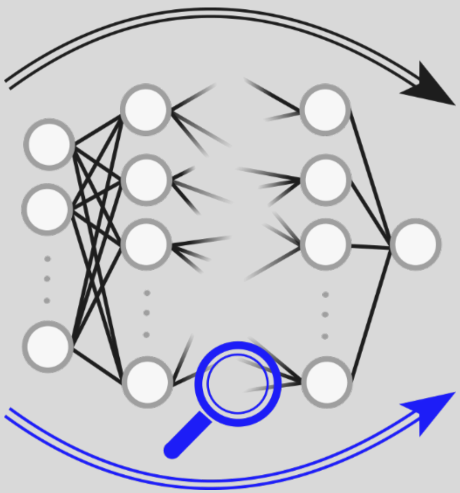
Replacing neural networks by optimal analytical predictors for the detection of phase transitions

Julian Arnold and Frank Schäfer
Department of Physics, University of Basel
Bruder group

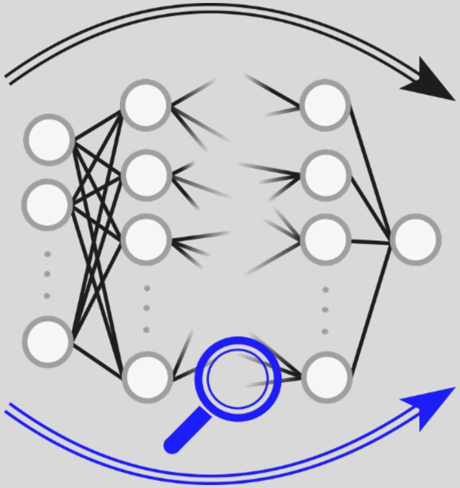


arXiv:2203.06084

FNSNF Swiss National Science Foundation
QSIT Quantum Science and Technology — National Centre of Competence in Research
University of Basel

Replacing neural networks ② by optimal analytical ③ predictors for the detection of phase transitions ①

Julian Arnold and Frank Schäfer
Department of Physics, University of Basel
Bruder group



arXiv:2203.06084

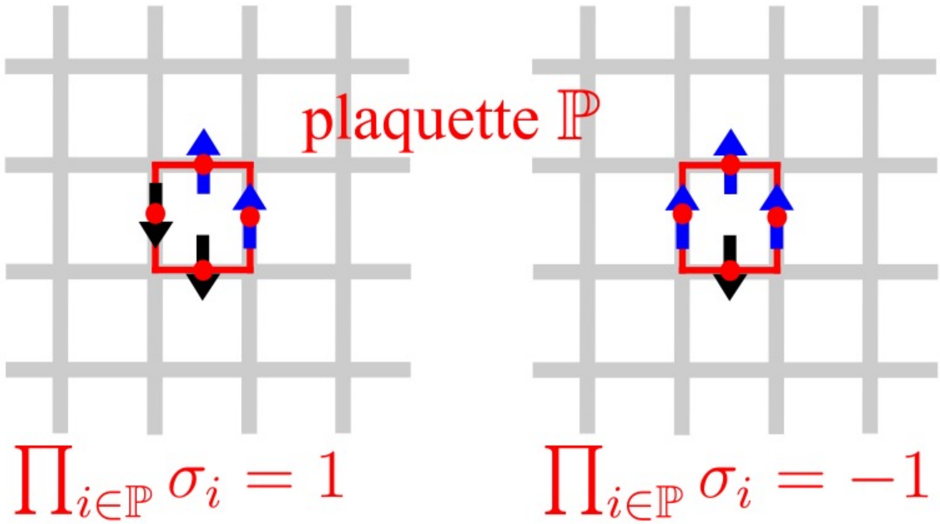
SWISS NATIONAL SCIENCE FOUNDATION

QSIT Quantum Science and Technology
National Centre of Competence in Research

University of Basel

Topological crossover in Ising gauge theory

- Hamiltonian $H(\boldsymbol{\sigma}) = -J \sum_{\mathbb{P}} \prod_{i \in \mathbb{P}} \sigma_i$



