The Inverse Renormalisation Group in Quantum Field Theories

Biagio Lucini

(work in collaboration with G. Aarts and D. Bachtis)

Overview

- Augmenting physical knowledge with Machine Learning
- Methods and models
- Quantitative characterization of phase transitions with Machine Learning
- Deriving new observables with Machine Learning
- Inverting the renormalization group flow with Machine Learning
- Summary and perspectives

Motivations

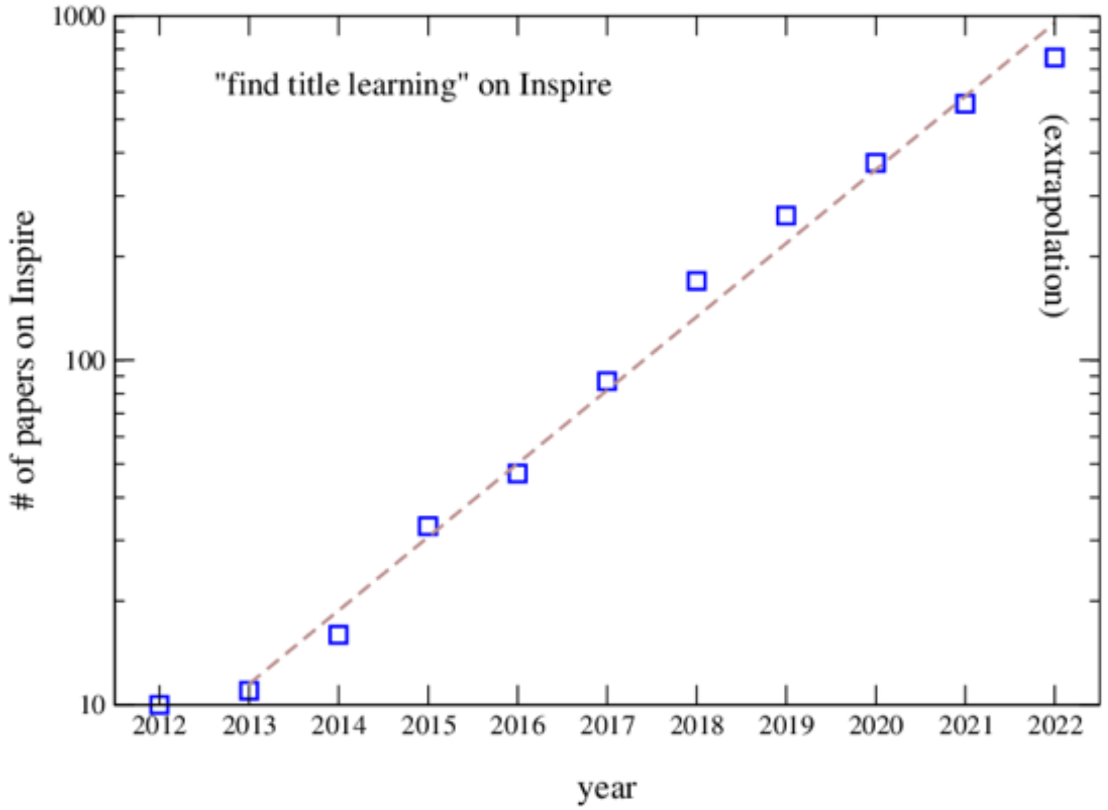
- First-principle studies of phase transitions (using the dynamics of an order parameter and the symmetry breaking pattern) is a well-established
- However, there are cases in which an order parameter is not known (e.g., QCD at finite quark mass), or the symmetry of the transition is debated (again, QCD)
- In other cases, we would like to understand whether the transition is driven by topological excitations (still debated for QCD)
- Topological phase transitions (currently heavily investigated in the Condensed Matter community) do not have an obvious parameter

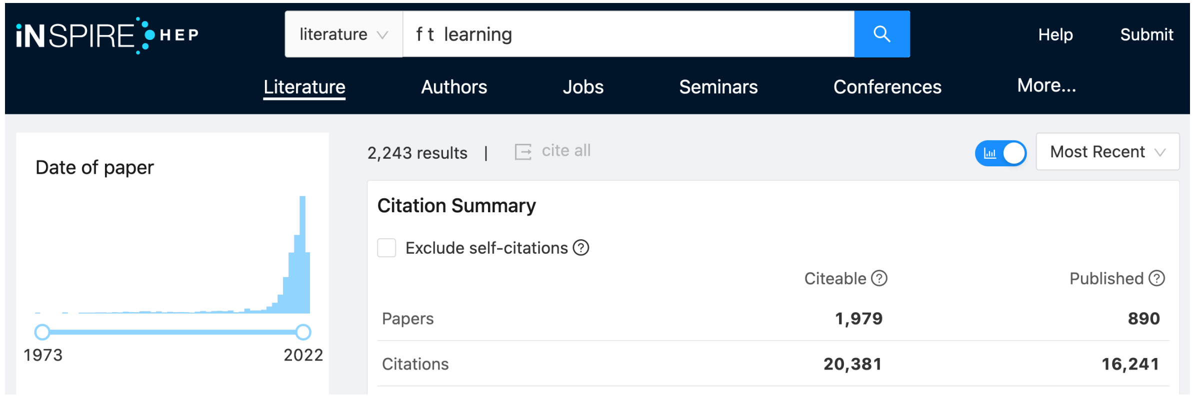
Can Machine Learning provide a universal tool to understand phase transitions?

Machine learning and the Physical Sciences

- Computational learning theory is a subfield of Artificial intelligence studies. Many algorithms available: (deep) neural networks, support vector machines, . . .
- Many ready-to-use libraries in a variety of programming languages: scikit-learn, tensorFlow, Theano, . . . [Chang, Chih-Chung and Lin, Chih-Jen, 2011]
- By now, Machine Learning used in various fields of Physics (High Energy experiments, Gravitational Waves, Astronomy, String Theory, Lattice, . . .) with various degrees of maturity
- Several investigations of Machine Learning applied to the study of phase transition are already present in the literature [following Melko and Carrasquilla]

Adoption of Machine Learning in HEP





Exponential growth !?!

Amara's law: We tend to overestimate the effect of a technology in the short run and to underestimate its effect in the long run.

Machine Learning for Phase Transitions

Prifysgol Abertawe Swansea University

Recent and current problems investigated include

- Can a Machine Learning algorithm detect a phase transition?
- Which algorithms are "better"?
- Can we find the order parameter?
- Can we reconstruct the symmetry that drives the transition?
- To which precision can we determine the transition temperature?
- With which accuracy can we measure quantities such as critical exponents?
- Can we see the *features* (e.g, topological excitations) that are relevant for the transition?
- Can machine learning invert the Renormalisation Group flow?

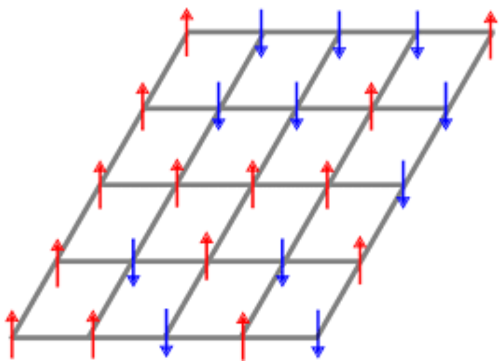
Overview

- Machine Learning for phase transitions
- Models and methods
- Precision calculations using Machine Learning
- Machine Learning derived observables
- The Inverse Renormalisation Group
- Summary and outlook

The Ising Model in D=2 dimensions

- Popular testbed for new numerical approaches, as it has analytic solution at $h = 0$
- Variables: spins $\sigma_i = \pm 1$ distributed on a L^2 grid
- Hamiltonian

$$\mathcal{H} = -\mathcal{J} \sum_{\langle i,j \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i \ , \qquad \mathcal{J} > 0$$



$\mathbb{Z}_2$ symmetry $\sigma_i \mapsto -\sigma_i$

- Partition function at temperature T

$$Z(\beta, h) = \sum_{\{\sigma_i = \pm 1\}} e^{-\beta \mathcal{H}} = e^{-\beta F} \ , \qquad \beta = (kT)^{-1}$$

For $h = 0$ phase transition at $T_c = \frac{2}{k\left(\log(1+\sqrt{2})\right)} = 2.2691853\ldots$

- Phase transition driven by spontaneous breaking of $\mathbb{Z}_2$ symmetry, with order parameter

$$m = \frac{1}{L^2 Z} \sum_{\{\sigma_i = \pm 1\}} \sigma_i e^{-\beta \mathcal{H}} = \frac{1}{L^2} \langle \sum_i \sigma_i \rangle$$

For $L \to \infty$, $m \neq 0$ for $T < T_c$, while $m = 0$ for $T > T_c$

The Ising critical point

- At $L = \infty$ the magnetic susceptibility has a divergence at T_c:

$$\chi = \frac{1}{L^2}\left(\langle\left(\sum_i \sigma_i\right)^2\rangle - \langle\sum_i \sigma_i\rangle^2\right) \underset{T\to T_c^\pm}{\propto} |T - T_c|^{-\gamma}$$

- At finite volume, the latter singularity gets smoothened down into a peak $\chi_{\text{max}}(T_c(L))$ and

$$|T_c(L) - T_c| \propto L^{-\frac{1}{\nu}}\ , \qquad \chi_{\text{max}}(T_c(L)) \propto L^{\frac{\gamma}{\nu}}$$

- Finite size scaling: extract γ and ν from the variation with L of $\chi_{\text{max}}(T_c(L))$

