Swansea University
Prifysgol Abertawe

Quantum field-theoretic machine learning and the inverse renormalization group

Dimitrios Bachtis

UKLFT annual meeting 2022

Outline

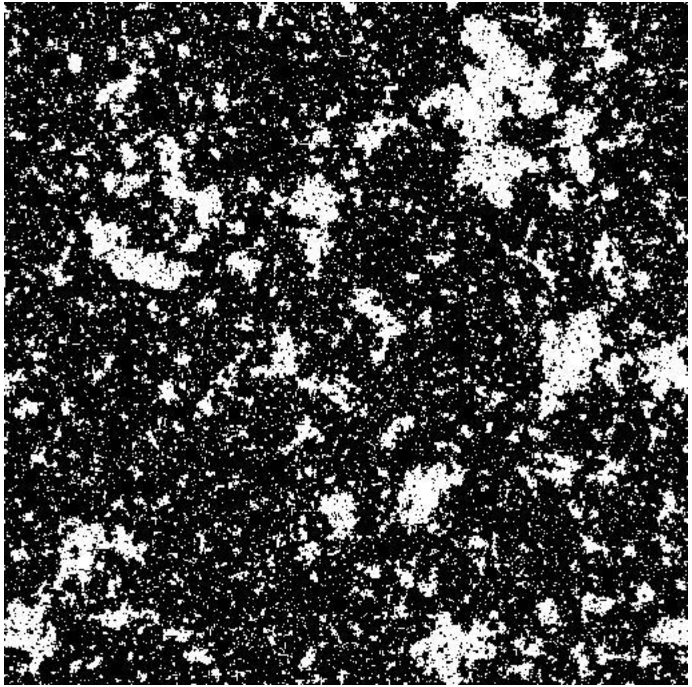
1. **(Standard) Renormalization group method**

2. **Inverse renormalization group with machine learning**

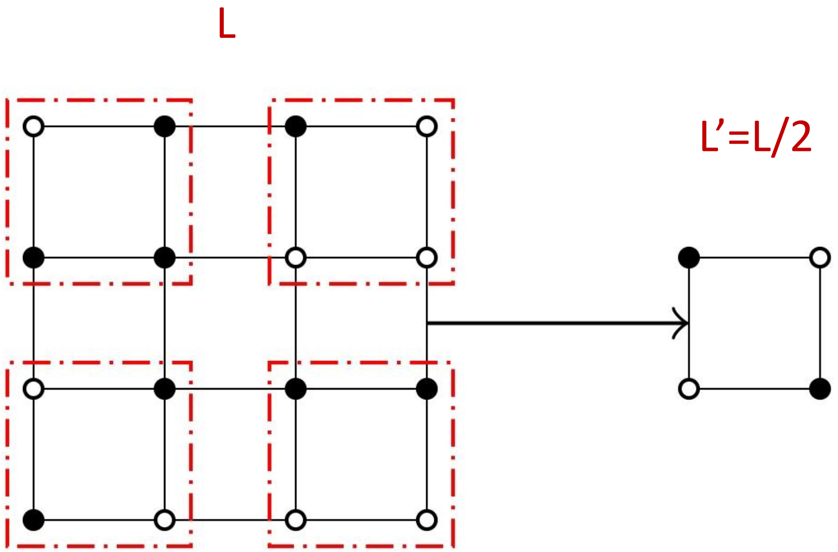
3. **Quantum field-theoretic machine learning**

The renormalization group

L



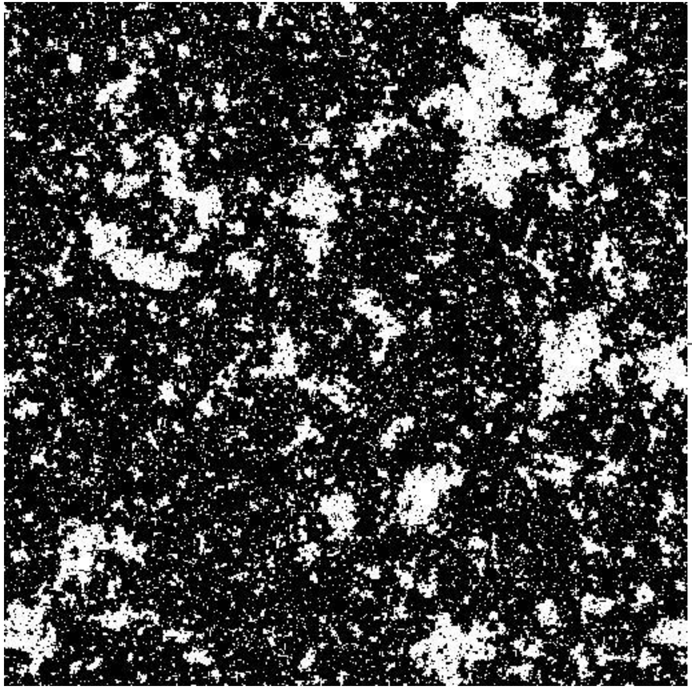
The renormalization group



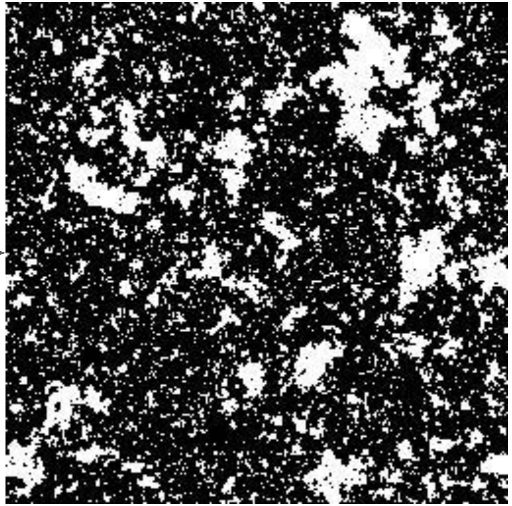
Spin blocking transformation with a rescaling factor of b=2 and the majority rule