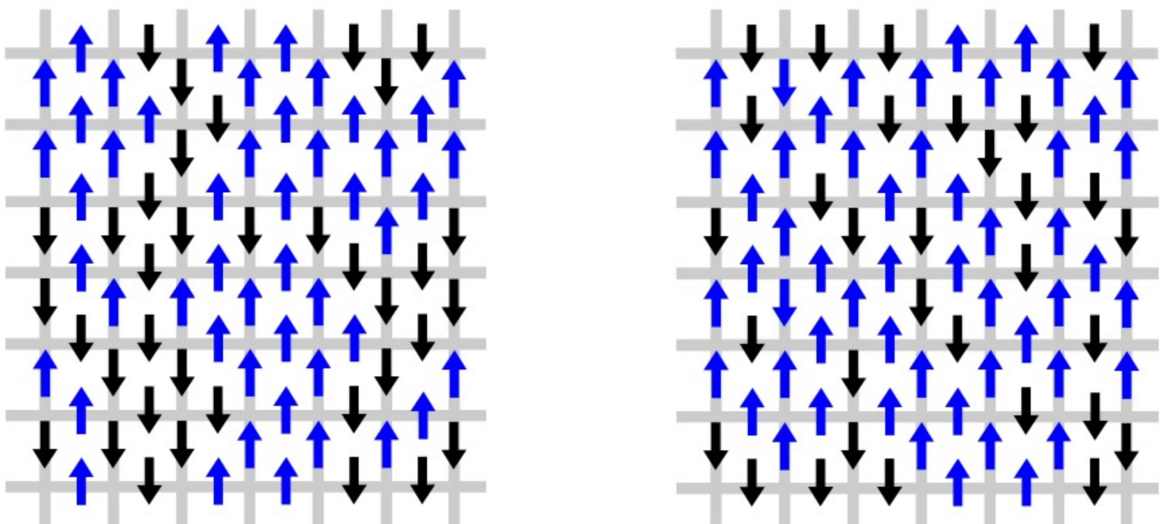
- spin configuration $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_{L\times L})$ with $\sigma_i \in \{+1, -1\}$
- Boltzmann distribution $\mathrm{P}(\boldsymbol{\sigma}) = \dfrac{e^{-H(\boldsymbol{\sigma})/k_{\mathrm{B}}T}}{Z}$

Topological crossover in Ising gauge theory

- Hamiltonian $H(\boldsymbol{\sigma}) = -J \sum_{\mathbb{P}} \prod_{i \in \mathbb{P}} \sigma_i$

ground state

excited state



$\prod_{i \in \mathbb{P}} \sigma_i = 1$

$\prod_{i \in \mathbb{P}} \sigma_i = \pm 1$

temperature T

Motivation

- detecting phase transitions from experimentally accessible data

  ⇒ does not require prior theoretical knowledge

  ⇒ could enable discovery of new phases of matter

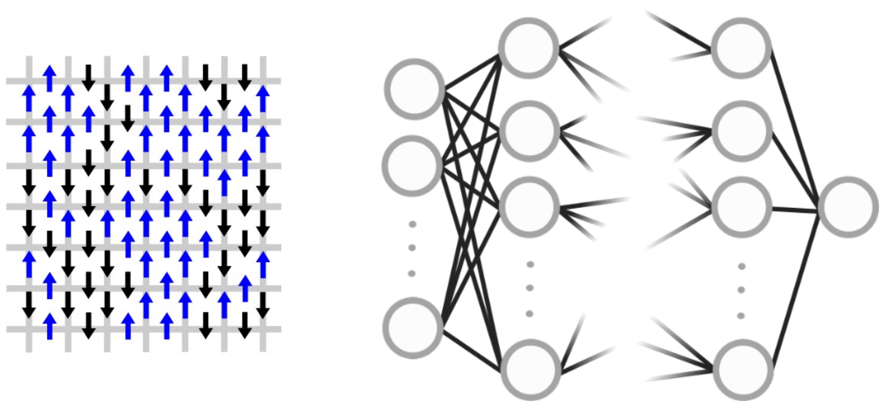
Motivation

- detecting phase transitions from experimentally accessible data

  $\Rightarrow$ does not require prior theoretical knowledge

  $\Rightarrow$ could enable discovery of new phases of matter

use (deep) neural networks which proved successful
in traditional image classification tasks



Detecting phase transitions using neural networks

- supervised learning (SL)

nature physics LETTERS
PUBLISHED ONLINE: 13 FEBRUARY 2017 | DOI: 10.1038/NPHYS4035

Machine learning phases of matter

Juan Carrasquilla[1]* and Roger G. Melko[1,2]

- learning by confusion (LBC)

nature physics LETTERS
PUBLISHED ONLINE: 13 FEBRUARY 2017 | DOI: 10.1038/NPHYS4037

Learning phase transitions by confusion

Evert P. L. van Nieuwenburg*, Ye-Hua Liu and Sebastian D. Huber

- prediction-based method (PBM)

