Study of the phase diagram of dense quark-gluon plasma from the first principles of the theory, enhanced by machine learning methods

Alexander Molochkov

Pacific Quantum Center, Vladivostok, Russia

in collaboration with

Vladimir Goy, Nikolai Gerasimenyuk, Sergei Liubimov

Pacific Quantum Center, Vladivostok, Russia

Maxim Chernodub

Institut Denis Poisson, Tours, France

Harold Erbin

Arnold Regge Center & University of Turin, Italy

Supported by RFBR grant 18-02-40121

Outline of the talk

— Motivation: mechanism of confinement with machine learning?

— Overview of machine learning and artificial neural networks

(here and in other places in the text)

— Recognizing confinement from dynamics of topological objects?

Example: compact U(1) gauge theory in (2+1)d as a toy model

[M.N. Chernodub, Harold Erbin, V.A. Goy, and A.V. Molochkov, Phys. Rev. D 102, 054501]

— ~~Guessing~~ Getting physics from unphysics with neural networks:

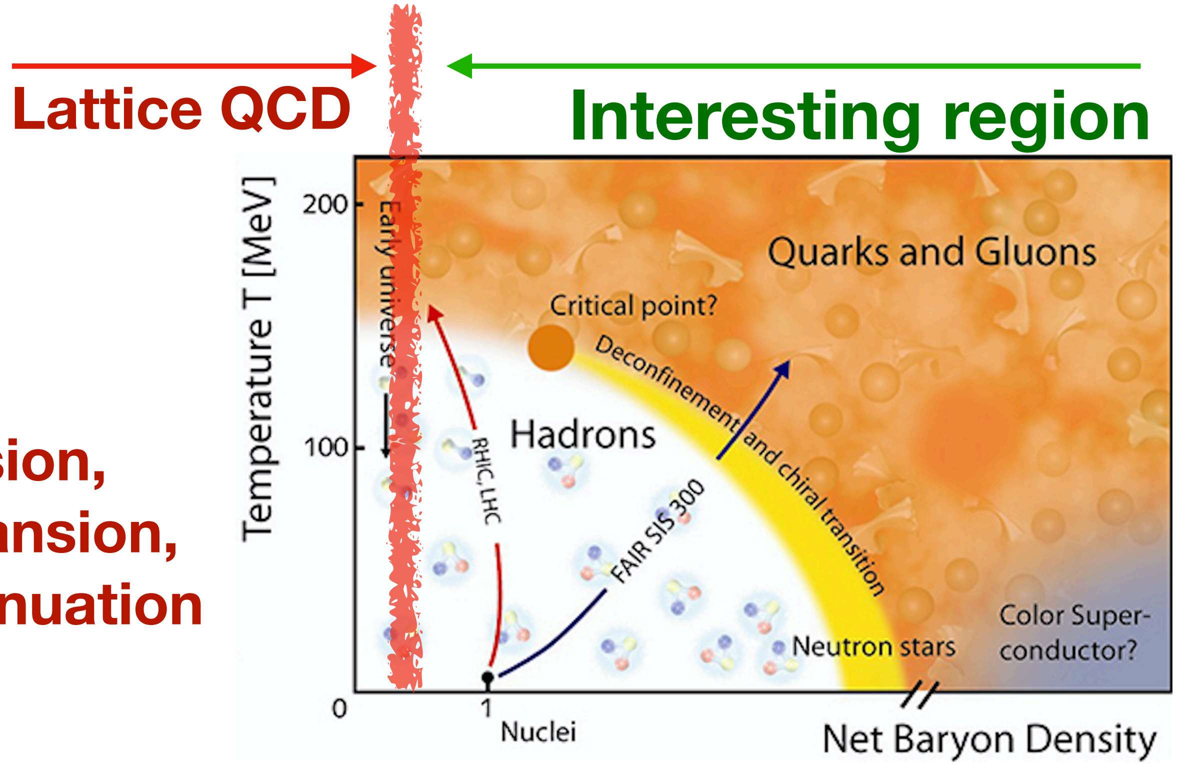
Finding physical deconfinement temperature in lattice Yang-Mills theories via a hint from outside the scaling window.

[D. Boyda, M. N. Chernodub, N. Gerasimenyuk, V. Goy, S. Lyubimov, and A. Molochkov, , arXiv:2009.10971 [hep-lat]]

— Future developments

Lattice QCD application to the finite density QGP study

Lattice QCD:
sign problem,
Taylor expansion,
Fugacity expansion,
analytic continuation



[from FAIR collaboration]

Patterns study via NN?

Analytic continuation via neural network?

In the machine learning field of computer science, artificial neural networks can successfully recognize and classify hidden patterns in (typically, huge) data sets.

NN feature: de-noising (technically) or renormalization (physically) large datasets

Many developments in the field (ML + QFT + …)

… +
1605.01735, Carrasquilla-Melko;
1608.07848, Broecker et al.;
1703.02435, Wetzel;
1705.05582, Wetzel-Scherzer;
1805.11058, Abe et al.;
1801.05784, Shanahan-Trewartha-Detmold;
1807.05971, Yoon-Bhattacharya-Gupta;
1810.12879, Zhou-Endrõdi-Pang;
1811.03533, Urban-Pawlowski;
1904.12072, Albergo-Kanwar-Shanahan;
1908.00281 Fukushima-Funai-Iida;
1909.06238, Matsumoto-Kitazawa-Kohno;
2004.14341 Bachtis-Aarts-Lucini
+ …

Machine learning and artificial neural networks

Features:

- pattern recognition (without explicit programming)
- flexible (wide range of applications)
- very general (no theory is needed, black box problem)
- for some, possesses some “creative” and “predictive” power

Applications, some of (General idea = pattern recognition):

- classification / clustering
- regression (prediction)
- transcription / translation
- anomaly detection
- de-noising
- synthesis and sampling

Applications in industry: computer vision, language processing, medical diagnosis, fraud detection, recommendation system, autonomous driving, …

Machine learning and artificial neural networks

Comparisons:

— results comparable and sometimes superior to human experts (cancer diagnosis, traffic sign recognition. . .)

— generally, performance is much better compared to any other machine algorithm

Drawbacks:

— black box

— magic

— numerical

(= how to extract analytical / exact results?)

For us, Machine Learning algorithms will upgrade the standard Monte Carlo

Deep neural network

Layered structure

(1) Input layer is the first layer of neurons which receives data (bits that encode colored pixels from an image, values of gauge fields in an MC-generated SU(N) field configuration)

(2) Hidden layer is a layer of neurons that gets information from the input layer, modifies it, and passes to the output (or, next) layer.

Terminology: A neural network that contains more than one hidden layer is a deep neural network.

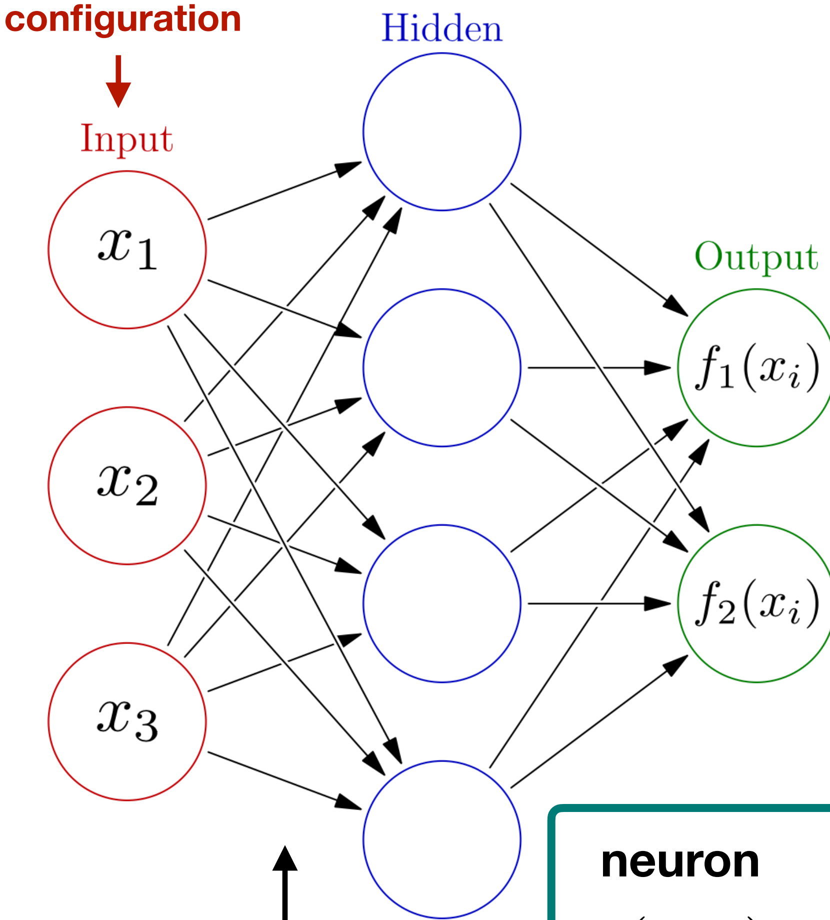
(3) The final layer of a neural network which contains the answer(s).

Sometimes the answer is “42”.

→ Link: is a matrix operation which acts on a vector from n^{th} layer and gives an input to the $(n+1)^{\text{th}}$ layer. The matrix elements are weights to be tuned.

for example, a gluon (QCD) configuration

hidden layer(s) of neurons



For example:

$$\begin{cases} 0 = \text{confinement} \\ 1 = \text{deconfinement} \end{cases}$$

value of $\langle \bar{\psi}\psi \rangle$

linear (matrix) operation

$$\vec{y}^{\,(n)} = \mathbb{W}^{(n)} \vec{x}^{\,(n)}$$

(out) (in)

neuron

$$x_i^{(n+1)} = G_i^{(n+1)}(y_i^{(n)})$$

(out) (in)

Activation function. Examples:

— ReLU

$$G(\xi) = \xi\Theta(\xi) \equiv \begin{cases} \xi, & \xi > 0 \\ 0, & \text{otherwise} \end{cases}$$

— Sigmoid

$$G(\xi) = \frac{1}{1+e^{-\xi}} \in (0,1)$$

Deep neural network

Operations:

— **convolution** = Reduces matrix (database) size.
Takes a convolution ("selective blurring") across an area.

— **pooling** = Reduces matrix (database) size.
Takes the maximum / average value across the pooled area.

