EoS-meter of QCD transition from deep learning

LongGang Pang
UC Berkeley & LBNL

With Kai Zhou, Nan Su, Hannah Petersen, Horst Stoecker
from Frankfurt Institute for Advanced Studies, Germany
and Xin-Nian Wang from CCNU and LBNL

arXiv:1612.04262 [hep-ph]

2017.06.16 in CERN, ML group

What is deep learning?

Artificial Intelligence (AI)

Machine Learning (ML)

- PCA, kNN, k-means
- SVM
- Bayesian analysis
- Decision Tree
- Random Forest
- Neural Network
- Ensemble method
- …

Deep Learning (DL)

Learning multiple levels of representations using hierarchical or recurrent structures

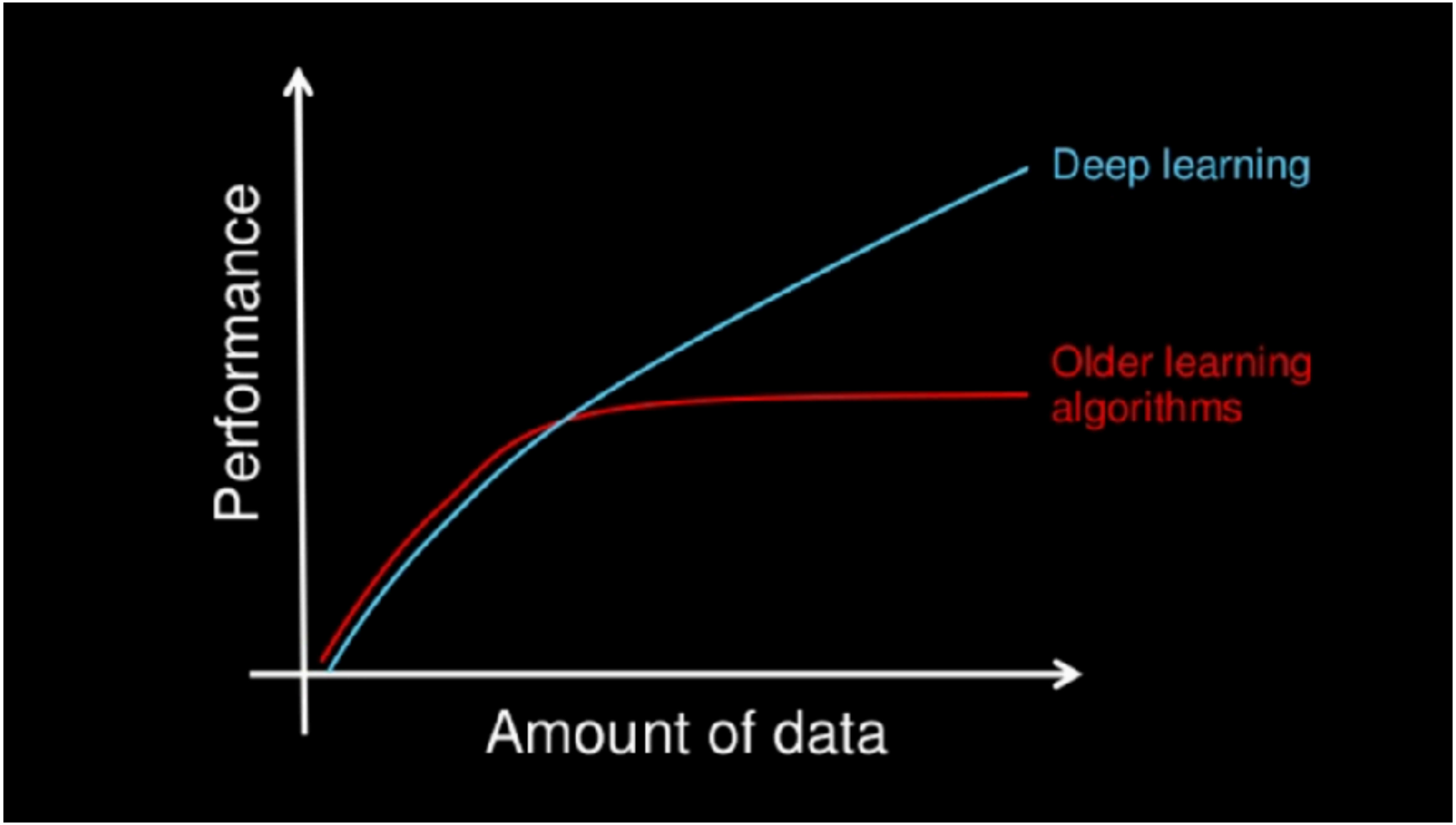
1. Big Data
2. GPU parallel
3. New architecture