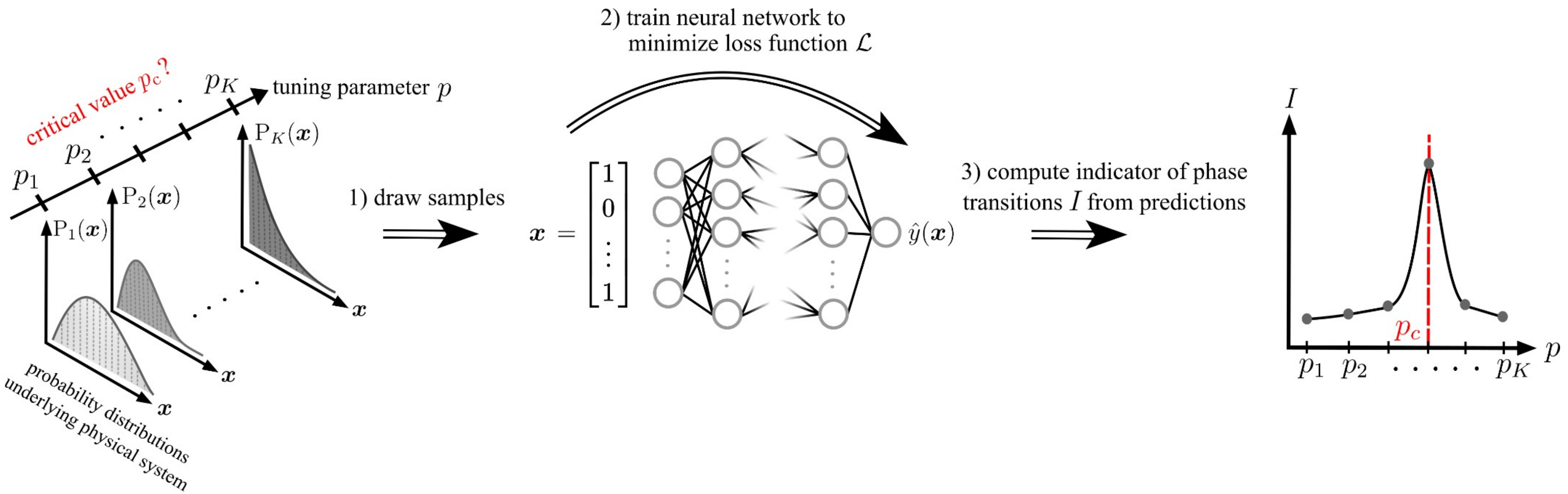
PHYSICAL REVIEW E **99**, 062107 (2019)

Vector field divergence of predictive model output as indication of phase transitions

Frank Schäfer and Niels Lörch*
Department of Physics, University of Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland

(Received 3 December 2018; revised manuscript received 16 May 2019; published 5 June 2019)

Detecting phase transitions using neural networks



What's the problem?

- methods were motivated in a heuristic fashion
- (deep) neural networks are difficult to interpret

  $\Rightarrow$ have limited understanding of their working principle

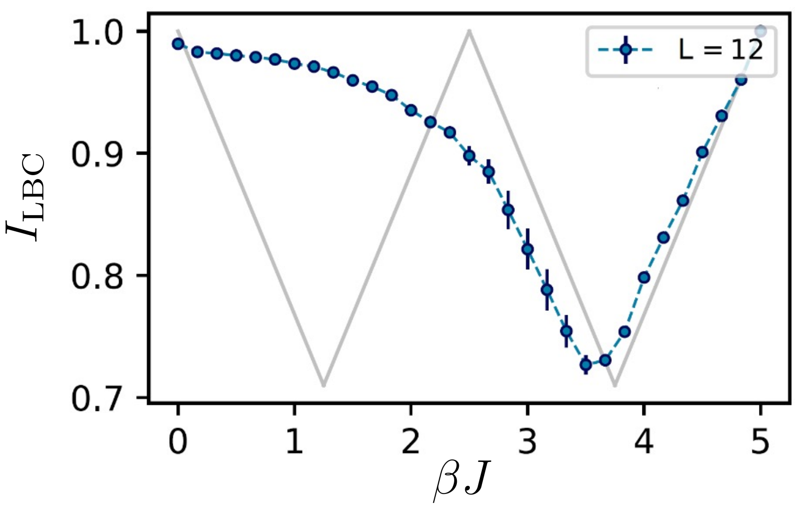
What's the problem?

- methods were motivated in a heuristic fashion
- (deep) neural networks are difficult to interpret

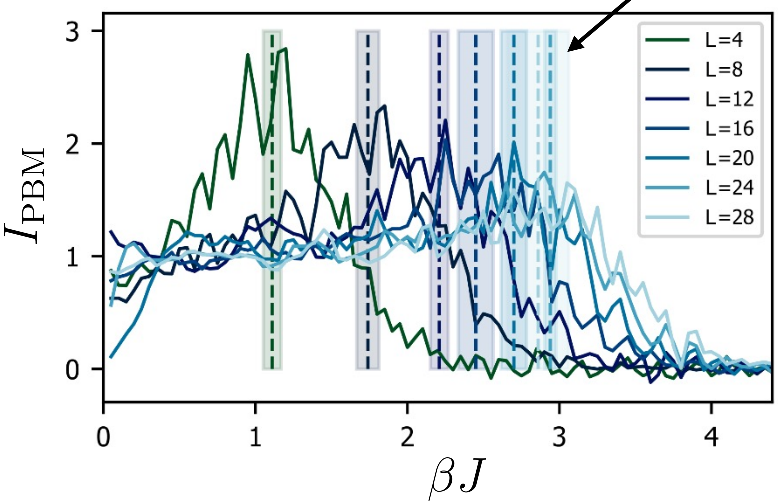
$\Rightarrow$ have limited understanding of their working principle

density-of-states model

unclear when and why they fail...



