Machine learning for lattice field theory and back

Gert Aarts

Swansea University Prifysgol Abertawe

FONDAZIONE BRUNO KESSLER

ECT* EUROPEAN CENTRE FOR THEORETICAL STUDIES IN NUCLEAR PHYSICS AND RELATED AREAS

Sejong University, Seoul, Nov 2022

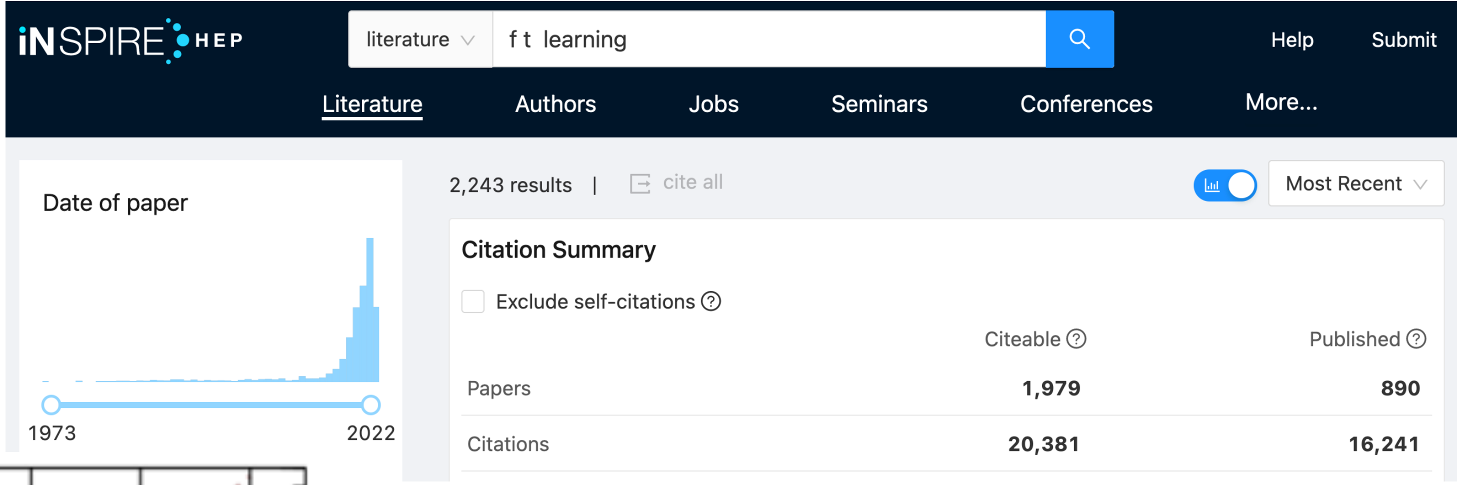
Introduction

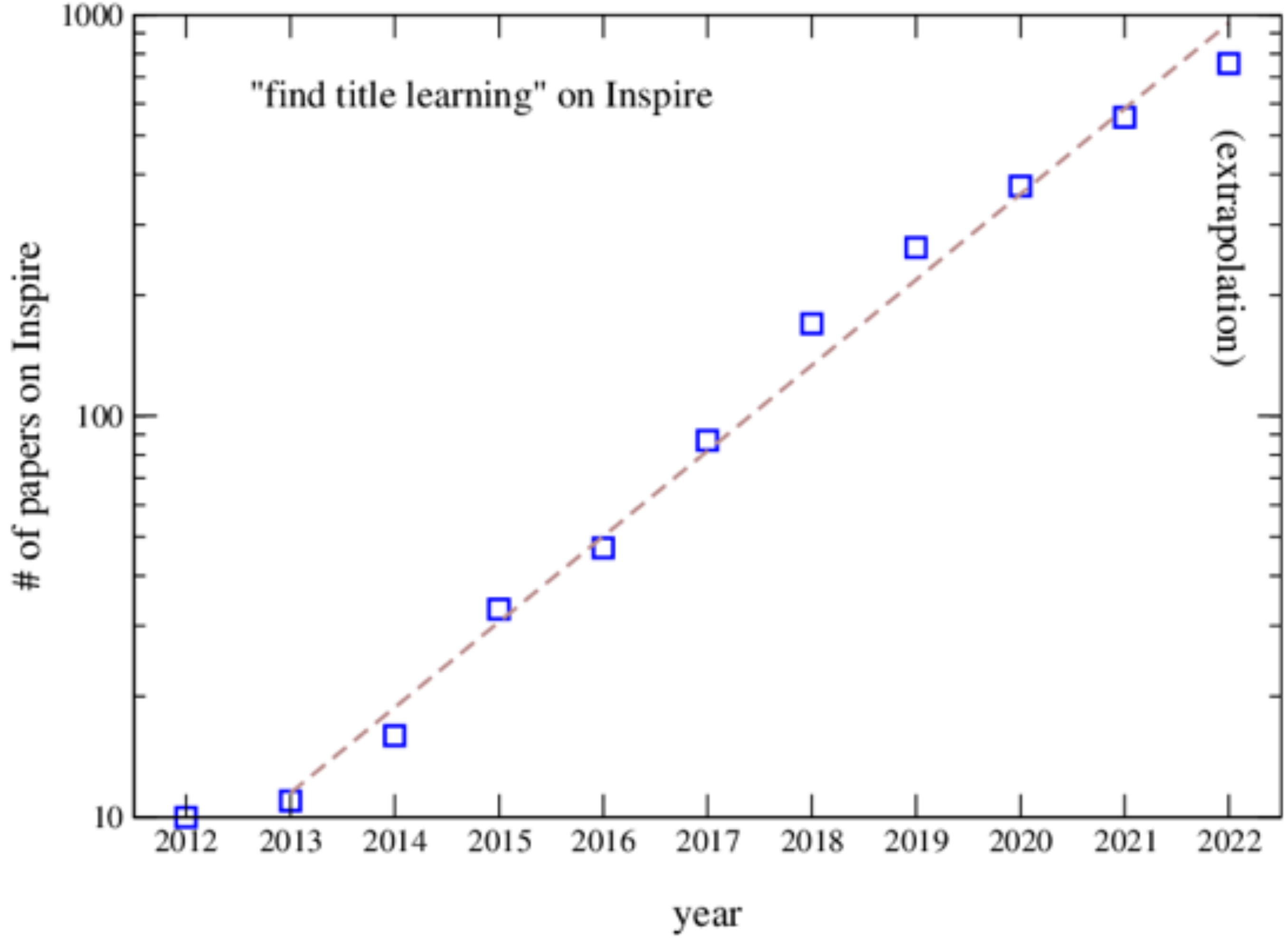
- past five years or so has seen a rapid rise of applications of machine learning (ML) in fundamental science, particle physics, theoretical physics

- of course ML has been around for quite some time, especially in experimental particle physics

- nevertheless, there is an **exponential** increase in activity

Introduction





- find title **learning** on the Inspire data base (high-energy physics)
- exponential growth!

Outline

a flavour of ideas and results in theoretical statistical physics
biased towards own work and interests in lattice field theory

- classification: order-disorder transition (by now classic application)
- transfer learning (widely used notion)
- inverse renormalisation group (new application and concepts)
- outlook

fruitful collaboration with Dimitrios Bachtis and Biagio Lucini

Based on the following papers:

- ✓ Extending machine learning classification capabilities with histogram reweighting
  Phys. Rev. E 102 (2020) 033303 [2004.14341 [cond-mat.stat-mech]]
- ✓ Mapping distinct phase transitions to a neural network
  Phys. Rev. E 102 (2020) 053306 [2007.00355 [cond-mat.stat-mech]]
- ✓ Adding machine learning within Hamiltonians: Renormalization group transformations, symmetry breaking and restoration
  Phys. Rev. Res. 3 (2021) 013134 [2010.00054 [hep-lat]]
- ✓ Quantum field-theoretic machine learning
  Phys. Rev. D 103 (2021) 074510 [2102.09449 [hep-lat]]
- ✓ Inverse renormalisation group in quantum field theory
  Phys. Rev. Lett. 128 (2022) 081603 [2107.00466 [hep-lat]]
  with Dimitrios Bachtis and Biagio Lucini and Francesco di Renzo