2006

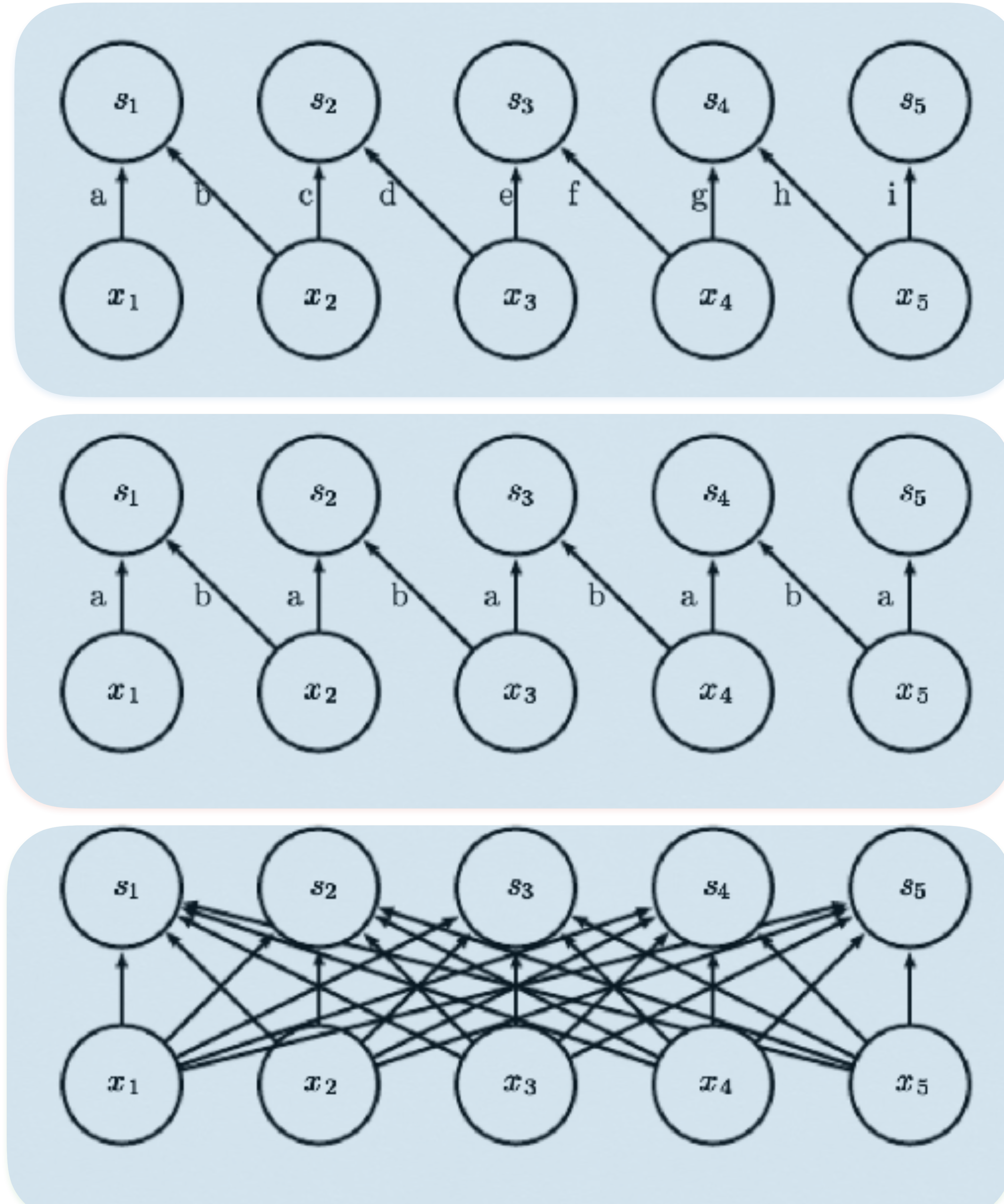
Geoffrey Hinton

Why deep learning?