...or succeed



Greplova *et al.,* New J. Phys. **22** 045003 (2020)

What's the problem?

- methods were motivated in a heuristic fashion
- (deep) neural networks are difficult to interpret

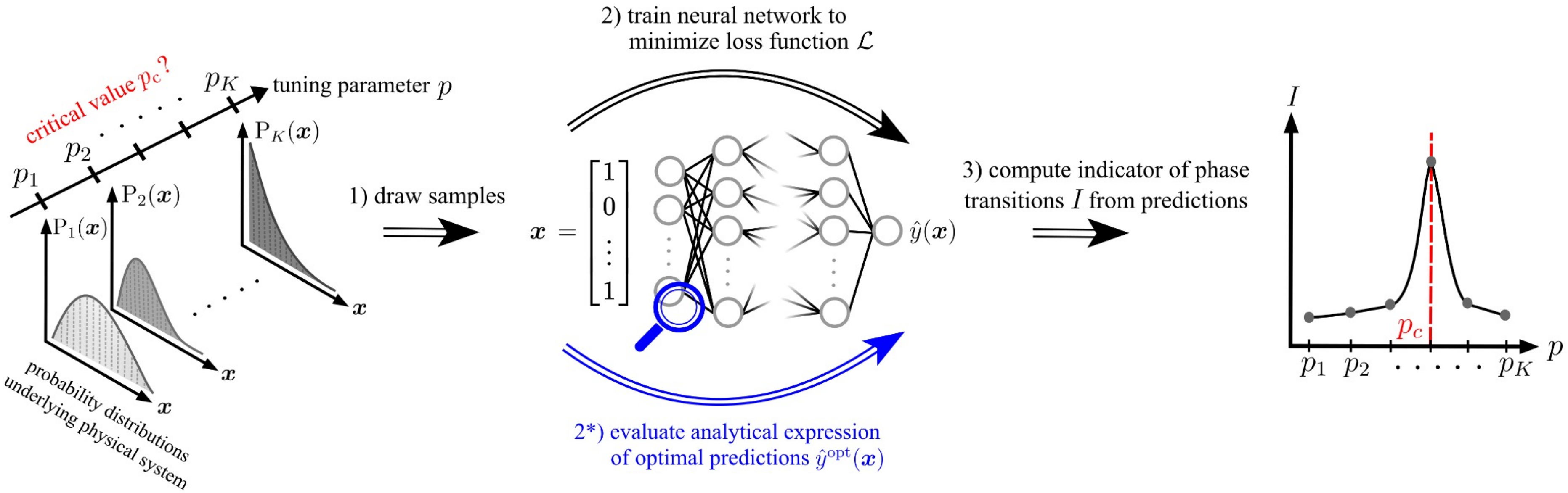
  $\Rightarrow$ have limited understanding of their working principle

high expressivity

$\Updownarrow$

low interpretability and high computational cost

Replacing neural networks



$$\arg\min_{\hat{y}(\boldsymbol{x})} \mathcal{L} = \hat{y}^{\text{opt}}(\boldsymbol{x})$$

$$\lim_{\text{expressive power} \to \infty} \hat{y}(\boldsymbol{x}) = \hat{y}^{\text{opt}}(\boldsymbol{x})$$

Optimal analytical predictors

- supervised learning $\hat{y}^{\text{opt}}_{\text{SL}}(\boldsymbol{x}) = \frac{\sum_{k\in \text{I}} \text{P}_k(\boldsymbol{x})}{\sum_{k=1}^{K} \text{P}_k(\boldsymbol{x})}$
- learning by confusion $\hat{y}^{\text{opt}}_{\text{LBC}}(\boldsymbol{x}) = \frac{\sum_{k\in \text{I}} \text{P}_k(\boldsymbol{x})}{\sum_{k=1}^{K} \text{P}_k(\boldsymbol{x})}$
- prediction-based method $\hat{y}^{\text{opt}}_{\text{PBM}}(\boldsymbol{x}) = \frac{\sum_{k=1}^{K} \text{P}_k(\boldsymbol{x})\, p_k}{\sum_{k=1}^{K} \text{P}_k(\boldsymbol{x})}$