Classification of phases of matter

Published: 13 February 2017

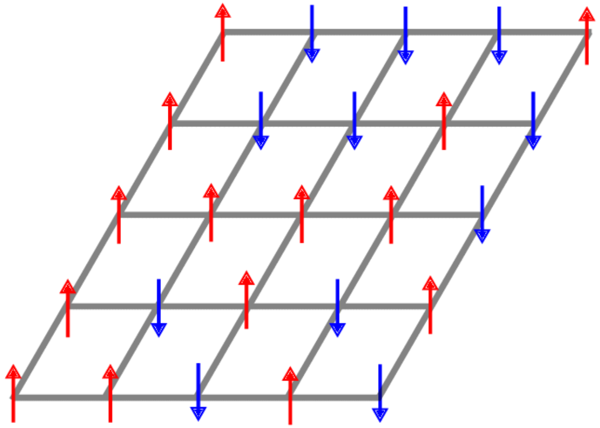
Machine learning phases of matter

Juan Carrasquilla & Roger G. Melko

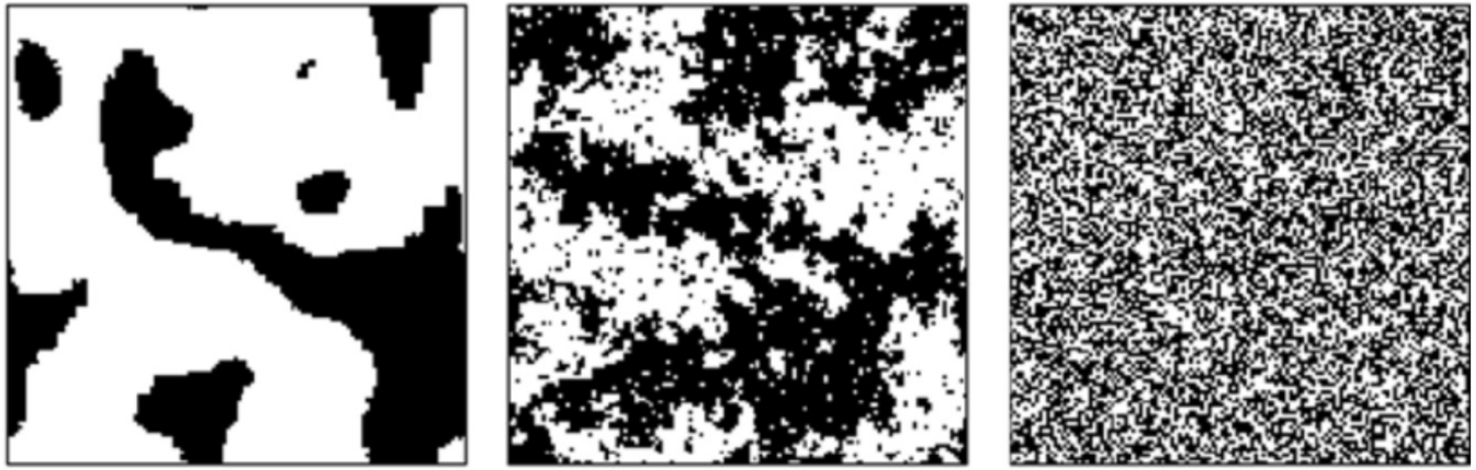
Nature Physics **13**, 431–434 (2017) | Cite this article

> 1100 citations since 2017

Classification of phases of matter



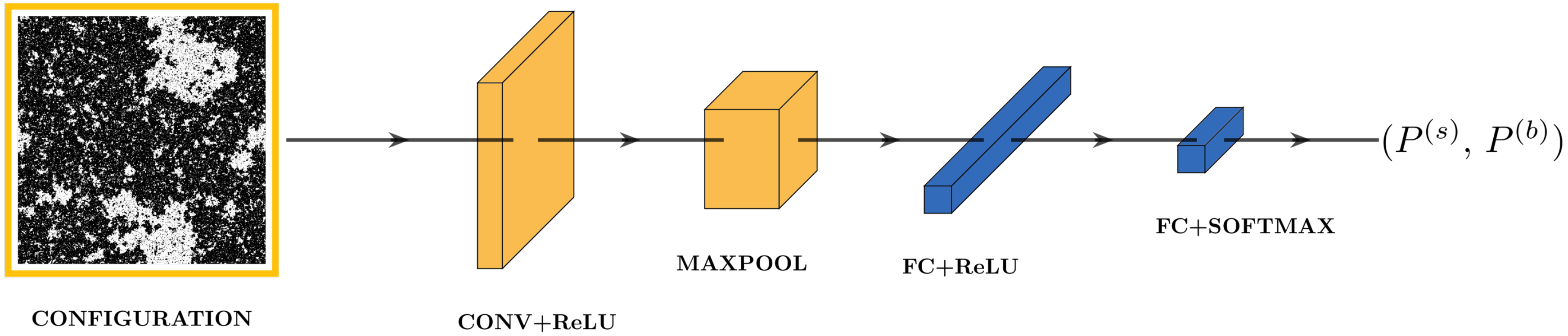
- matter can exist in different phases
- prototype: 2d Ising model -> ordered/disordered or cold/hot phases
- task: determine phase a configuration is in, determine critical coupling or temperature



Ordered -- ? -- Disordered

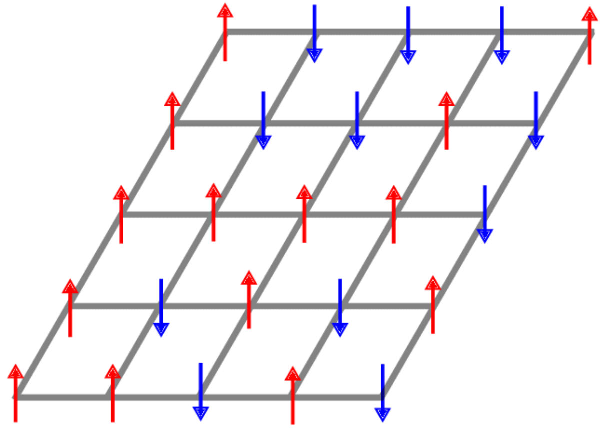
ML excels in pattern finding

- supervised learning problem: use sets of configurations deep in the ordered and in the disordered phase
- input: configurations < -- > output: ordered/disordered
- "train the ML algorithm", i.e. adjust parameters in the neural network so that it reproduces the correct classification for the training set



- new, unseen configurations -- > determine **probability** to be (dis)ordered

2d Ising model



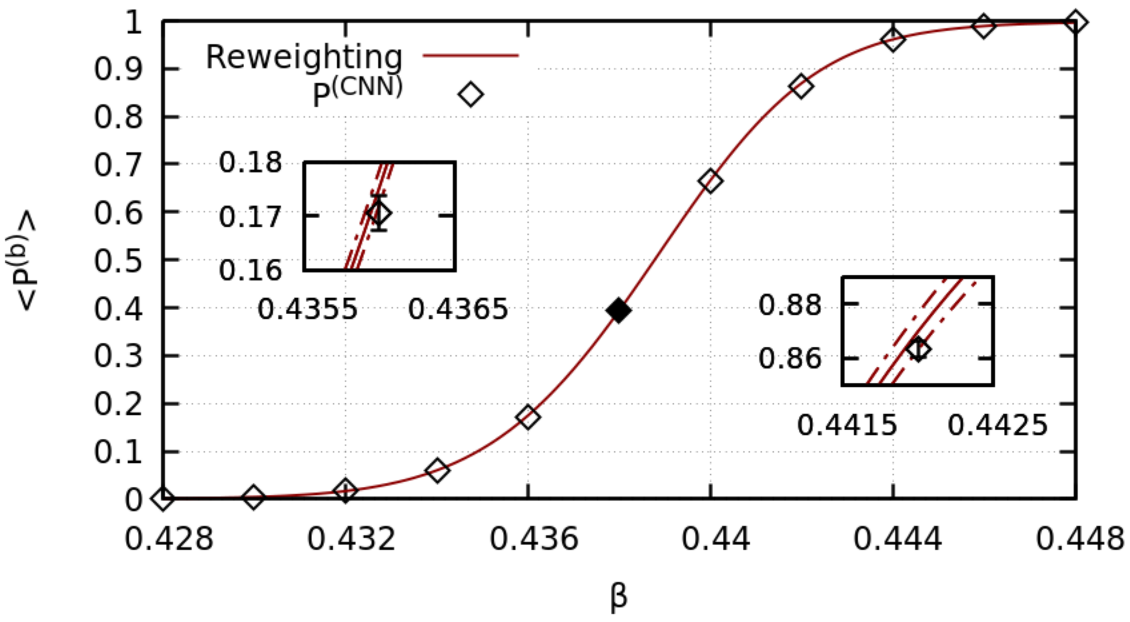
- $Z = \mathrm{Tr}\, e^{-\beta E}$ with $E = -\sum_{<ij>} s_i s_j$ $\quad (s_i = \pm 1)$
- critical coupling or inverse temperature β_c
- correlation length ξ, magnetic susceptibility χ diverge at transition
- critical exponents $\quad \xi \sim |t|^{-\nu} \quad \chi \sim |t|^{-\gamma}$ $\quad$ reduced temperature $t = \dfrac{\beta_c - \beta}{\beta_c}$
- $\nu = 1, \quad \gamma/\nu = 7/4, \quad \beta_c = \frac{1}{2}\ln(1+\sqrt{2}) \approx 0.440687$
- finite-size scaling $\quad |t| = \left|\dfrac{\beta_c - \beta_c(L)}{\beta_c}\right| \sim \xi^{-\frac{1}{\nu}} \sim L^{-\frac{1}{\nu}} \qquad \chi \sim L^{\gamma/\nu}$