The renormalization group

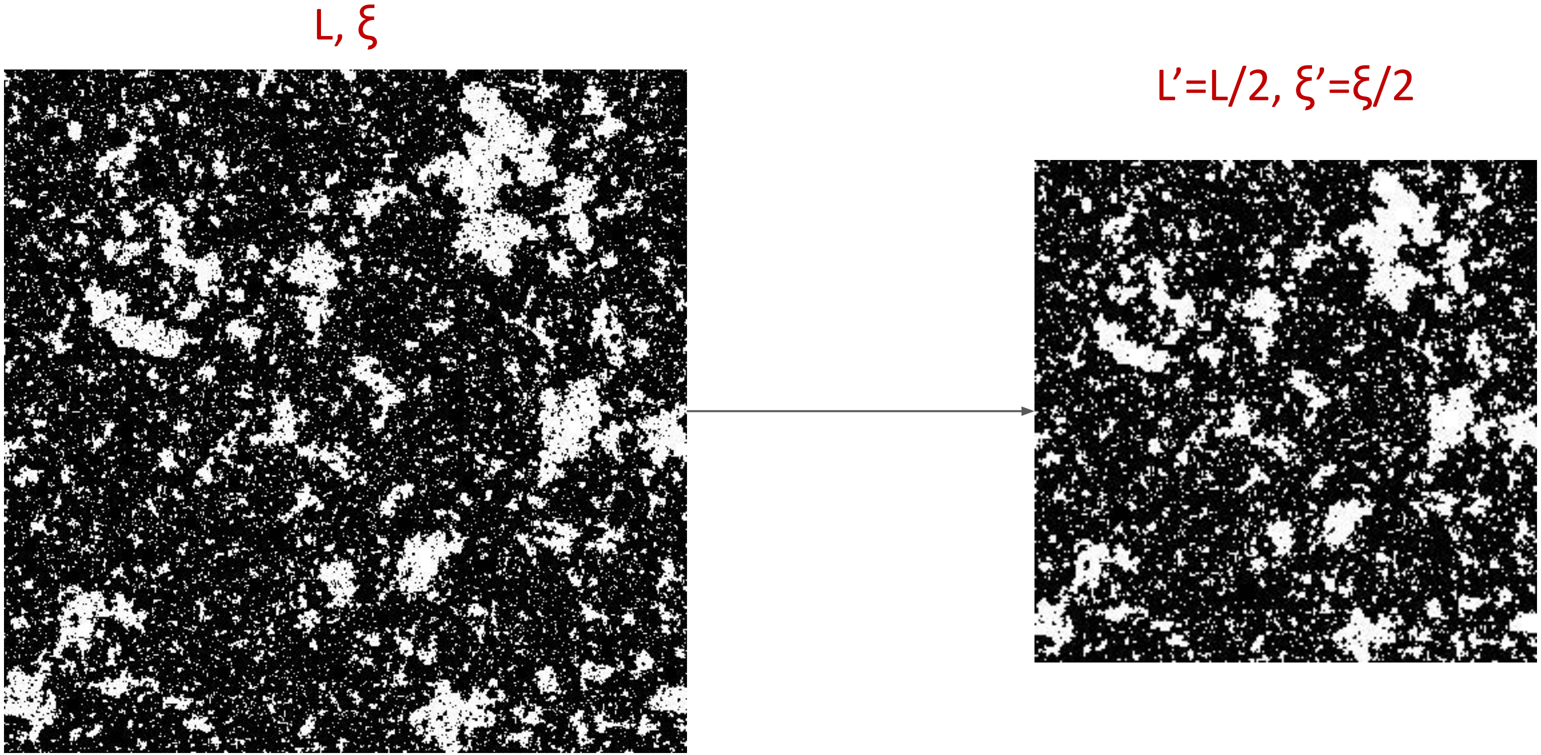
L



L'=L/2

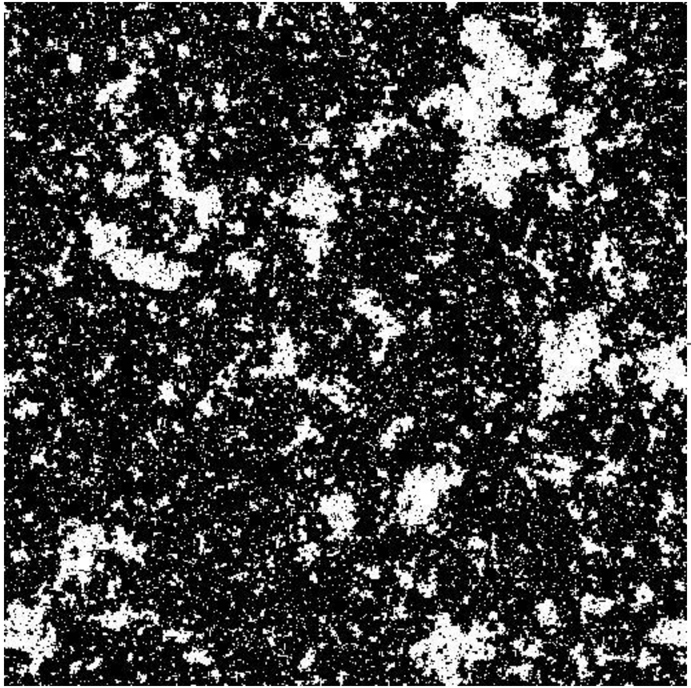


The renormalization group

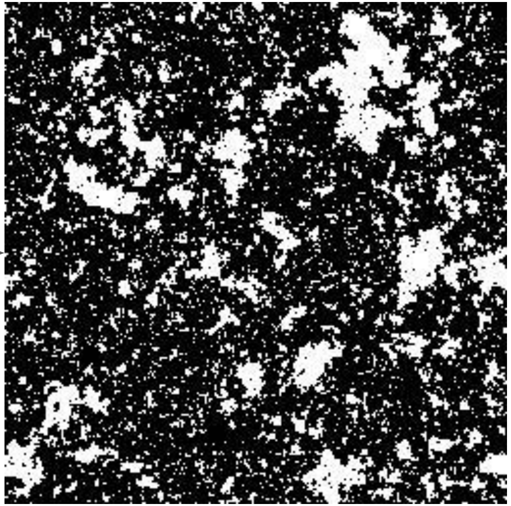


The renormalization group

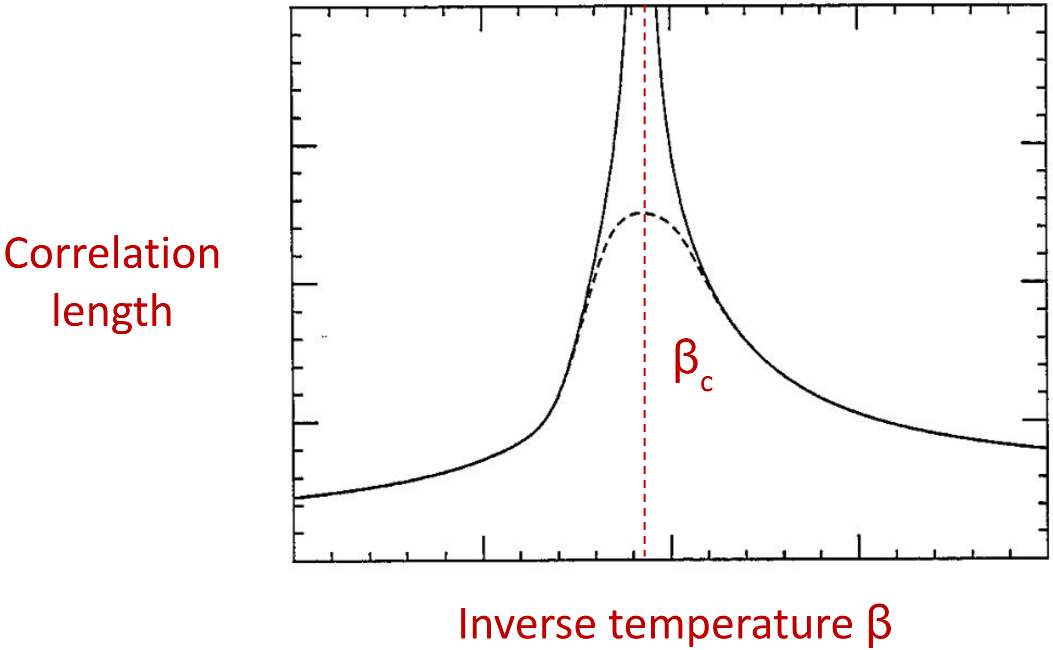
L, ξ, β



L’=L/2, ξ’=ξ/2, β’