$\Rightarrow$ gauge changes in underlying probability distributions $\{\text{P}_k\}_{k=1}^{K}$

analytical expressions reveal dependence of output on input data

Optimal analytical predictors

- supervised learning $\hat{y}_{\mathrm{SL}}^{\mathrm{opt}}(\boldsymbol{x}) = \frac{\sum_{k\in \mathrm{I}} \mathrm{P}_k(\boldsymbol{x})}{\sum_{k=1}^{K} \mathrm{P}_k(\boldsymbol{x})}$
- learning by confusion $\hat{y}_{\mathrm{LBC}}^{\mathrm{opt}}(\boldsymbol{x}) = \frac{\sum_{k\in \mathrm{I}} \mathrm{P}_k(\boldsymbol{x})}{\sum_{k=1}^{K} \mathrm{P}_k(\boldsymbol{x})}$
- prediction-based method $\hat{y}_{\mathrm{PBM}}^{\mathrm{opt}}(\boldsymbol{x}) = \frac{\sum_{k=1}^{K} \mathrm{P}_k(\boldsymbol{x})\, p_k}{\sum_{k=1}^{K} \mathrm{P}_k(\boldsymbol{x})}$

$\Rightarrow$ gauge changes in underlying probability distributions $\{\mathrm{P}_k\}_{k=1}^{K}$

analytical expressions reveal dependence of output on input data

$\Rightarrow$ use analytical expressions for probability distributions or estimate it based on drawn samples

can compute optimal indicators I^{opt} directly from data

Optimal analytical predictors

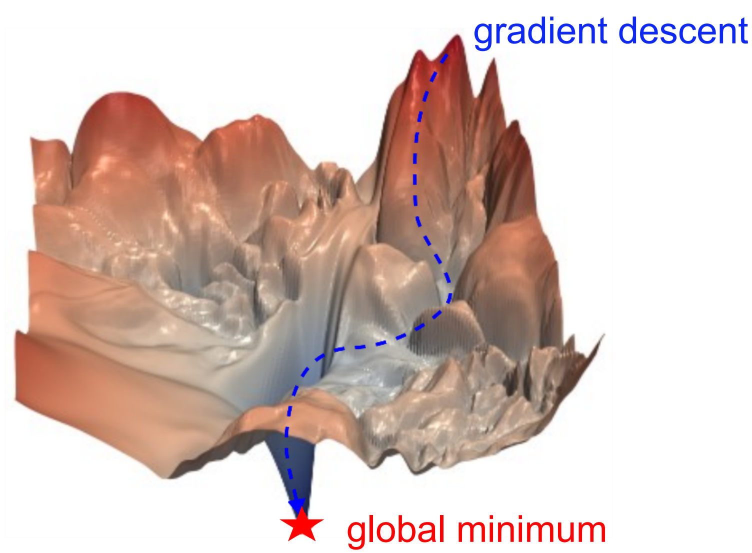
- require less computational resources compared to training neural networks

  ⇒ single training epoch $\gtrsim$ evaluation time of optimal indicators

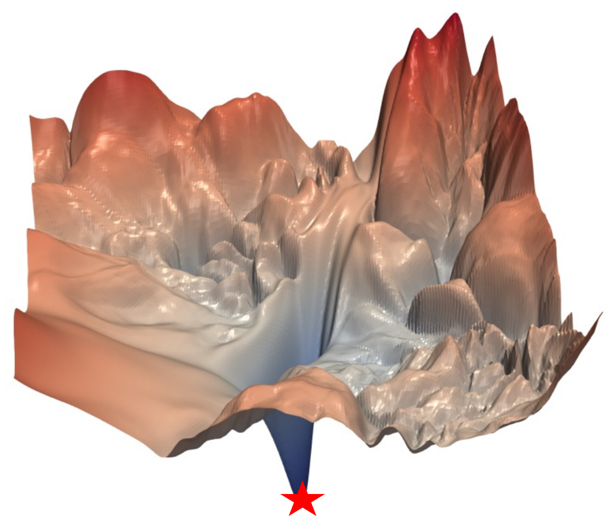
  ⇒ no need to tune hyperparameters

  ⇒ convergence is explicitly guaranteed

train neural networks



evaluate optimal predictors

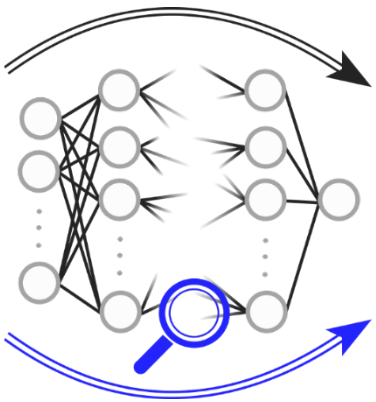


Li *et al.*, Adv. Neural Inf. Process. Syst. 31 (2018)