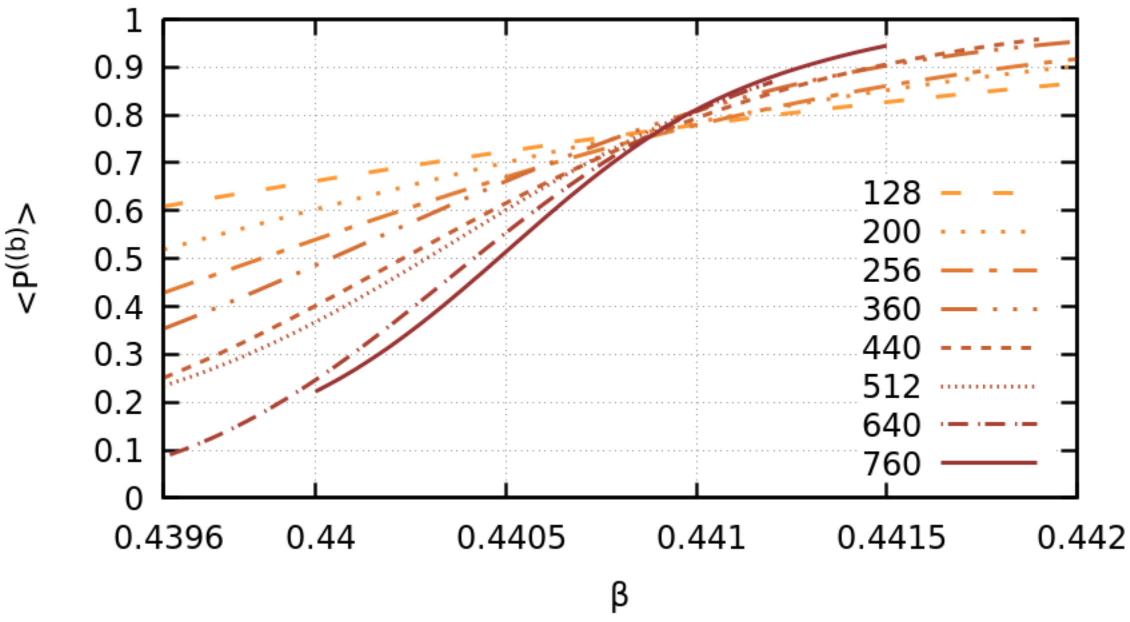
$\beta_c \sim 0.440687$

Probability to be in ordered phase

train NN away from the phase transition: $\beta \leq 0.41$ and $\beta \geq 0.47$

investigate unseen configurations at intermediate β on lattices of different sizes

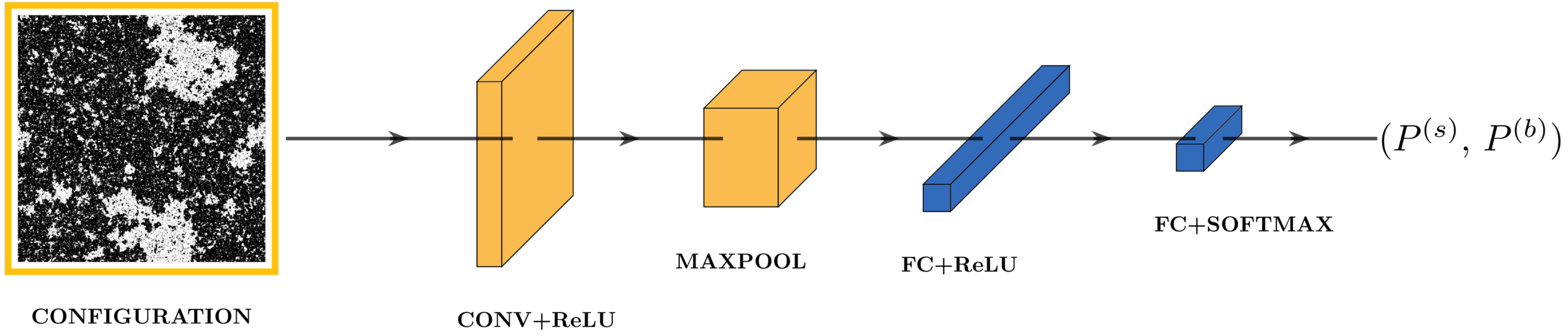




probability behaves as an approximate order parameter

What's new?

- by now well-established procedure, what can we add?
- interpret output from a NN as an observable in a statistical system
- input: configurations, distributed according to Boltzmann weight
- output: observable, "order parameter" in statistical system



$$\Rightarrow \quad < P > = \frac{1}{Z} \sum_i P_i \, e^{-\beta E_i}$$

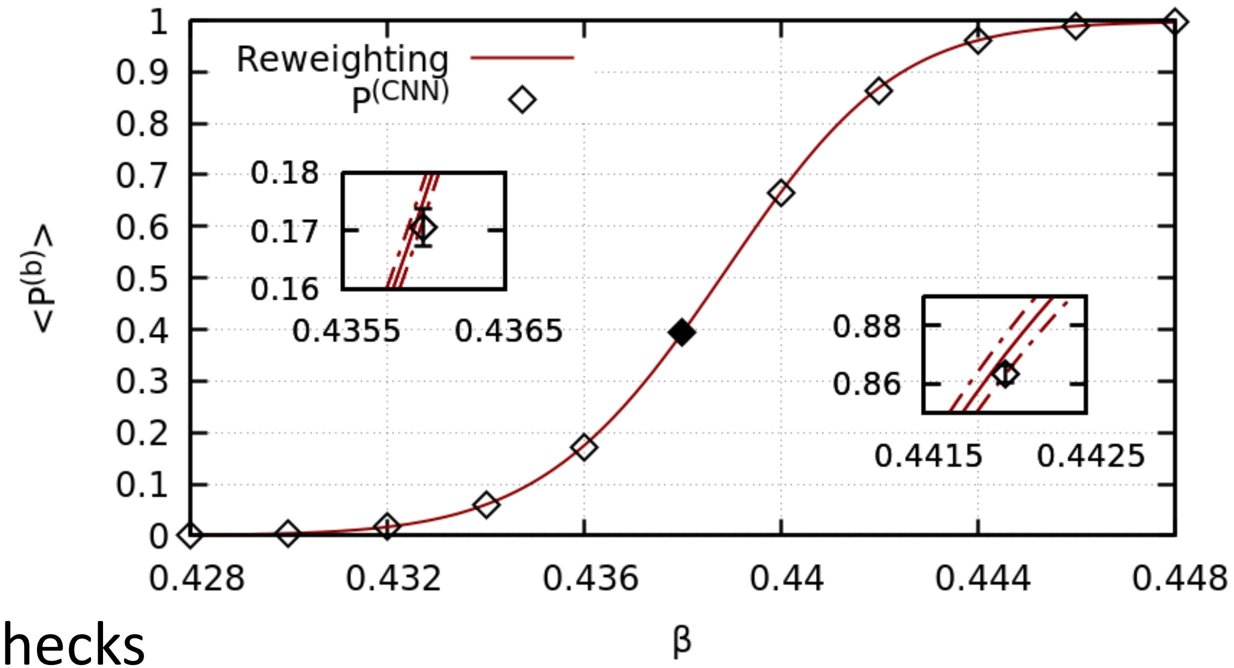
Output of NN as physical observable

- opens up possibility to use “standard” numerical/statistical methods

→ histogram reweighting: extrapolation to other parameter values

- starting from computation at given β_0: extrapolate to other β values

$$< P > (\beta) = \frac{\sum P_i e^{-(\beta-\beta_0)E_i}}{\sum e^{-(\beta-\beta_0)E_i}}$$