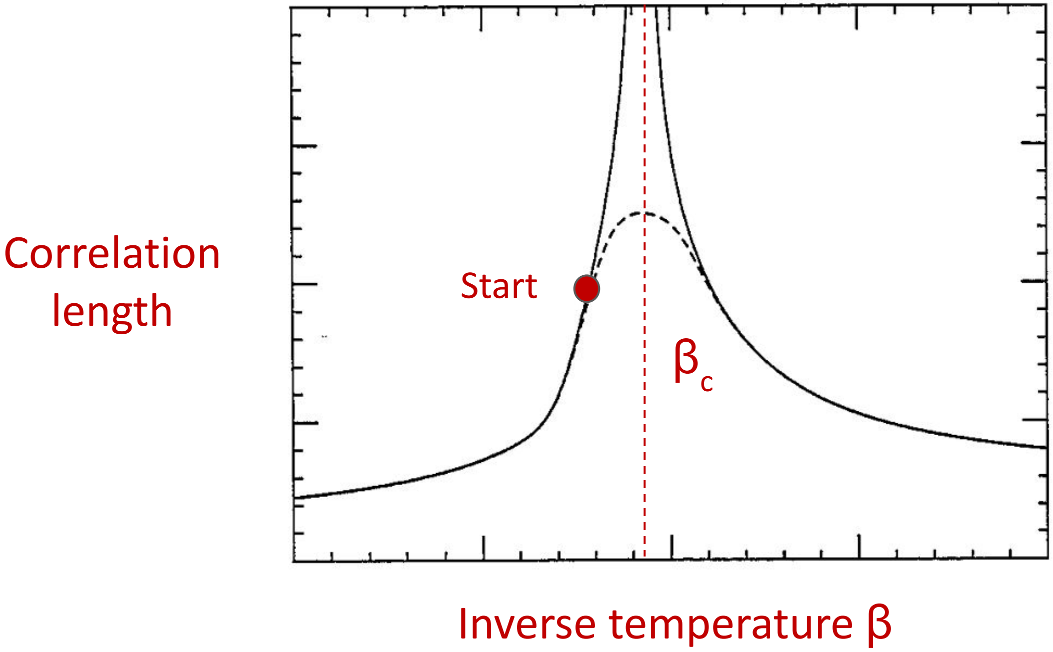


The renormalization group



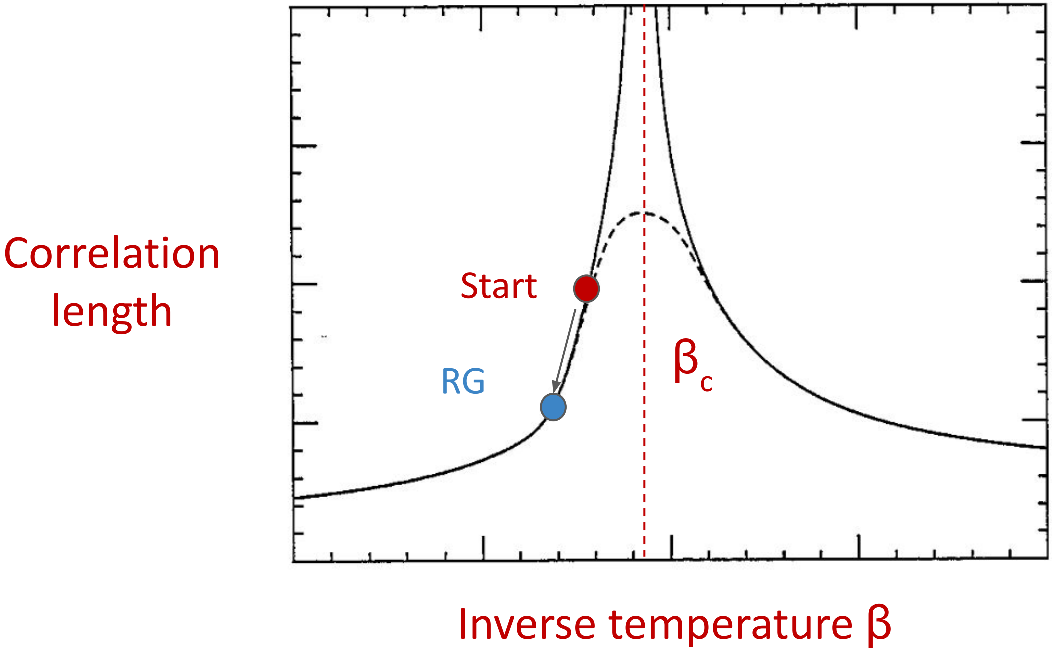
Altered figure from (Newman, Barkema) book (Fig 4.1)

The renormalization group



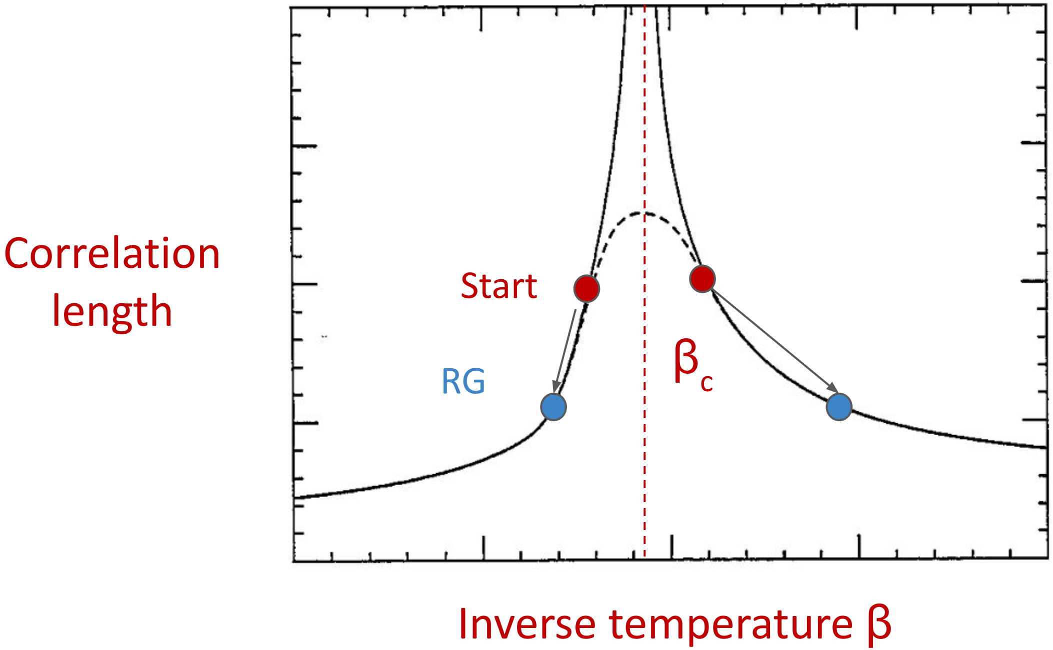
Altered figure from (Newman, Barkema) book (Fig 4.1)

The renormalization group



Altered figure from (Newman, Barkema) book (Fig 4.1)

The renormalization group



Altered figure from (Newman, Barkema) book (Fig 4.1)

Neural Networks as Physical Observables

The original and the rescaled systems have a different distance from the critical point.

This distance can be measured by defining the reduced inverse temperature for the original and the rescaled system:

$$t = \frac{\beta_c - \beta}{\beta_c}$$

Original

$$t' = \frac{\beta_c - \beta'}{\beta_c}$$

Rescaled

The renormalization group

There is one inverse temperature where the original and the rescaled systems have the same correlation length: the inverse critical temperature β_c=0.440687.

At the inverse critical temperature β_c the correlation length diverges, it becomes infinite, and intensive observable quantities of the two systems will become equal.

$$O'(\beta_c)=O(\beta_c)$$

Inverse renormalization group

Can we devise an **inverse renormalization group** approach that can be applied for an arbitrary number of steps to iteratively increase the lattice size of the system?

Inverse Renormalization Group in Quantum Field Theory, D. Bachtis, G. Aarts, F. Di Renzo, B. Lucini, Phys. Rev. Lett. 128, 081603 (2022)

Inverse renormalization group

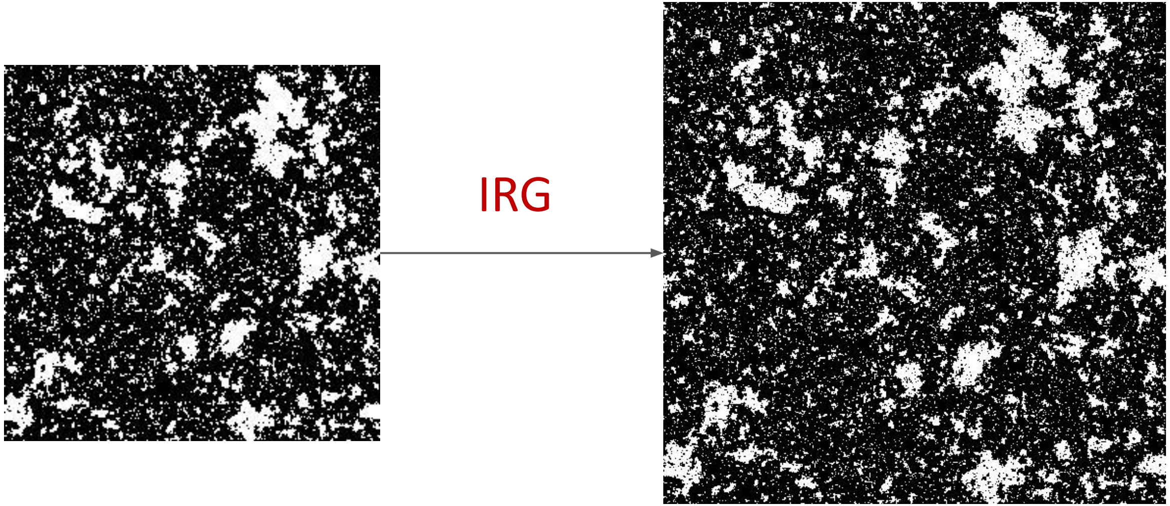
Can we devise an **inverse renormalization group** approach that can be applied for an arbitrary number of steps to iteratively increase the lattice size of the system?

If yes, then we can obtain configurations of systems with larger lattice size without simulating them, hence evading the critical slowing down effect.

Inverse Renormalization Group in Quantum Field Theory, D. Bachtis, G. Aarts, F. Di Renzo, B. Lucini, Phys. Rev. Lett. 128, 081603 (2022)

Inverse renormalization group

In the inverse renormalization group new degrees of freedom will be introduced within the system.



Inverse Renormalization Group in Quantum Field Theory, D. Bachtis, G. Aarts, F. Di Renzo, B. Lucini, Phys. Rev. Lett. 128, 081603 (2022)

Inverse renormalization group

Inversion of a majority rule in the **Ising** model

Original degree of freedom

+1

Possible rescaled degrees of freedom

+1	+1
+1	-1

-1	+1
+1	+1

-1	+1
+1	-1

-1	+1
+1	-1

…

For the inverse renormalization group in the Ising model, see:

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

Inverse renormalization group

Inversion of a summation in the $\boldsymbol{\varphi}^4$ model

Original degree of freedom

0.40

Possible rescaled degrees of freedom

0.01	0.36
0.02	0.01

-421.1	90.1
0.5	330.9

…

Inverse Renormalization Group in Quantum Field Theory, D. Bachtis, G. Aarts, F. Di Renzo, B. Lucini, Phys. Rev. Lett. 128, 081603 (2022)

Inverse renormalization group

We can learn a set of transformations that can mimic the inversion of a standard renormalization group transformation.

