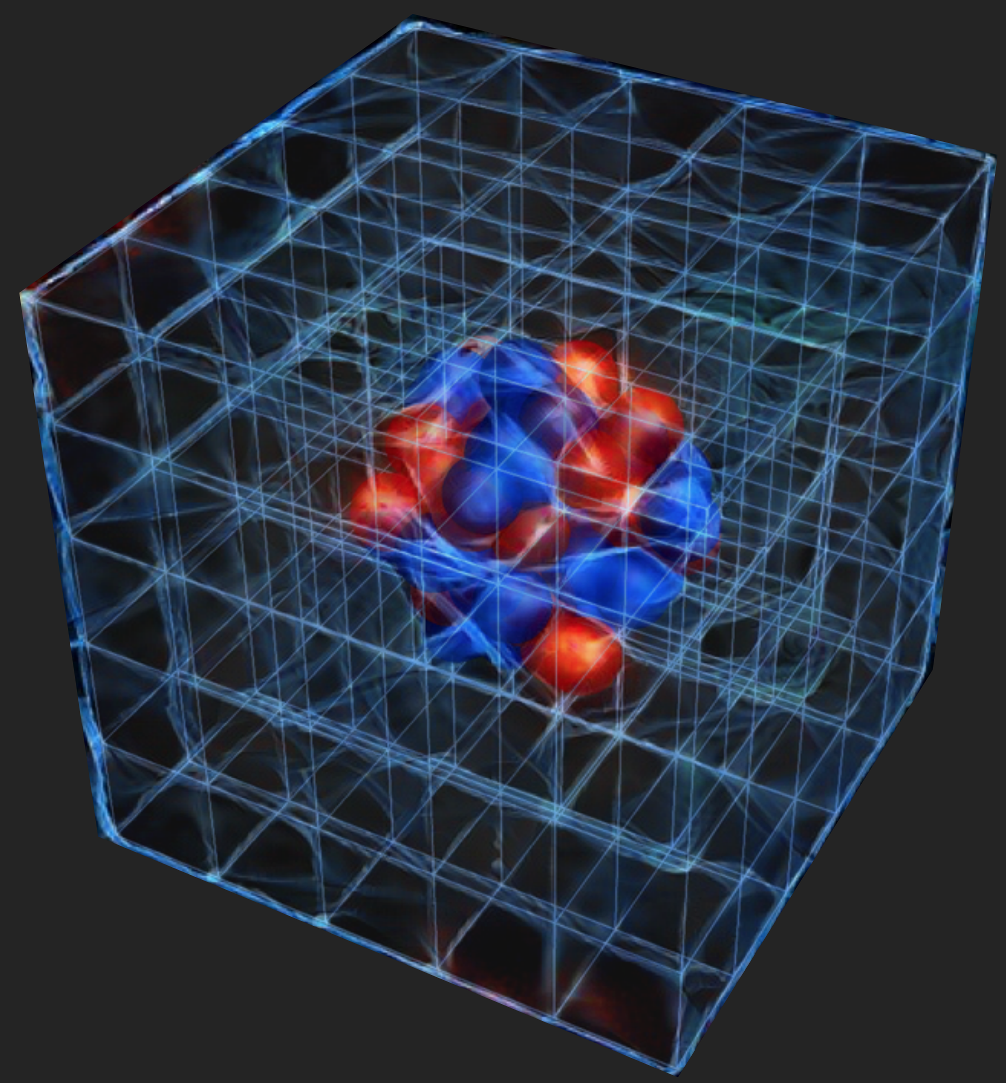
Machine Learning Matched Action Parameters



Phiala Shanahan

Massachusetts Institute of Technology

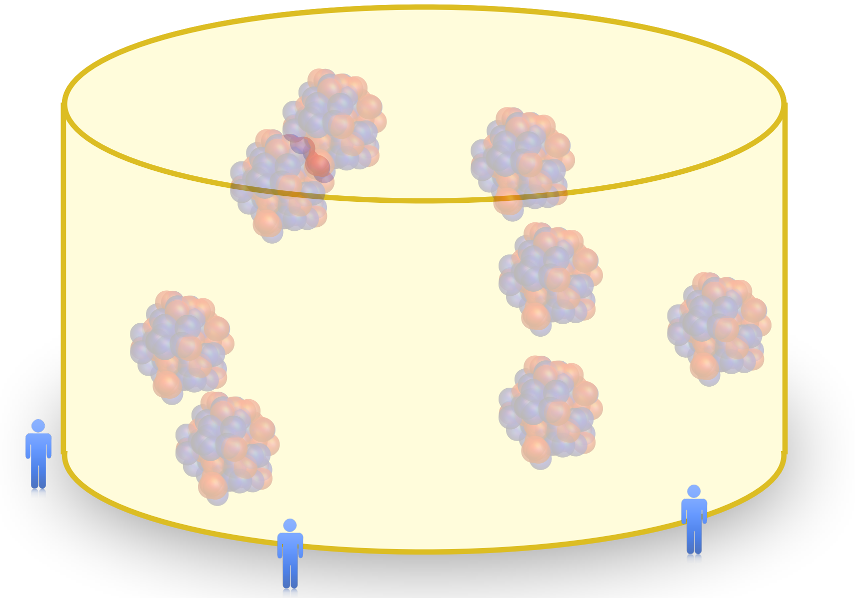
Motivation: ML for LQCD

First-principles nuclear physics beyond A=4

How finely tuned is the emergence of nuclear structure in nature?

Interpretation of intensity-frontier experiments

- Scalar matrix elements in A=131
  XENON1T dark matter direct detection search
- Axial form factors of Argon A=40
  DUNE long-baseline neutrino expt.
- Double-beta decay rates of Calcium A=48



Exponentially harder problems → Need exponentially improved algorithms

Machine learning for LQCD

APPROACH

Machine learning as ancillary tool for lattice QCD

- Accelerate gauge-field generation
- Optimise extraction of physics from gauge field ensemble

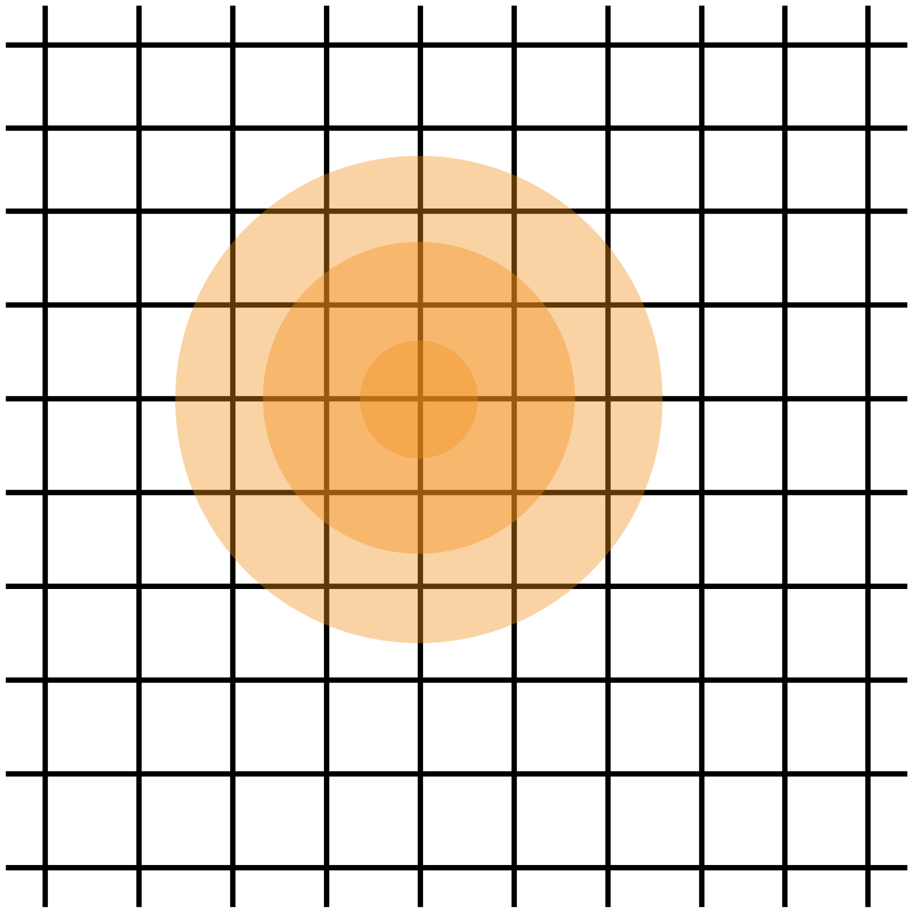
Will need to accelerate all stages of lattice QCD workflow to achieve physics goals