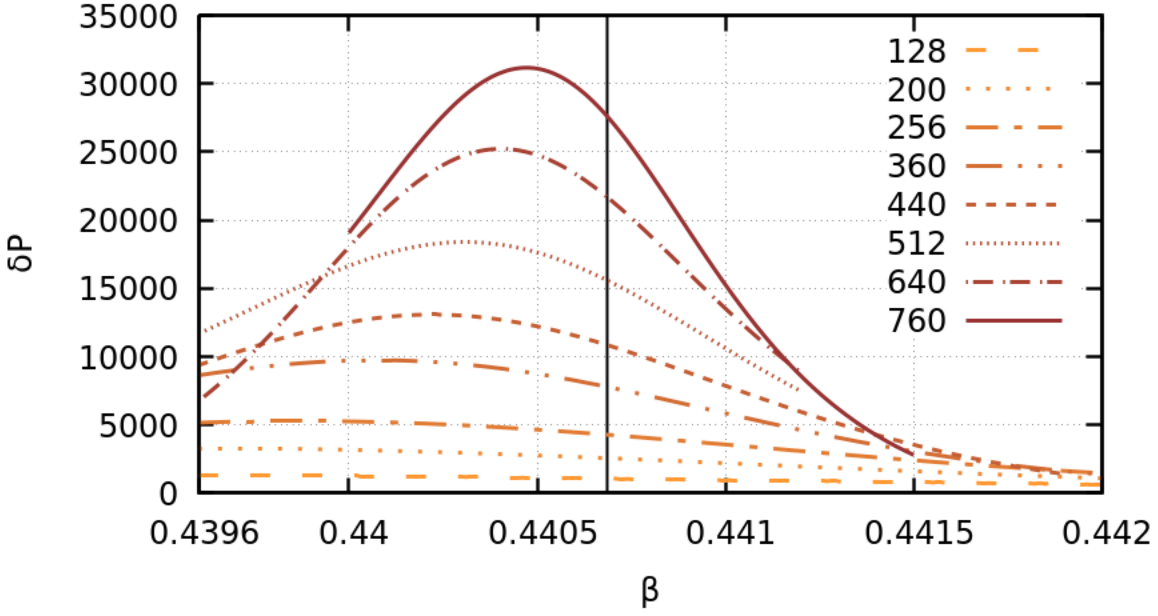
✓ filled diamond at β_0

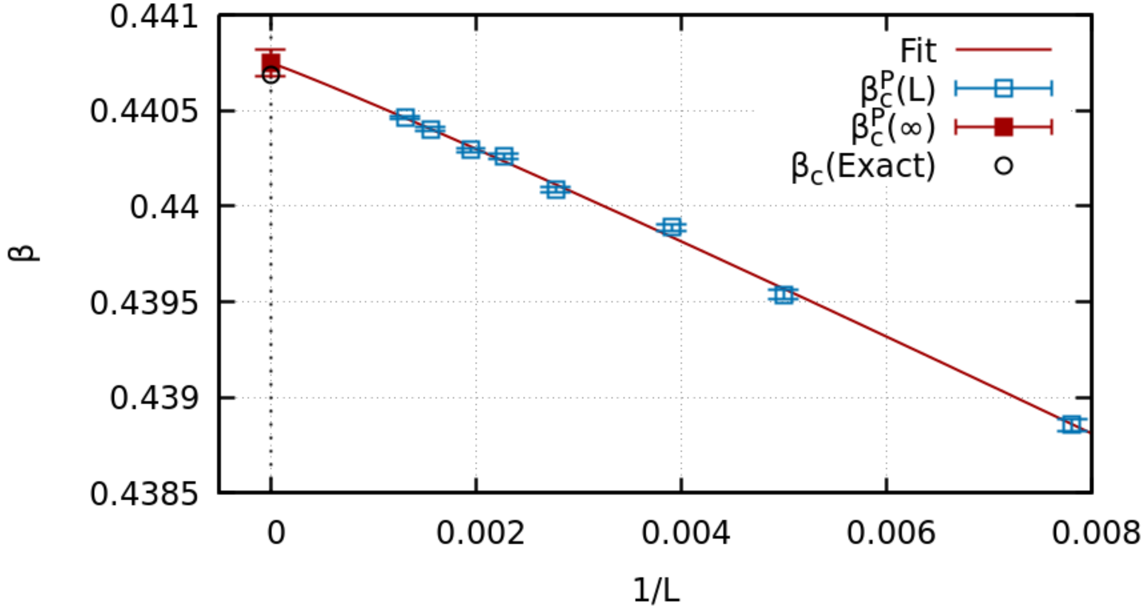
✓ line obtained by reweighting in β

✓ open diamonds are independent cross checks

Critical behaviour from NN observables

- Determine L dependent susceptibility δP and its maximum at $\beta_c(L)$





Extract critical properties from NN observables only

	β_c	ν	γ/ν
CNN+Reweighting	0.440749(68)	0.95(9)	1.78(4)
Exact	$\ln(1+\sqrt{2})/2$ ≈ 0.440687	1	7/4 =1.75

Intermediate summary

- ✓ proposed to identify NN outputs as observables in statistical physics
- ✓ introduced histogram reweighting to employ in supervised machine learning
- ✓ critical properties obtained from a finite-size scaling analysis using quantities derived from NN alone (no need for explicit order parameter, knowledge of symmetries)
- ✓ quantitative studies of phase transitions based on a synergistic relation between machine learning and statistical mechanics

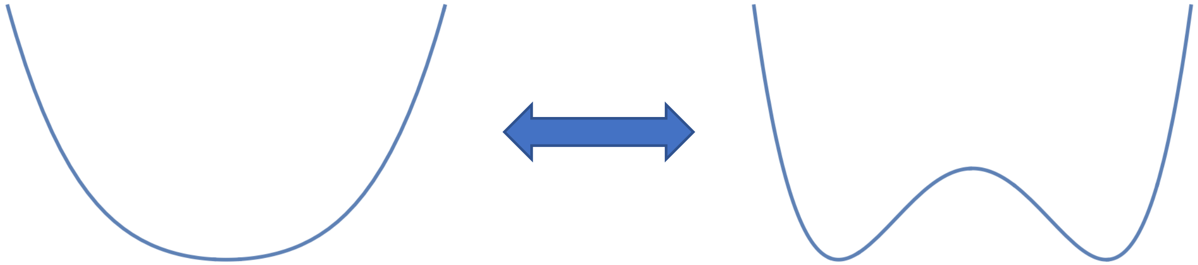
Transfer learning with histogram reweighting

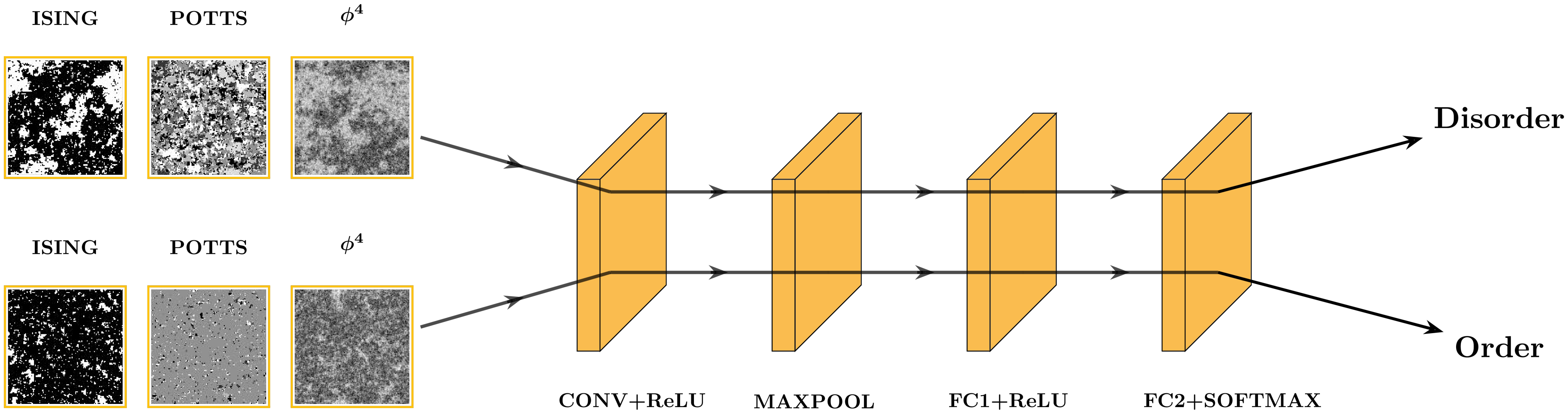
- NN has learned patterns, or *features*, in 2d Ising model
- are these sufficiently universal to predict the structure of phase transitions in other systems?
- what about universality class, order of transition, type of degrees of freedom?