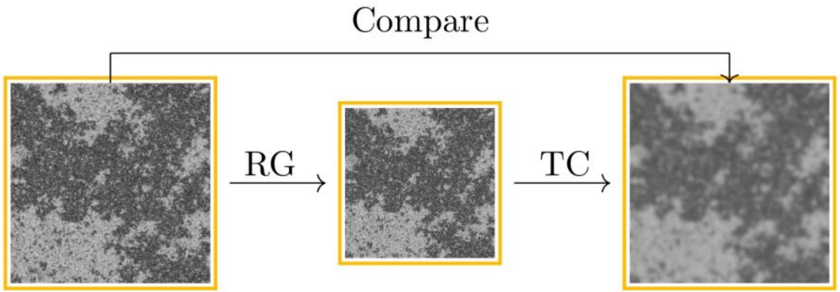


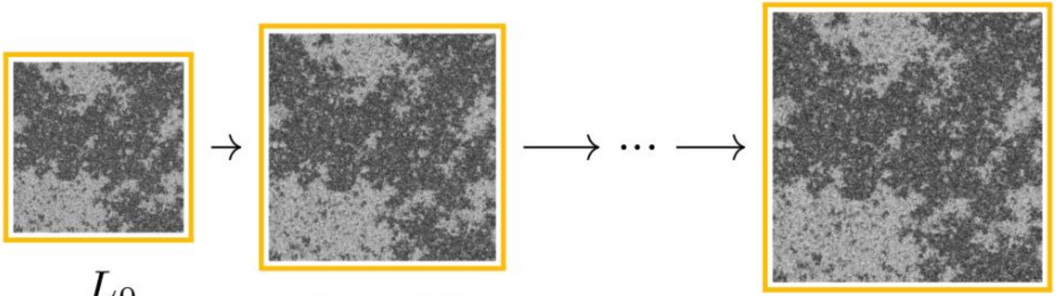
FIG. 3. Illustration of the optimization approach. Transposed convolutions (TC) are applied on configurations produced with the renormalization group (RG) to construct a set of configuration which is compared with the original.

Inverse Renormalization Group in Quantum Field Theory, D. Bachtis, G. Aarts, F. Di Renzo, B. Lucini, Phys. Rev. Lett. 128, 081603 (2022)

Inverse renormalization group

The benefit:

Once learned, we can apply this set of inverse transformations iteratively to <u>arbitrarily increase the size of the system</u>.



L_0 $\quad$ $L_1 = bL_0$ $\quad$ $L_j = b^{(j-i)} L_i$

Inverse Renormalization Group in Quantum Field Theory, D. Bachtis, G. Aarts, F. Di Renzo, B. Lucini, Phys. Rev. Lett. 128, 081603 (2022)

Inverse renormalization group

The set of transformations can be applied iteratively to arbitrarily increase the lattice size:

$$L_j = b^{(j-i)} L_i \qquad j > i \geq 0, \text{ and } L_0 = L$$

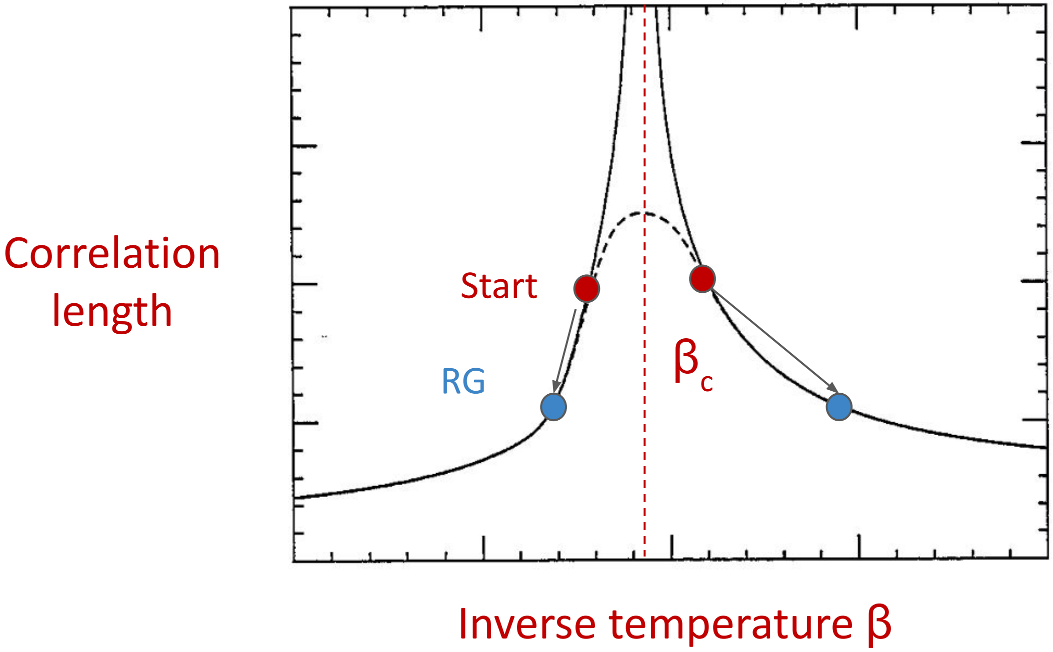
However the increase in the lattice size will induce an analogous increase in the correlation length of the system:

$$\xi_j = b^{(j-i)} \xi_i$$

What are the implications?

Inverse Renormalization Group in Quantum Field Theory, D. Bachtis, G. Aarts, F. Di Renzo, B. Lucini, Phys. Rev. Lett. 128, 081603 (2022)

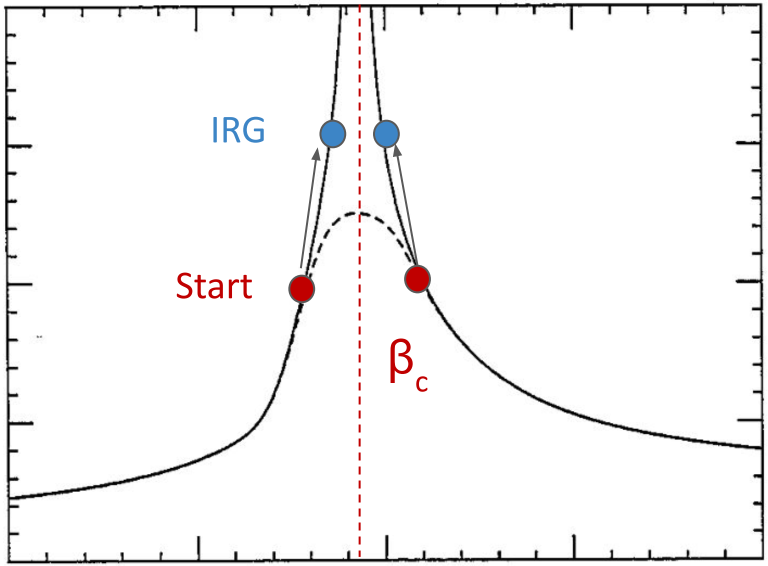
Inverse renormalization group



Inverse Renormalization Group in Quantum Field Theory, D. Bachtis, G. Aarts, F. Di Renzo, B. Lucini, Phys. Rev. Lett. 128, 081603 (2022)

Inverse renormalization group

Correlation length

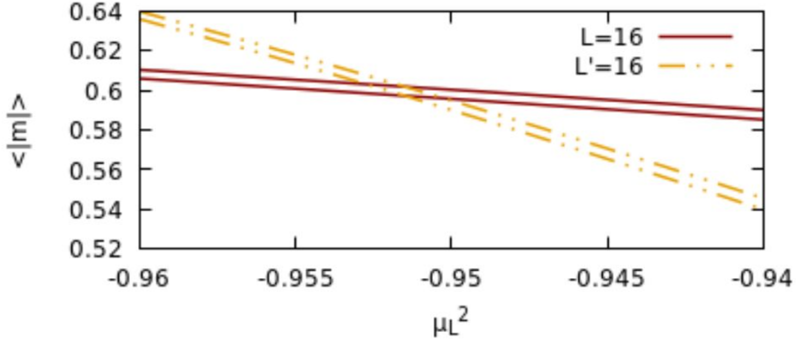


Inverse temperature β

Inverse Renormalization Group in Quantum Field Theory, D. Bachtis, G. Aarts, F. Di Renzo, B. Lucini, Phys. Rev. Lett. 128, 081603 (2022)

Inverse renormalization group

First, we verify that the **standard MC** renormalization group method works in the φ^4 theory:

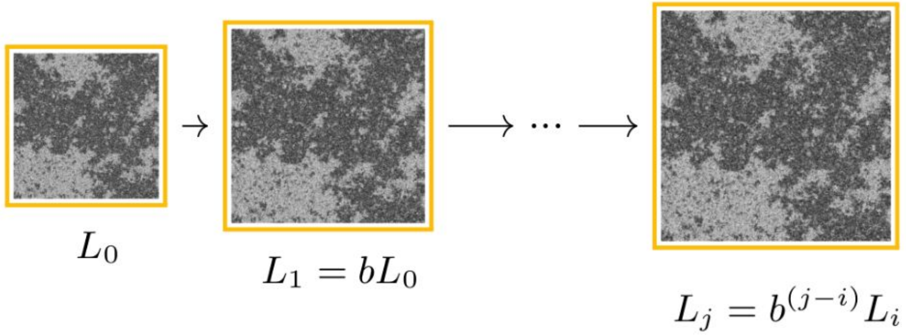


Then we invert the standard transformation that we verified as being successful.