- **ONLY apply where quantum field theory can be rigorously preserved**

Accelerating HMC: action matching

QCD gauge field configurations sampled via

Hamiltonian dynamics + Markov Chain Monte Carlo



Updates diffusive

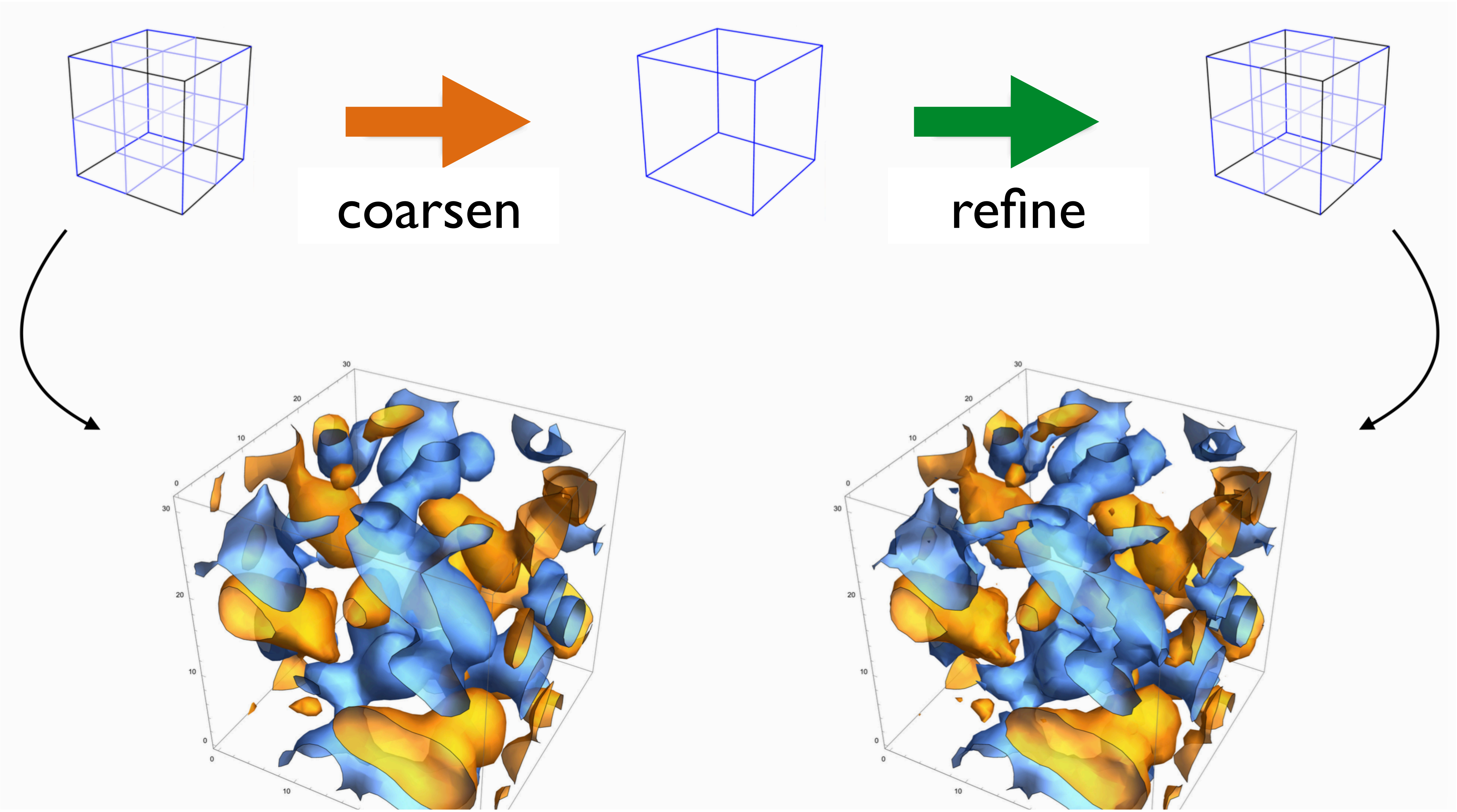
Lattice spacing → 0

Number of updates to change fixed physical length scale → ∞

"Critical slowing-down" of generation of uncorrelated samples

Multi-scale HMC updates

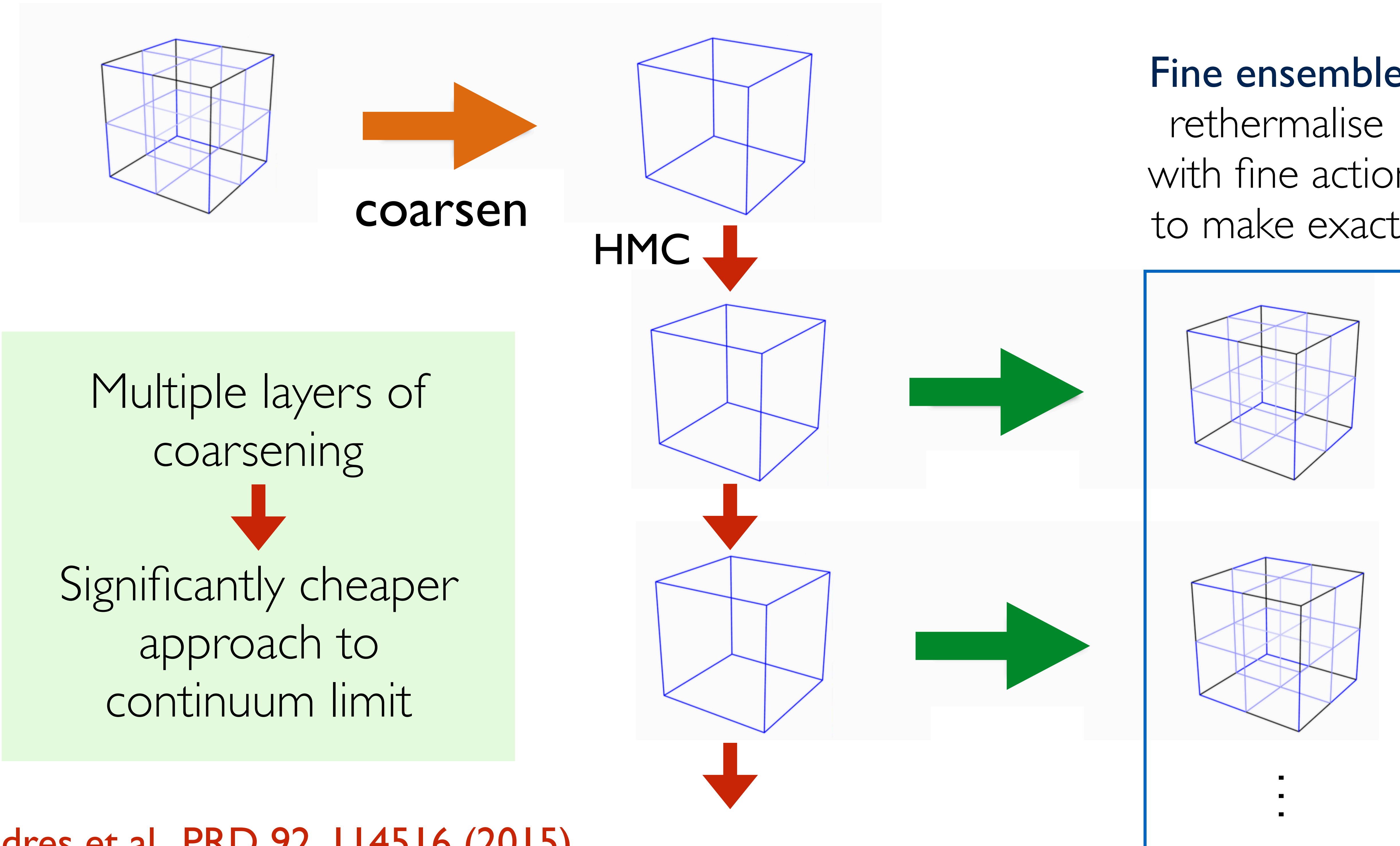
Given coarsening and refinement procedures…



Endres et al., PRD 92, 114516 (2015)

Multi-scale HMC updates

Perform HMC updates at coarse level



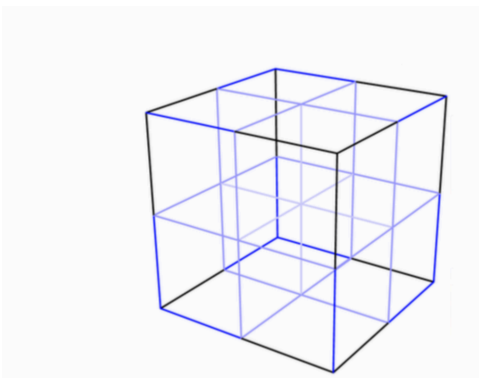
Multiple layers of coarsening

Significantly cheaper approach to continuum limit

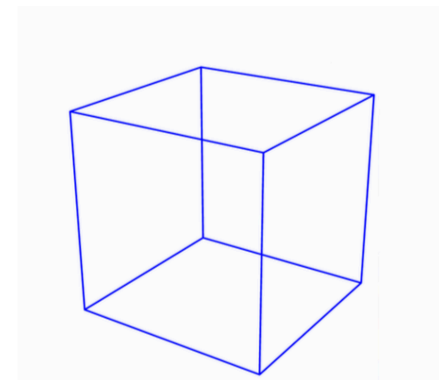
Endres et al., PRD 92, 114516 (2015)

Multi-scale HMC updates

Perform HMC updates at coarse level



encode same long-distance physics



MUST KNOW parameters of coarse QCD action that reproduce ALL physics parameters of fine simulation

Map a subset of physics parameters in the coarse space and match to coarsened ensemble

OR

Solve regression problem directly: "Given a coarse ensemble, what parameters generated it?"