Credit: Andrew NG



Most popular DL method in physics: DCNN



Locally Connected

Locally Connected + Share Weights

Convolution ✓

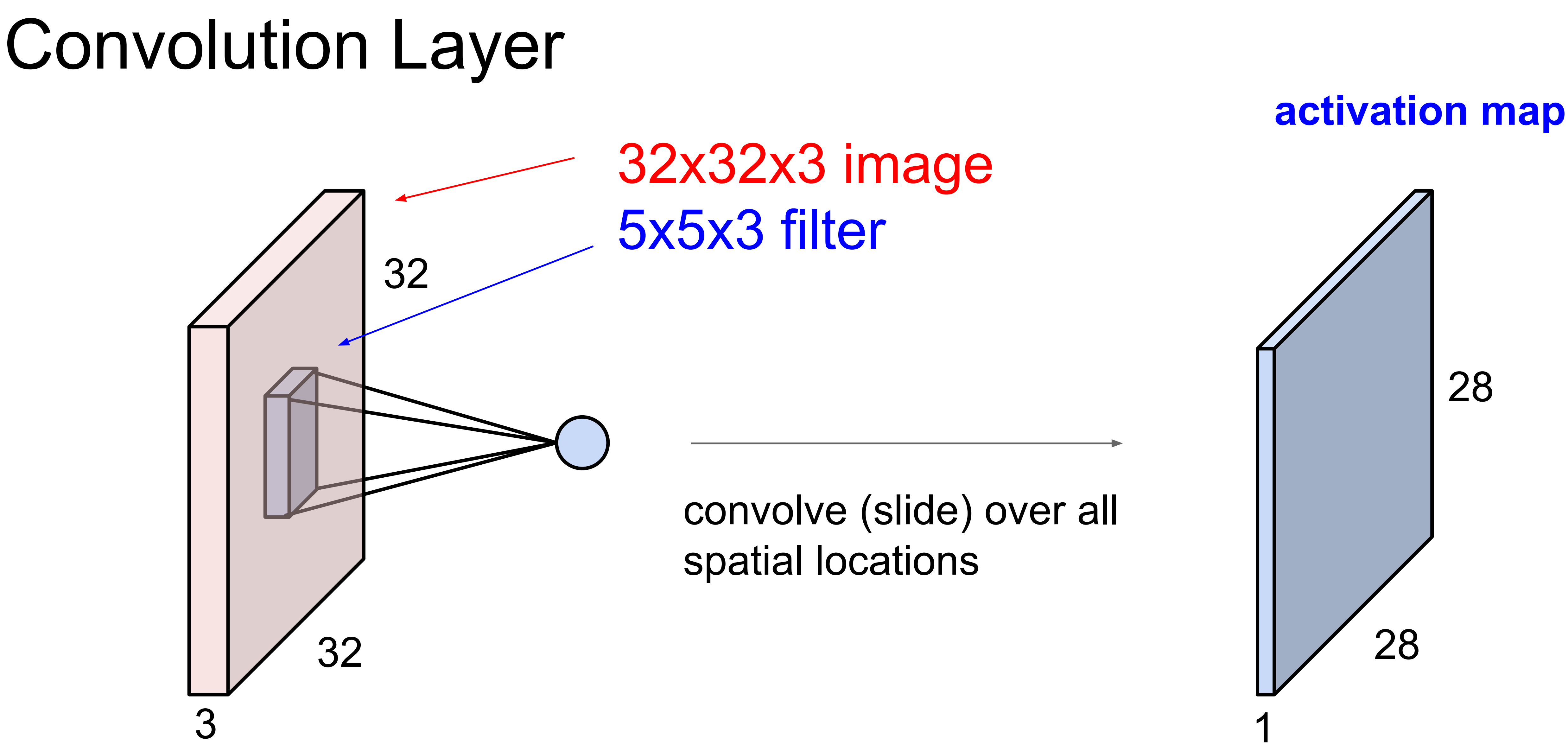
Fully Connected

DCNN = Deep Convolution Neural Network

"Deep Learning" Book

Most popular DL method in physics: DCNN

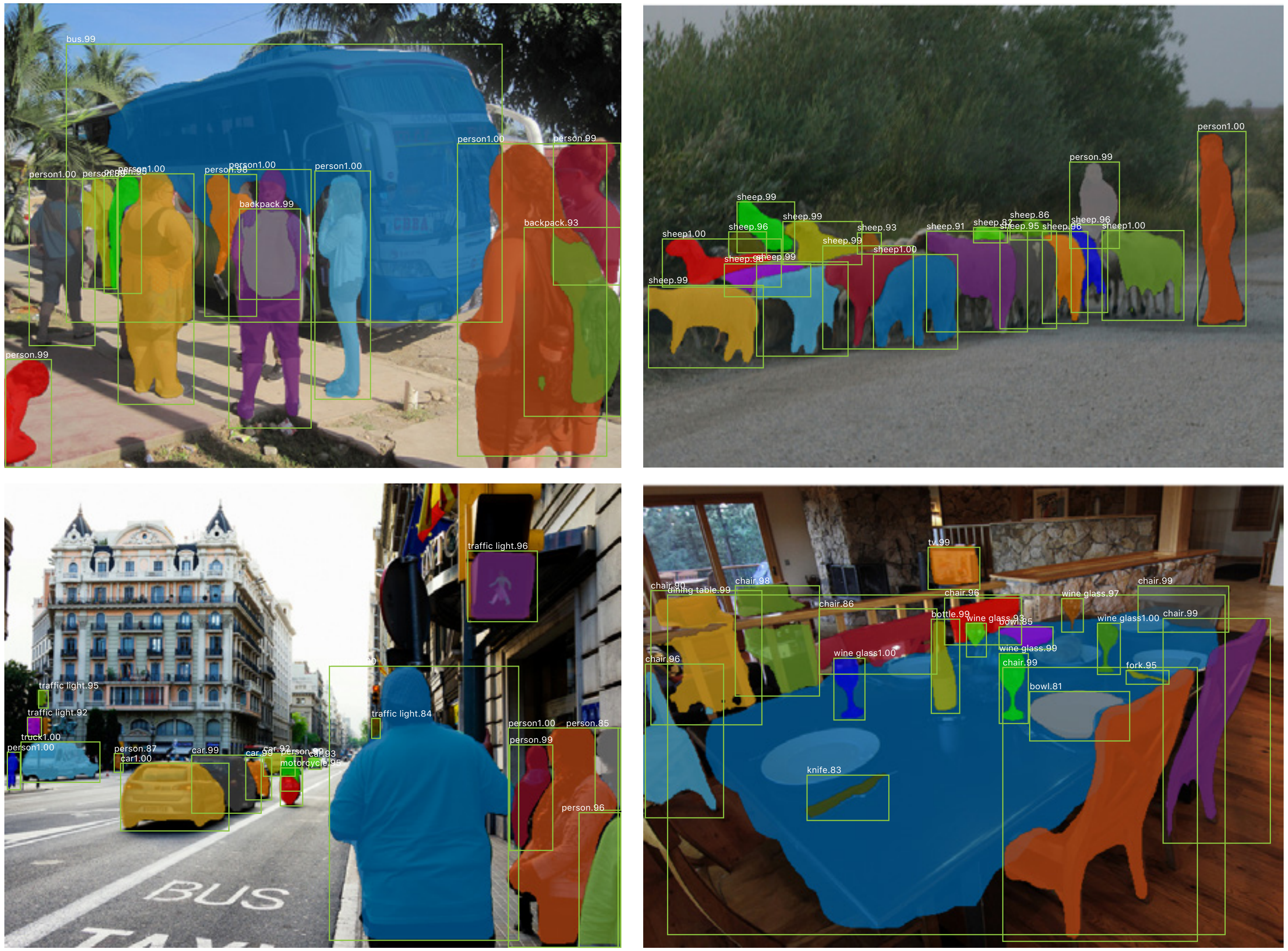
Fig from CS231N, Stanford



1. Much fewer parameters (local connection, share weights)
2. Translating, rotating, scaling invariance

Deep Convolution Neural Network (DCNN)

Mask R-CNN Kaiming He, Georgia Gkioxari, Piotr Dollár, Ross Girshick, 2017.04.05



- Extract abstract concepts (what is a dog?)
- High-quality segmentation mask, pixel level precision