Bypassing the trade-off

high expressivity
⇕
low interpretability and high computational cost

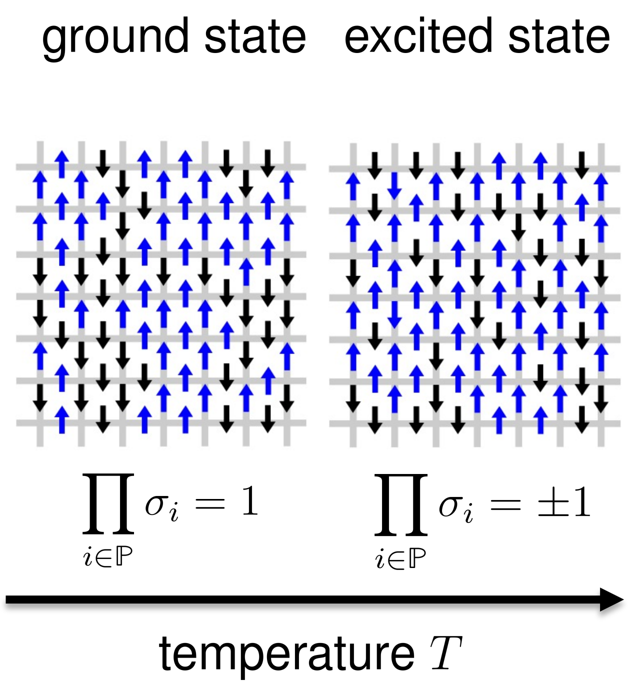
neural networks

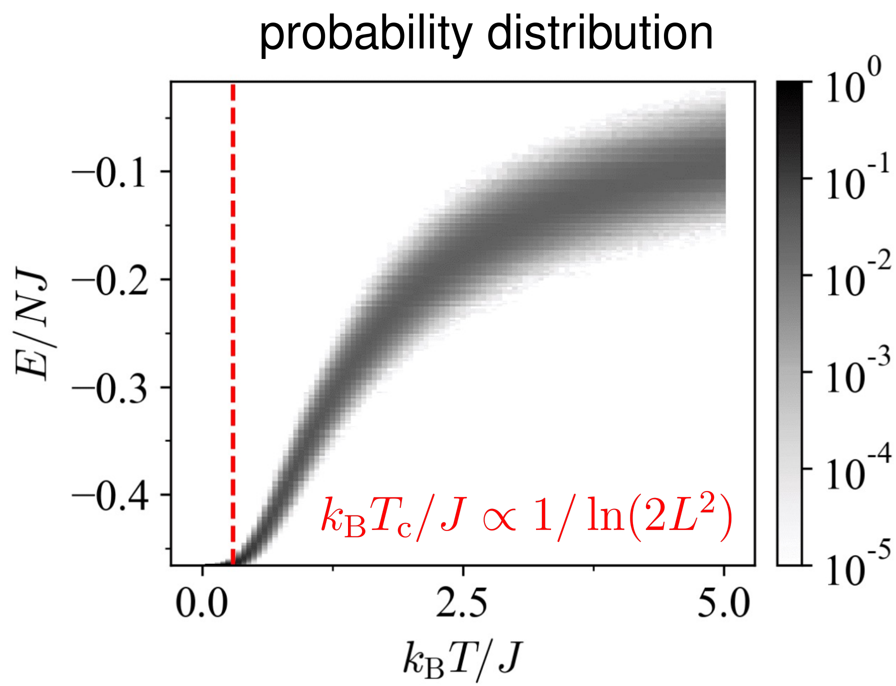


optimal analytical predictors

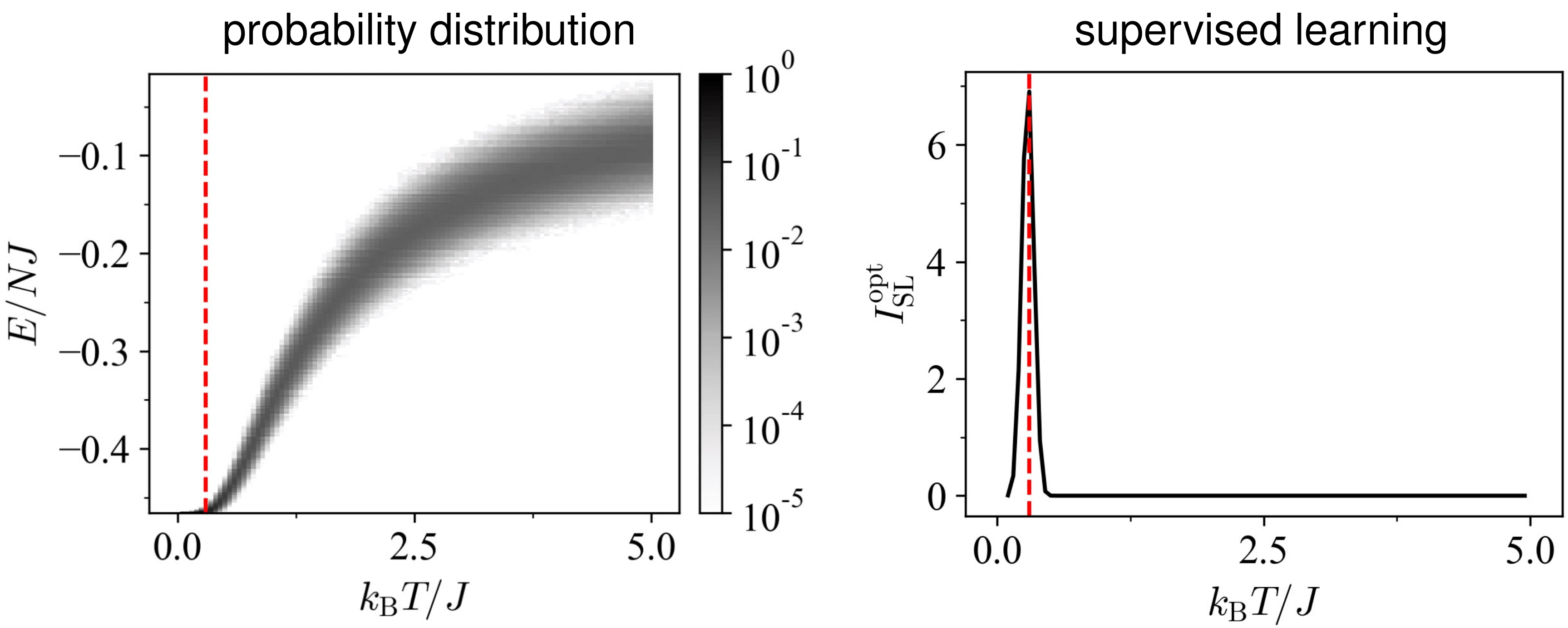
high expressivity
with
high interpretability at low computational cost