Inverse Renormalization Group in Quantum Field Theory, D. Bachtis, G. Aarts, F. Di Renzo, B. Lucini, Phys. Rev. Lett. 128, 081603 (2022)

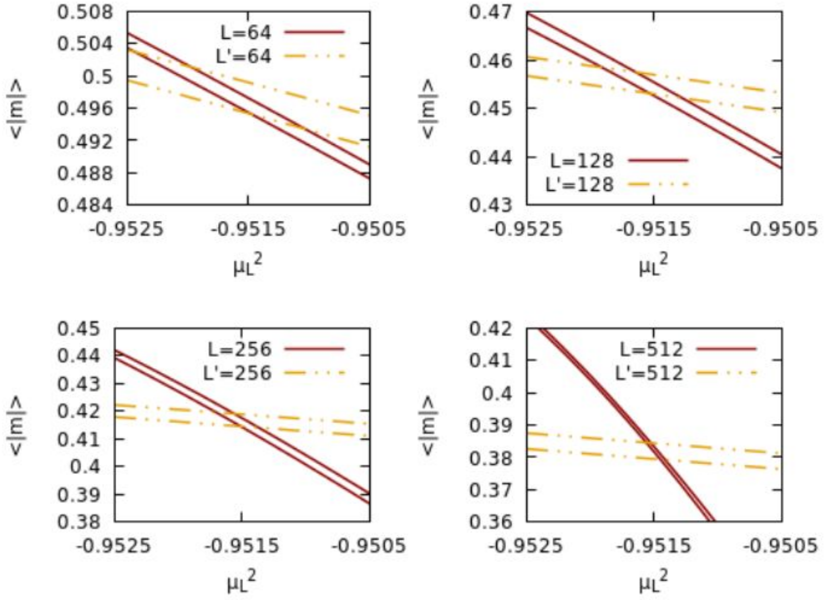
Inverse renormalization group

Now, we start from a lattice size L_0=32 in each dimension and apply the inverse transformations to obtain systems of lattice sizes L_1=64, L_2=128, L_3=256, L_4=512.



Inverse Renormalization Group in Quantum Field Theory, D. Bachtis, G. Aarts, F. Di Renzo, B. Lucini, Phys. Rev. Lett. 128, 081603 (2022)

Inverse renormalization group



Inverse Renormalization Group in Quantum Field Theory, D. Bachtis, G. Aarts, F. Di Renzo, B. Lucini, Phys. Rev. Lett. 128, 081603 (2022)

Inverse renormalization group

Can we now use the inverse renormalization group approach to calculate critical exponents?

The relations that govern the critical behaviour of the magnetization for an original (i) and a rescaled (j) system are

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

They can be equivalently expressed in terms of the correlation length as

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

where v is the correlation length exponent

Inverse Renormalization Group in Quantum Field Theory, D. Bachtis, G. Aarts, F. Di Renzo, B. Lucini, Phys. Rev. Lett. 128, 081603 (2022)

Inverse renormalization group

By dividing the magnetizations (or magnetic susceptibilities), taking the natural logarithm, and applying L'Hôpital's rule, we obtain

$$\frac{\beta}{\nu} = -\frac{\ln \frac{dm_j}{dm_i}\Big|_{K_c}}{\ln \frac{\xi_j}{\xi_i}} = -\frac{\ln \frac{dm_j}{dm_i}\Big|_{K_c}}{(j-i)\ln b}. \qquad \frac{\gamma}{\nu} = \frac{\ln \frac{d\chi_j}{d\chi_i}\Big|_{K_c}}{\ln \frac{\xi_j}{\xi_i}} = \frac{\ln \frac{d\chi_j}{d\chi_i}\Big|_{K_c}}{(j-i)\ln b}.$$

We can use the expressions above to calculate the critical exponents without ever experiencing a critical slowing down effect.

Inverse Renormalization Group in Quantum Field Theory, D. Bachtis, G. Aarts, F. Di Renzo, B. Lucini, Phys. Rev. Lett. 128, 081603 (2022)

Inverse renormalization group

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 $\mu_L^2 = -0.9515$, $\lambda_L = 0.7$, $\kappa_L = 1$. 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)

TABLE II. Values of the critical exponents γ/ν and β/ν. The original system has lattice size $L = 8$ in each dimension and its action has coupling constants $\mu_L^2 = -1.2723$, $\lambda_L = 1$, $\kappa_L = 1$. The rescaled systems are obtained through inverse renormalization group transformations.

L_i/L_j	8/16	8/32	8/64	8/128	8/256	8/512	16/32	16/64	16/128	16/256	16/512
γ/ν	1.694(6)	1.708(6)	1.717(6)	1.723(6)	1.727(6)	1.730(6)	1.721(6)	1.728(6)	1.732(6)	1.735(6)	1.737(6)
β/ν	0.154(2)	0.147(2)	0.142(2)	0.139(2)	0.137(2)	0.135(2)	0.140(2)	0.136(2)	0.134(2)	0.132(2)	0.131(2)

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(6)	1.738(6)	1.740(6)	1.740(6)	1.741(6)	1.742(6)	1.742(7)	1.743(6)	1.743(7)	1.743(7)
β/ν	0.133(2)	0.131(2)	0.130(2)	0.129(2)	0.129(2)	0.129(2)	0.128(2)	0.128(2)	0.127(2)	0.127(2)

Ising universality class: γ/v=1.75, β/v=0.125.

Inverse Renormalization Group in Quantum Field Theory, D. Bachtis, G. Aarts, F. Di Renzo, B. Lucini, Phys. Rev. Lett. 128, 081603 (2022)

Quantum field-theoretic machine learning

Can we view machine learning as part of quantum field theory?

And why?

Construction of quantum fields from Markoff fields, E. Nelson, J. Funct. Anal. 12, 97 (1973)

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

A probability distribution is a product of strictly positive and appropriately normalized factors (or potential functions) ψ:

$$p(\phi) = \frac{\prod_{c \in C} \psi_c(\phi)}{\int_{\boldsymbol{\phi}} \prod_{c \in C} \psi_c(\boldsymbol{\phi}) d\boldsymbol{\phi}},$$

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

A probability distribution is a product of strictly positive and appropriately normalized factors (or potential functions) ψ:

$$p(\phi) = \frac{\prod_{c \in C} \psi_c(\phi)}{\int_{\boldsymbol{\phi}} \prod_{c \in C} \psi_c(\boldsymbol{\phi}) d\boldsymbol{\phi}},$$

1. Factors are the fundamental building blocks of probability distributions.

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

A probability distribution is a product of strictly positive and appropriately normalized factors (or potential functions) ψ:

$$p(\phi) = \frac{\prod_{c \in C} \psi_c(\phi)}{\int_{\boldsymbol{\phi}} \prod_{c \in C} \psi_c(\boldsymbol{\phi}) d\boldsymbol{\phi}},$$

1. Factors are the fundamental building blocks of probability distributions.

2. By controlling the factors we are able to control the probability distribution.

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

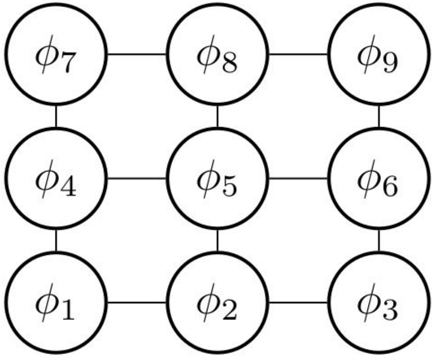
We require some form of representation to construct the probability distribution. We are going to use a finite set Λ that we express as a graph $G(\Lambda,e)$ where e is the set of edges in G.

A clique c is a subset of Λ where the points are pairwise connected. A maximal clique is a clique where we cannot add another point that is pairwise connected with ***all*** the points in the subset.