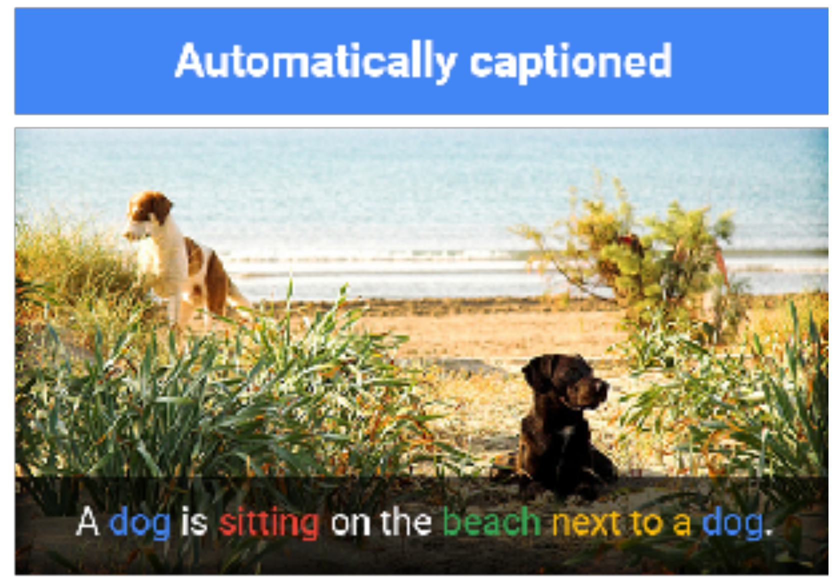
Deep Convolution Neural Network (DCNN)

Perceptual losses for real-time style transfer and super-resolution, by Johnson, Justin and Alahi, Alexandre and Fei-Fei, Li

- Extract/apply artistic style to new FIAS building.

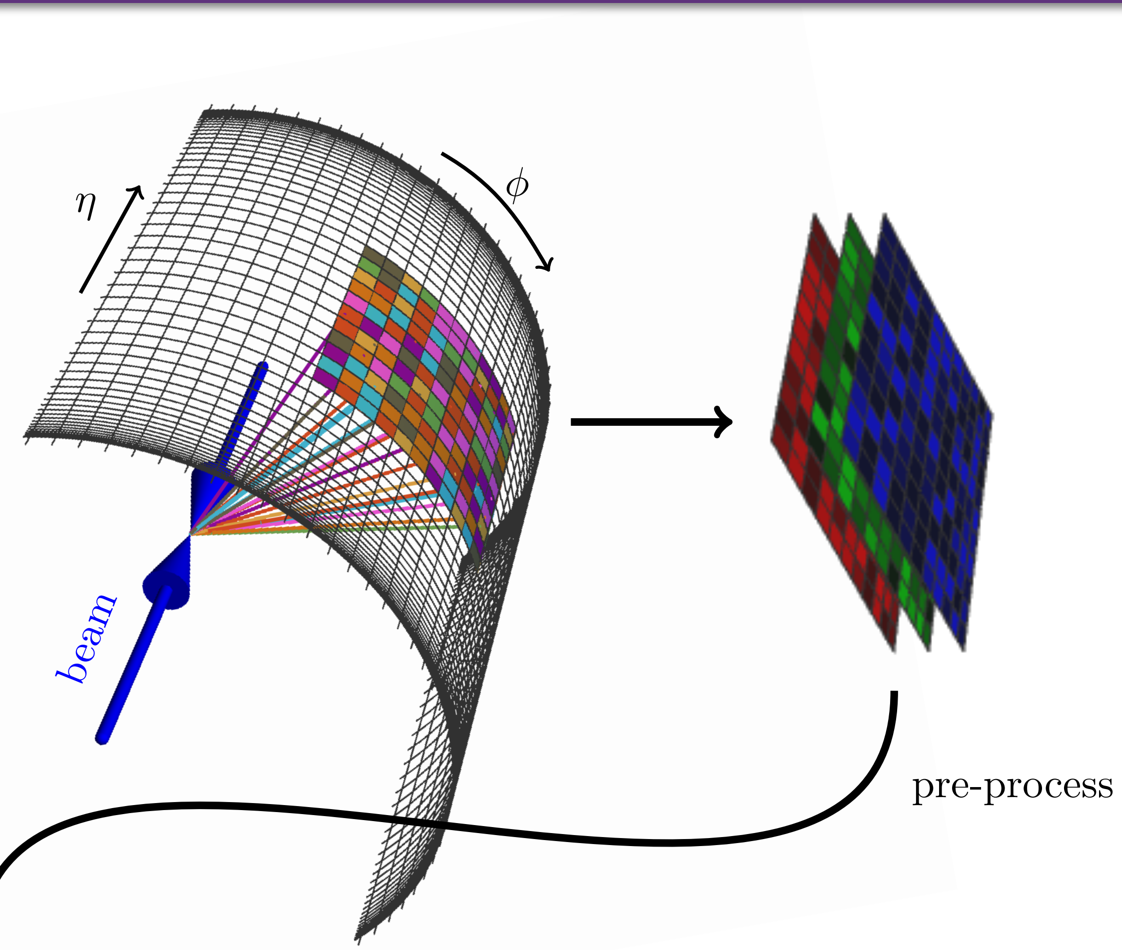
Deep Convolution Neural Network (DCNN)