— **dropout** = A form of regularization useful in training neural networks.
Removes a random selection of a fixed number of the units

Supervision:

— supervised network: teach network at a set of examples using mean squared error as a validation (= examination)

— un-supervised network: leave network alone and let it classify features of the system (of the database) itself.

Question: can we “learn confinement” in QCD?

— Big aim: Assuming that some (topological, semi-classical) structures are responsible for confinement, can we find (determine, describe) them in gluon configurations?

— Modest (intermediate) aim: first take a well-known toy model; make test, tune the artificial network, and compare with (expected) results.

Compact U(1) gauge theory in (2+1)d

Similarities with SU(N) Yang-Mills theories:

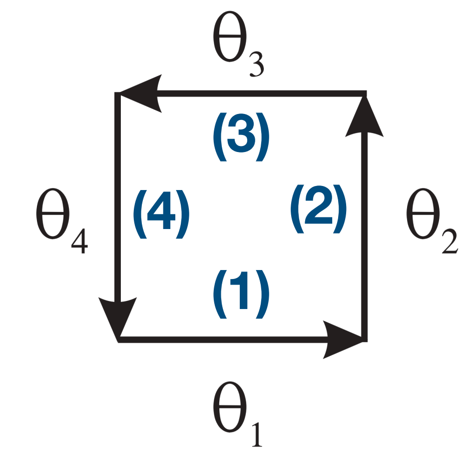
- Mass gap generation
- Linear confinement
- Presence of instantons (instanton-like objects)
- Finite–temperature deconfinement

cU(1) gauge theory on the lattice

- **Lattice action**

gauge coupling $\beta = \dfrac{1}{g^2 a}$

$$S[\theta] = \beta \sum_P (1 - \cos\theta_P)$$

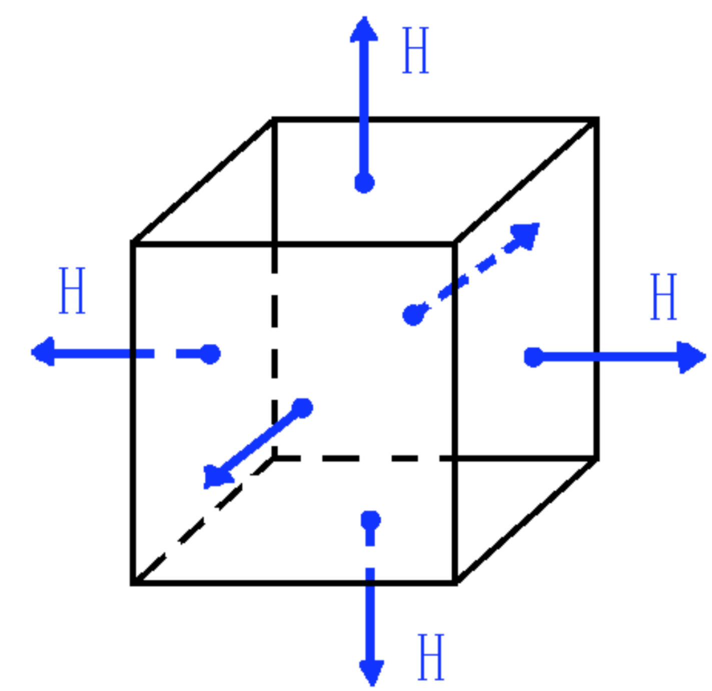


- **Plaquette angle**

$$\theta_{P_{x,\mu\nu}} = \underset{(1)}{\theta_{x,\mu}} + \underset{(2)}{\theta_{x+\hat{\mu},\nu}} - \underset{(3)}{\theta_{x+\hat{\nu},\mu}} - \underset{(4)}{\theta_{x,\nu}}$$

- **Compactness**

$$\theta_{x,\mu} \in [-\pi, +\pi)$$



- **Density of abelian monopoles**

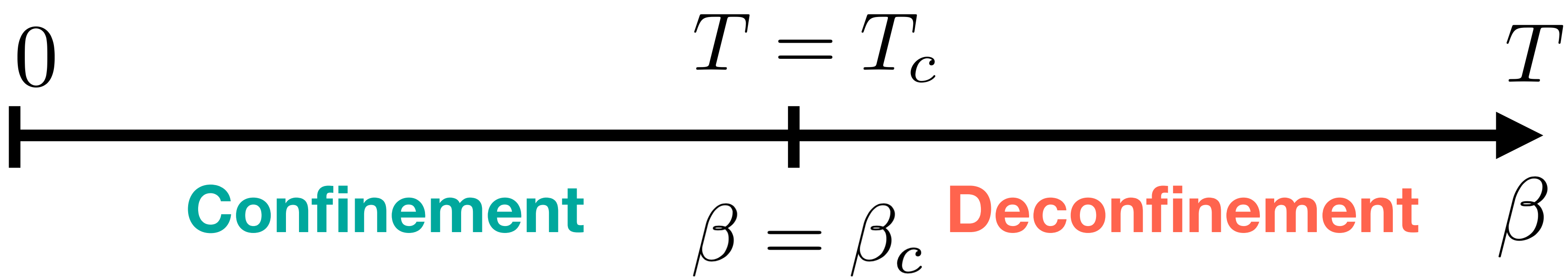
$$\rho_x = \frac{1}{2\pi} \sum_{P \in \partial C_x} \bar{\theta}_P$$

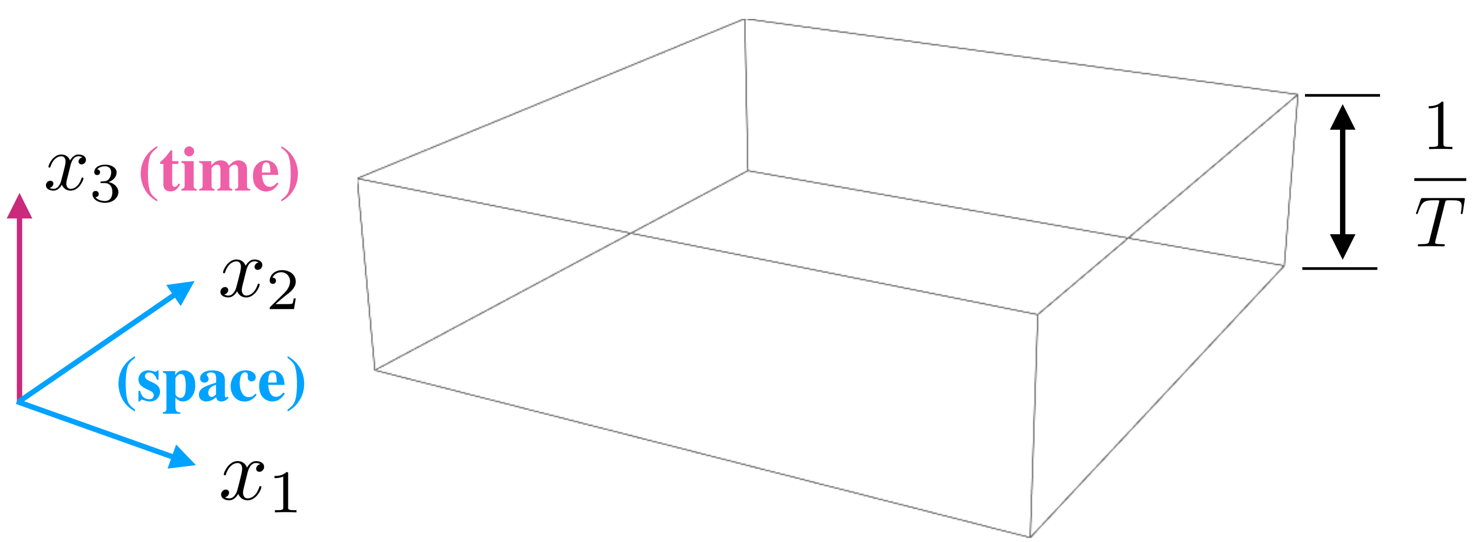
$$\bar{\theta}_P = \theta_P + 2\pi k_P \in [-\pi, \pi), \qquad k_P \in \mathbb{Z}$$