Machine learning LQCD

Neural networks excel on problems where

Basic data unit has little meaning

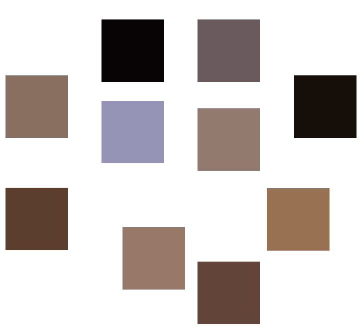
Combination of units is meaningful

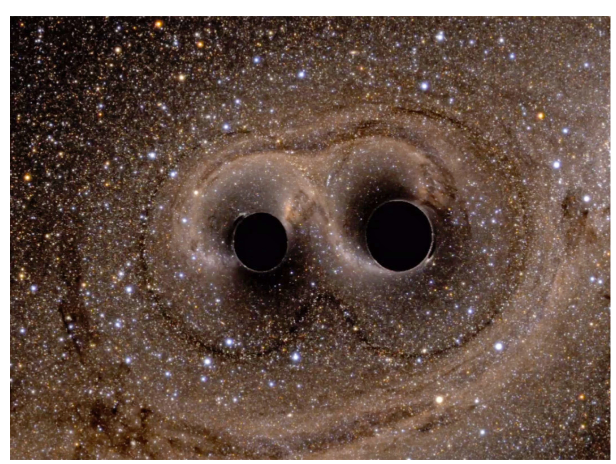
Image recognition

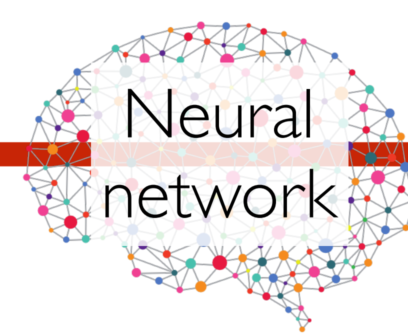
Pixel

Image

Label







"Colliding black holes"

Machine learning LQCD

Neural networks excel on problems where

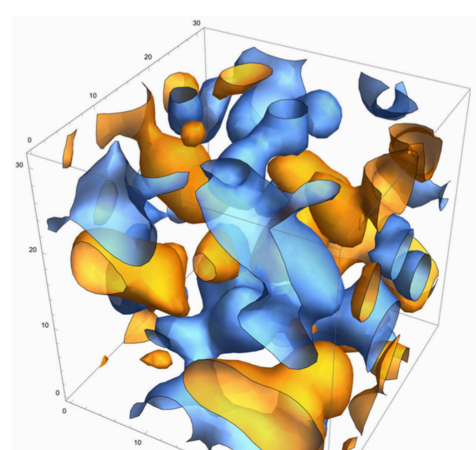
Basic data unit has little meaning ⟷ Combination of units is meaningful

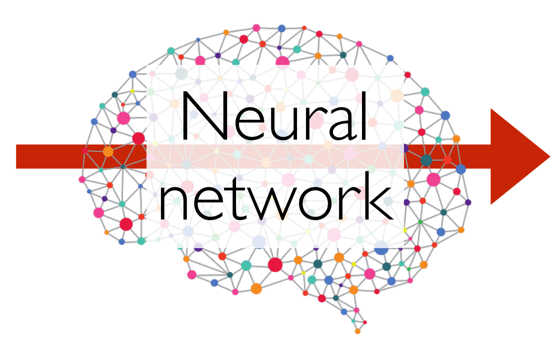
Parameter identification

Element of a colour matrix at one discrete space-time point

0 6 7 3 5
2 8 4 1
6

Ensemble of lattice QCD gauge field configurations





Label

Parameters of action

Machine learning LQCD

CIFAR benchmark image set for machine learning

- 32×32 pixels $\times$ 3 cols $\simeq 3000$ numbers
- 60000 samples
- Each image has meaning
- Local structures are important
- Translation-invariance within frame

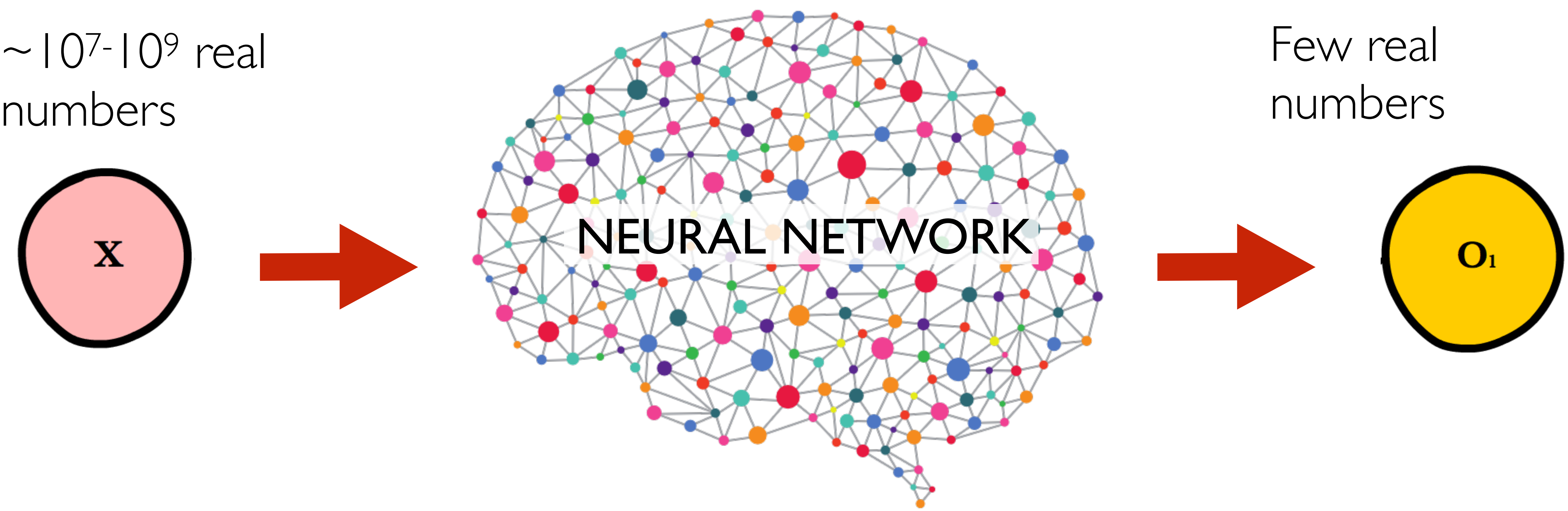
Ensemble of lattice QCD gauge fields

- $64^3 \times 128 \times 4 \times N_c^2 \times 2 \simeq 10^9$ numbers
- ~ 1000 samples
- Ensemble of gauge fields has meaning
- Long-distance correlations are important
- Gauge and translation-invariant with periodic boundaries

Regression by neural network

Lattice QCD gauge field

Parameters of lattice action



- **Complete:** not restricted to affordable subset of physics parameters
- **Instant:** once trained over a parameter range

Naive neural network

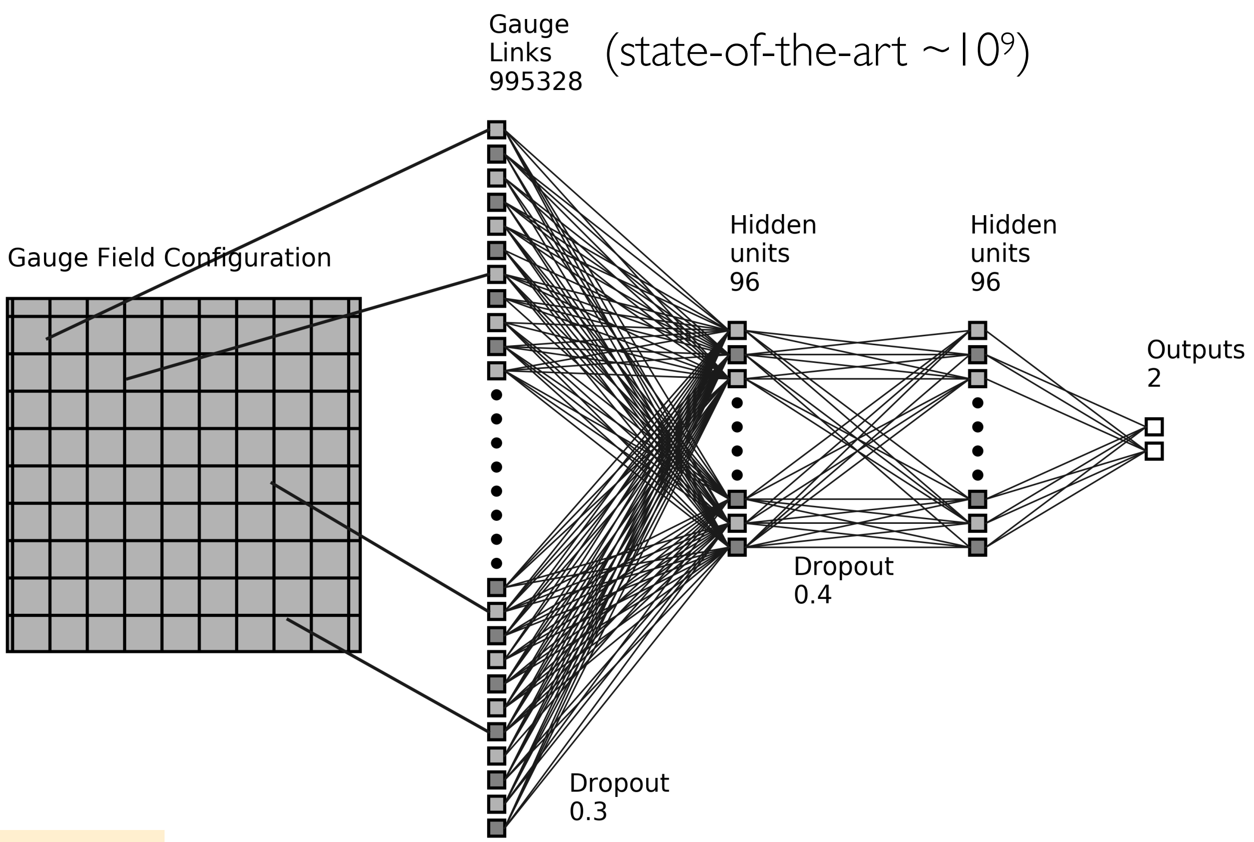
Simplest approach ⟷ Ignore physics symmetries

Train simple neural network on regression task

- Fully-connected structure
- Far more degrees of freedom than number of training samples available

"Inverted data hierarchy" → Recipe for overfitting!