The other critical exponents can be derived from scaling relations

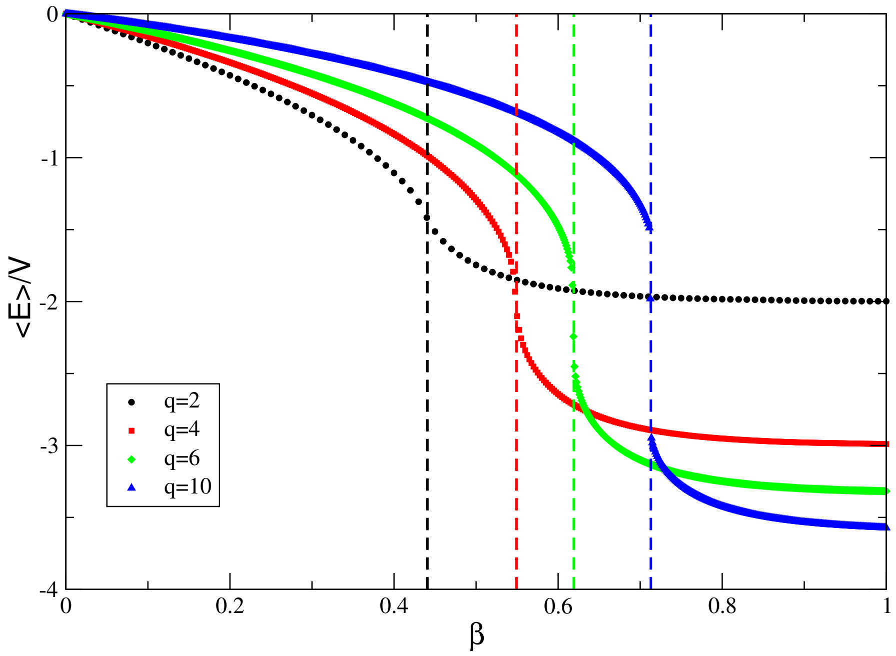
Fisher Law:	$\gamma = \nu(2 - \eta)$,
Widom Law:	$\gamma = \beta(\delta - 1)$,
Rushbrooke Law:	$\alpha + 2\beta + \gamma = 2$,
Josephson Law:	$\nu d = 2 - \alpha$,

The Potts model in D=2

Hamiltonian $$H = 2\beta \sum_{\langle ij \rangle} \left(\frac{1}{q} - \delta_{\sigma_i, \sigma_j} \right)$$

Second order phase transition for q < 5, first order phase transition otherwise

$$\beta_c = \frac{1}{2} \log \left(1 + \sqrt{q}\right)$$



The self-interacting scalar field in D=2

- Action

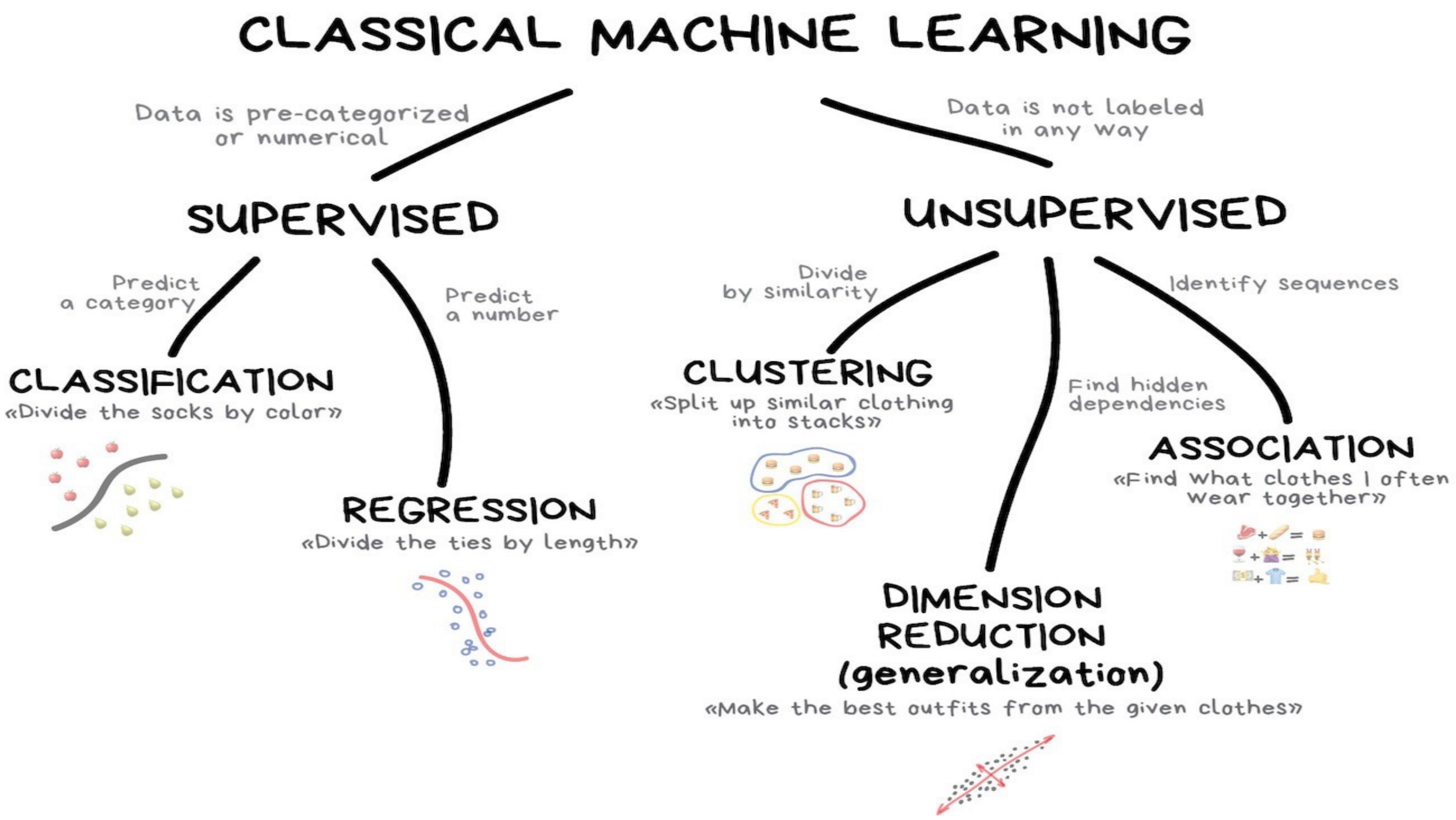
$$S = -\kappa_L \sum_{\langle ij \rangle} \phi_i \phi_j + \frac{(\mu_L^2 + 4\kappa_L)}{2} \sum_i \phi_i^2 + \frac{\lambda_L}{4} \sum_i \phi_i^4$$

- We fix $\kappa_L = 1$ and find a line of critical points, depending on the ratio λ_L / μ_L^2
- We consider the reference critical values

$$\lambda_L = 0.7 \,, \qquad \mu_L^2 = -0.95153(16)$$

[D. Schaich, W. Loinaz, arXiv:0902.0045]

What is machine learning?

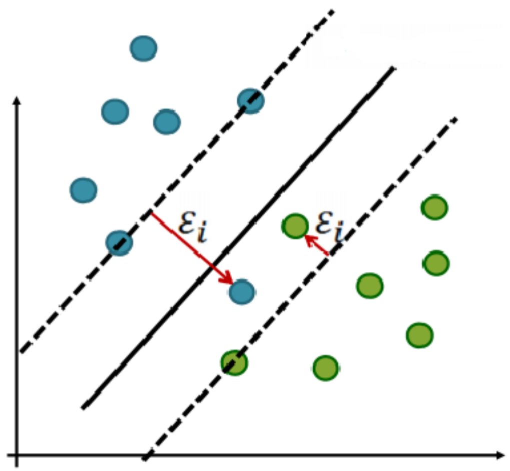


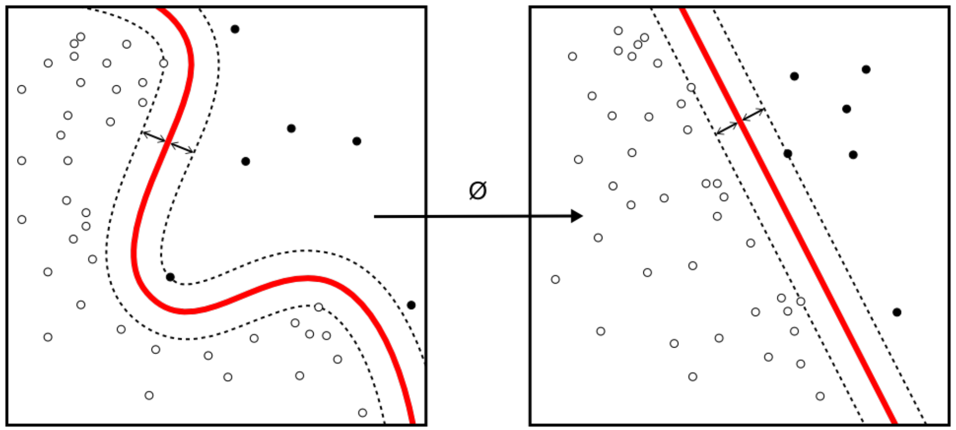
[Credits: https://vas3k.com/blog/machine_learning]

The Support Vector Machine

Problem: separate two classes of data through a maximally separating hyperplane

Transformations (*kernels*) can be requested in order to find the maximally separating hyperplane