Example: Ising gauge theory





Example: Ising gauge theory

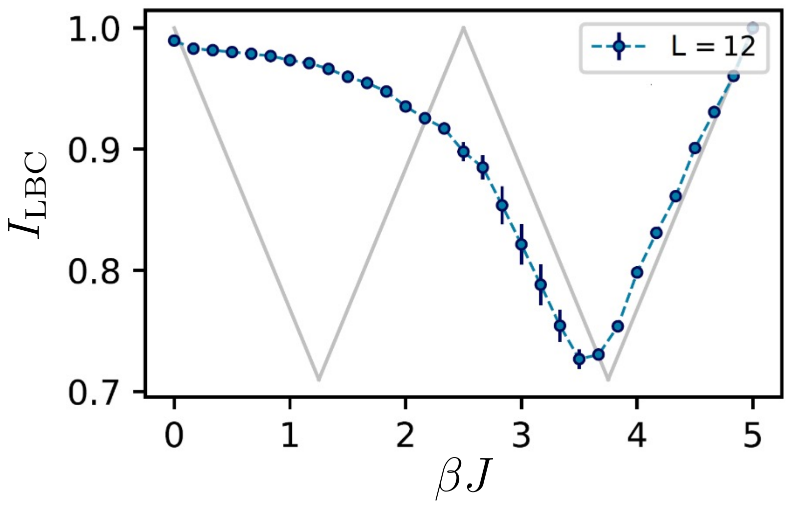


- for Boltzmann distribution: $\hat{y}_{\mathrm{SL}}^{\mathrm{opt}}(p_k) \propto \mathrm{P}_k(E_{\mathrm{gs}})$

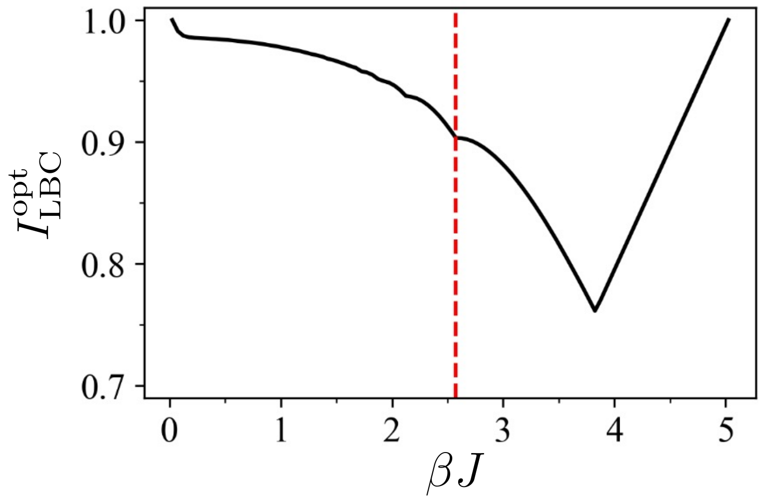
$\Rightarrow$ supervised learning tracks the relevant physical quantity