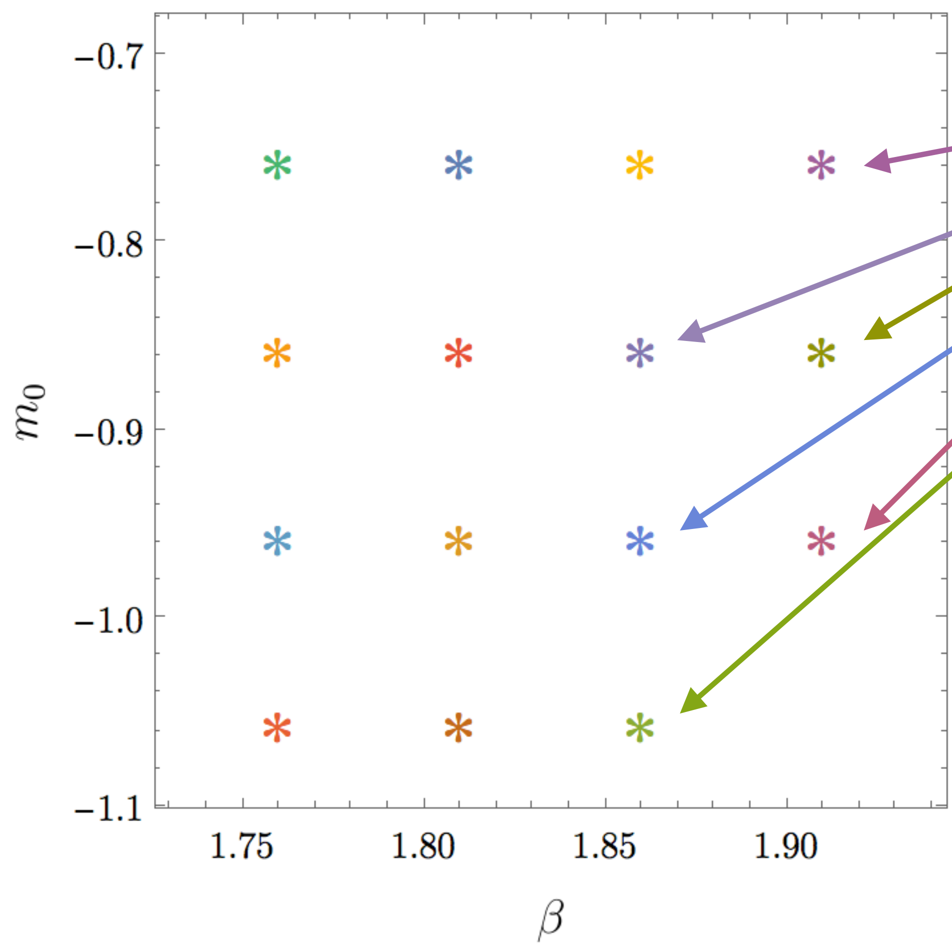
(state-of-the-art ~10^9)



Naive neural network

Training and validation datasets

Quark mass parameter



Parameter related to lattice spacing

* Parameters of training and validation datasets

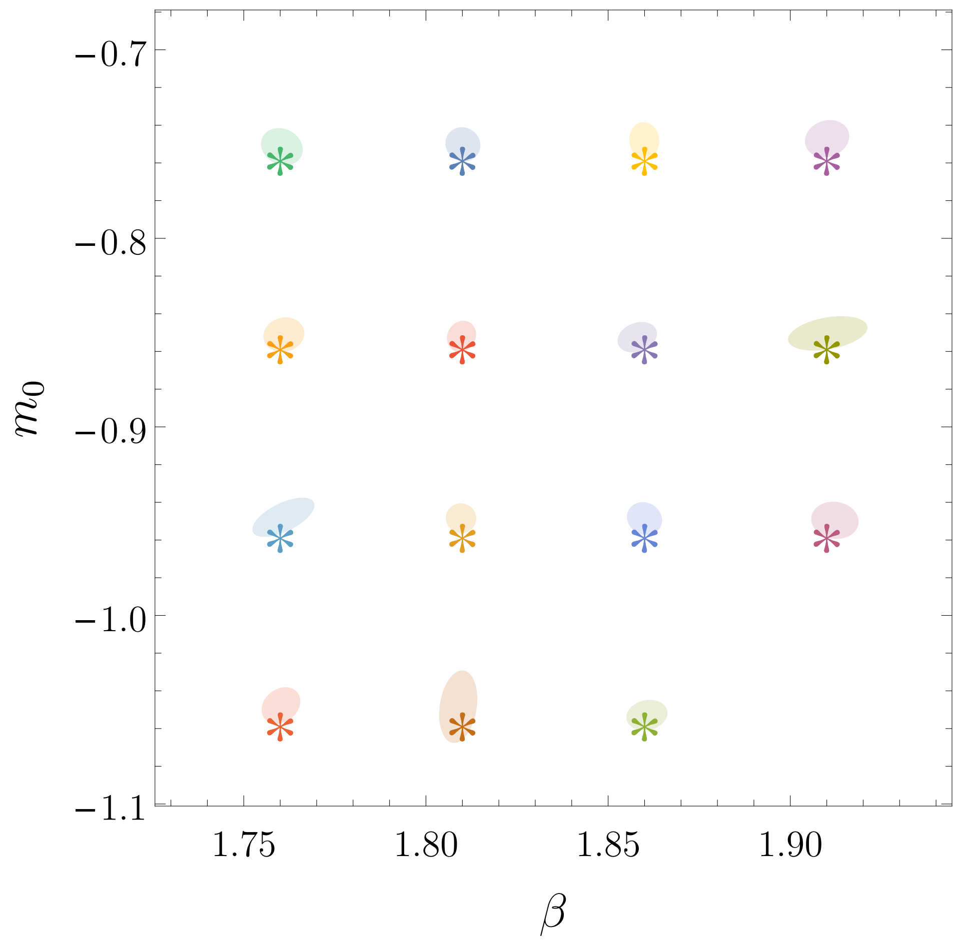
- O(10,000) independent configurations generated at each point
- Validation configurations randomly selected from generated streams

Spacing in evolution stream >> correlation time of physics observables

Naive neural network

Neural net predictions on validation data sets

Quark mass parameter



Parameter related to lattice spacing

* True parameter values

Confidence interval from ensemble of gauge fields

SUCCESS?

No sign of overfitting

- Training and validation loss equal
- Accurate predictions for validation data

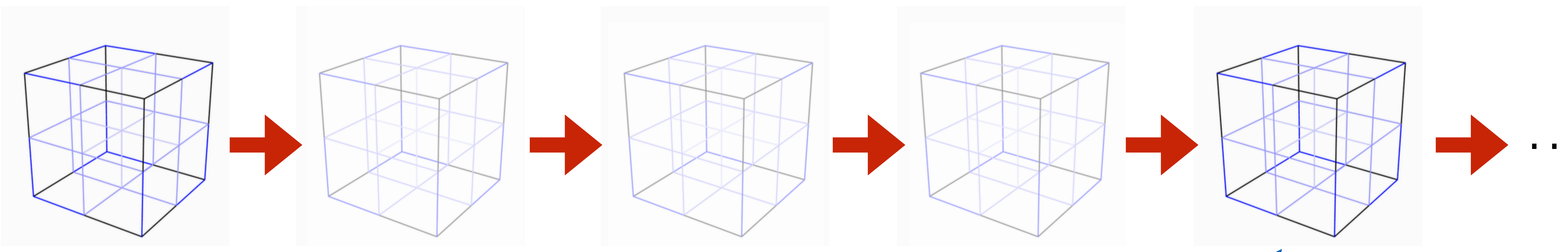
BUT fails to generalise to

- Ensembles at other parameters
- New streams at same parameters

NOT POSSIBLE IF CONFIGS ARE UNCORRELATED

Naive neural network

Stream of generated gauge fields at given parameters



Training/validation data selected from configurations spaced to be decorrelated (by physics observables)

- Network succeeds for validation configs from same stream as training configs
- Network fails for configs from new stream at same parameters