Monopoles are defined on the cubes of the lattice

cU(1) gauge theory on the lattice

– Finite-temperature deconfinement transition





lattice size $N_t \times N_s \times N_s$

lattice spacing a

lattice coupling

$$\beta = \frac{1}{g^2 a}$$

temperature

$$T = \frac{1}{N_t a}$$

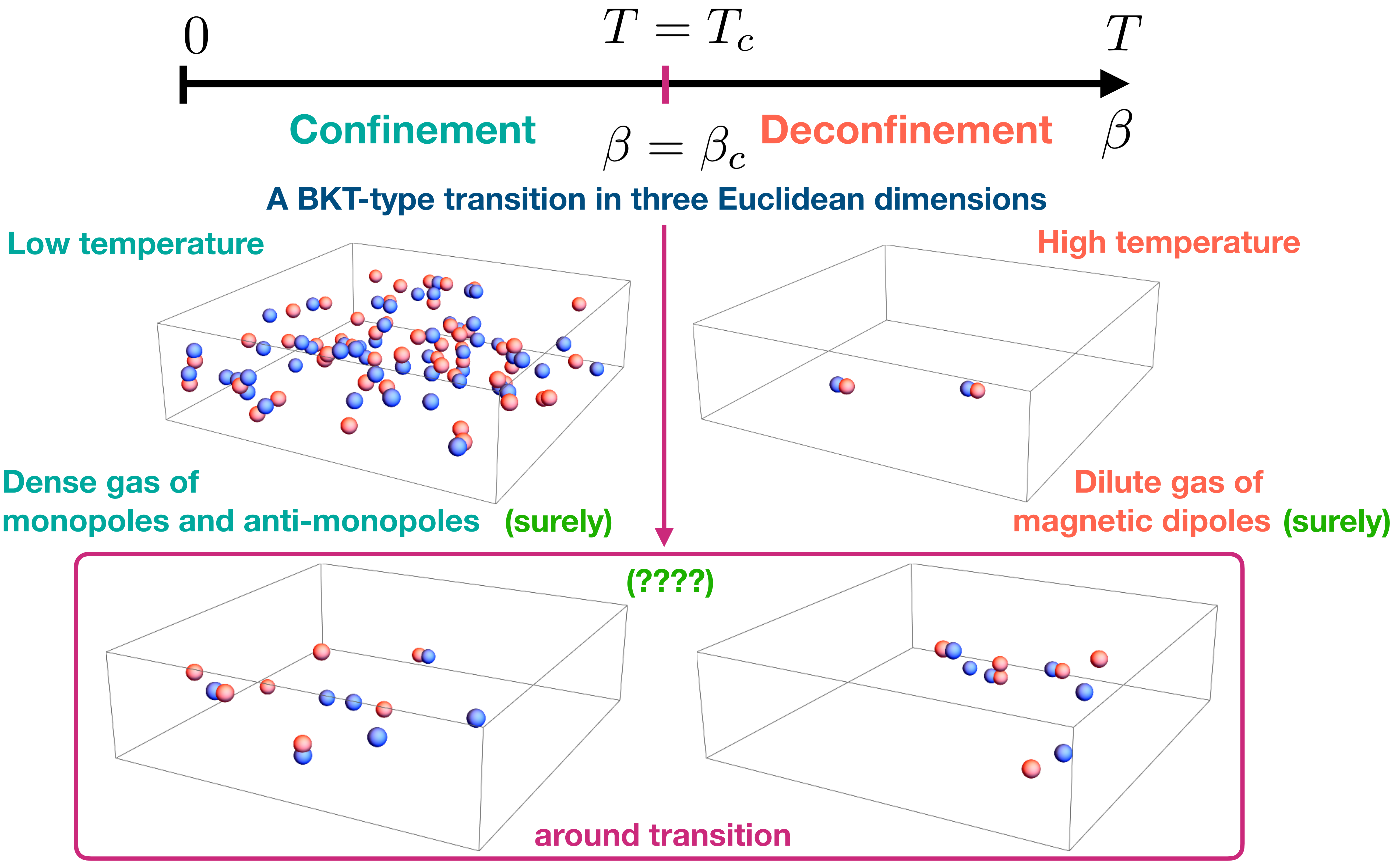
temperature

$$\frac{T}{g^2} = \frac{\beta}{N_t}$$

continuum | lattice

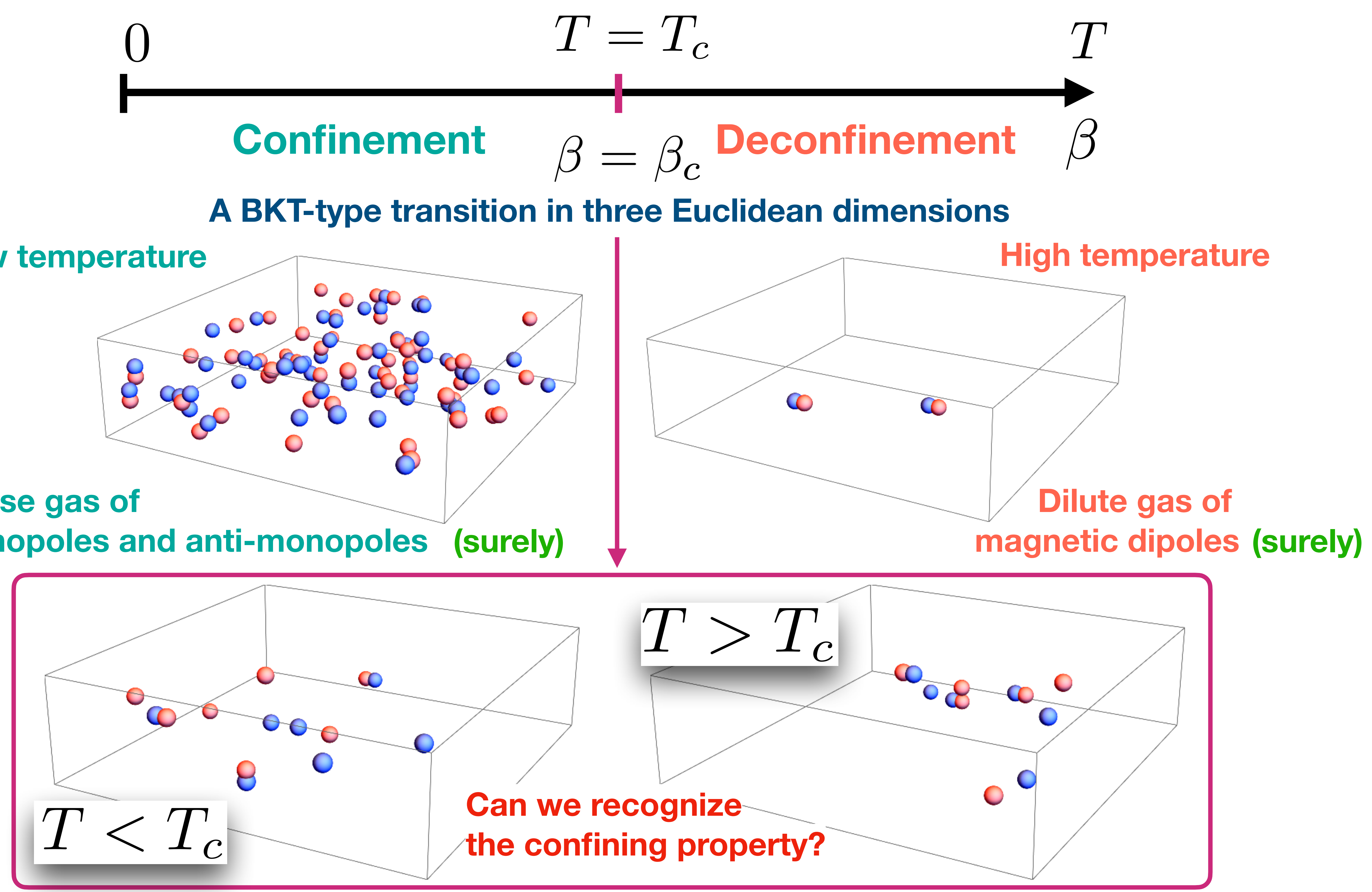
cU(1) gauge theory on the lattice

— Finite-temperature deconfinement transition seen by/via monopoles



cU(1) gauge theory on the lattice

— Finite-temperature deconfinement transition seen by/via monopoles



Machine learning monopoles and confinement

We should learn how to machine learn physical effects:

— vast diversity of different architectures

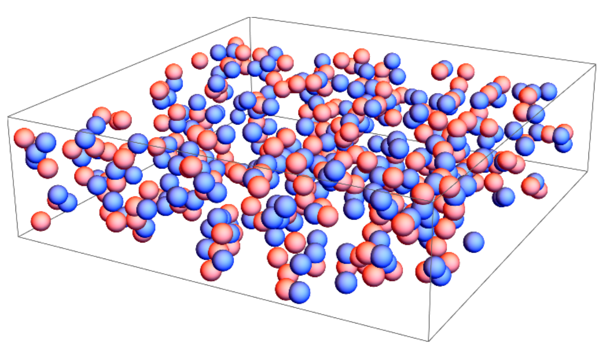
— choice of architecture depends on a specific problem (empirical)

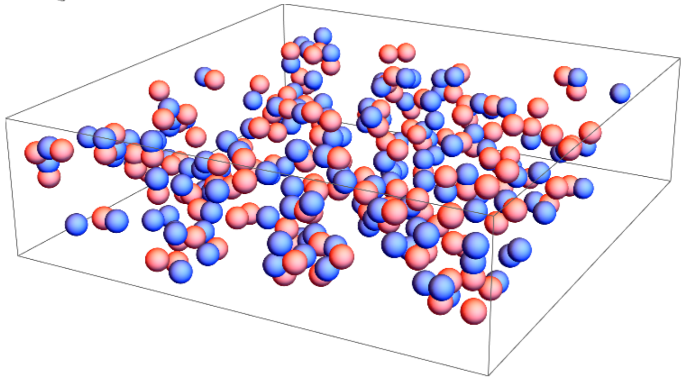
Our objectives:

— train ANN at small lattice volumes (Lt,Ls) = (4,16)

— predict for larger lattices (Lt = 4,6,8, Ls = 16,32):

- the nature of the phase (confinement/deconfinement)
- values of Polyakov loop, $|L|$
- the critical temperature, T_c





Input:

— Lattice monopole configurations

Neural network:

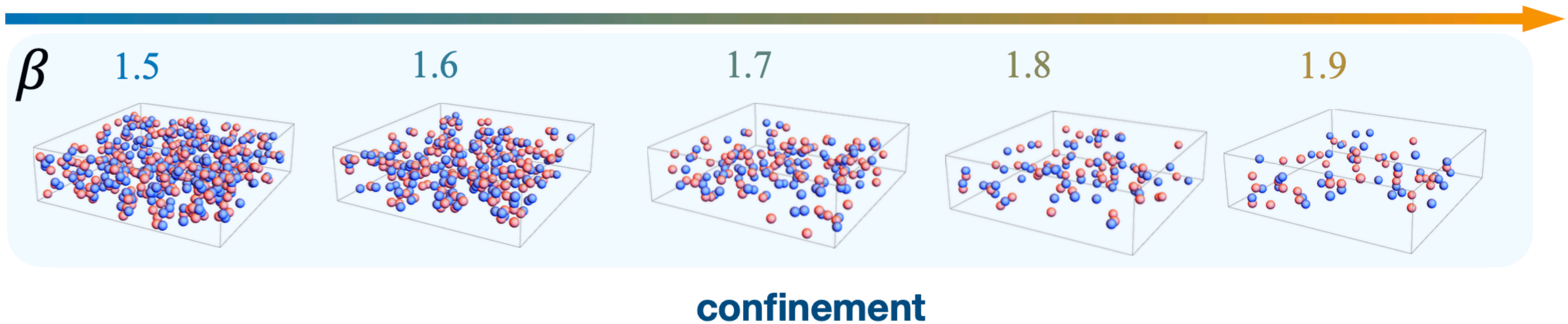
— network: convolution + dense layers

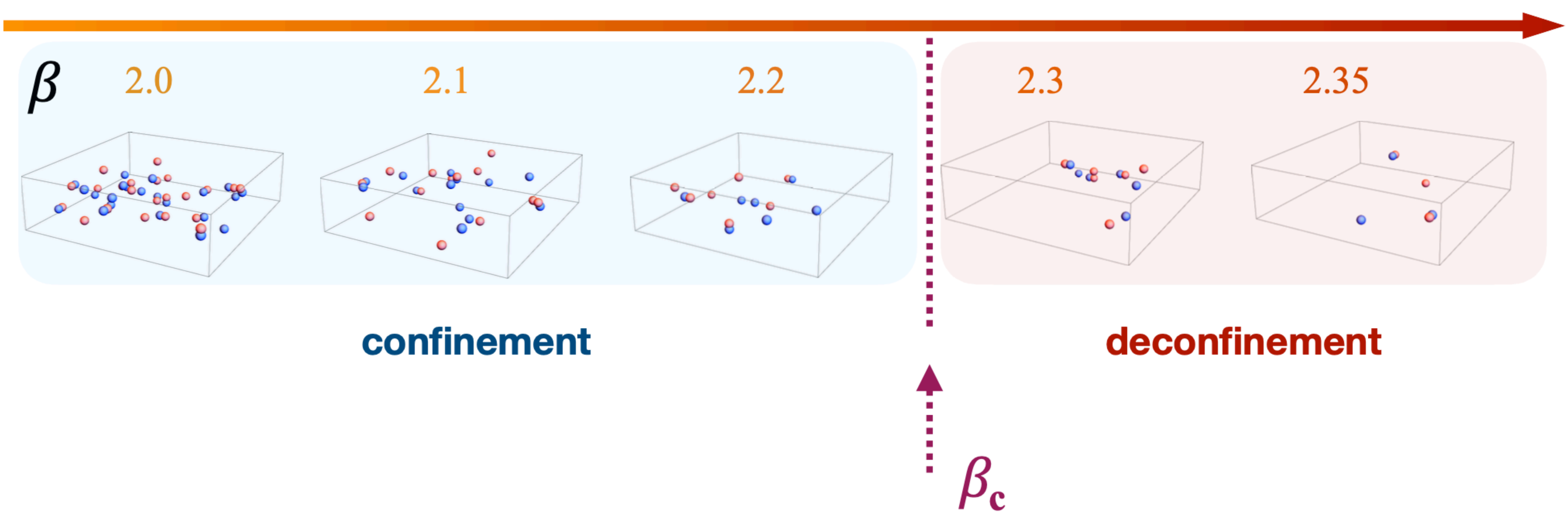
— supervised learning technique

— 1.28M parameters … more than just fitting

Machine learning monopoles and confinement

Classifying the phases using the monopole configurations

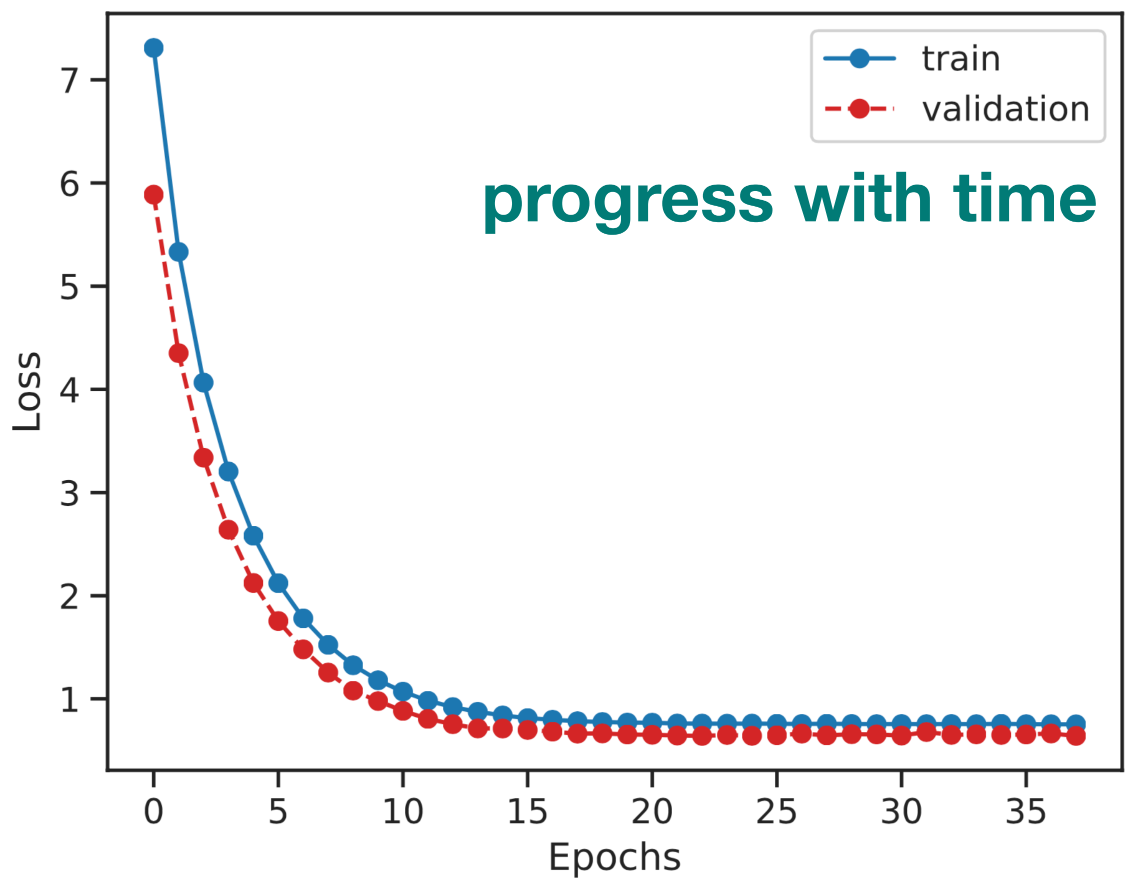




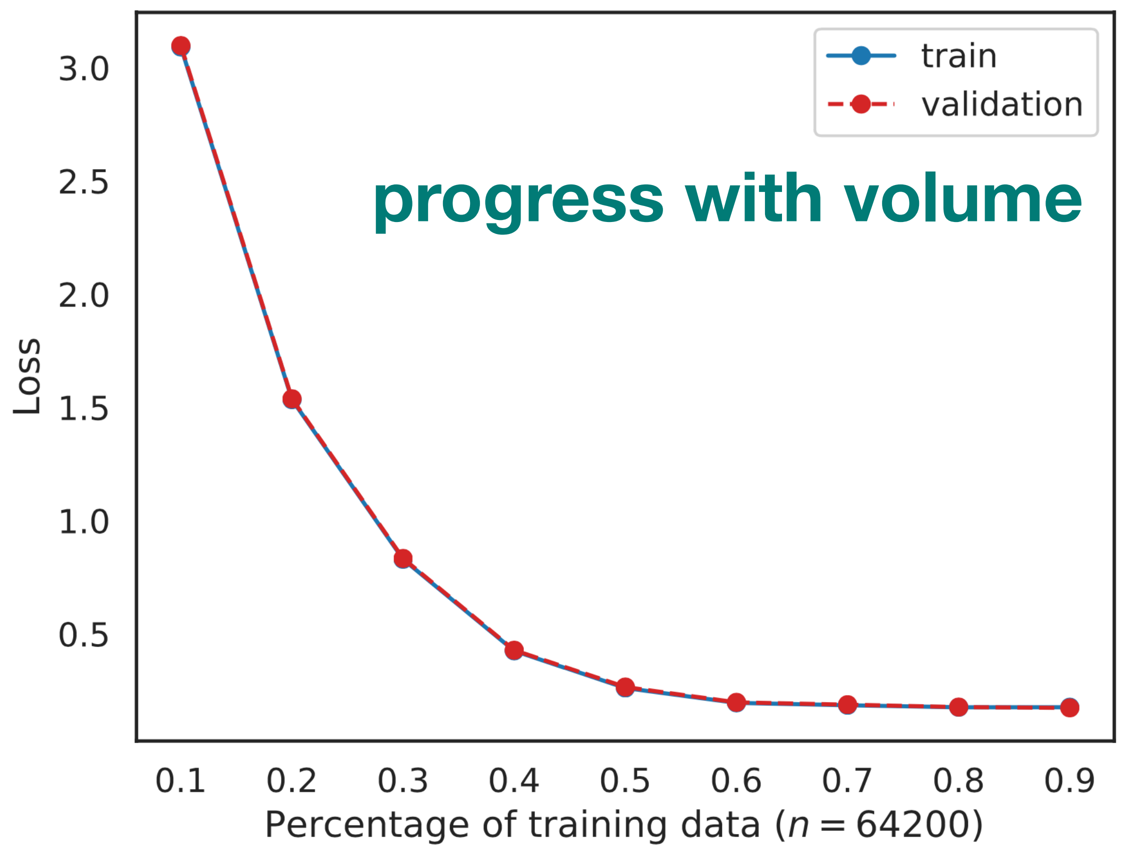
Machine learning monopoles and confinement

How fast does this ANN learn?

Training curve



Learning curve



Epoch = A full "training course" using the full dataset

Uses the **Adam optimization** algorithm (an extension to stochastic gradient descent)

training: about 2000 configurations
validation: 200 configurations

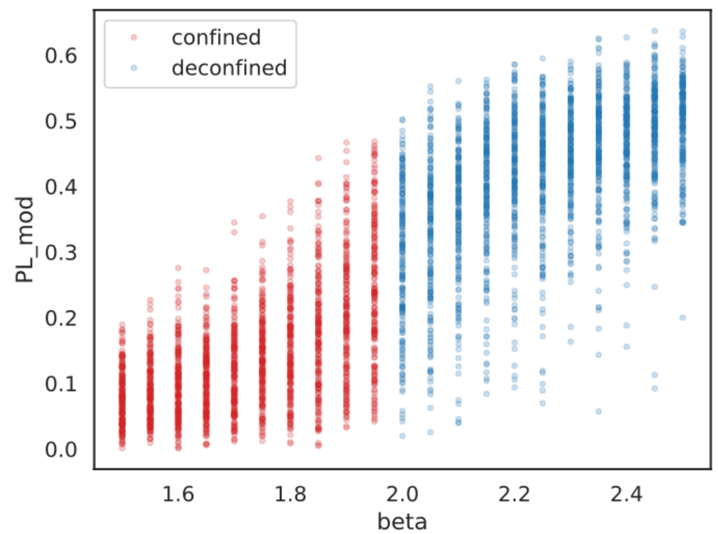
Machine learning monopoles and confinement

The neural network was trained at a small volume and asked to make predictions at larger volumes (never seen configurations).