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

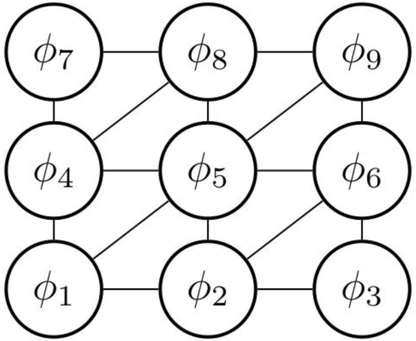
Quantum field-theoretic machine learning

On the square lattice a maximal clique is an edge.



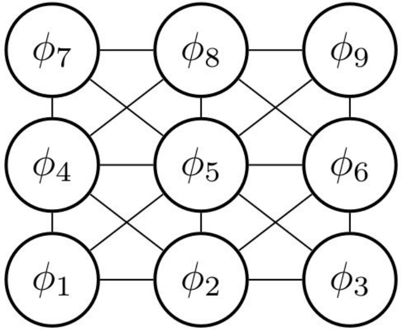
(a)

On a triangular lattice a maximal clique is a triangle.

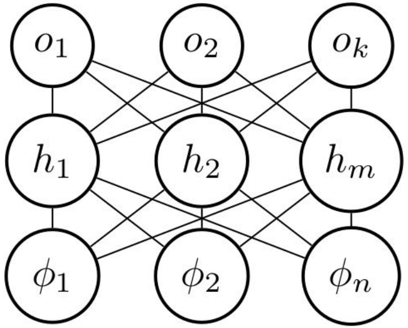


(b)

On the square lattice with both diagonals a maximal clique is a square.



(c)



(d)

On the bipartite graph, which represents standard neural network architectures a maximal clique is an edge.

Quantum field-theoretic machine learning

Hammersley-Clifford theorem

A strictly positive distribution p satisfies the local Markov property of an undirected graph G:

$$p(\phi_i | (\phi_j)_{j\in\Lambda-i}) = p(\phi_i | (\phi_j)_{j\in\mathcal{N}_i})$$

if and only if p can be represented as a product of strictly positive potential functions ψ_c over G, one per maximal clique c, i.e.

$$p(\phi) = \frac{1}{Z}\prod_{c\in C}\psi_c(\phi), \quad Z = \int_{\boldsymbol{\phi}} \prod_{c\in C}\psi_c(\boldsymbol{\phi})d\boldsymbol{\phi}$$

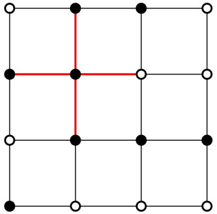
where Z is the partition function and **φ** are all possible states of the system.

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

The φ^4 lattice field theory is, by definition, formulated on a square lattice which is equivalent to a graph $G(\Lambda,e)$. A non-unique choice of potential function per each maximal clique is:

$$\psi_c = \exp\left[-w_{ij}\phi_i\phi_j + \frac{1}{4}(a_i\phi_i^2 + a_j\phi_j^2 + b_i\phi_i^4 + b_j\phi_j^4)\right],$$



The probability distribution is expressed as a product of strictly positive potential functions ψ, over each maximal clique:

$$p(\phi;\theta) = \frac{\exp\left[\sum_{c\in C}\ln\psi_c(\phi)\right]}{\int_\phi \exp\left[\sum_{c\in C}\ln\psi_c(\phi)\right]d\phi} = \frac{1}{Z}\prod_{c\in C}\psi_c(\phi).$$

The φ^4 theory satisfies Markov properties and it is therefore a Markov random field.

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

Having established that certain physical systems are Markov random fields, how do we use them for machine learning?

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

Having established that certain physical systems are Markov random fields, how do we use them for machine learning?

Exactly in the same way as any other machine learning algorithm...

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

The φ^4 theory has a probability distribution $p(\varphi;\theta)$ with action $S(\varphi;\theta)$:

$$p(\phi;\theta) = \frac{\exp\left[-S(\phi;\theta)\right]}{\int_{\boldsymbol{\phi}} \exp[-S(\boldsymbol{\phi},\theta)]d\boldsymbol{\phi}}.$$

We now consider a quantum field theory with action A and a target probability distribution $q(\varphi)$:

$$q(\phi) = \exp[-\mathcal{A}]/Z_{\mathcal{A}}$$

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

We can then define an asymmetric distance between the probability distributions p(φ;θ) and q(φ), which is called the Kullback-Leibler divergence:

$$KL(p||q) = \int_{-\infty}^{\infty} p(\phi;\theta) \ln \frac{p(\phi;\theta)}{q(\phi)} d\phi \geq 0.$$

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

We can then define an asymmetric distance between the probability distributions p(φ;θ) and q(φ), which is called the Kullback-Leibler divergence:

$$KL(p||q) = \int_{-\infty}^{\infty} p(\phi;\theta) \ln \frac{p(\phi;\theta)}{q(\phi)} d\phi \geq 0.$$

We want to minimize the Kullback-Leibler divergence.

By minimizing it we would make the two probability distributions equal. We can then use the probability distribution p(φ;θ) of the φ^4 theory to draw samples from the target distribution q(φ) of action A.

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

We substitute the two probability distributions in the Kullback-Leibler divergence to obtain:

$$F_{\mathcal{A}} \leq \langle \mathcal{A} - S \rangle_{p(\phi;\theta)} + F \equiv \mathcal{F},$$

Bogoliubov Inequality

<> denotes expectation value

There are two important observations on the above equation:

1. It sets a rigorous upper bound to the calculation of the free energy of the system with action A.
2. The bound is dependent entirely on samples drawn from the distribution p(φ;θ) of the φ^4 theory.

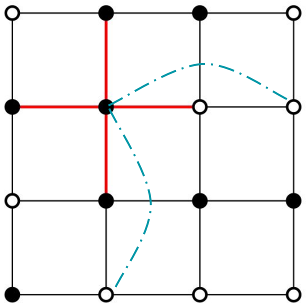
Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

We have conducted a variety of proof-of-principle applications to demonstrate that the inhomogeneous action

$$S(\phi;\theta) = -\sum_{\langle ij\rangle} w_{ij}\phi_i\phi_j + \sum_i a_i\phi_i^2 + \sum_i b_i\phi_i^4,$$

is able to represent more intricate actions, such as actions that include longer range interactions



$$\mathcal{A}_{\{4\}}(\phi) = -\sum_{\langle ij\rangle}\phi_i\phi_j + 1.52425\sum_i\phi_i^2 + 0.175\sum_i\phi_i^4 - \sum_{\langle ij\rangle_{nnn}}\phi_i\phi_j$$

See Ref.

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

What if the target probability distribution $q(\varphi)$ is unknown?

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

Earlier we defined the Kullback-Leibler divergence as:

$$KL(p||q) = \int_{-\infty}^{\infty} p(\phi;\theta) \ln \frac{p(\phi;\theta)}{q(\phi)} d\phi \geq 0.$$

We will now consider the opposite divergence:

$$KL(q||p) = \int_{-\infty}^{\infty} q(\phi) \ln \frac{q(\phi)}{p(\phi;\theta)} d\phi.$$

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

We are searching for the optimal values of the coupling constants in the φ^4 action that are able to reproduce the data as configurations in the equilibrium distribution.

$$S(\phi;\theta) = -\sum_{\langle ij \rangle} w_{ij}\phi_i\phi_j + \sum_i a_i \phi_i^2 + \sum_i b_i \phi_i^4,$$

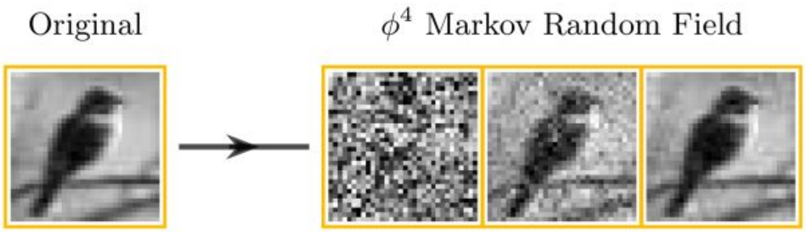
Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

We are searching for the optimal values of the coupling constants in the φ^4 action that are able to reproduce the data as configurations in the equilibrium distribution.

$$S(\phi;\theta) = -\sum_{\langle ij\rangle} w_{ij}\phi_i\phi_j + \sum_i a_i\phi_i^2 + \sum_i b_i\phi_i^4,$$

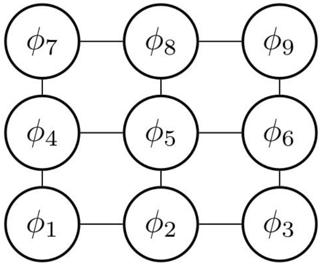
Case of an image:



Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

φ^4 Markov random field

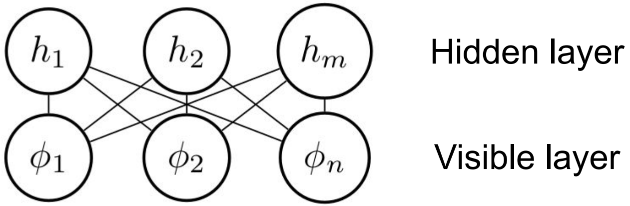


$$S(\phi;\theta) = -\sum_{\langle ij\rangle} w_{ij}\phi_i\phi_j + \sum_i a_i\phi_i^2 + \sum_i b_i\phi_i^4,$$

$$\theta = \{w_{ij}, a_i, b_i\}$$

$$p(\phi;\theta) = \frac{\exp\left[-S(\phi;\theta)\right]}{\int_{\boldsymbol{\phi}} \exp[-S(\boldsymbol{\phi},\theta)]d\boldsymbol{\phi}}.$$

φ^4 neural network



$$S(\phi,h;\theta) = -\sum_{i,j} w_{ij}\phi_i h_j + \sum_i r_i\phi_i + \sum_i a_i\phi_i^2 + \sum_i b_i\phi_i^4 + \sum_j s_j h_j + \sum_j m_j h_j^2 + \sum_j n_j h_j^4,$$

$$\theta = \{w_{ij}, r_i, a_i, b_i, s_j, m_j, n_j\}$$

$$p(\phi,h;\theta) = \frac{\exp[-S(\phi,h;\theta)]}{\int_{\boldsymbol{\phi},\boldsymbol{h}} \exp[-S(\boldsymbol{\phi},\boldsymbol{h};\theta)]d\boldsymbol{\phi}d\boldsymbol{h}}.$$

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

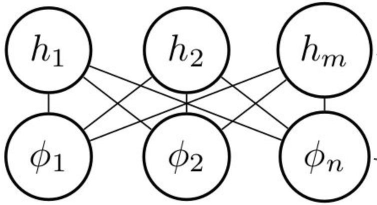
Quantum field-theoretic machine learning

From the joint probability distribution of the φ^4 neural network

$$p(\phi, h; \theta) = \frac{\exp[-S(\phi, h; \theta)]}{\int_{\boldsymbol{\phi}, \boldsymbol{h}} \exp[-S(\boldsymbol{\phi}, \boldsymbol{h}; \theta)] d\boldsymbol{\phi} d\boldsymbol{h}}.$$

We are able to marginalize out variables and derive marginal probability distributions $p(\varphi;\theta)$ and $p(h;\theta)$:

Hidden layer



$$p(h; \theta) = \int_{\boldsymbol{\phi}} p(\boldsymbol{\phi}, h; \theta) d\boldsymbol{\phi} = \frac{\int_{\boldsymbol{\phi}} \exp[-S(\boldsymbol{\phi}, h; \theta)] d\boldsymbol{\phi}}{\int_{\boldsymbol{\phi}, \boldsymbol{h}} \exp[-S(\boldsymbol{\phi}, \boldsymbol{h}; \theta)] d\boldsymbol{\phi} d\boldsymbol{h}},$$

$$p(\phi; \theta) = \int_{\boldsymbol{h}} p(\phi, \boldsymbol{h}; \theta) d\boldsymbol{h} = \frac{\int_{\boldsymbol{h}} \exp[-S(\phi, \boldsymbol{h}; \theta)] d\boldsymbol{h}}{\int_{\boldsymbol{\phi}, \boldsymbol{h}} \exp[-S(\boldsymbol{\phi}, \boldsymbol{h}; \theta)] d\boldsymbol{\phi} d\boldsymbol{h}},$$

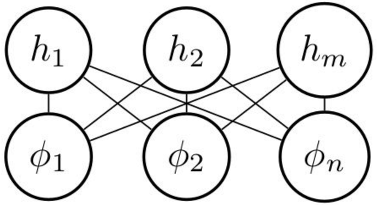
Visible layer

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

We now want to minimize the asymmetric distance between the empirical probability distribution q(φ) and the marginal probability distribution p(φ;θ):

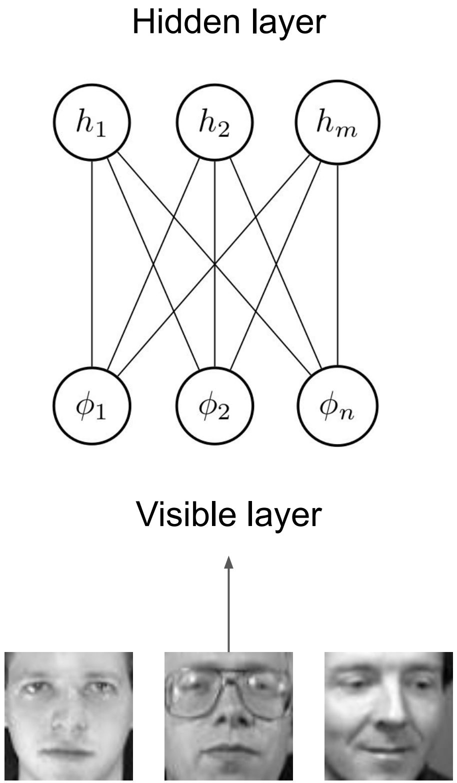
$$KL(q||p) = \int_{-\infty}^{\infty} q(\phi) \ln \frac{q(\phi)}{p(\phi;\theta)} d\phi.$$



In other words, we want to reproduce the dataset in the visible layer. The hidden layer will then uncover dependencies on the data.

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

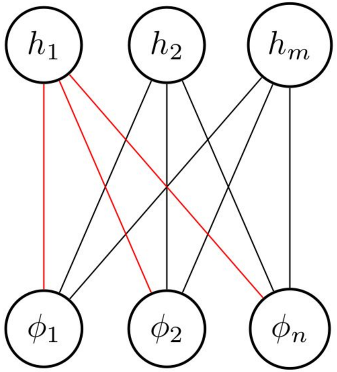
Quantum field-theoretic machine learning



Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

Hidden layer



Visible layer

Examples of the coupling constants w_{ij} with j fixed

$$S(\phi, h; \theta) = -\sum_{i,j} w_{ij}\phi_i h_j + \sum_i r_i \phi_i + \sum_i a_i \phi_i^2$$
$$+ \sum_i b_i \phi_i^4 + \sum_j s_j h_j + \sum_j m_j h_j^2 + \sum_j n_j h_j^4,$$

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Quantum field-theoretic machine learning

The φ^4 neural network:

$$S(\phi, h; \theta) = -\sum_{i,j} w_{ij}\phi_i h_j + \sum_i r_i \phi_i + \sum_i a_i \phi_i^2 + \sum_i b_i \phi_i^4 + \sum_j s_j h_j + \sum_j m_j h_j^2 + \sum_j n_j h_j^4,$$

is a generalization of other neural network architectures:

Gaussian-Gaussian restricted Boltzmann machine:

$b_i = n_j = 0$

Gaussian-Bernoulli restricted Boltzmann machine:

$b_i = n_j = m_j = 0$
h_j binary

Bernoulli-Bernoulli restricted Boltzmann machine:

$b_i = n_j = m_j = a_i = 0$
$\varphi_{I,}$ h_j binary

φ^4-Bernoulli restricted Boltzmann machine:

$m_j = n_j = 0$
h_j binary

φ^4 equivalence with the Ising model (under an appropriate limit)

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).

Summary

1) **Inverse renormalization group with machine learning**:
   a) How to generate configurations of systems with larger lattice size without having to simulate these systems and without critical slowing down effect.
   b) How do inverse renormalization group flows emerge.
   c) How to calculate multiple critical exponents with the inverse renormalization group.
2) **Quantum field-theoretic machine learning:**
   a) How to derive machine learning algorithms and neural networks from quantum field theories

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 813942

Thank you for your attention!