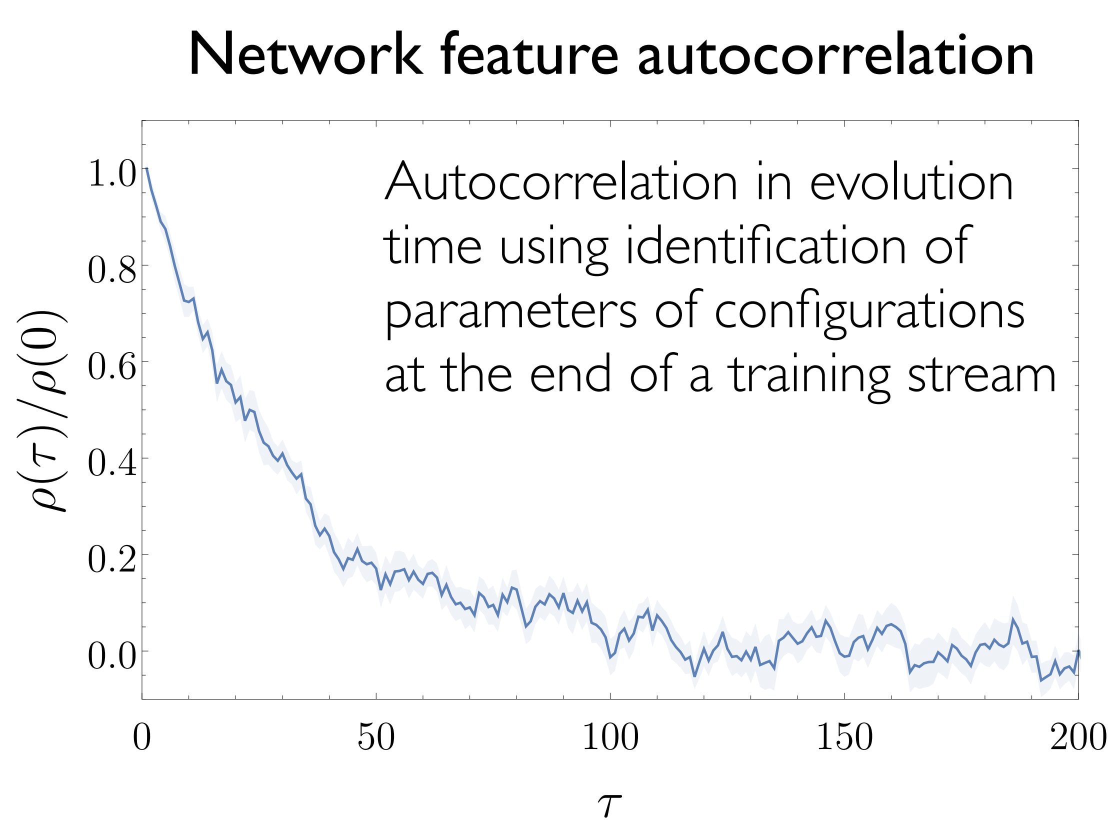
Network has identified feature with a longer correlation length than any known physics observable

Naive neural network

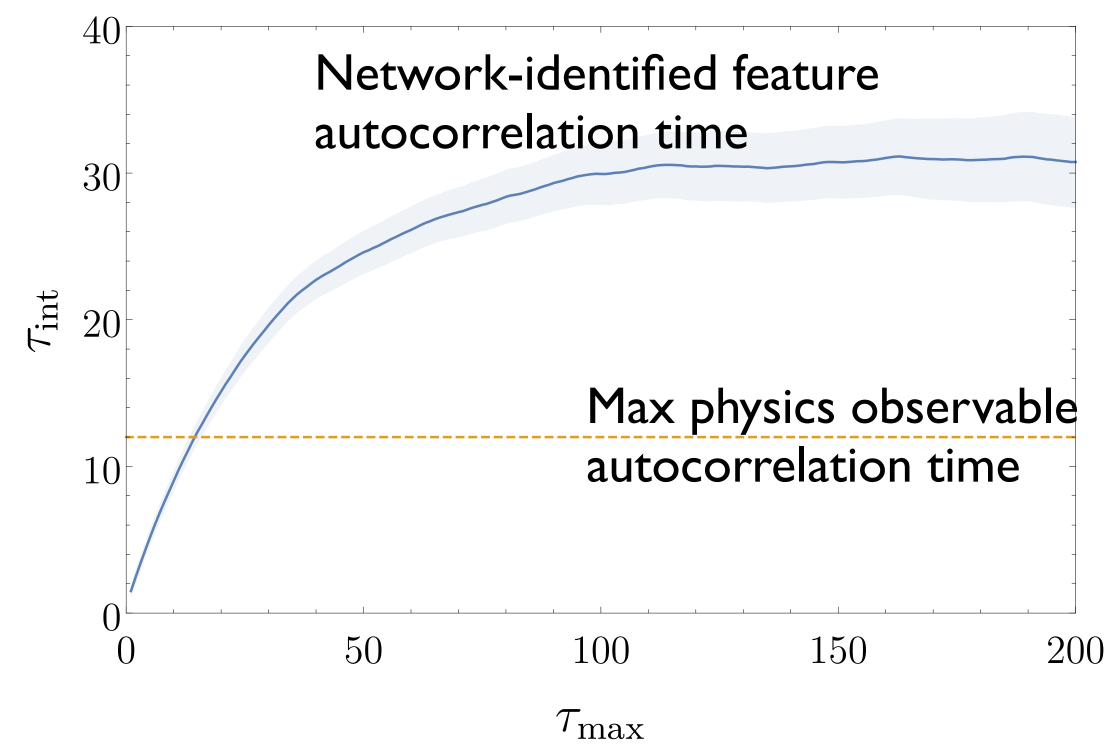
- Naive neural network that does not respect symmetries fails at parameter regression task

BUT

- Identifies unknown feature of gauge fields with a longer correlation length than any known physics observable



$$\tau_{\text{int}} = \frac{1}{2} + \lim_{\tau_{\text{max}}\to\infty} \frac{1}{\rho(0)} \sum_{\tau=0}^{\tau_{\text{max}}} \rho(\tau)$$



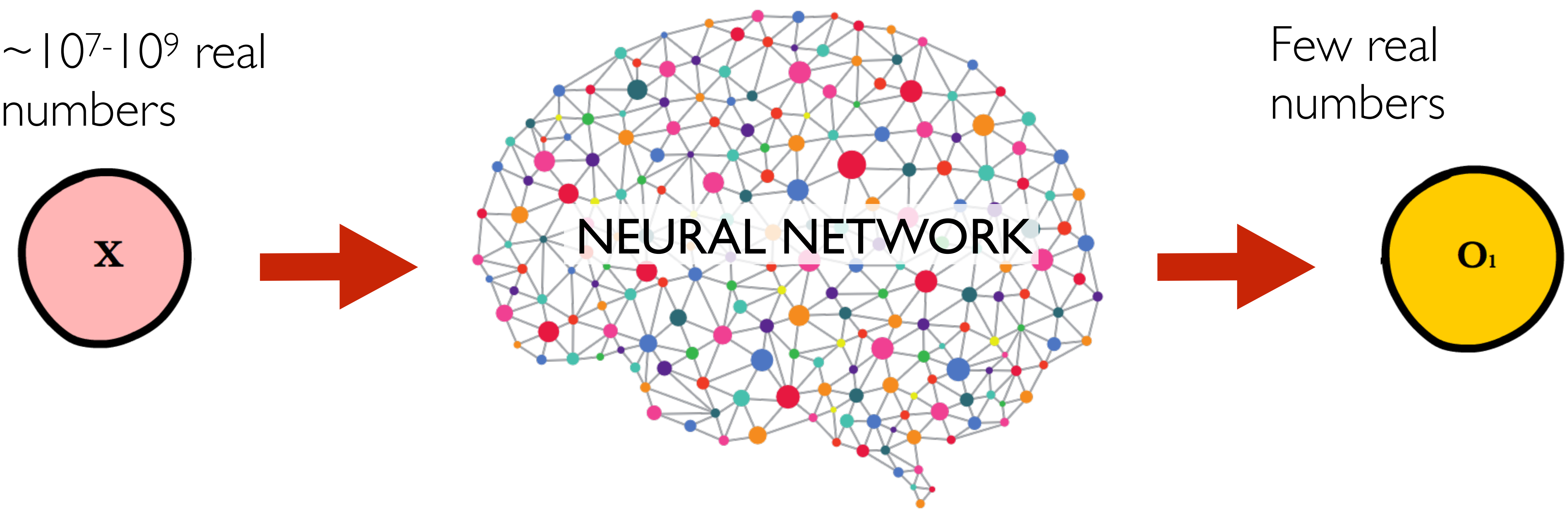
Regression by neural network

Lattice QCD gauge field

~10^7-10^9 real numbers

Parameters of lattice action

Few real numbers

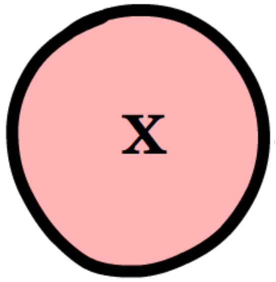


- **Complete:** not restricted to affordable subset of physics parameters
- **Instant:** once trained over a parameter range

Regression by neural network

Lattice QCD gauge field

~10^7-10^9 real numbers

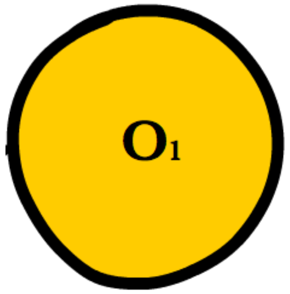


Custom network structures

- Respects gauge-invariance, translation-invariance, boundary conditions
- Emphasises QCD-scale physics
- Range of neural network structures find same minimum