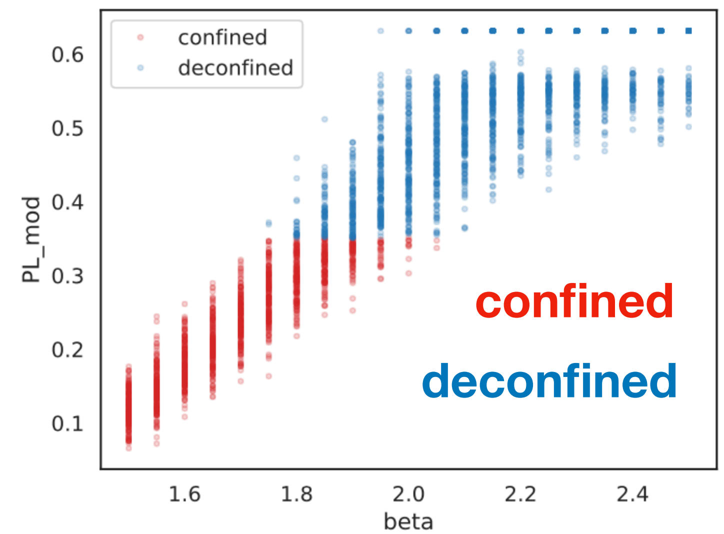
Predicting the value of the Polyakov loop and the phase

Monte Carlo (original)

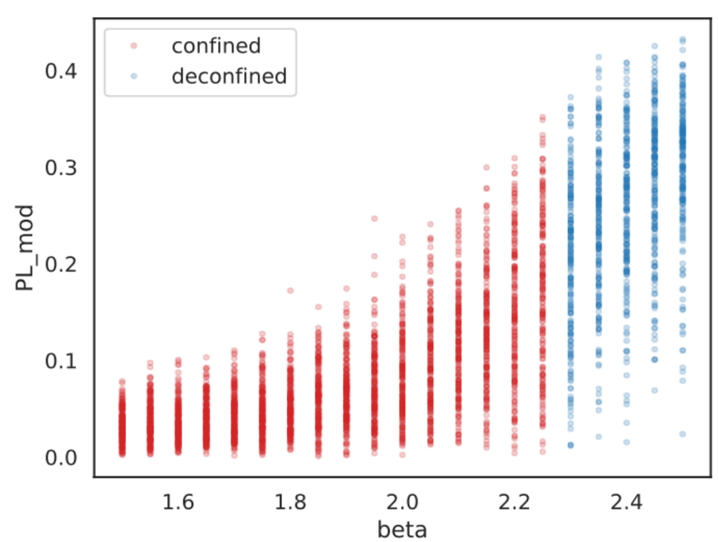
Artificial Neural Network (predicted)



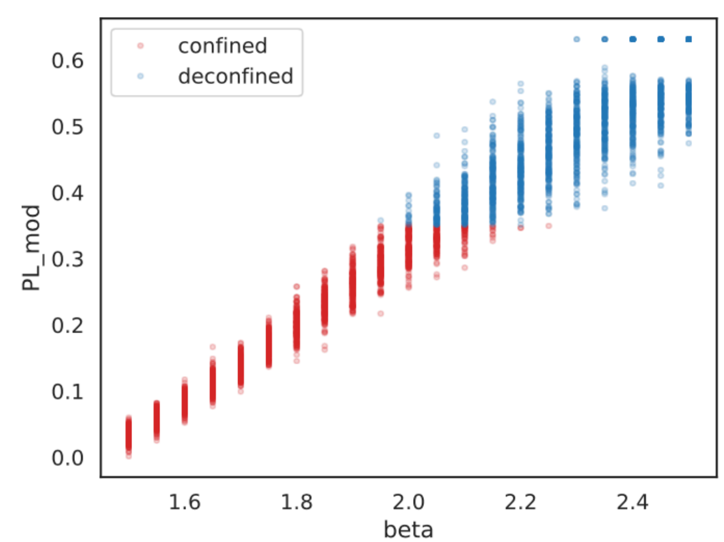
(a) 6×16^2 – real



(b) 6×16^2 – predicted



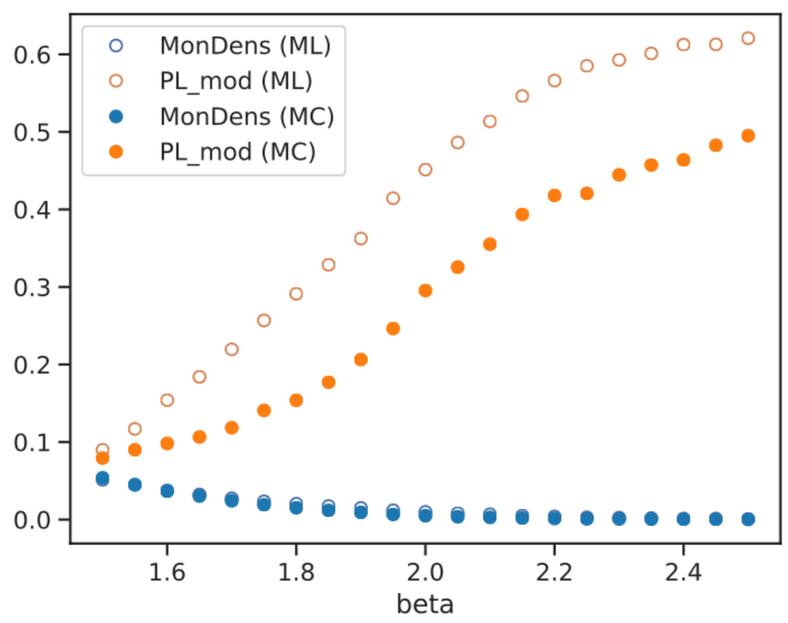
(c) 8×32^2 – real



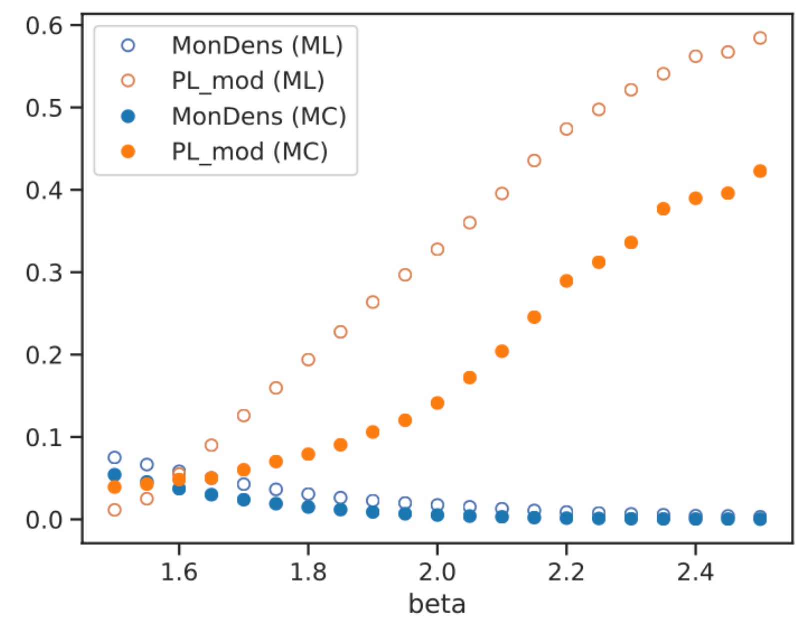
(d) 8×32^2 – predicted

Machine learning monopoles and confinement

Predicting the value of the Polyakov loop and the monopole density

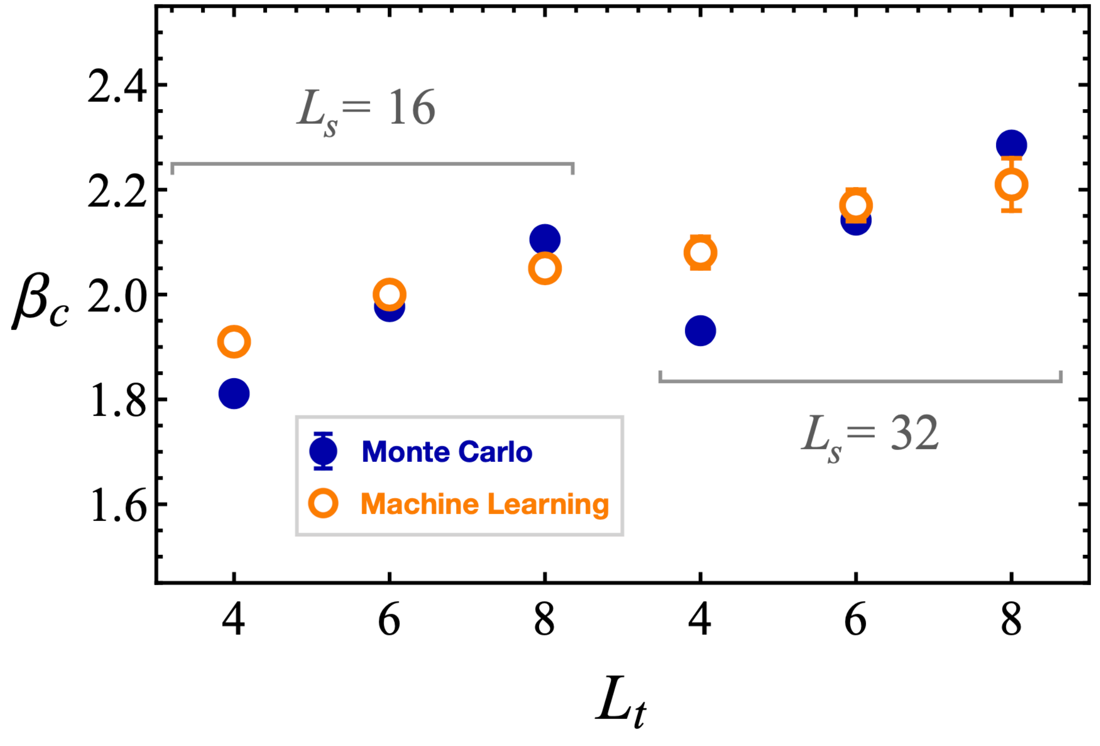


(c) 6×16^2



(d) 6×32^2

Predicting the critical temperature



Locating the phase transition by the degree of the confusion experienced by the neural network

Machine learning monopoles and confinement

Summary. The artificial neural network (ANN):

— uses the supervised learning technique to acquire knowledge about monopole configurations generated by the standard Monte-Carlo technique

— processes the monopole configurations as 3d holograms

— studies how to associate these monopole holograms with the vacuum expectation value of the Polyakov loop

— after training, uses the monopole configurations at larger-volume lattices to distinguish confinement and deconfinement phases

— neglects the renormalization effects: while the predicted Polyakov loop differs from the original order parameter, the critical inflection points are close to each other.