Example: Ising gauge theory

training neural networks



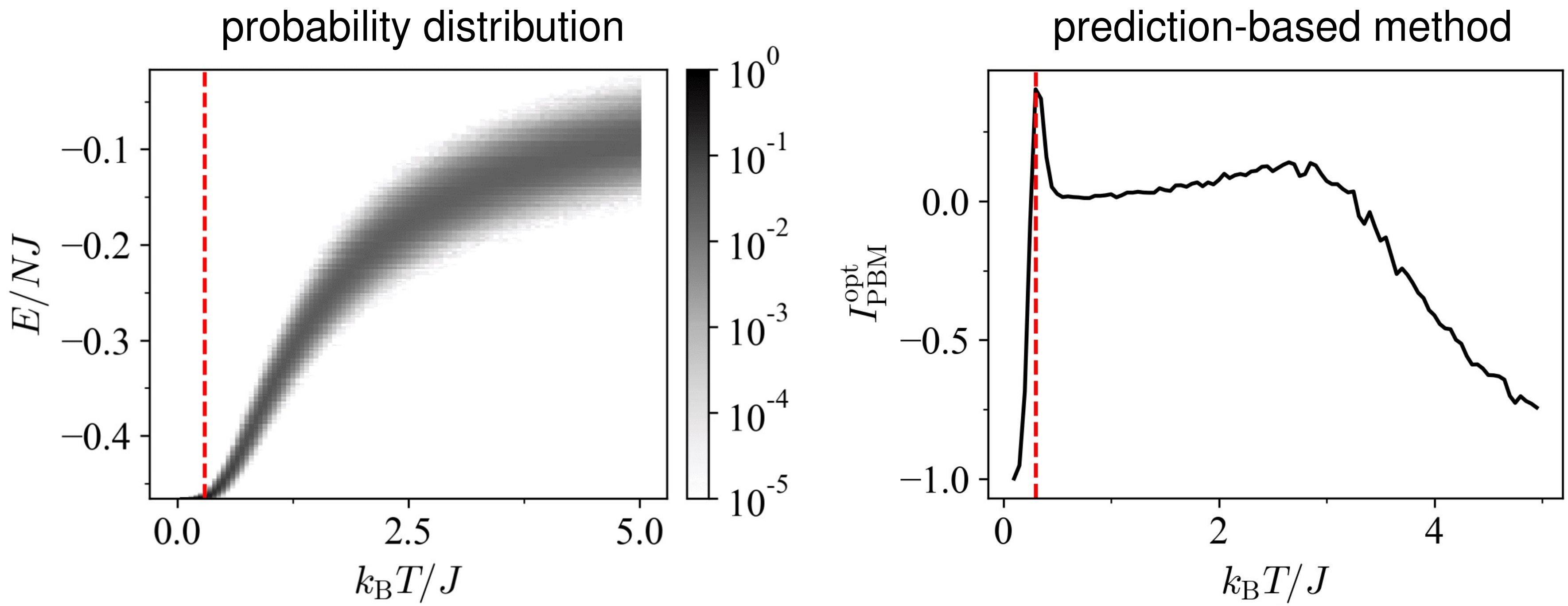
evaluating optimal predictors



$\Rightarrow$ learning by confusion fails in this setting

Greplova *et al.*, New J. Phys. **22** 045003 (2020)

Example: Ising gauge theory



- for Boltzmann distribution: $\hat{y}_{\mathrm{PBM}}^{\mathrm{opt}} \Leftrightarrow \hat{y}_{\mathrm{DOS}}$

$\Rightarrow$ prediction-based method is equivalent to density-of-states (DOS) model

Greplova *et al.,* New J. Phys. **22** 045003 (2020)

...and many more

- classical many-body systems

  $\Rightarrow$ symmetry-breaking phase transition in Ising model

  $\Rightarrow$ Berezinskii-Kosterlitz-Thouless transition in XY model

- quantum many-body systems

  $\Rightarrow$ first-order phase transition in XXZ chain

  $\Rightarrow$ topological phase transition in Kitaev chain

  $\Rightarrow$ Mott-insulator to superfluid transition in Bose-Hubbard model

  $\Rightarrow$ many-body localization transition in Bose-Hubbard model

Thank you
for your attention.

University of Basel

Bruder group

Code:

https://github.com/arnoldjulian/Replacing-neural-networks-by-optimal-analytical-predictors-for-the-detection-of-phase-transitions

Paper:

J. Arnold and F. Schäfer, Replacing neural networks by optimal analytical predictors for the detection of phase transitions, arXiv:2203.06084 (2022).