Google, DeepMind



- Understand actions and relations.
- Treat objects of the images as irrelevant features.

Deep learning in Physics (Hadron colliders)



P. Baldi, K. Bauer, C. Eng, P. Sadowski, and D. Whiteson, *Jet Substructure Classification in High-Energy Physics with Deep Neural Networks*, *Phys. Rev.* **D93** (2016), no. 9 094034, [arXiv:1603.09349].

D. Guest, J. Collado, P. Baldi, S.-C. Hsu, G. Urban, and D. Whiteson, *Jet Flavor Classification in High-Energy Physics with Deep Neural Networks*, arXiv:1607.08633.

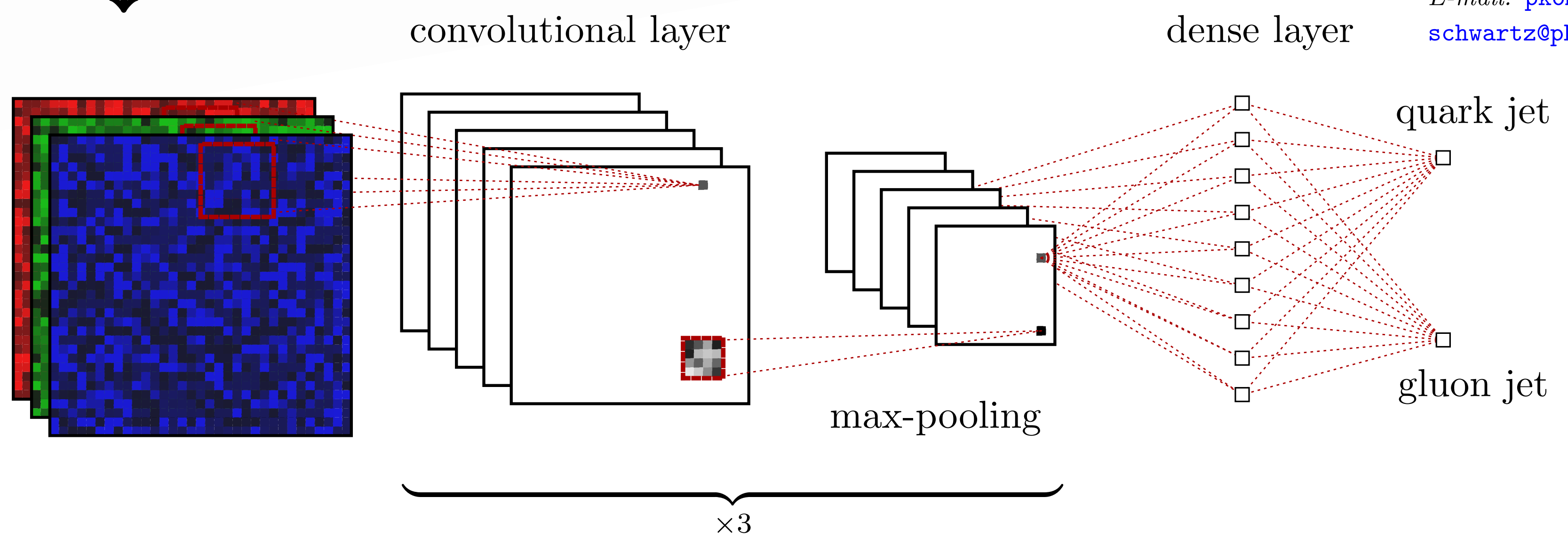
J. S. Conway, R. Bhaskar, R. D. Erbacher, and J. Pilot, *Identification of High-Momentum Top Quarks, Higgs Bosons, and W and Z Bosons Using Boosted Event Shapes*, arXiv:1606.06859.