— the best criterion for locating the phase transition: the degree of the confusion experienced by the neural network.

the last point agrees with [1610.02048]

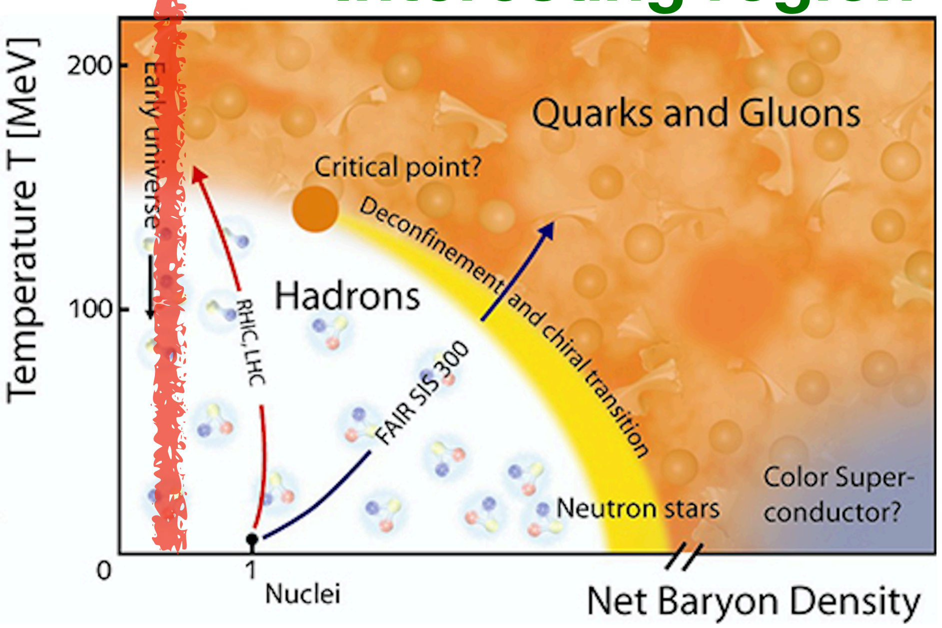
Learning physics from “unphysics”

Motivation: we know very-well a result at an unphysical or not-so-interesting point, but we want to get some information about a physical/interesting point.

Example: finite baryonic chemical potential in QCD

Lattice QCD → ← **Interesting region**

Lattice QCD: sign problem, Taylor expansion, analytic continuation



[from FAIR collaboration]

Analytic continuation via neural network?

Important feature: de-noising (technically) or renormalization (physically). The neural network removes uninteresting UV fluctuations and leaves only non-perturbative physics
→ the Polyakov loop in compact electrodynamics

Learning physics from “unphysics”

Simplified example: lattice Yang-Mills at strong coupling

[D. Boyda, N. Gerasimenyuk, V. Goy, S. Lyubimov, A. Molochkov, M.C., to appear]

Input: We know the value of the order parameter at unphysical coupling(s)

- bad version No. 1: lattice Yang-Mills at very strong coupling,
  no relation to the continuum limit

- bad version No. 2: lattice Yang-Mills at very weak coupling,
  purely perturbative finite-volume physics

Output: We want to know the value of the order parameter for any value of the coupling including the scaling region in lattice Yang-Mills theory

We take the bad choice No. 2:
SU(N_c) gauge theory on the lattice with β=4 (for N_c=2) and β=10 (for N_c=3);
on the lattices with spatial extensions N_s=8,16,32

Our choices are really very bad: physical spatial lattice size is ~0.01 fm.

P.S. We also checked that a good ANN can treat the bad choice No. 1 very well as well.

Learning physics from “unphysics”

Architecture of the neural network

Layer		Structure
InputLayer	In	$(N_t = 4,\ N_s{\times}N_s,\ N_s,\ \mathrm{Dim}{\times}\mathrm{U})$
	Out	$(N_t = 4,\ N_s{\times}N_s,\ N_s,\ \mathrm{Dim}{\times}\mathrm{U})$
Conv3D	In	$(4,\ N_s{\times}N_s,\ N_s,\ \mathrm{Dim}{\times}\mathrm{U})$
	Out	$(2,\ N_s{\times}N_s,\ N_s,\ 256)$
Conv3D	In	$(2,\ N_s{\times}N_s,\ N_s,\ 256)$
	Out	$(1,\ N_s{\times}N_s,\ N_s,\ 32)$
AveragePooling3D	In	$(1,\ N_s{\times}N_s,\ N_s,\ 32)$
	Out	(1, 1, 1, 32)
Flatten	In	(1, 1, 1, 32)
	Out	(32)
Dense	In	(32)
	Out	(1)

Architecture of the neural network for the prediction of the Polyakov Loop in the SU(N) gauge theory with the temporal size of the lattice $N_t = 4$. Here Dim is dimension of theory, U is dimension of vector representation

Input: raw configuration of SU(N) gauge field.

Hidden layers

Output: expectation value of the Polyakov loop

Train in the unphysical region:
known value of the Polyakov loop for a set of SU(N) configurations at unphysically small lattices ~$(0.01\ \mathrm{fm})^3$

Lattice configuration representation for ANN

SU(2):

$$U = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \equiv \begin{pmatrix} a_1 + ia_2 & a_3 + ia_4 \\ -a_3 + ia_4 & a_1 - ia_2 \end{pmatrix} \rightarrow \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}$$

$$\{U_\mu(t,x,y,z)\} \rightarrow [N_t, N_s, N_s, N_s, Dim, 4]$$

SU(3):

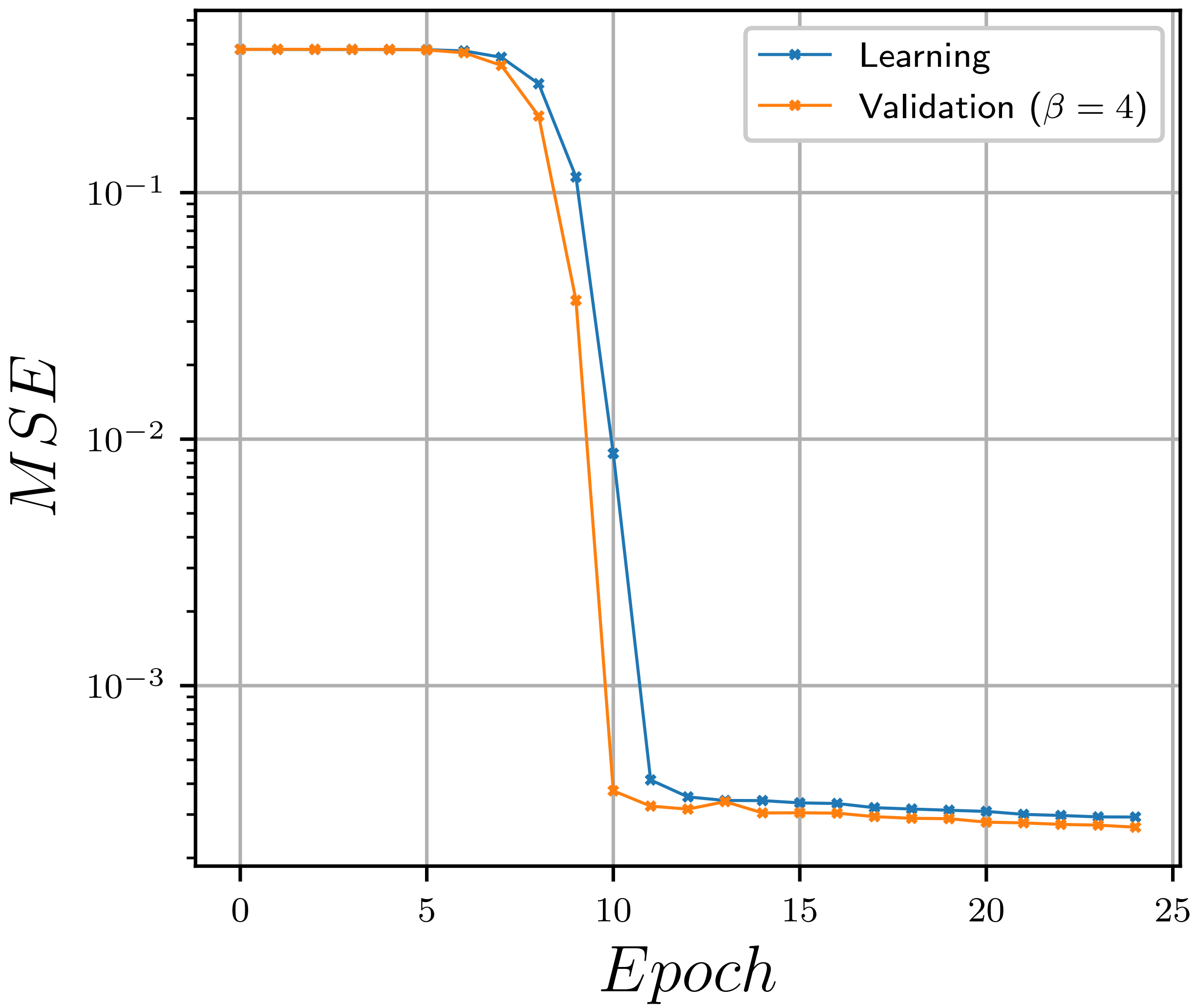
$$U = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix} \equiv \begin{pmatrix} a_1 + ia_2 & a_3 + ia_4 & a_5 + ia_6 \\ a_7 + ia_8 & a_9 + ia_{10} & a_{11} + ia_{12} \\ a_{13} + ia_{14} & a_{15} + ia_{16} & a_{17} + ia_{18} \end{pmatrix} \rightarrow \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_{18} \end{pmatrix}$$

$$\{U_\mu(t,x,y,z)\} \rightarrow [N_t, N_s, N_s, N_s, Dim, 18]$$

Dimension of the set reduction:

$$[N_t, N_s, N_s, N_s, Dim, 18] \rightarrow [N_t, N_s \times N_s, N_s, N_s, Dim \times U]$$