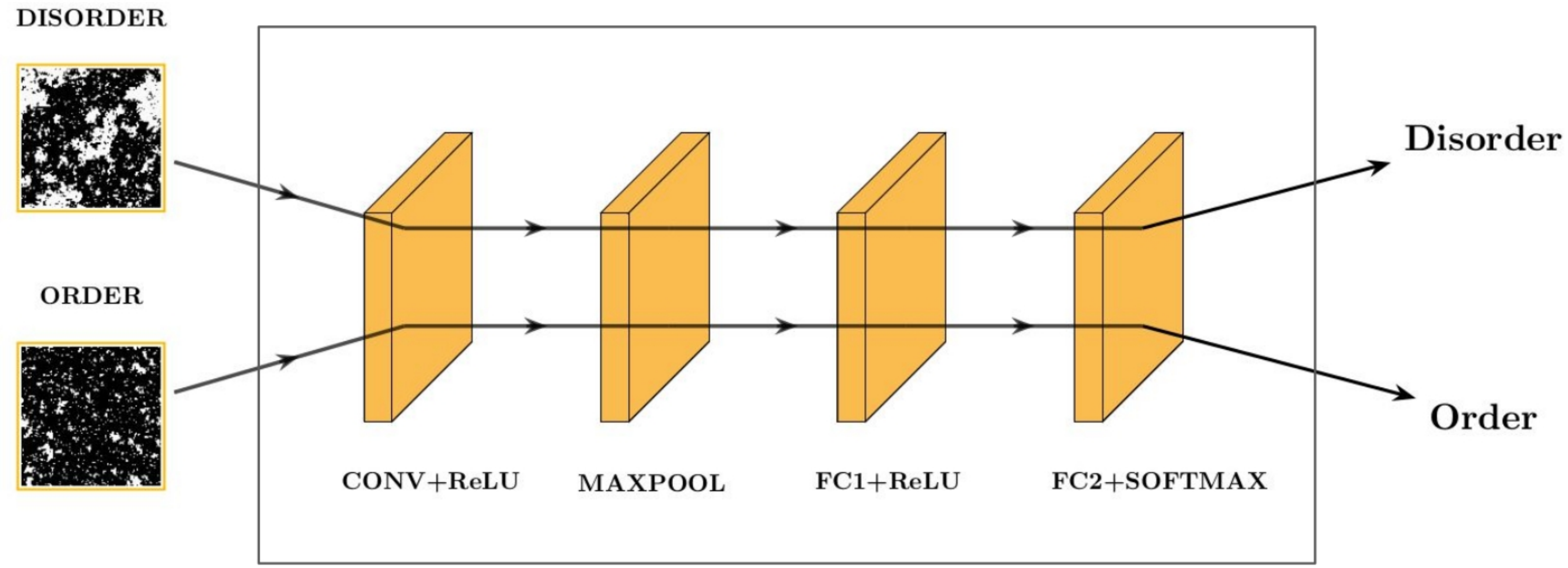




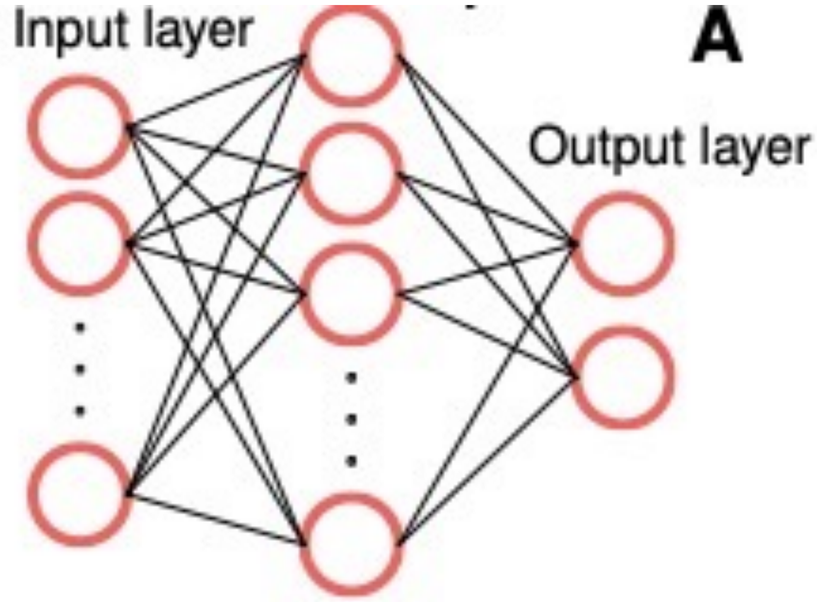
Hyperplane (or hypersurface) identified by decision function d, whose sign identifies the class

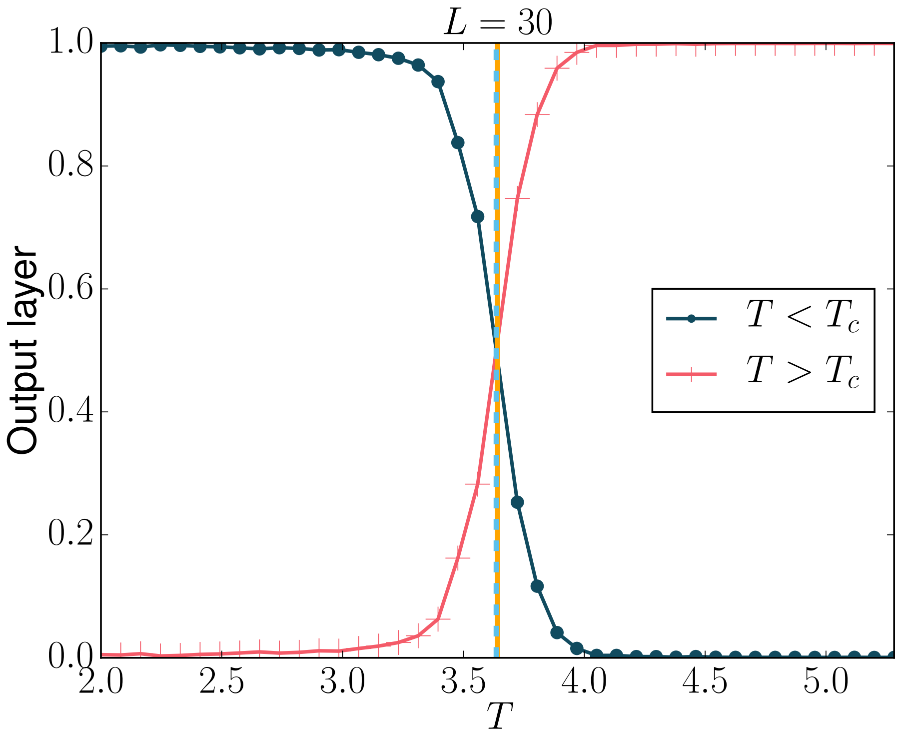
Convolutional Neural Networks

$f(\sigma_i)$



Exposing the phase structure





- Neural Network trained on a square lattice
- Critical temperature on the triangular lattice determined at the permille level (finite size shift?)

[Carrasquilla and Melko, Nature Physics volume 13, pages 431–434 (2017), arXiv:1605.1735]

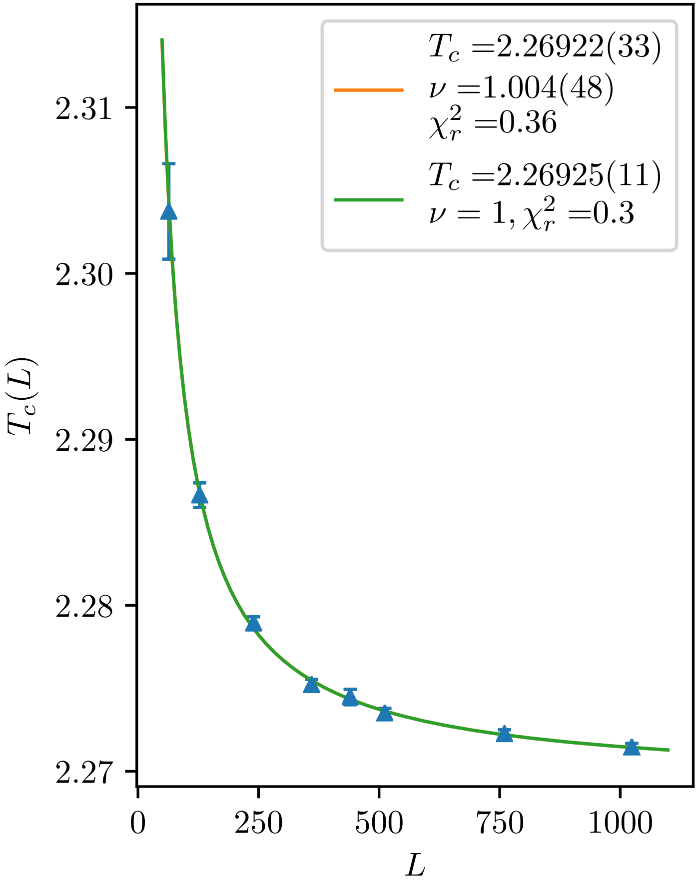
Overview

- Machine Learning for phase transitions
- Models and methods
- Precision calculations using Machine Learning
- Machine Learning derived observables
- The Inverse Renormalisation Group
- Summary and outlook

Determination of ν (Ising)