→ Transfer learning

- apply to q-state Potts model (with $q = 3, \ldots, 7$), φ^4 scalar field theory

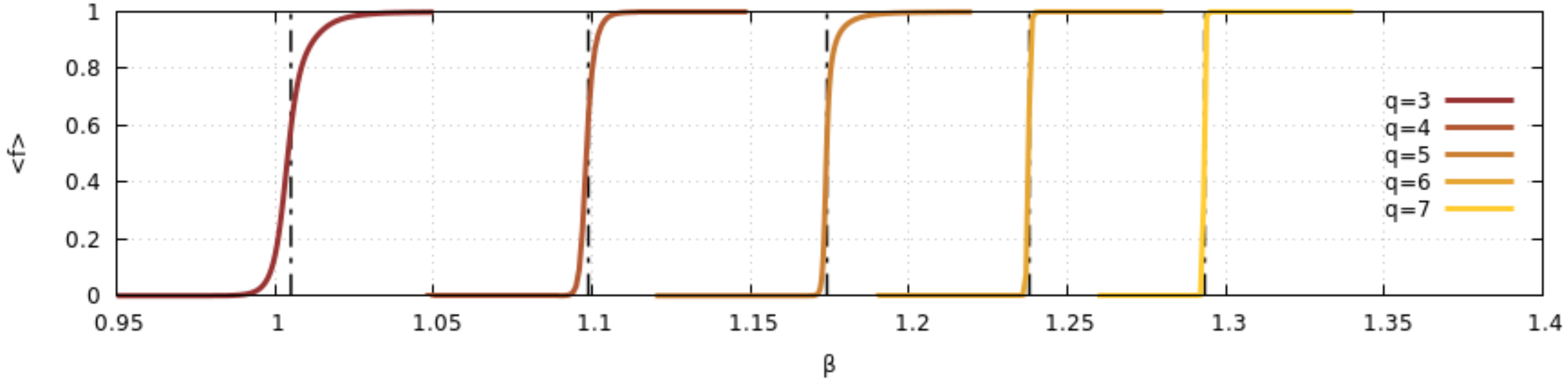
Transfer learning





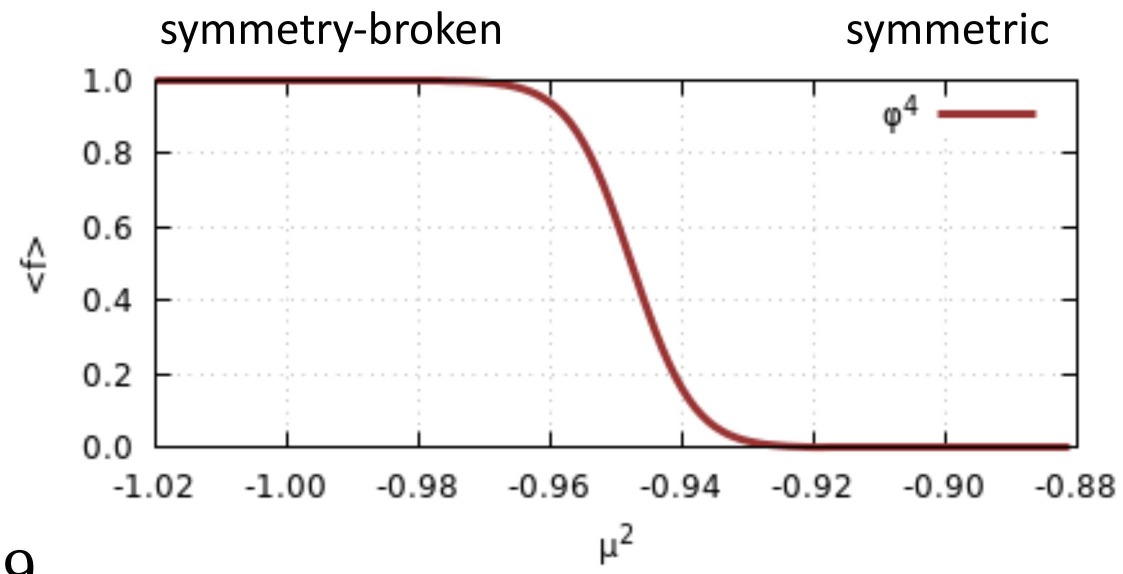
✓ Mapping distinct phase transitions to a neural network
Phys. Rev. E 102 (2020) 053306 [2007.00355 [cond-mat.stat-mech]]

Transfer learning: q-state Potts model

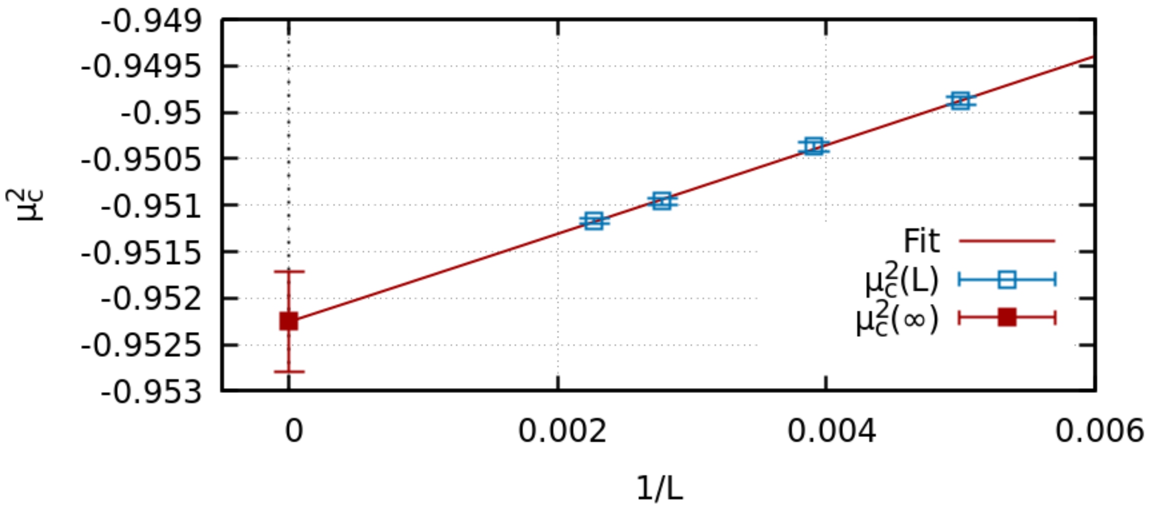


- training on Ising model, not Potts model
- continuous lines using histogram reweighting
- vertical dashed lines indicate expected transition at $\beta_c = \ln(1 + \sqrt{q})$
- $q = 3, 4$: 2nd order transition, $q = 5, 6, 7$: 1st order transition

φ^4 scalar field theory



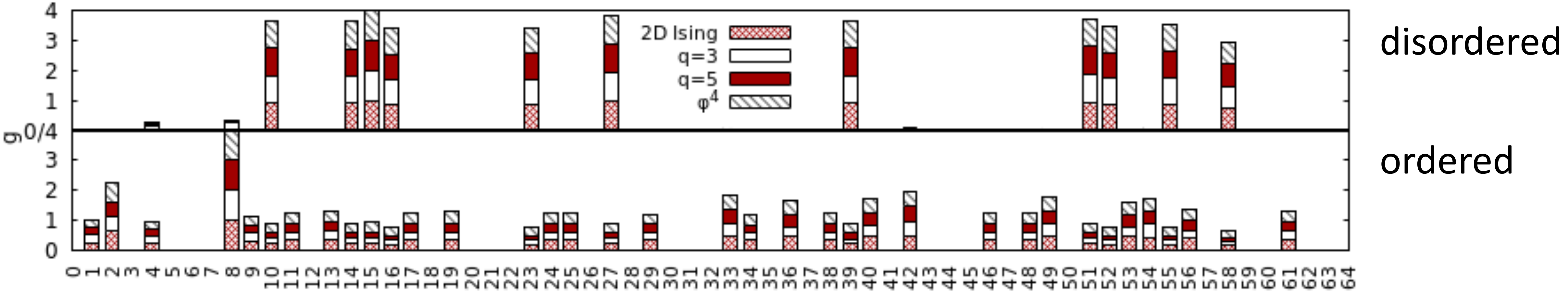
- reweight in mass parameter, μ^2
- identify regions where phase is clear
- retrain NN using $\mu^2 < -1.0$ and $\mu^2 > -0.9$
- repeat finite-size scaling analysis as in 2d Ising model



