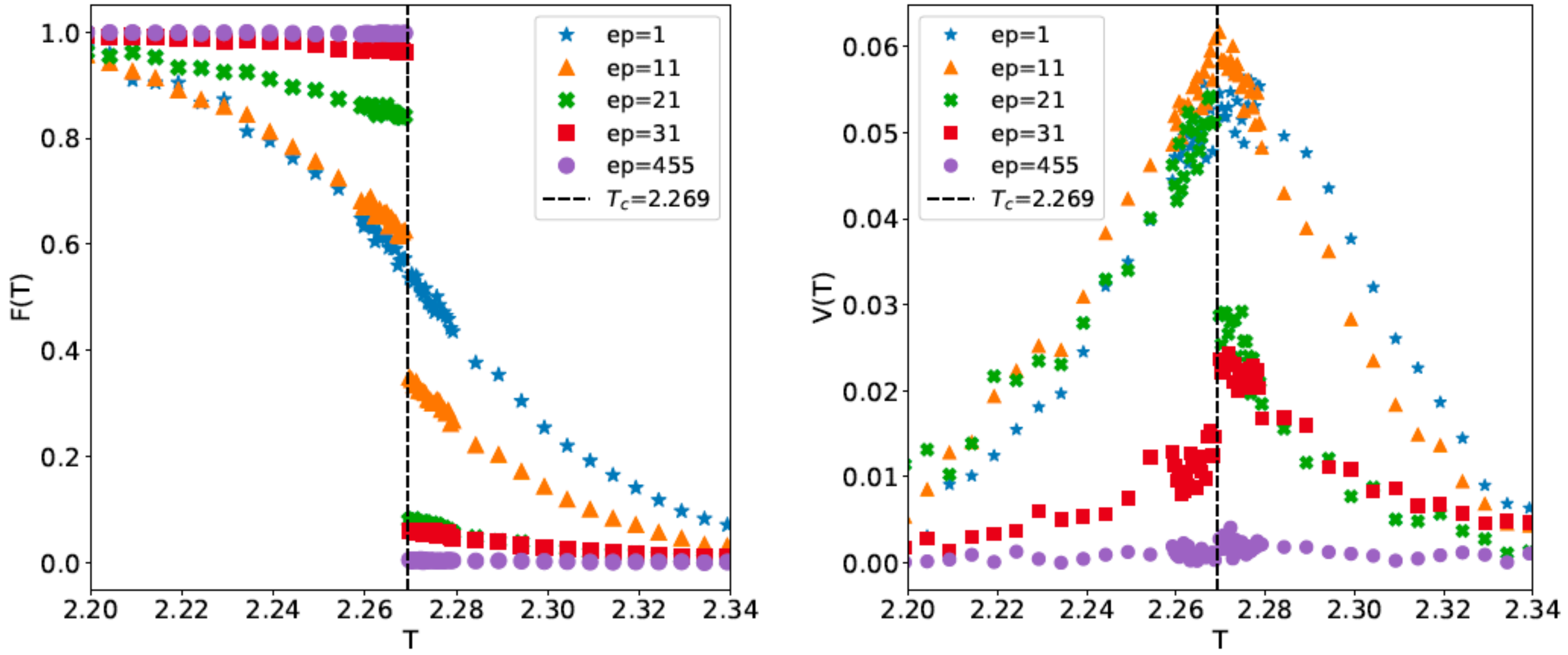


Fig. 6. (left) $F(T)$, the average of the output function, and the variance of output function, $V(T)$, (right) for different numbers of epochs 1, 11, 21, 31, 455. Dashed black vertical line indicates the critical temperature T_c.

Pseudo-critical behaviour

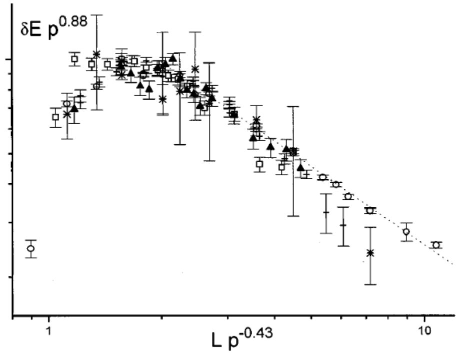
PHYSICAL REVIEW E VOLUME 55, NUMBER 5 MAY 1997

Cluster Monte Carlo: Scaling of systematic errors in the two-dimensional Ising model

Lev N. Shchur[1] and Henk W. J. Blöte[2]

$x_n = x_{n-p} \oplus x_{n-q}$ Shift-register RNG

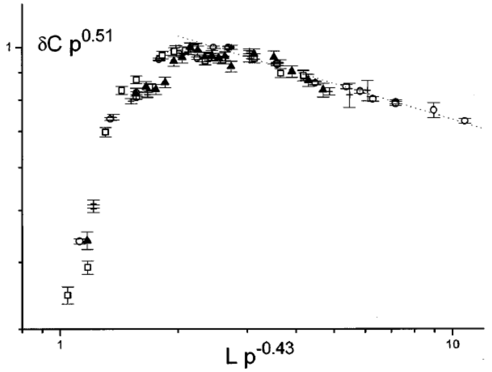
Wolff cluster algorithm – cluster size at Tc = magn. susceptibility



“Resonance” of two length

$$p \propto \chi \propto L_{\max}^{\gamma/\nu}$$

Leads to scaling iof systematic errors!



$$\delta E \lesssim 0.3 L^{-0.84} p^{-0.52}.$$

$$-\delta C \lesssim 0.85 L^{-0.21} p^{-0.42}$$

Leads to scaling iof systematic errors!

Exponents are specific, but! - the width of their distribution scales according with Ferdinand-Fisher, with the inverse correlation length exponent!

3D Ising scaling of errors $-\delta C \approx L^{-0.24}\, p^{-0.19}$ $\delta Q \approx L^{-1.15}\, p^{0.09}$