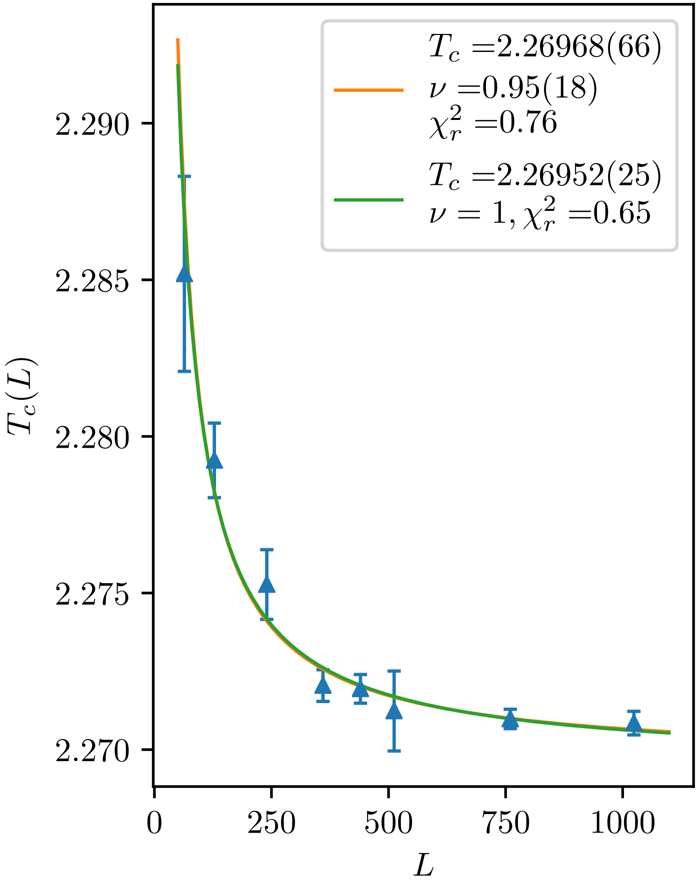
[C. Giannetti, B. Lucini and D. Vadacchino, Nucl.Phys.B 944 (2019) 114639, arXiv:1812.06726]

Scaling of peak position of χ



Exact value:
T_c = 2.2691865...

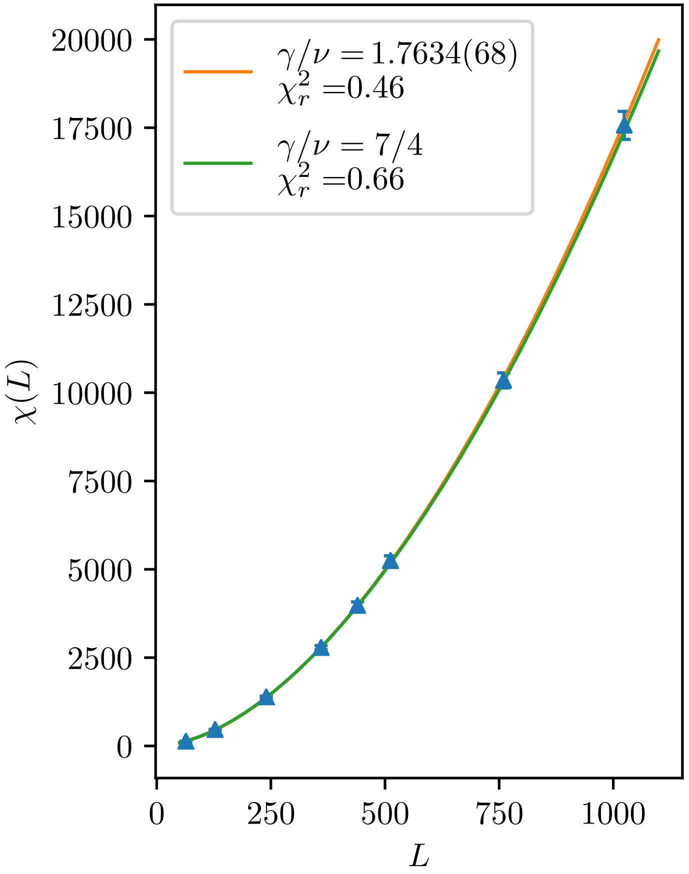
Scaling of peak position of d



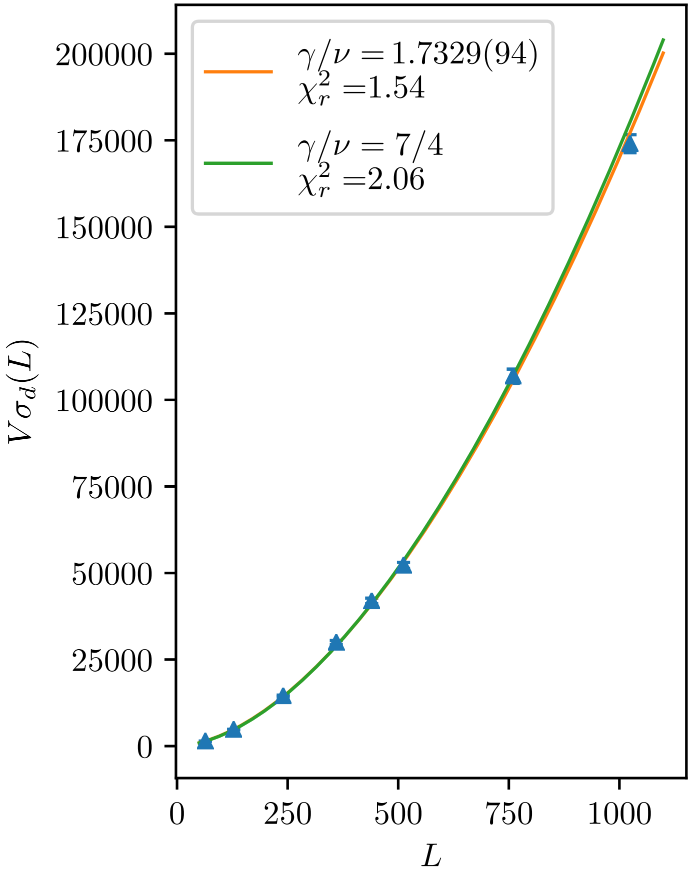
Determination of γ/ν (Ising)

[C. Giannetti, B. Lucini and D. Vadacchino, Nucl.Phys.B 944 (2019) 114639, arXiv:1812.06726]

Scaling of peak height of χ



Scaling of peak height of d



Summary of other results

[C. Giannetti, B. Lucini and D. Vadacchino, Nucl.Phys.B 944 (2019) 114639, arXiv:1812.06726]

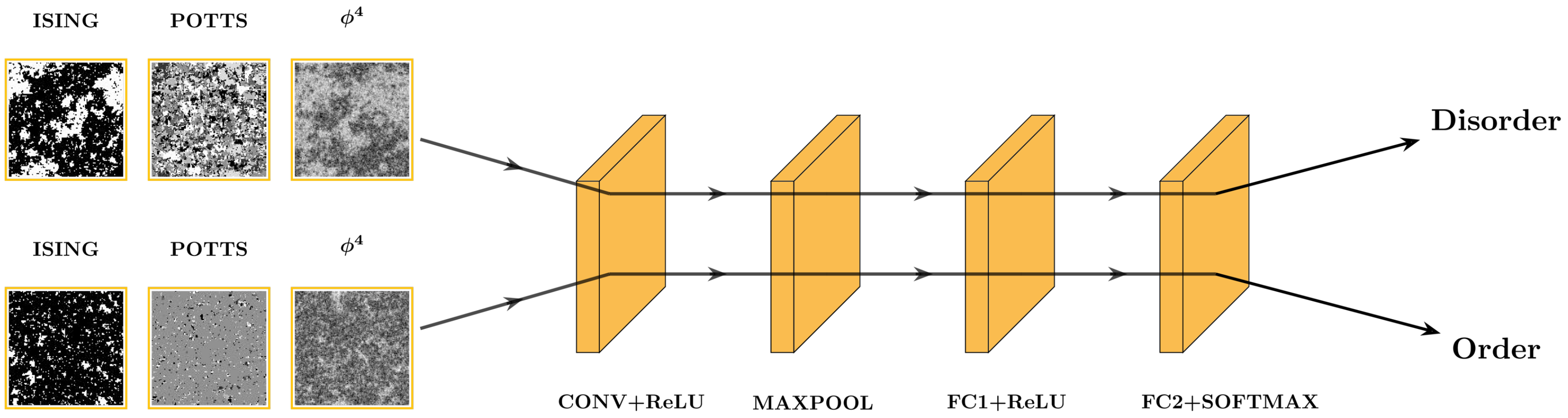
- Critical exponents reproduced with very good accuracy

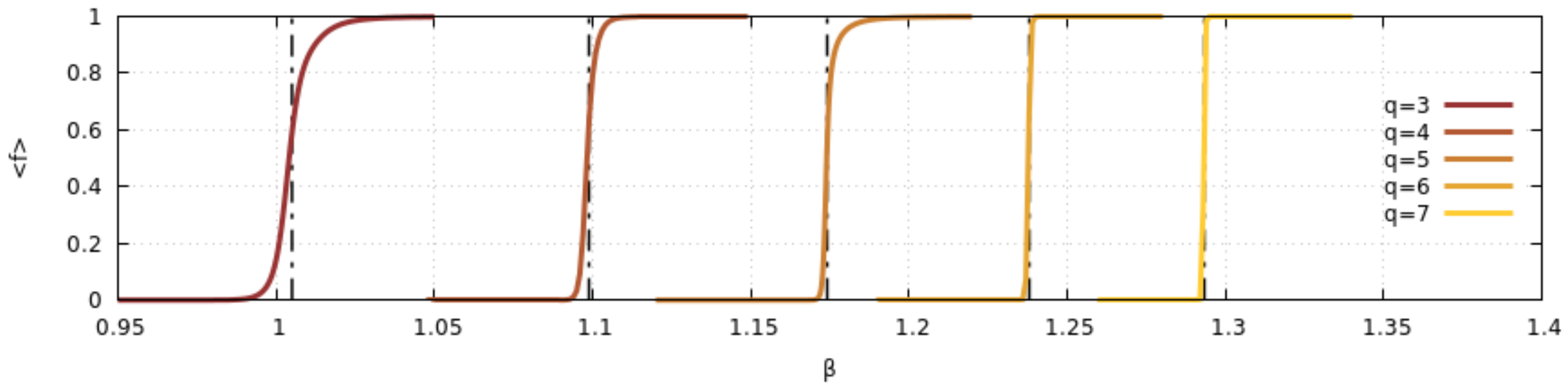
Method	T_c	ν	χ_r^2	γ/ν	χ_r^2
Reweighting	2.26922(33)	1.004(48)	0.36	1.7634(68)	0.46
	2.26925(11)	1 (exact)	0.3	7/4 (exact)	0.66
SVM	2.26968(66)	0.95(18)	0.79	1.733(10)	1.54
	2.26954(25)	1 (exact)	0.65	7/4 (exact)	2.06