Learning physics from “unphysics”

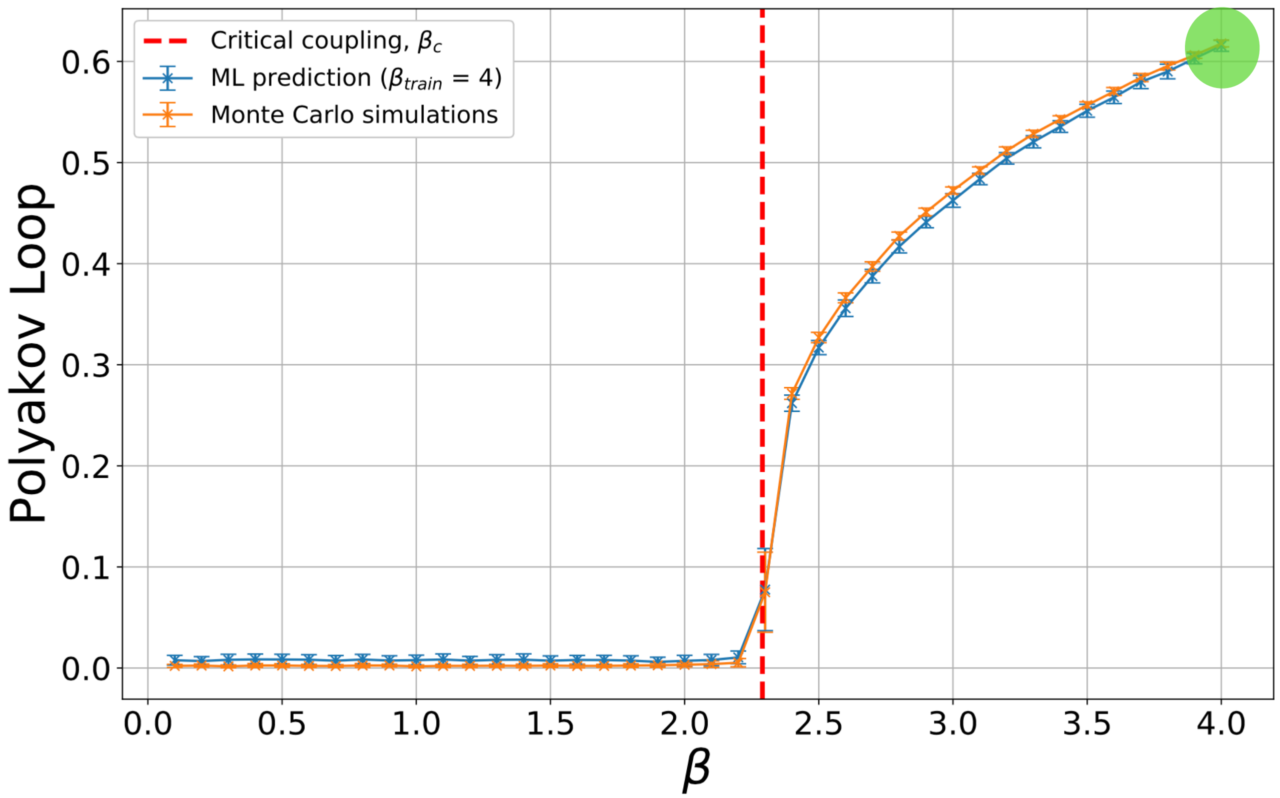


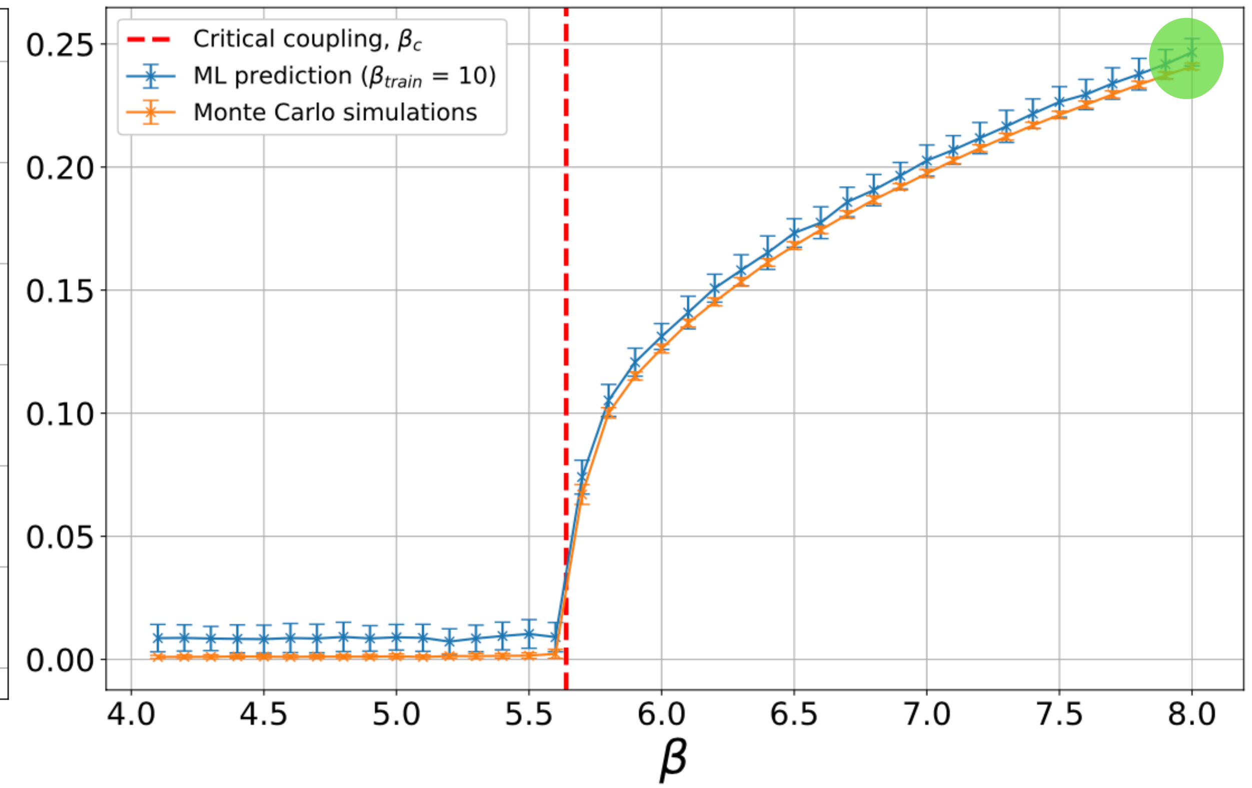
Learning curves for training and validation at the point β = 4 of the SU(2) gauge theory on $16^3 \times 4$ lattice with the mean squared error (MSE) used as a loss function. The MSE normalized on the value of the order parameter squared, $\langle |L| \rangle^2$, gives qualitatively the same picture.

The machine-learning algorithm finds a function of the lattice configuration parameters that correlates with the Polyakov loop.

Learning physics from "unphysics"

Results:*



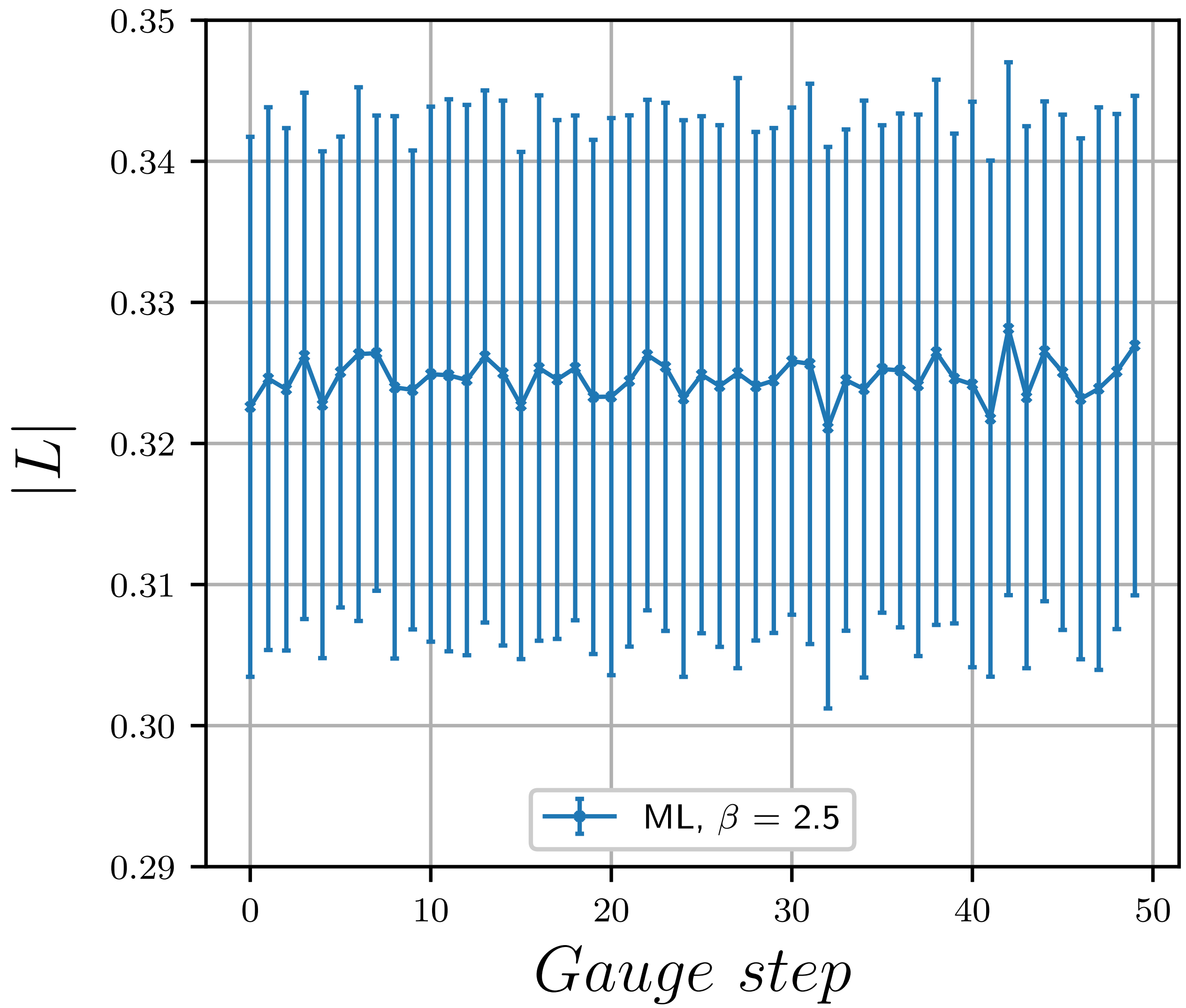


SU(2) **Lattice size: 4×32^3** **SU(3)**

The neural network is able to find a "proper expression" of the order parameter in the unphysical point. We restore it in the whole parameter space, including the transition region.