Inverse renormalization group

$$\begin{array}{|c|c|}\hline 3 & 1 \\ \hline 2 & 0 \\ \hline\end{array} * \begin{array}{|c|c|}\hline w_{11} & w_{12} \\ \hline w_{21} & w_{22} \\ \hline\end{array}$$

Example

$$\begin{array}{|c|c|}\hline 3 & 1 \\ \hline 2 & 0 \\ \hline\end{array} * \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 0 & 1 \\ \hline\end{array} =$$

$$\begin{array}{|c|c|c|}\hline 6 & 9 & \\ \hline 0 & 3 & \\ \hline & & \\ \hline\end{array} + \begin{array}{|c|c|c|}\hline & 2 & 3 \\ \hline & 0 & 1 \\ \hline & & \\ \hline\end{array} + \begin{array}{|c|c|c|}\hline & & \\ \hline 4 & 6 & \\ \hline 0 & 2 & \\ \hline\end{array} + \begin{array}{|c|c|c|}\hline & & \\ \hline & 0 & 0 \\ \hline & 0 & 0 \\ \hline\end{array} =$$

$$\begin{array}{|c|c|c|}\hline 6 & 11 & 3 \\ \hline 4 & 9 & 1 \\ \hline 0 & 2 & 0 \\ \hline\end{array}$$

Inverse Renormalization Group in Quantum Field Theory, D. Bachtis, G. Aarts, F. Di Renzo, B. Lucini, Phys. Rev. Lett. 128, 081603 (2022)

Learning

A first proof-of-principle demonstration is to use the inhomogeneous action S:

$$S(\phi;\theta) = -\sum_{\langle ij\rangle} w_{ij}\phi_i\phi_j + \sum_i a_i\phi_i^2 + \sum_i b_i\phi_i^4,$$

to learn a homogeneous action A:

$$\mathcal{A}(\phi) = -\sum_{\langle ij\rangle} \phi_i\phi_j + 1.52425\sum_i \phi_i^2 + 0.175\sum_i \phi_i^4$$

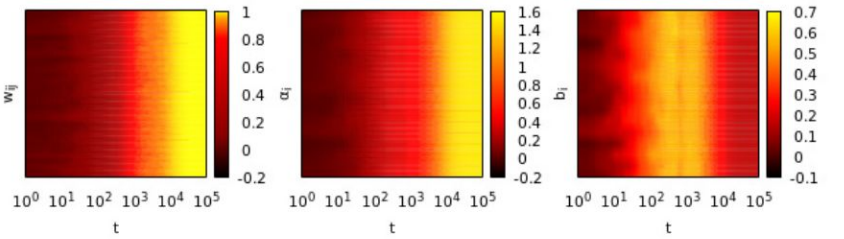


FIG. 2. Variational parameters $\theta = \{w_{ij}, a_i, b_i\}$ versus epochs t on logarithmic scale. The figures depict the evolution of the parameters θ towards the expected values of the coupling constants in the target homogeneous action.

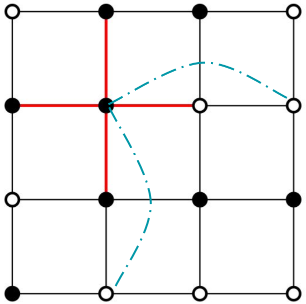
Learning

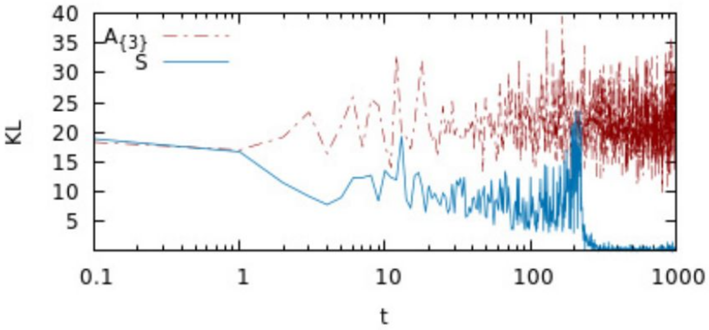
Another proof-of-principle demonstration is to use the inhomogeneous action S:

$$S(\phi;\theta) = -\sum_{\langle ij \rangle} w_{ij}\phi_i\phi_j + \sum_i a_i\phi_i^2 + \sum_i b_i\phi_i^4,$$

to learn an action that includes longer-range interactions:

$$\mathcal{A}_{\{4\}}(\phi) = -\sum_{\langle ij \rangle}\phi_i\phi_j + 1.52425\sum_i\phi_i^2 + 0.175\sum_i\phi_i^4 - \sum_{\langle ij \rangle_{nnn}}\phi_i\phi_j$$





Learning

Three reweighting (simultaneous) steps: Make the (already trained) inhomogeneous action S:

$$S(\phi;\theta) = -\sum_{\langle ij \rangle} w_{ij}\phi_i\phi_j + \sum_i a_i\phi_i^2 + \sum_i b_i\phi_i^4,$$

Equal to the target action A (acts as a correction step):

$$\mathcal{A}_{\{4\}}(\phi) = -\sum_{\langle ij \rangle}\phi_i\phi_j + 1.52425\sum_i \phi_i^2 + 0.175\sum_i \phi_i^4 - \sum_{\langle ij \rangle_{nnn}} \phi_i\phi_j$$

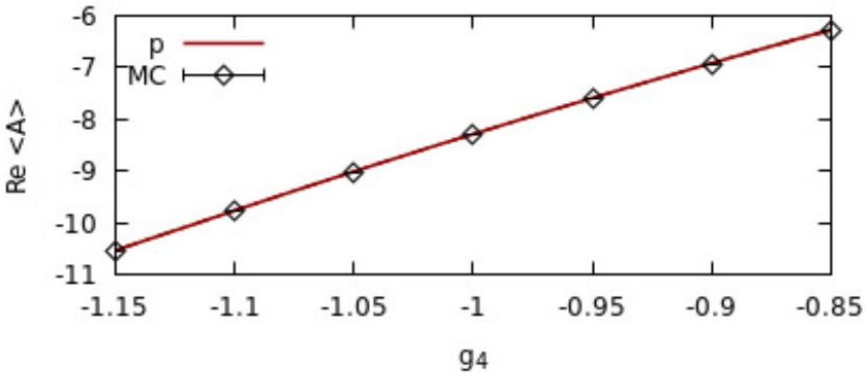
Extrapolate in the parameter space along the trajectory of a coupling constant g' of A

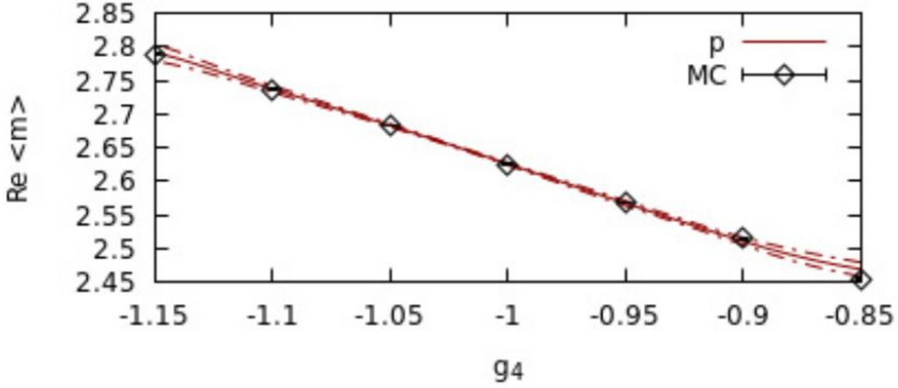
$$\mathcal{A}_{\{4\}}(\phi) = -\sum_{\langle ij \rangle}\phi_i\phi_j + 1.52425\sum_i \phi_i^2 + 0.175\sum_i \phi_i^4 - \sum_{\langle ij \rangle_{nnn}} g'\phi_i\phi_j$$

Extrapolate to an imaginary term

$$\mathcal{A}_{\{5\}}(\phi) = -\sum_{\langle ij \rangle}\phi_i\phi_j + 1.52425\sum_i \phi_i^2 + 0.175\sum_i \phi_i^4 - \sum_{\langle ij \rangle_{nnn}} g'\phi_i\phi_j + i0.15\sum_i \phi_i^2$$

Learning





The results include reweighting to a complex-valued coupling constant on the mass term and extrapolations in parameter space along the trajectory of the coupling constant g_4 in the longer-range interaction.

$$\mathcal{A}_{\{5\}}(\phi) = -\sum_{\langle ij \rangle} \phi_i \phi_j + 1.52425 \sum_i \phi_i^2 + 0.175 \sum_i \phi_i^4 - \sum_{\langle ij \rangle_{nnn}} g' \phi_i \phi_j + i0.15 \sum_i \phi_i^2$$

Quantum field-theoretic machine learning

To minimize the variational free energy we implement a gradient-based approach:

$$\frac{\partial \mathcal{F}}{\partial \theta_i} = \langle \mathcal{A} \rangle \left\langle \frac{\partial S}{\partial \theta_i} \right\rangle - \left\langle \mathcal{A} \frac{\partial S}{\partial \theta_i} \right\rangle + \left\langle S \frac{\partial S}{\partial \theta_i} \right\rangle - \langle S \rangle \left\langle \frac{\partial S}{\partial \theta_i} \right\rangle,$$

We then update the coupling constants θ at each step t until convergence.

$$\theta^{(t+1)} = \theta^{(t)} - \eta * \mathcal{L}, \qquad \mathcal{L} = \partial \mathcal{F} / \partial \theta^{(t)}$$

After training we expect that, practically:

$$\mathcal{F} \approx F_A \qquad p(\phi; \theta) \approx q(\phi).$$

Quantum field-theoretic machine learning, D. Bachtis, G. Aarts and B. Lucini, Phys. Rev. D 103, 074510, (arXiv:2102.09449).