- The SVM finds the (square of the) magnetization as d
- The symmetry is encoded in the kernel transformation
- Independence of the (sensibly chosen) training temperatures

Transfer learning

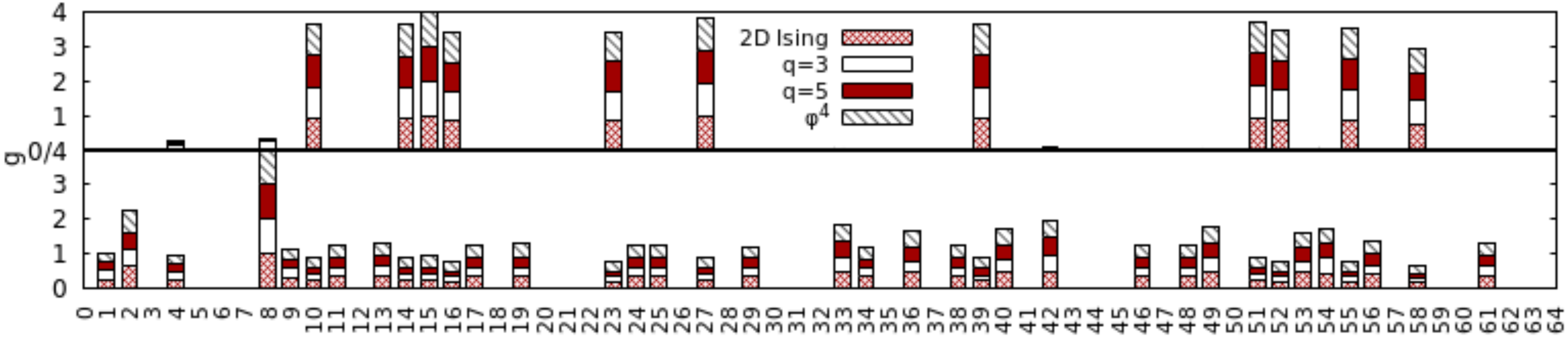
[D. Bachtis, G. Aarts and B. Lucini, Phys.Rev.E 102 (2020) 5, 053306, arXiv:2007.00355]





A Convolutional Neural Network trained on Ising 2D can locate the order-disorder transition in other spin models

Towards interpretability: activation functions in NN



disordered

ordered

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

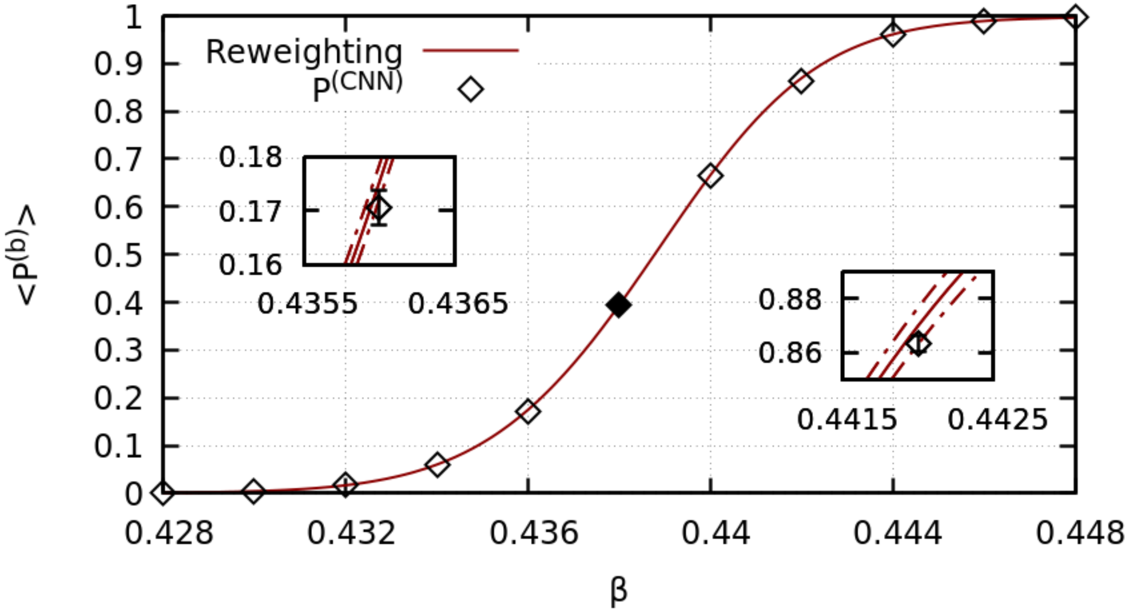
Overview

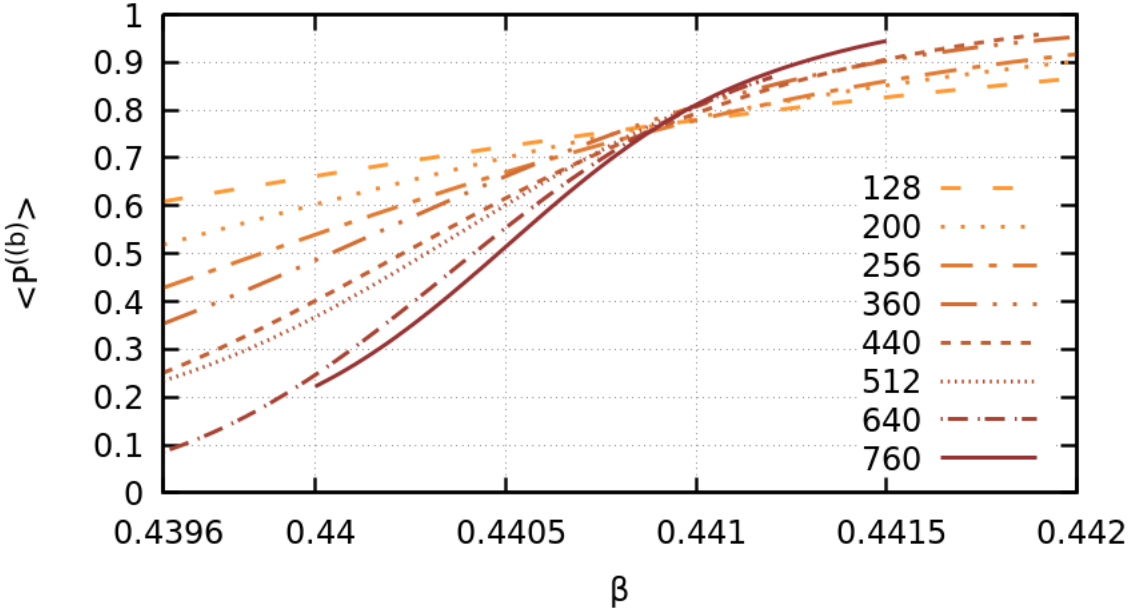
- Machine Learning for phase transitions
- Models and methods
- Precision calculations using Machine Learning
- Machine Learning derived observables
- The Inverse Renormalisation Group
- Summary and outlook

Probability of classification as an observable

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

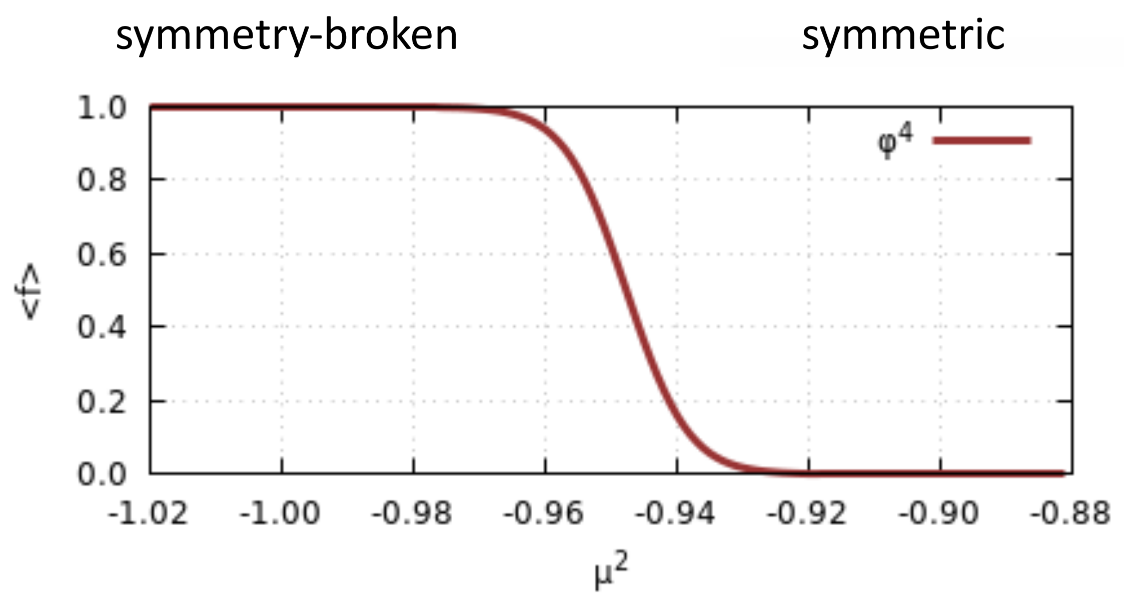
The probability of classification reweighted using a single point agrees with direct measurement
➡ this probability is a thermodynamic observable!