*) preliminary, to be beautified later

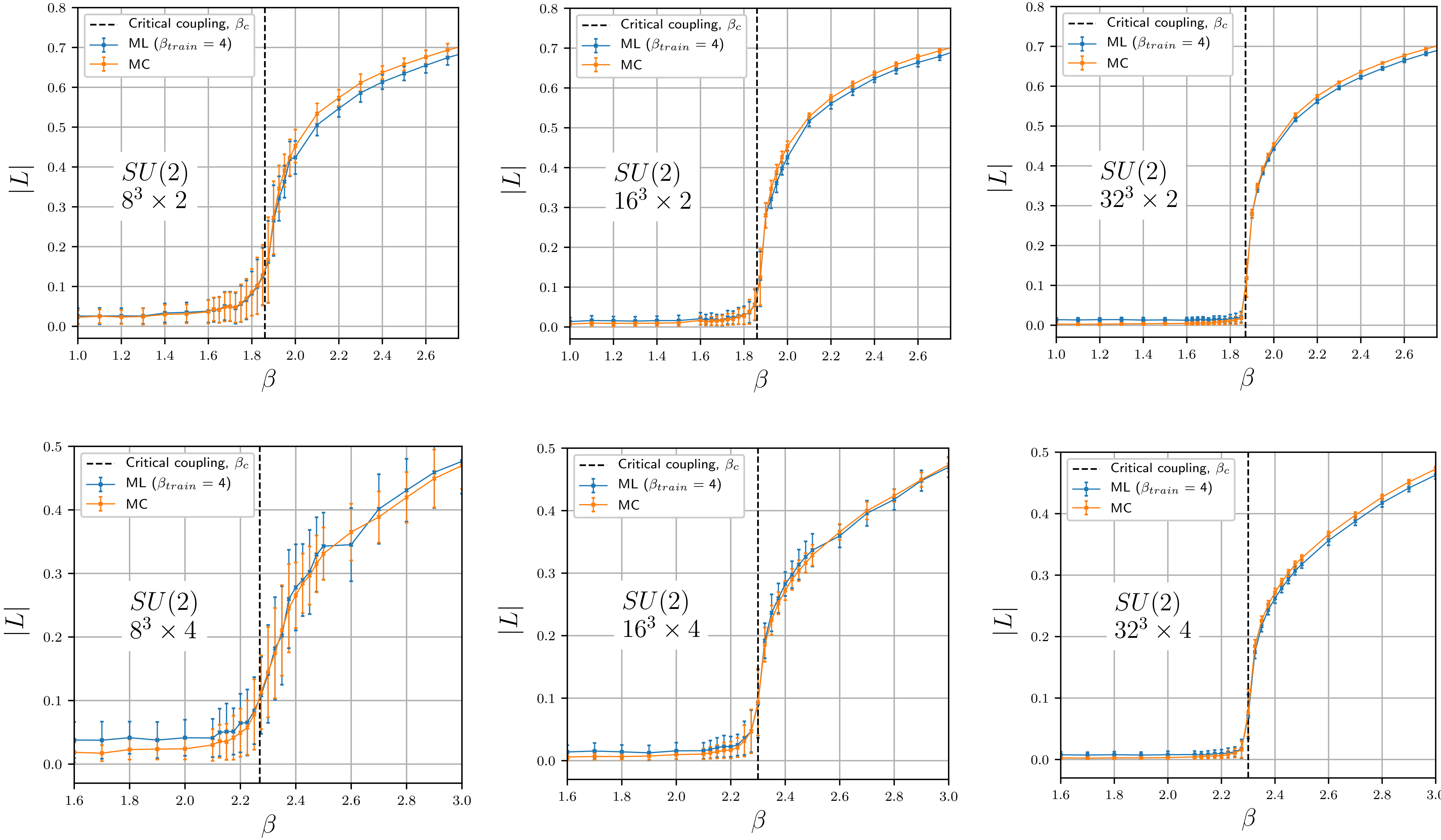
Learning physics from "unphysics"



The degree of the gauge dependence in the prediction of the order parameter by the ML algorithm. The predicted order parameter along with the prediction uncertainty vs. the number of the gauge randomization steps of the initial $16^3 \times 4$ gluon configuration at β = 2.5.

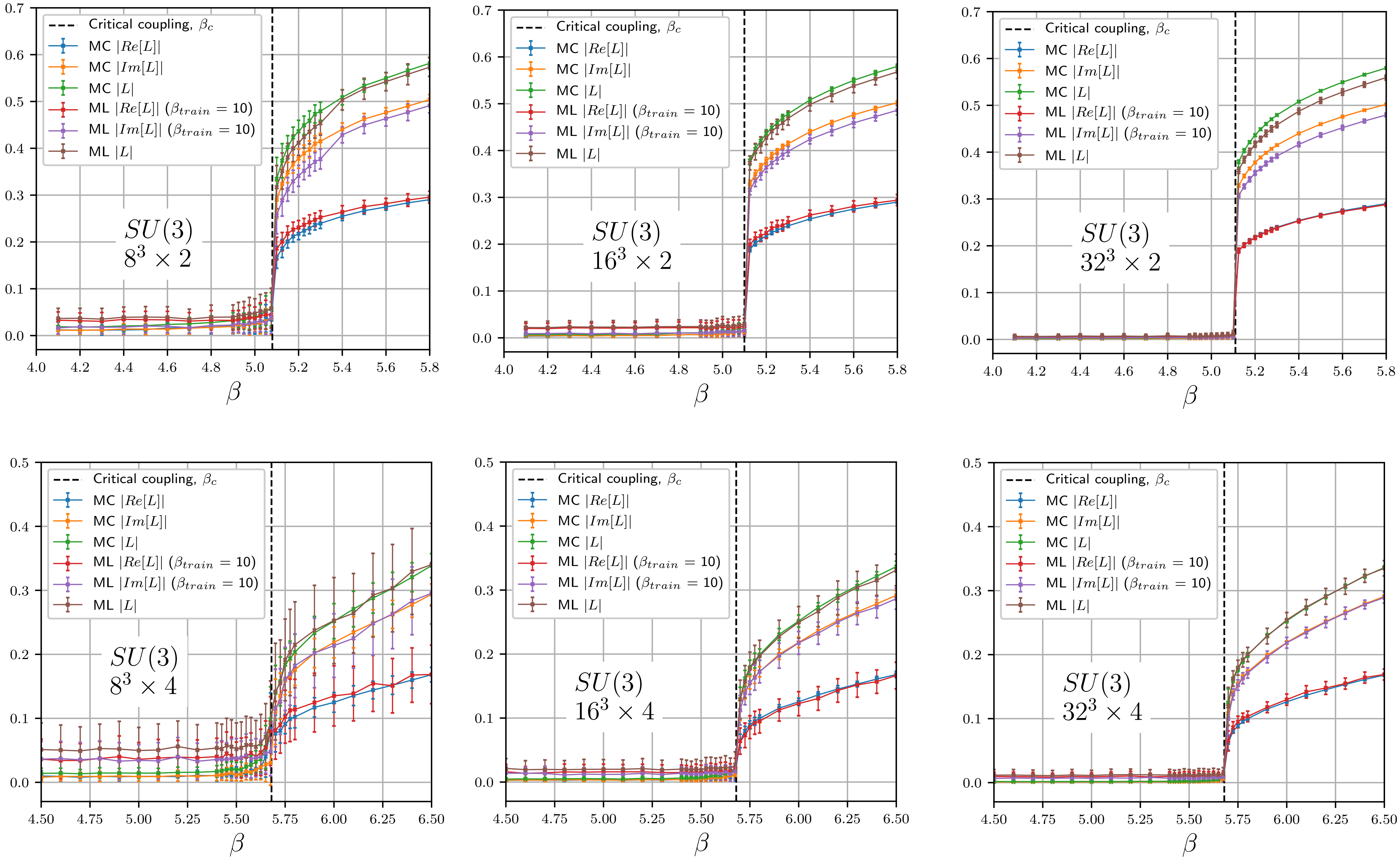
The ML algorithm's prediction is a gauge- invariant quantity that does not depend on the strength of the gluonic configuration's randomization in the gauge group's space transformations.

Learning physics from “unphysics”



The Polyakov loop in SU(2) gauge theory at the N_t = 2,4 and N_s = 8, 16, 32 lattices. The Monte-Carlo (MC) simulation, shown by the blue line, and the prediction of the machine-learning (ML) algorithm, shown by the orange line, overlap within the error bars. We use 100 configurations for all three lattice sizes.

Learning physics from “unphysics”



Polyakov loop for SU(3) gauge theory at Nt = 2,4 obtained with the Monte Carlo simulations compared to the neural network prediction. The absolute value, the real and imaginary parts of the loop are shown. The value of ML |L| restored from ML predictions of |Re[L]| and |Im[L]|.

Machine learning confinement

Summary again:

— the neural network uses the supervised learning technique to make reasonable predictions about the phase diagram;

— the neural network may learn essential properties of the lattice field theory: group symmetry, gauge invariance etc.;

— the machine-learning algorithm allows us to restore a gauge-invariant order parameter in the whole physical region of the parameter space after being trained on lattice configurations at one unphysical point in the lattice parameter space;

— the neural network may serve as an efficient numerical counterpart of an ``analytical continuation'' of physical observable. as a function of lattice configuration.

Future developments (very briefly)

Use artificial neural network,
together with Monte Carlo to find

— gluonic field configurations responsible for color confinement in Yang-Mills theory.

— QCD endpoint at real baryonic chemical potential.