	μ_c^2	ν	γ/ν
CNN+Reweighting	-0.95225(54)	0.99(34)	1.78(7)

- same universality class as 2d Ising model
- critical mass in agreement with results obtained with standard methods (Binder cumulant, susceptibility)

Under the hood: activation functions in NN



mean activation functions in the 64 neurons in the fully connected (FC1) layer of 2d Ising-trained neural network, for:

- 2d Ising model
- $q = 3$ and $q = 5$ Potts model
- φ^4 scalar field theory

universal features distinguish ordered and disordered phases, irrespective of e.g. order of transition

Intermediate summary: transfer learning

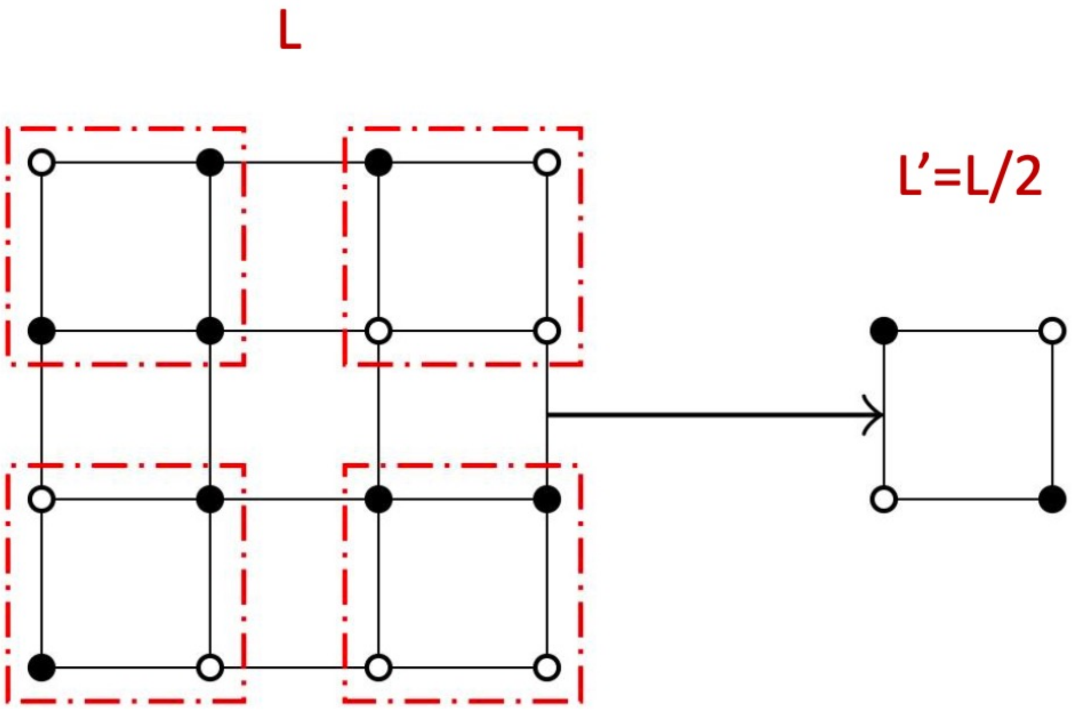
- ✓ enables the use of simplistic systems to study complicated models with partially known behaviour
- ✓ combine with reweighting to scan parameter space and reconstruct effective order parameters
- ✓ locate (unknown) phase transitions
- ✓ given this knowledge, train a new NN on configurations of the target system labelled according to previous step
- ✓ study infinite-volume limit in this new NN to make accurate predictions

Renormalisation Group (RG)

- standard renormalisation group: coarse-graining,
  blocking transformation, integrating out degrees of freedom, …

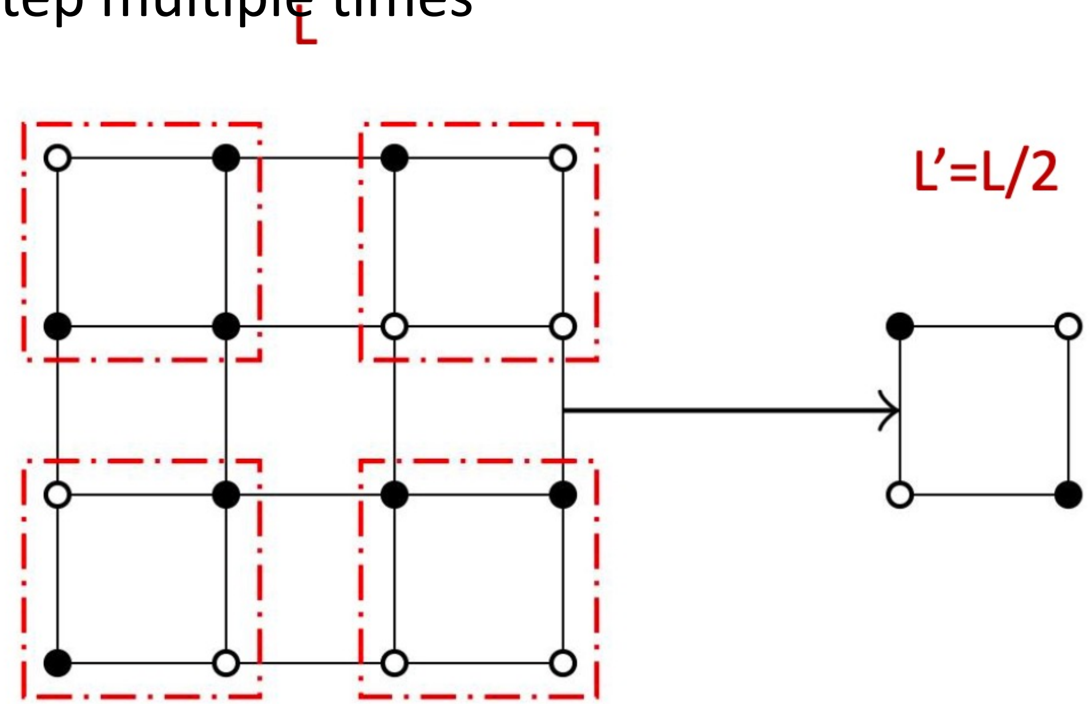
- Ising model: Kadanoff block spin
- majority rule
- reduction of degrees of freedom
- study critical scaling

- not invertible: semi-group



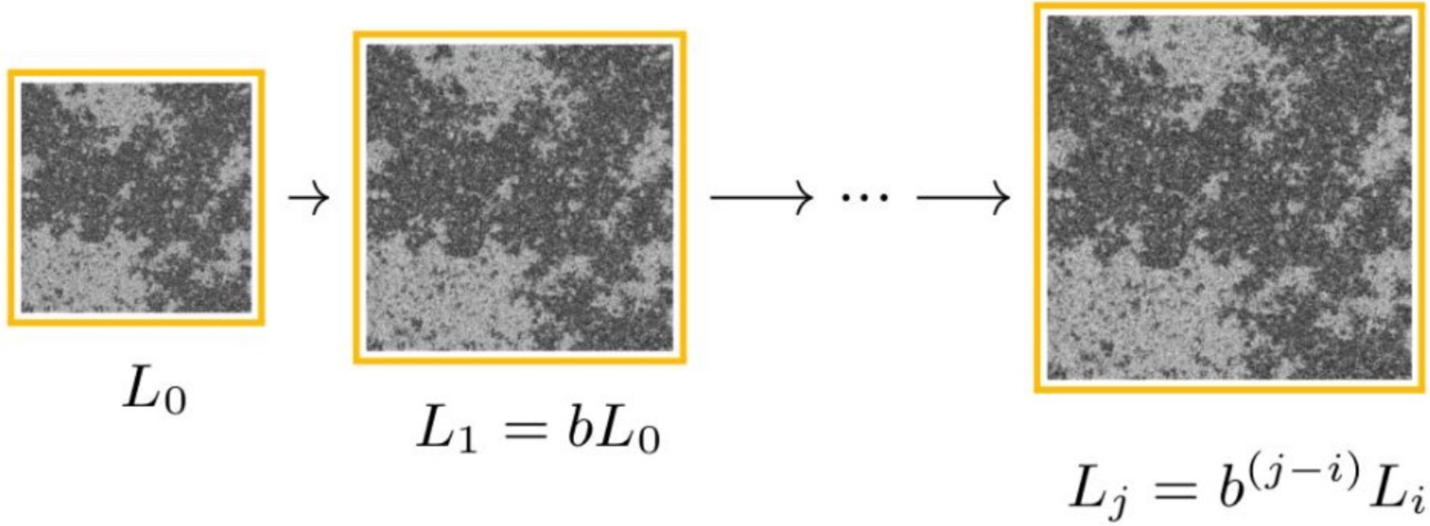
Renormalisation group

- generates flow in parameter space
- due to repeated blocking: run out of degrees of freedom
- need to start with large system to apply RG step multiple times
- large systems, close to a transition,
  suffer from critical slowing down