Parameters of lattice action

Few real numbers

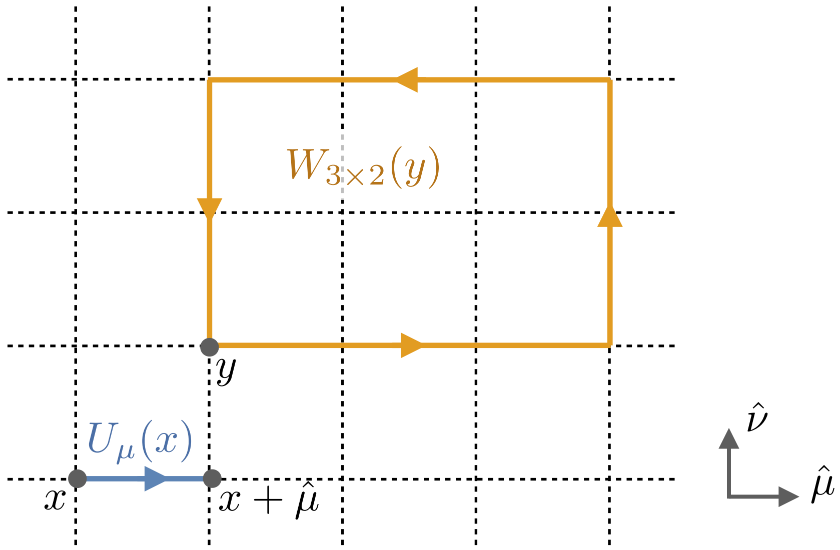


- **Complete:** not restricted to affordable subset of physics parameters
- **Instant:** once trained over a parameter range

Symmetry-preserving network

Network based on symmetry-invariant features

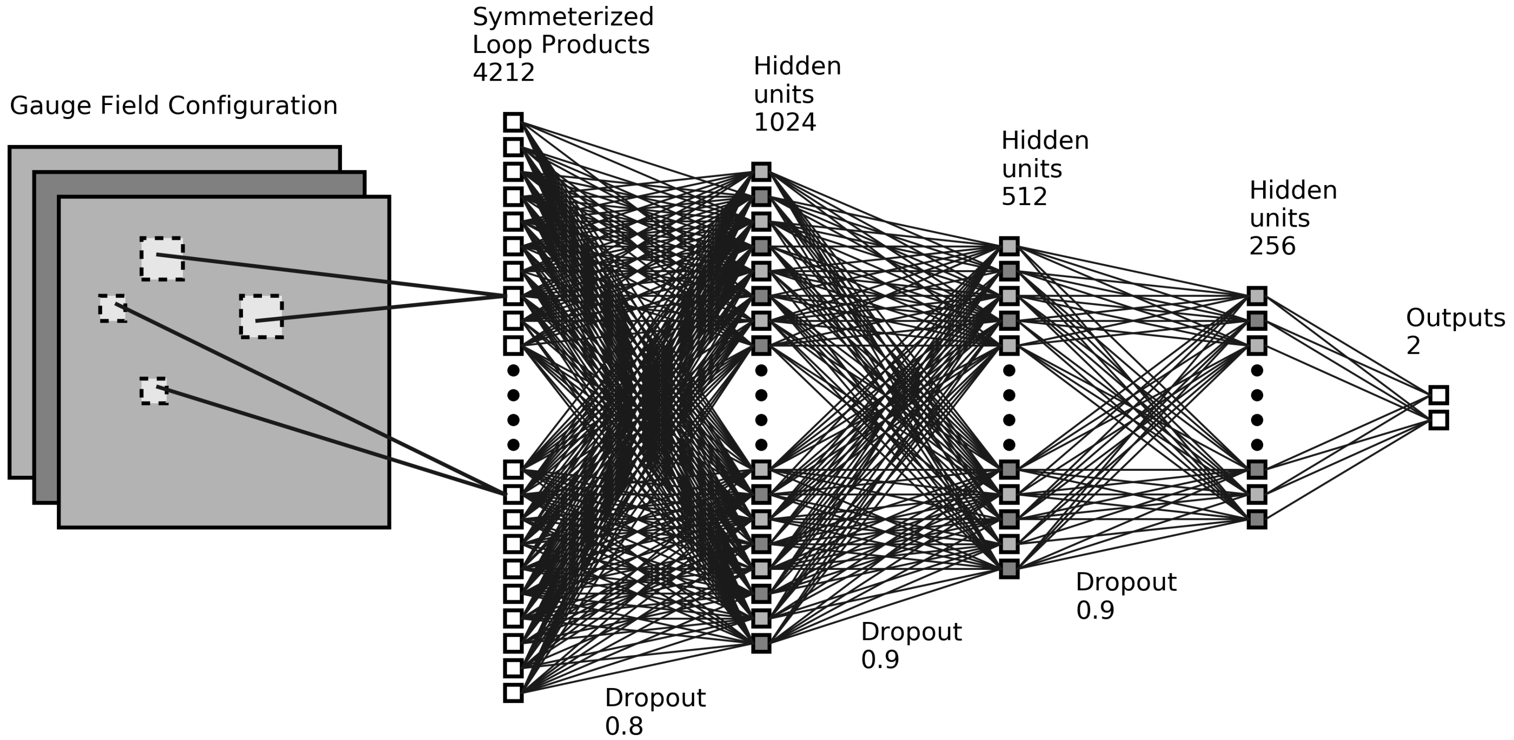
Closed Wilson loops
(gauge-invariant)



- Loops
- Correlated products of loops at various length scales
- Volume-averaged and rotation-averaged

Symmetry-preserving network

Network based on symmetry-invariant features

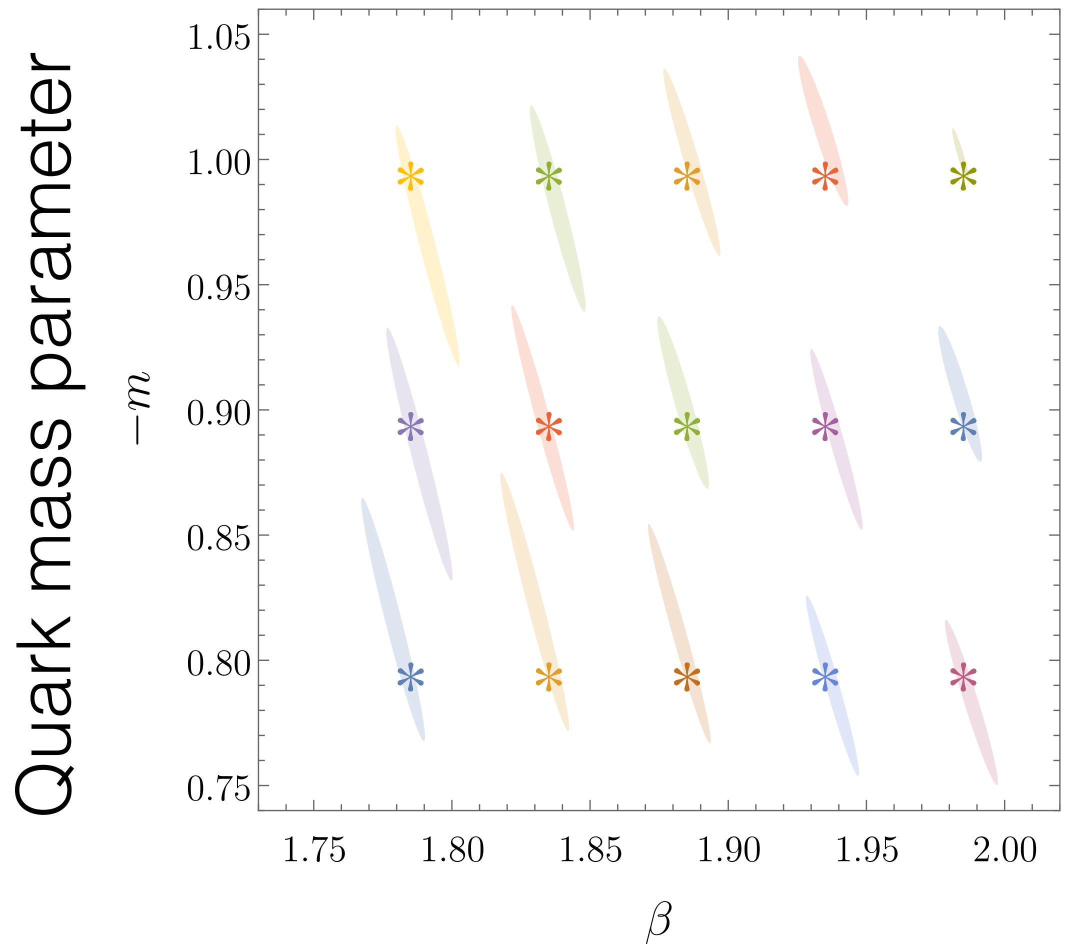


- Fully-connected network structure
- First layer samples from set of possible symmetry-invariant features