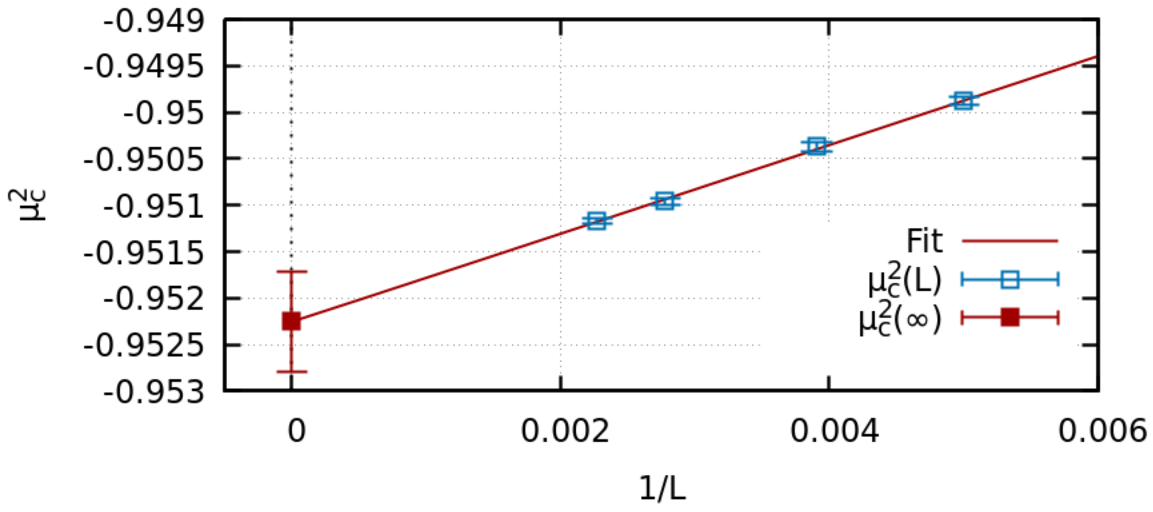


[D. Bachtis, G. Aarts and B. Lucini, Phys.Rev.E 102 (2020) 3, 033303, arXiv:2004.14341]

ϕ^4 scalar field theory



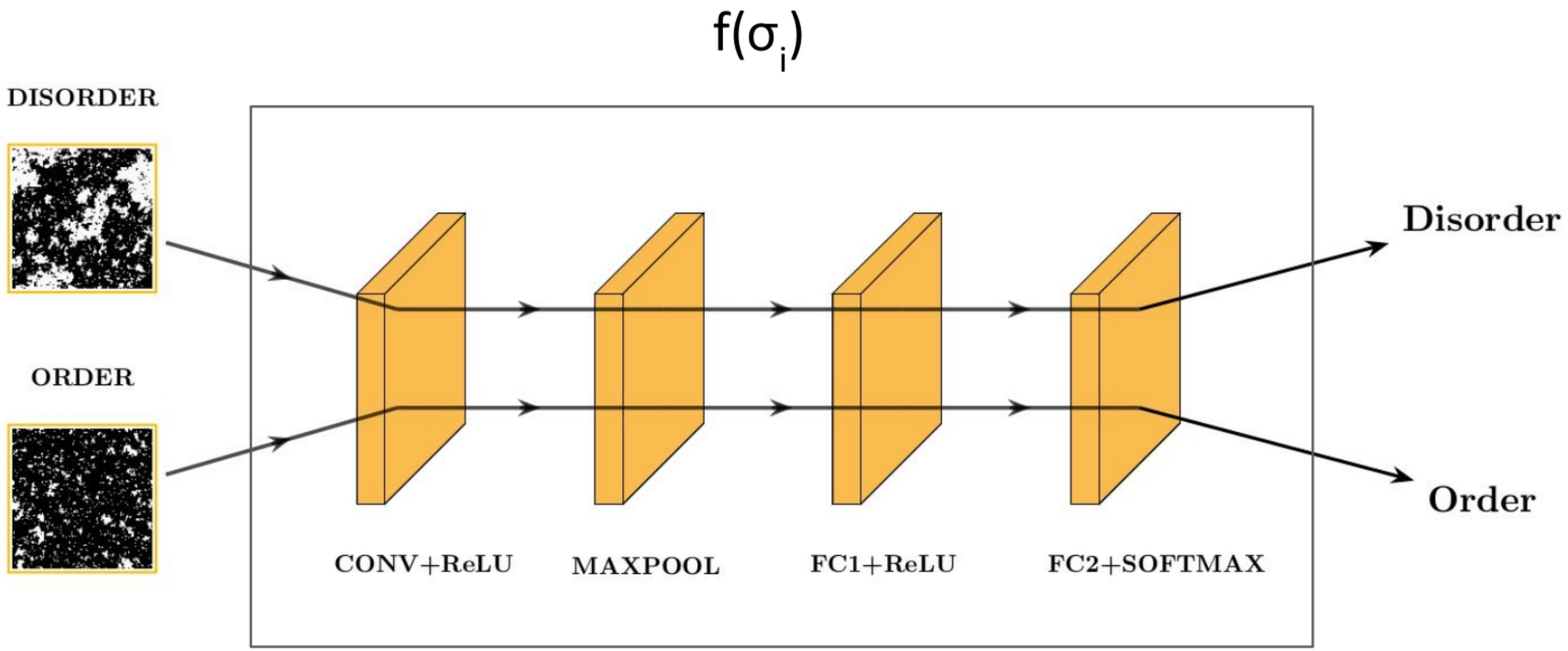
- reweight in mass parameter, μ^2
- identify regions where phase is clear
- retrain NN using $\mu^2 < -1.0$ and $\mu^2 > -0.9$
- repeat finite-size scaling analysis as in 2d Ising model



	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)

Order parameters from machine learning



Can the function f act as an order parameter?

[D. Bachtis, G. Aarts and B. Lucini, Phys.Rev.E 102 (2020) 3, 033303, arXiv:2004.14341]

Coupling f to the Hamiltonian

Define an observable variable Y conjugated to f and write an extended Hamiltonian

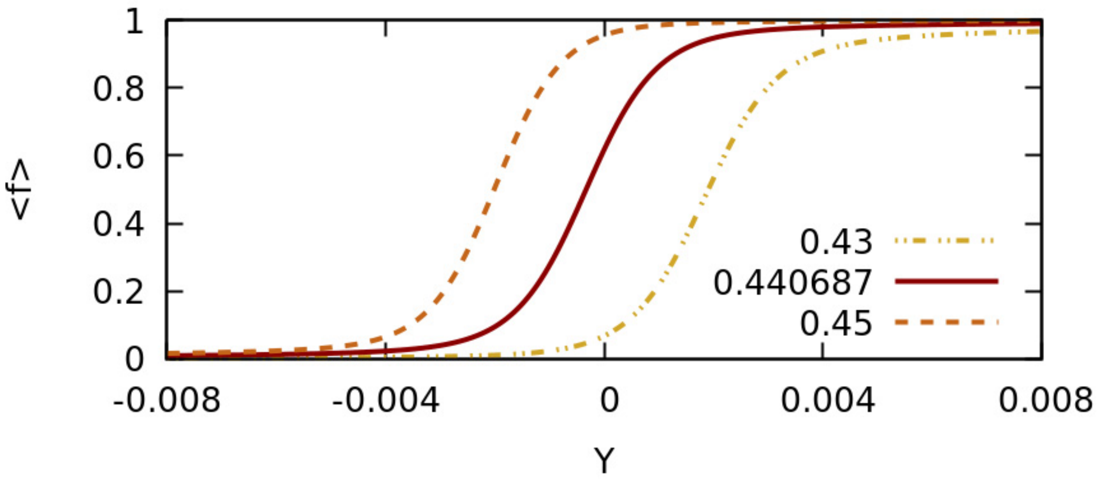
$$E_Y = E - V f Y.$$

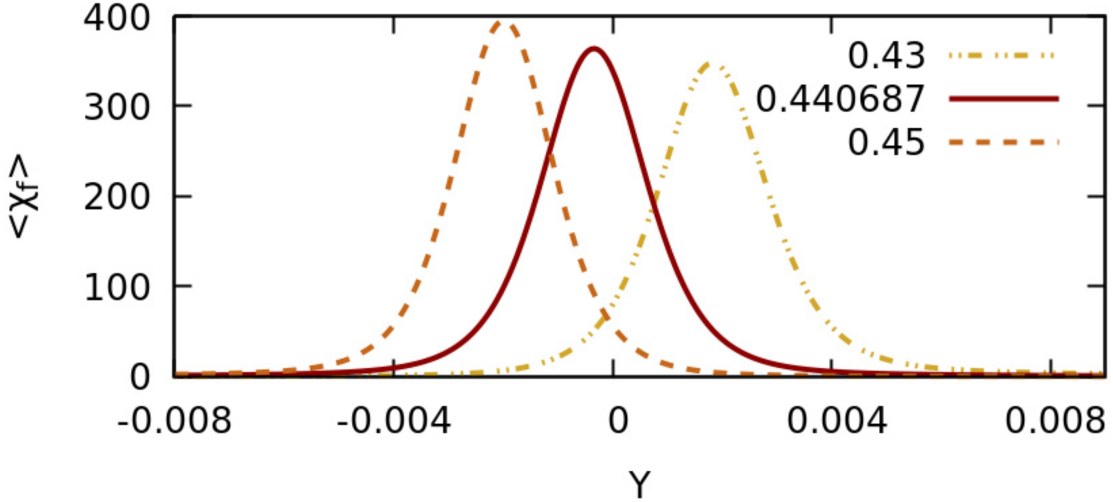
Now f can be computed using path integral methods

$$\langle f \rangle = \frac{1}{\beta V} \frac{\partial \ln Z_Y}{\partial Y} = \frac{\sum_\sigma f_\sigma \exp[-\beta E_\sigma + \beta V f_\sigma Y]}{\sum_\sigma \exp[-\beta E_\sigma + \beta V f_\sigma Y]}$$