Inverse renormalisation group

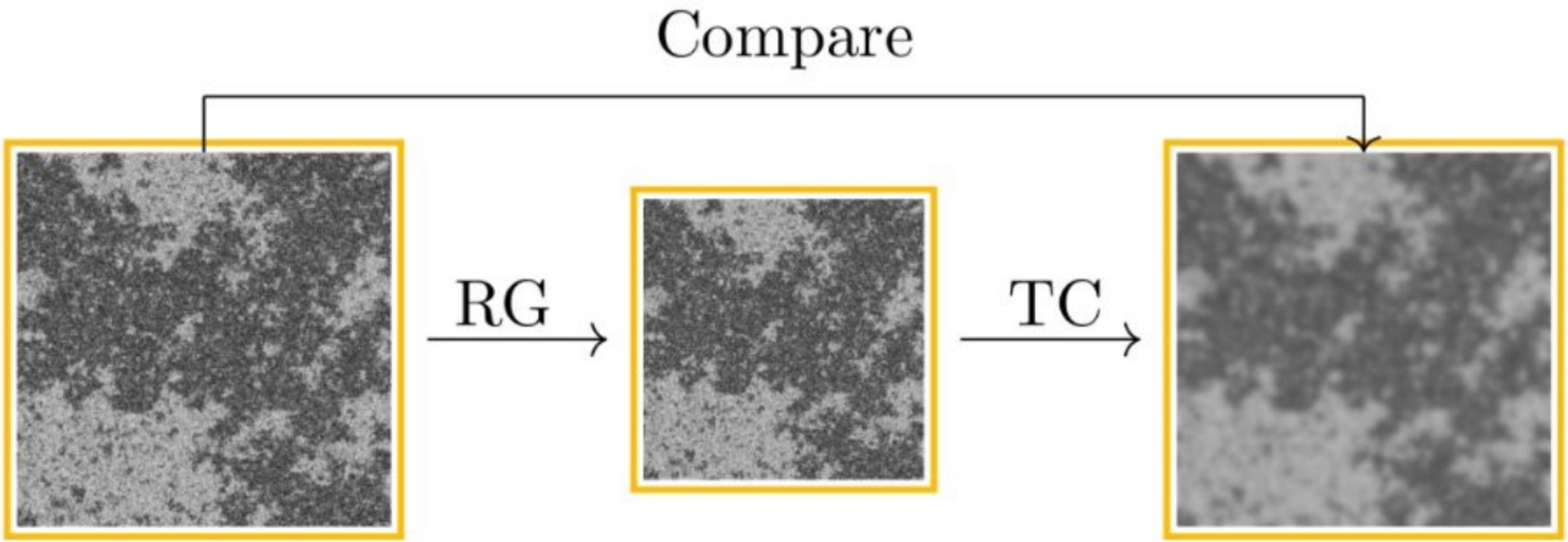
- what if we could invert the RG?
- add degrees of freedom, fill in the ‘details’
- inverse flow in parameter space
- can be applied arbitrary number of steps
- evade critical slowing down



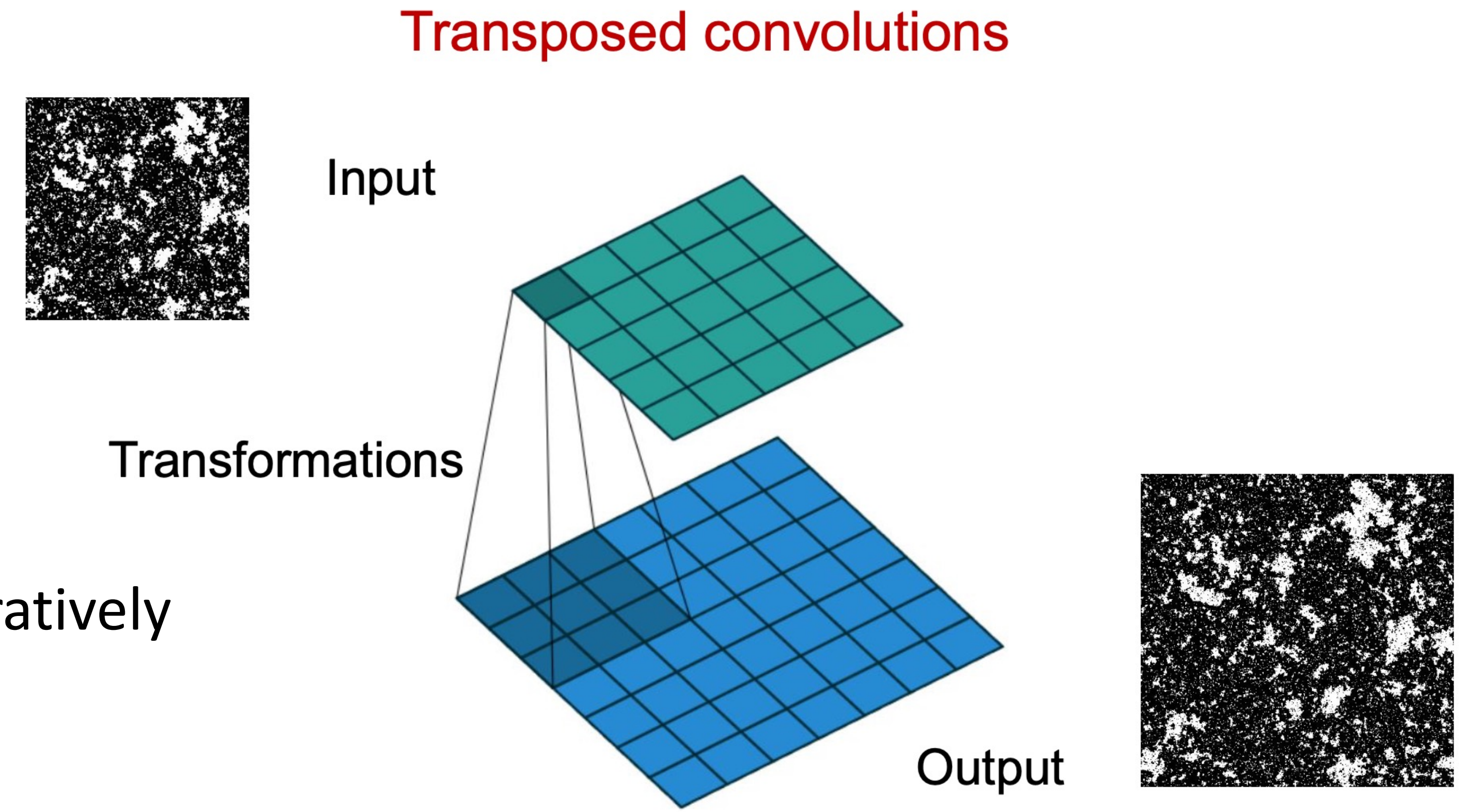
for Ising model: Inverse Monte Carlo Renormalization Group Transformations for Critical Phenomena, D. Ron, R. Swendsen, A. Brandt, Phys. Rev. Lett. 89, 275701 (2002)

How to devise an inverse transformation?