Number of degrees of freedom of network comparable to size of training dataset

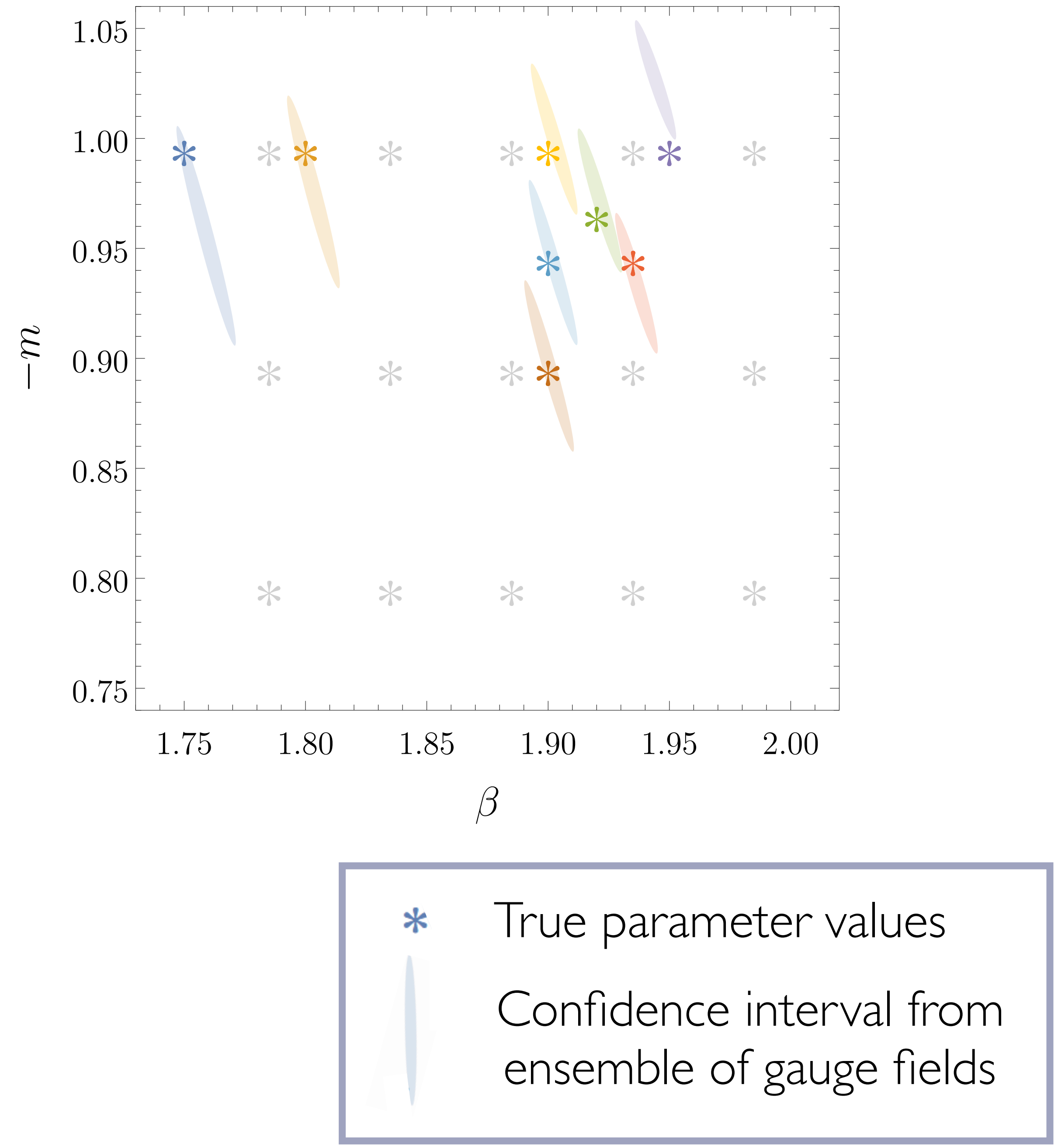
Gauge field parameter regression

Neural net predictions on validation data sets



Parameter related to lattice spacing

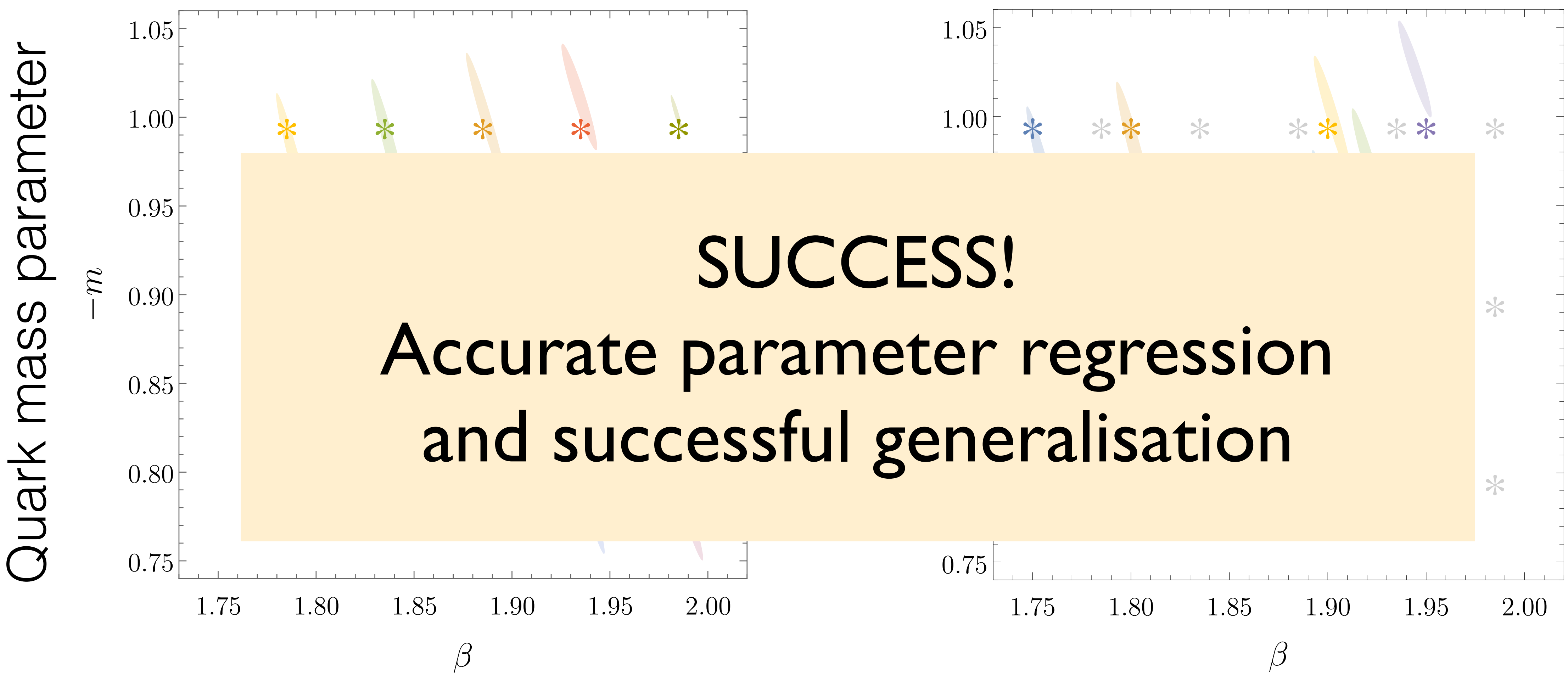
Predictions on new datasets



Gauge field parameter regression

Neural net predictions on validation data sets

Predictions on new datasets



SUCCESS!

Accurate parameter regression and successful generalisation

Parameter related to lattice spacing

* True parameter values

Confidence interval from ensemble of gauge fields

Gauge field parameter regression

PROOF OF PRINCIPLE

Step towards fine lattice generation at reduced cost

1. Generate one fine configuration
2. Find matching coarse action
3. HMC updates in coarse space
4. Refine and rethermalise

Accurate matching minimises cost of updates in fine space

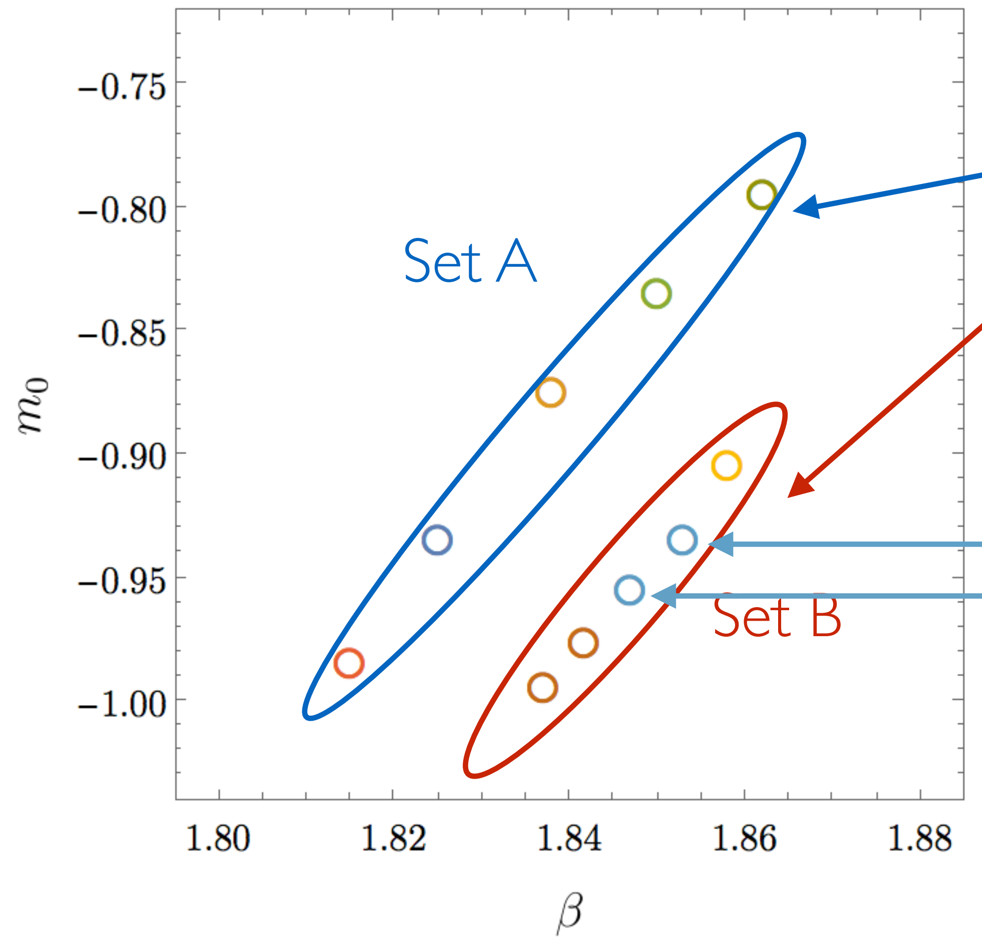
Guarantees correctness ✓

Shanahan, Trewartha, Detmold, PRD (2018) [1801.05784]

Tests of network success

How does neural network regression perform compared with other approaches?