Note that Y define a new direction for reweighting and that reweighting in this direction does not require the knowledge of E_Y

Induced phase transition





Critical exponents calculated with Renormalisation Group methods

	β_c	ν	θ_Y, θ
RG+NN	0.44063(21)	1.01(2)	$\theta_Y = 0.534(3)$
Exact	$\ln(1+\sqrt{2})/2$	1	$\theta = 8/15$

f allows access to the magnetic critical exponent θ

Overview

- Machine Learning for phase transitions
- Models and methods
- Precision calculations using Machine Learning
- Machine Learning derived observables
- The Inverse Renormalisation Group
- Summary and outlook

The Inverse Renormalisation Group

[D. Bachtis, G. Aarts, F. Di Renzo, and B. Lucini, Phys. Rev. Lett., 128:081603 (2022)]

Prifysgol Abertawe
Swansea University

Purpose: generating configurations on larger lattices starting from smaller ones near criticality with negligible computational cost

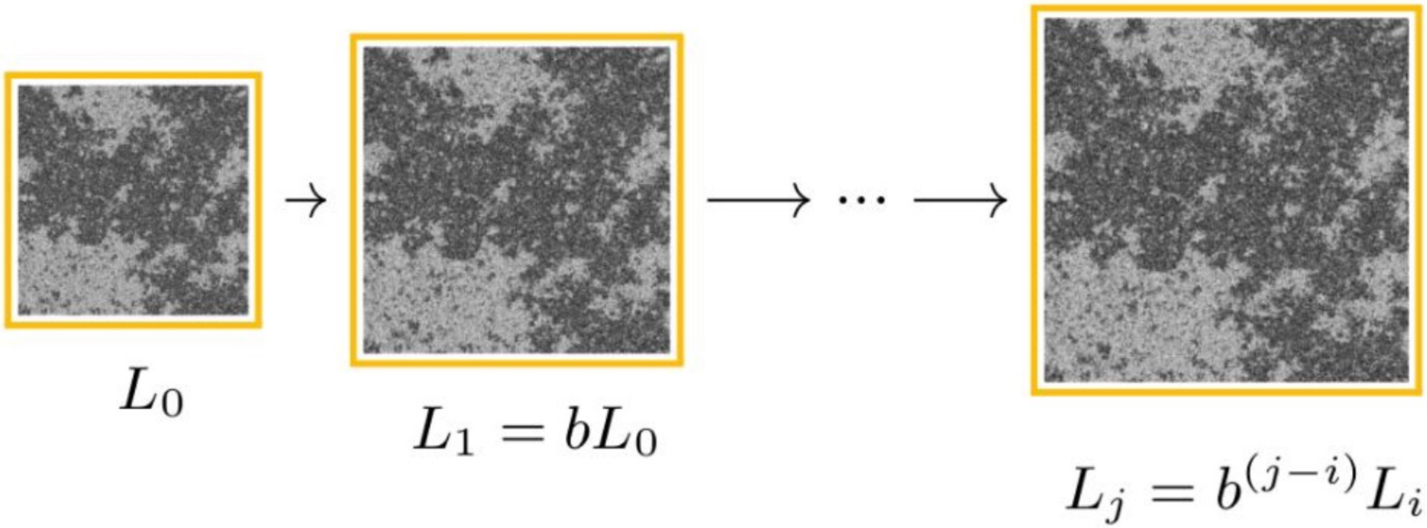
Not a new idea, e.g.

- R.H. Swendsen, Phys. Rev. Lett., 42:859–861 (1979)
- D. Ron, R.H. Swendsen, and A. Brandt, Phys. Rev. Lett., 89:275701 (2002)
- S. Efthymiou, M.J.S. Beach, and R.G. Melko, Phys. Rev. B, 99:075113, (2019)
- S.-H. Li and L. Wang, Phys. Rev. Lett., 121:260601 (2018)
- K. Shiina, H. Mori, Y. Tomita, H.K. Lee, and Y. Okabe, Scientific Reports, 11(1):9617 (2021)

Our work presents the first IRG calculation for a Quantum Field Theory

Benefits of the IRG

- Overcome critical slowing down $\tau \propto \xi^z$
- More precise calculations of observables at criticality
- Better insights on the infrared dynamics of the model
- Can grow the lattice size indefinitely



Known problem: the RG is not invertible

To invert the RG, we would need to grow the number of degrees of freedom, but the process is not unique

E.g., for a blocked spin equal to +1 possibilities (majority rules) include

+1	+1
+1	-1

-1	+1
+1	+1

-1	+1
+1	-1

-1	+1
+1	-1

…

Even worse for the scalar field, e.g.

0.01	0.36
0.02	0.01

-421.1	90.1
0.5	330.9

…

compatible with a blocked spin value 0.4

What we mean by inverting the RG then?

- We start from a set of configuration generated via a Monte Carlo on a lattice of size L
- Using a Machine Learning algorithm, from those we derive a set of configurations on a lattice L' = b L (typically, b=2)
- We assume that the ensemble at L' as distributed according to the Boltzmann measure at L
- This enables us to compute (and to reweight!) observables at L'
- Using crossing of curves, we compute critical quantities

Advantage: numerical effort done on small lattices, hence relatively cheap

Critical to the process: blocking method, ML algorithm and assumption of Boltzmann distribution

The blocking method

- Given a block B with generic point i, consider

$$\phi_B^+ = \frac{\sum_{i\in B} \phi(i)\theta(\phi(i))}{\sum_{i\in B} \theta(\phi(i))} \quad \text{and} \quad \phi_B^- = \frac{\sum_{i\in B} \phi(i)\theta(-\phi(i))}{\sum_{i\in B} \theta(-\phi(i))}$$

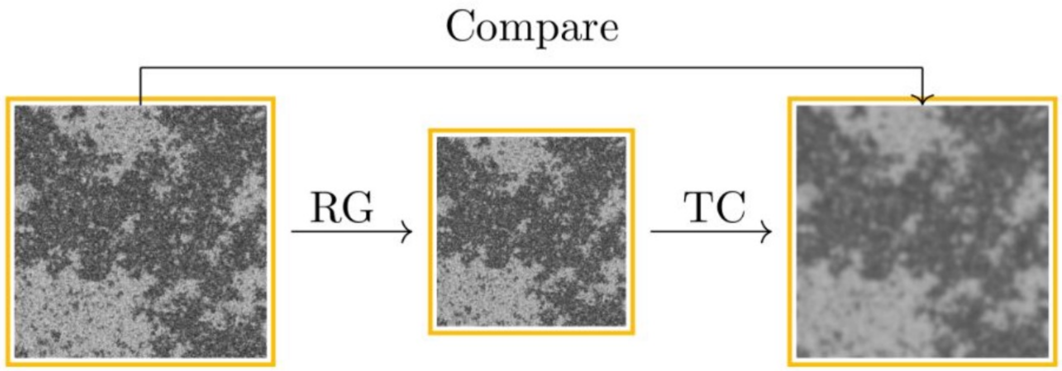
- Now, set

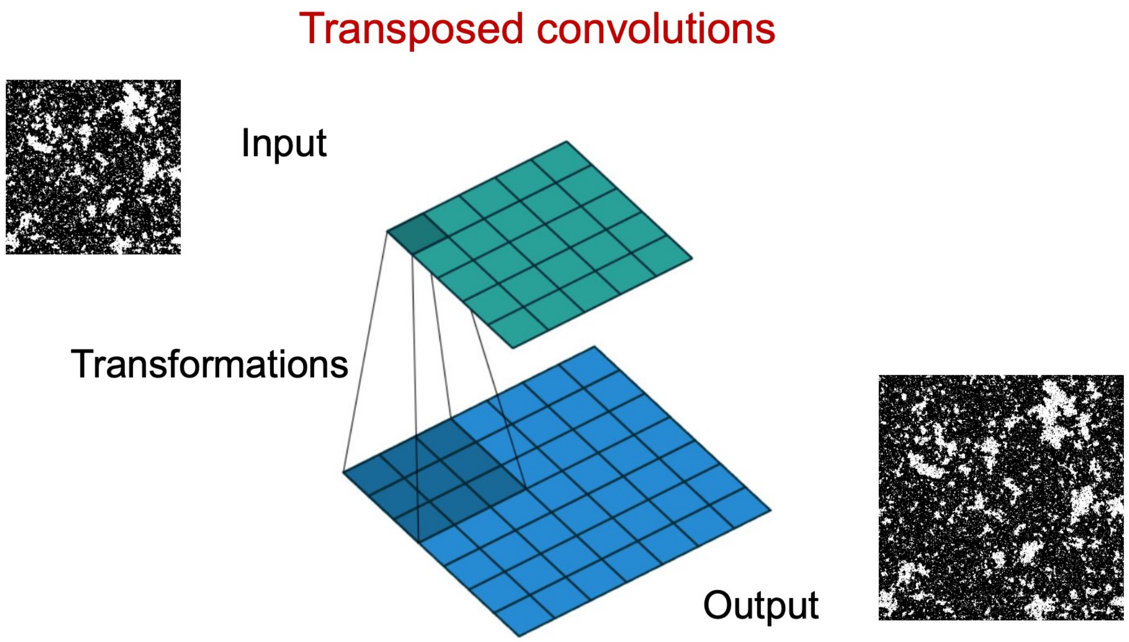
$$\phi_B = \phi_B^+\theta(\phi_B^+ + \phi_B^-) + \phi_B^-\theta(-\phi_B^+ - \phi_B^-)$$