- new degrees of freedom should be introduced
- learn a set of transformations (*transposed convolutions*) to invert a standard RG step
- minimise difference between original and constructed configuration



Inverse renormalisation group



- local transformation
- apply inverse transformations iteratively
- evade critical slowing down
- generate flow in parameter space
- invariance at critical point

Application to φ^4 scalar field theory

- repeated steps
- locking in on critical point

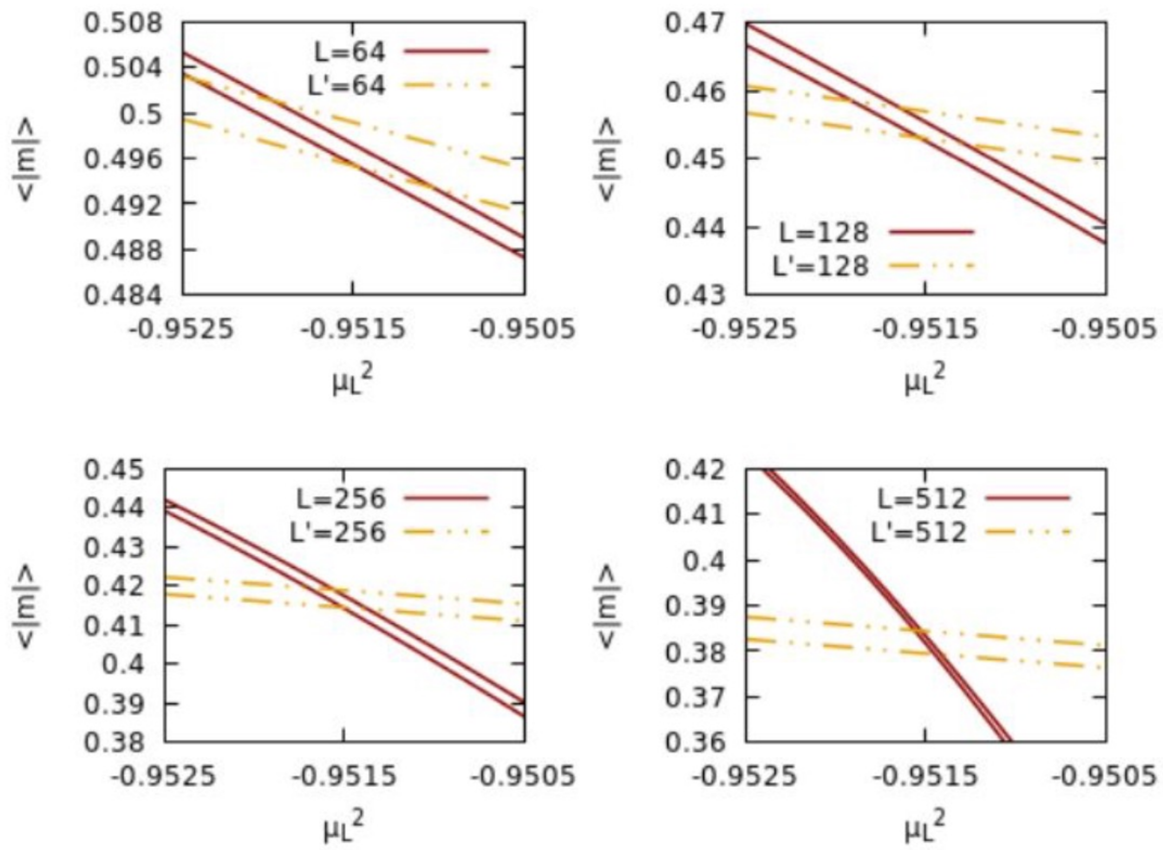


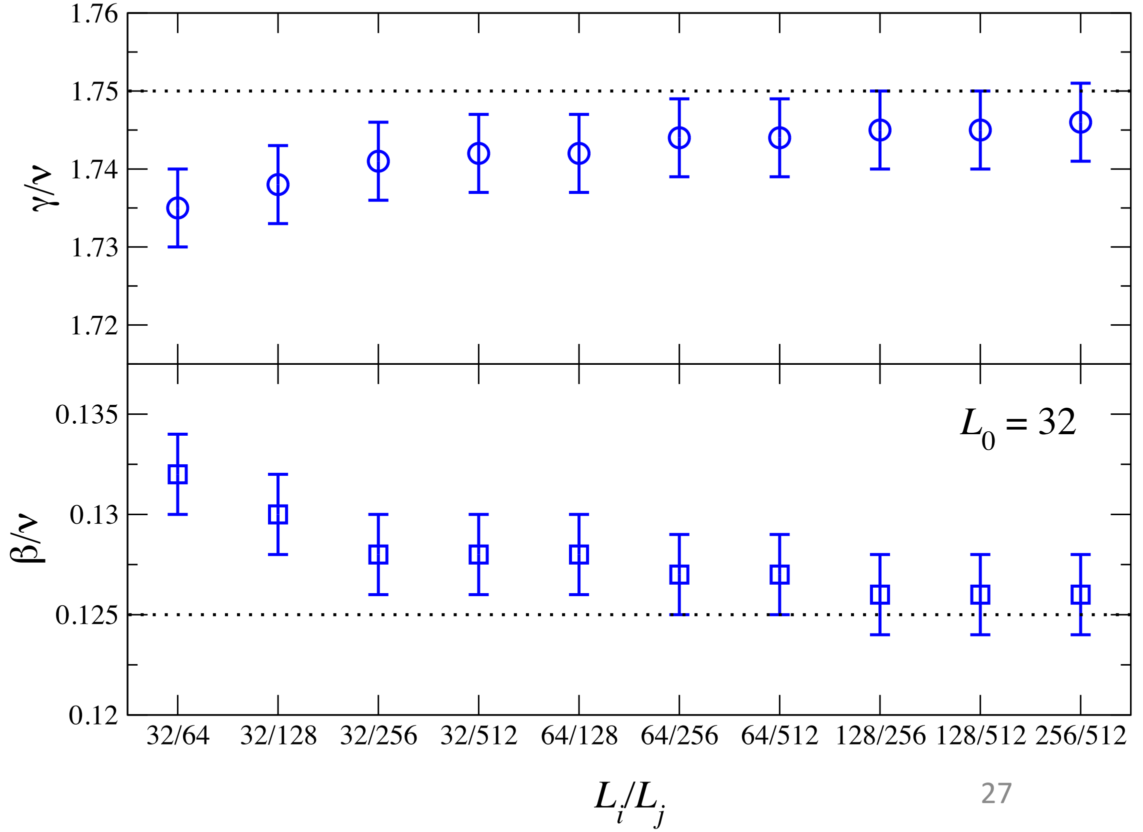
TABLE I. Values of the critical exponents γ/ν and β/ν. The original system has lattice size $L = 32$ in each dimension and its action has coupling constants $\mu_L^2 = -0.9515$, $\lambda_L = 0.7$, and $\kappa_L = 1$. The rescaled systems are obtained through inverse renormalization group transformations.

L_i/L_j	32/64	32/128	32/256	32/512	64/128	64/256	64/512	128/256	128/512	256/512
γ/ν	1.735(5)	1.738(5)	1.741(5)	1.742(5)	1.742(5)	1.744(5)	1.744(5)	1.745(5)	1.745(5)	1.746(5)
β/ν	0.132(2)	0.130(2)	0.128(2)	0.128(2)	0.128(2)	0.127(2)	0.127(2)	0.126(2)	0.126(2)	0.126(2)

GA, Bachtis, Lucini, di Renzo
PRL 128 (2022) 081603

Application to φ^4 scalar field theory

- start with lattice of size 32^2 and apply IRG steps repeatedly
- $32^2 \rightarrow 64^2 \rightarrow 128^2 \rightarrow 256^2 \rightarrow 512^2$
- IRG flow towards critical point
- extract critical exponents
  γ/υ and β/υ from comparison
  between two volumes
- constructed a large (512^2) lattice
  very close to criticality
  without critical slowing down



Intermediate summary: inverse RG

- ✓ flow to critical point without critical slowing down
- ✓ reach large lattices from easy-to-simulate lattice sizes
- ✓ relies on ‘reliable’ blocking step (nontrivial: scalar field majority rule is new)
- ✓ new concept for continuous field theories

Outlook

- ✓ machine learning has seen major boost in physical sciences
- ✓ largely underexplored in statistical/lattice field theory
- ✓ new concepts introduced
- ✓ more progress can be found here:

arXiv > hep-lat > arXiv:2202.05838

Search...

Help | Advanced S

High Energy Physics – Lattice

[Submitted on 10 Feb 2022]

Applications of Machine Learning to Lattice Quantum Field Theory

Denis Boyda, Salvatore Calì, Sam Foreman, Lena Funcke, Daniel C. Hackett, Yin Lin, Gert Aarts, Andrei Alexandru, Xiao-Yong Jin, Biagio Lucini, Phiala E. Shanahan

There is great potential to apply machine learning in the area of numerical lattice quantum field theory, but full exploitation of that potential will require new strategies. In this white paper for the Snowmass community planning process, we discuss the unique requirements of machine learning for lattice quantum field theory research and outline what is needed to enable exploration and deployment of this approach in the future.