Consider very closely-spaced validation ensembles at new parameters



Sets along lines of **constant 1x1 Wilson loop** (most precise feature allowed by network)

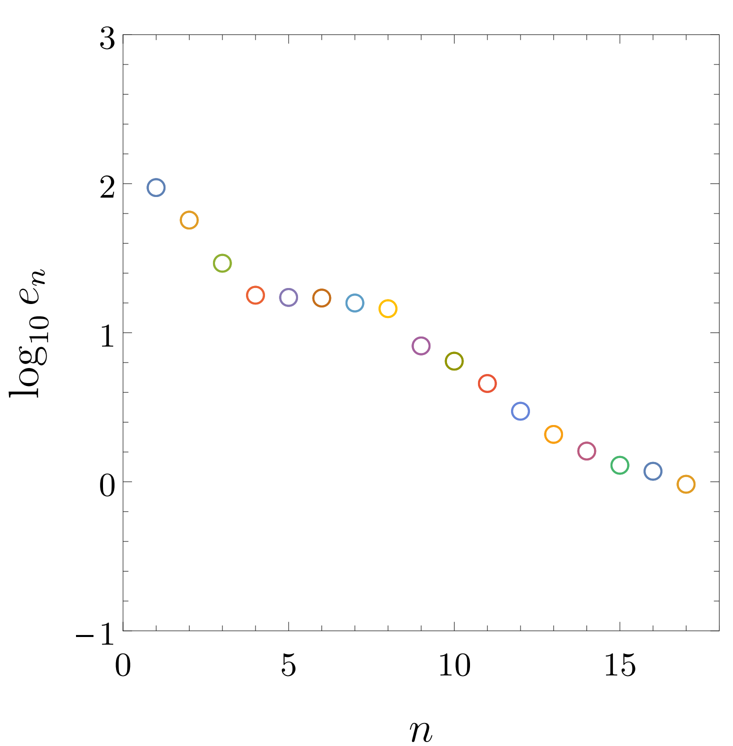
Much closer spacing than separation of training ensembles

Tests of network success

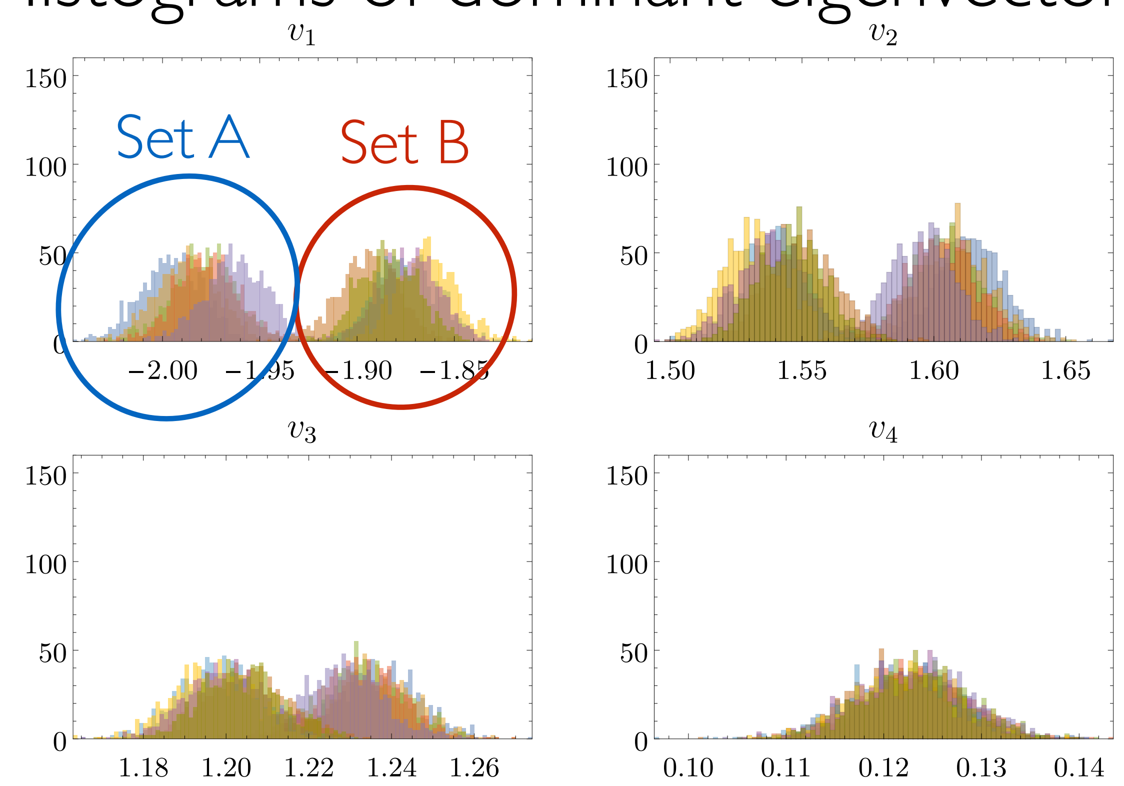
How does neural network regression perform compared with other approaches?

Consider very closely-spaced validation ensembles at new parameters: **not distinguishable to principal component analysis in loop space**

Eigenvalues



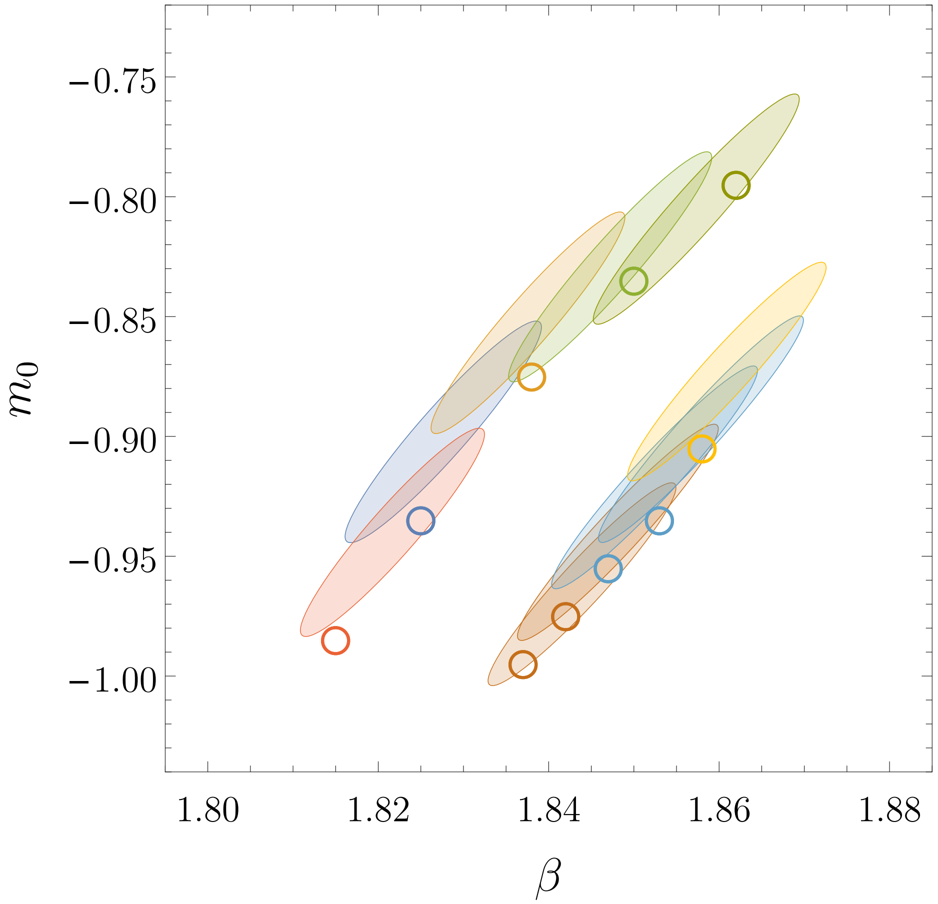
Histograms of dominant eigenvectors



Tests of network success

How does neural network regression perform compared with other approaches?

Consider very closely-spaced validation ensembles at new parameters: **distinguishable to trained neural network**



- Correct ordering of central values
- Accurate regression differences even at very fine resolution