- This is equivalent to the majority rule in the Ising model

Lattice augmentation with Machine Learning

Central concept: transposed convolution





More on transposed convolutions

$$\begin{bmatrix}3 & 1\\ 2 & 0\end{bmatrix} * \begin{bmatrix}w_{11} & w_{12}\\ w_{21} & w_{22}\end{bmatrix}$$

Example

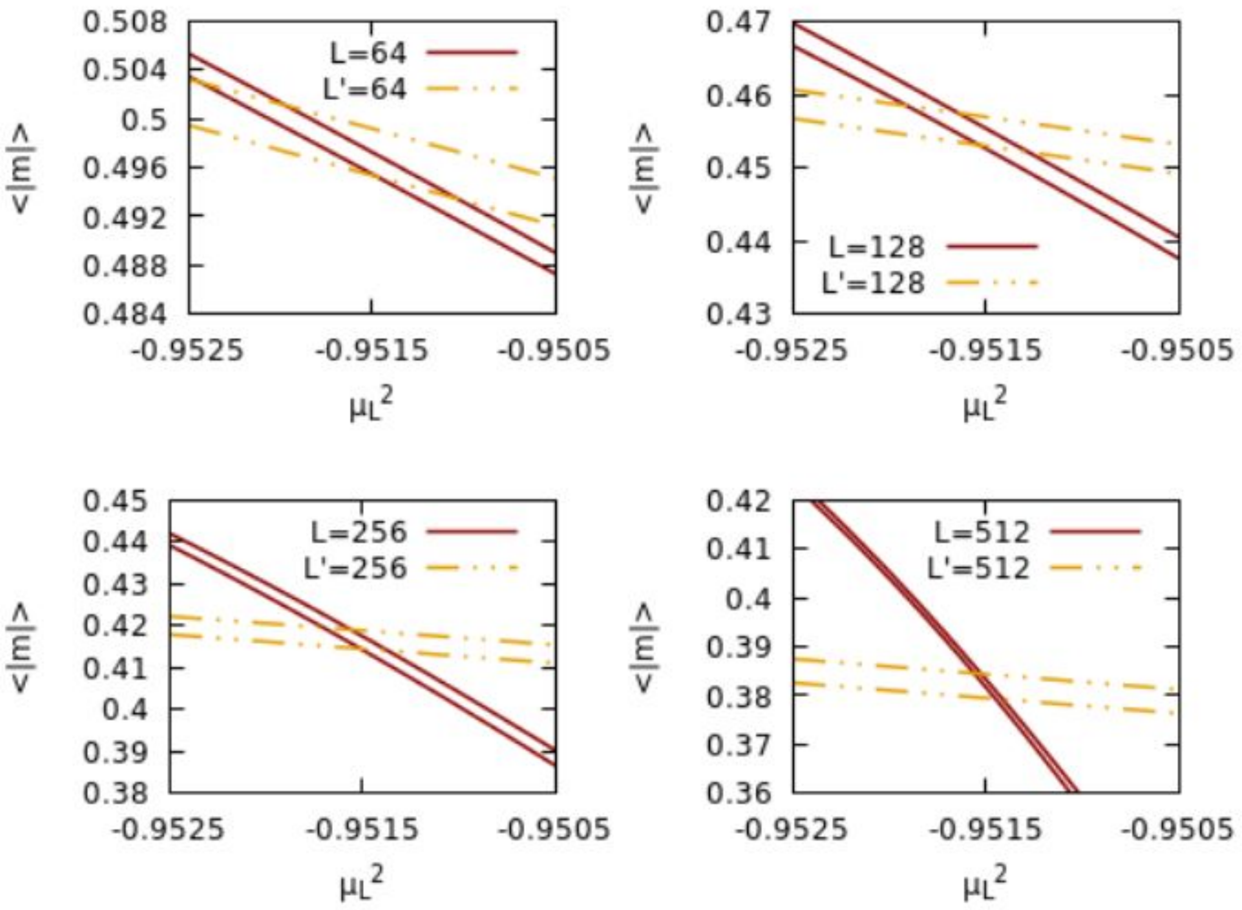
$$\begin{bmatrix}3 & 1\\ 2 & 0\end{bmatrix} * \begin{bmatrix}2 & 3\\ 0 & 1\end{bmatrix} =$$

$$\begin{bmatrix}6 & 9 & \\ 0 & 3 & \\ & & \end{bmatrix} + \begin{bmatrix} & 2 & 3\\ & 0 & 1\\ & & \end{bmatrix} + \begin{bmatrix} & & \\ 4 & 6 & \\ 0 & 2 & \end{bmatrix} + \begin{bmatrix} & & \\ & 0 & 0\\ & 0 & 0\end{bmatrix} =$$

$$\begin{bmatrix}6 & 11 & 3\\ 4 & 9 & 1\\ 0 & 2 & 0\end{bmatrix}$$

Determining the direction of the RG flow

- Comparison with directly simulated lattices show that in the augmented system the coupling flows towards the critical point
- Plotting two different lattice sizes (no need for direct simulation!) the crossing identifies an estimate for the critical coupling



Determining critical quantities

We can rewrite the scaling relationships for the magnetisation

$$m_i \sim |t_i|^\beta \qquad m_j \sim |t_j|^\beta$$

in terms of the correlation length

$$m_i \sim \xi_i^{-\beta/\nu} \qquad m_j \sim \xi_j^{-\beta/\nu}$$

to obtain the operational definition of the critical exponent ratio

$$\frac{\beta}{\nu} = -\frac{\ln \frac{dm_j}{dm_i}\Big|_{K_c}}{\ln \frac{\xi_j}{\xi_i}} = -\frac{\ln \frac{dm_j}{dm_i}\Big|_{K_c}}{(j-i)\ln b}$$

Similarly, from χ we get γ/ν

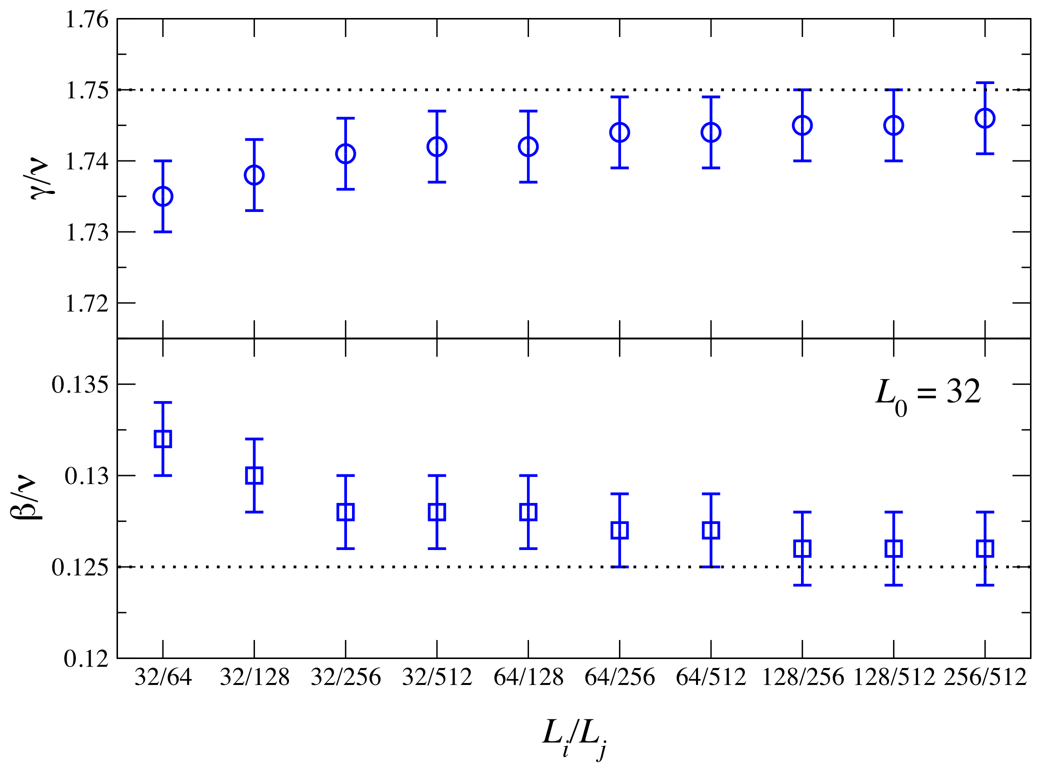
Critical exponents

Method: configurations obtained with a simulation for L=32 and IRG augmentation up to L=512

Ratios of critical exponents extracted for pairs of lattices

Expected asymptotic approach to Ising values clearly observed

All with no critical slowing down!



Conclusions and Outlook

Prifysgol Abertawe Swansea University

- Machine Learning offers a novel angle to look at phase transitions
- It enables precise calculations of critical properties with no assumed knowledge on the underlying symmetry
- Machine Learning exposes novel observables, whose behaviour can offer insights on the dynamics of the phase transition
- A powerful demonstrator of the potential of Machine Learning is the Inverse Renormalisation Group
- Future work focusing on interpretability
- Related work ongoing to derive more efficient and interpretable Machine Learning methods from Quantum Field Theories

[e.g., D. Bachtis, G. Aarts and B. Lucini, Phys.Rev.D 103 (2021) 7, 074510, arXiv:2107.00466]