J. Barnard, E. N. Dawe, M. J. Dolan, and N. Rajcic, *Parton Shower Uncertainties in Jet Substructure Analyses with Deep Neural Networks*, arXiv:1609.00607.

Deep learning in color: towards automated quark/gluon jet discrimination

Patrick T. Komiske,[a] Eric M. Metodiev,[a] and Matthew D. Schwartz[b]

[a] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[b] *Department of Physics, Harvard University, Cambridge, MA 02138, USA*
E-mail: pkomiske@mit.edu, metodiev@mit.edu, schwartz@physics.harvard.edu

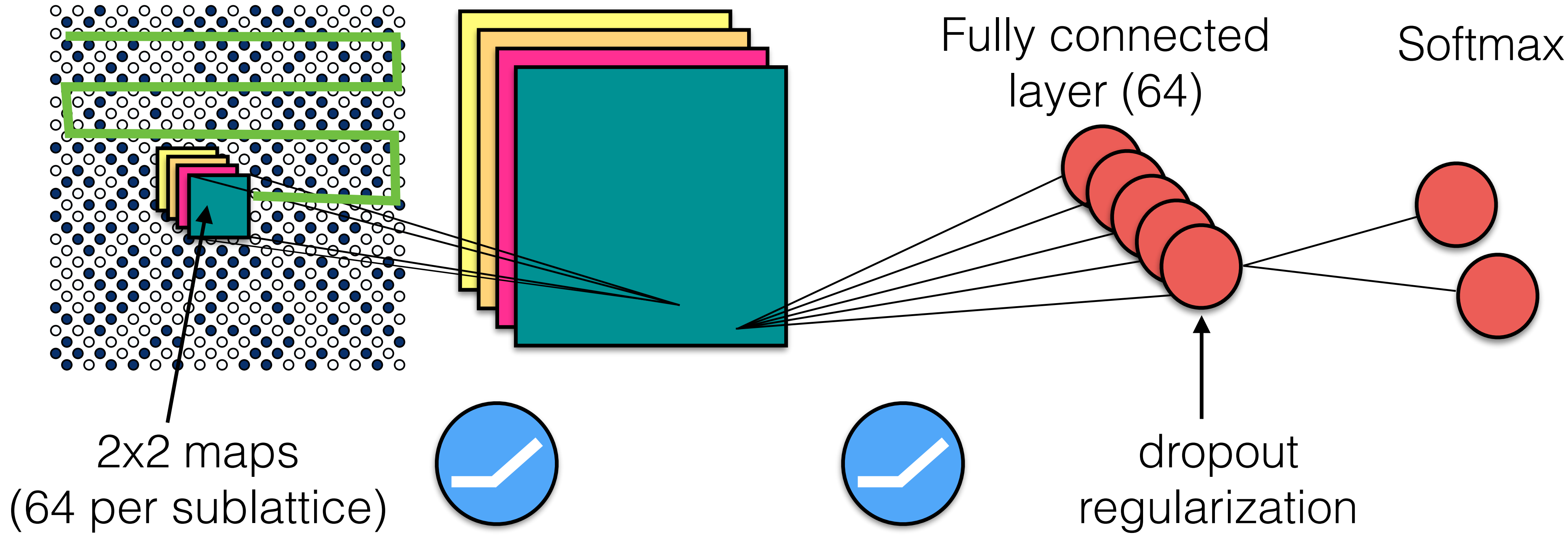


Deep CNN match or outperform traditional jet observables.

Figure 2: An illustration of the deep convolutional neural network architecture. The first layer is the input jet image, followed by three convolutional layers, a dense layer and an output layer.

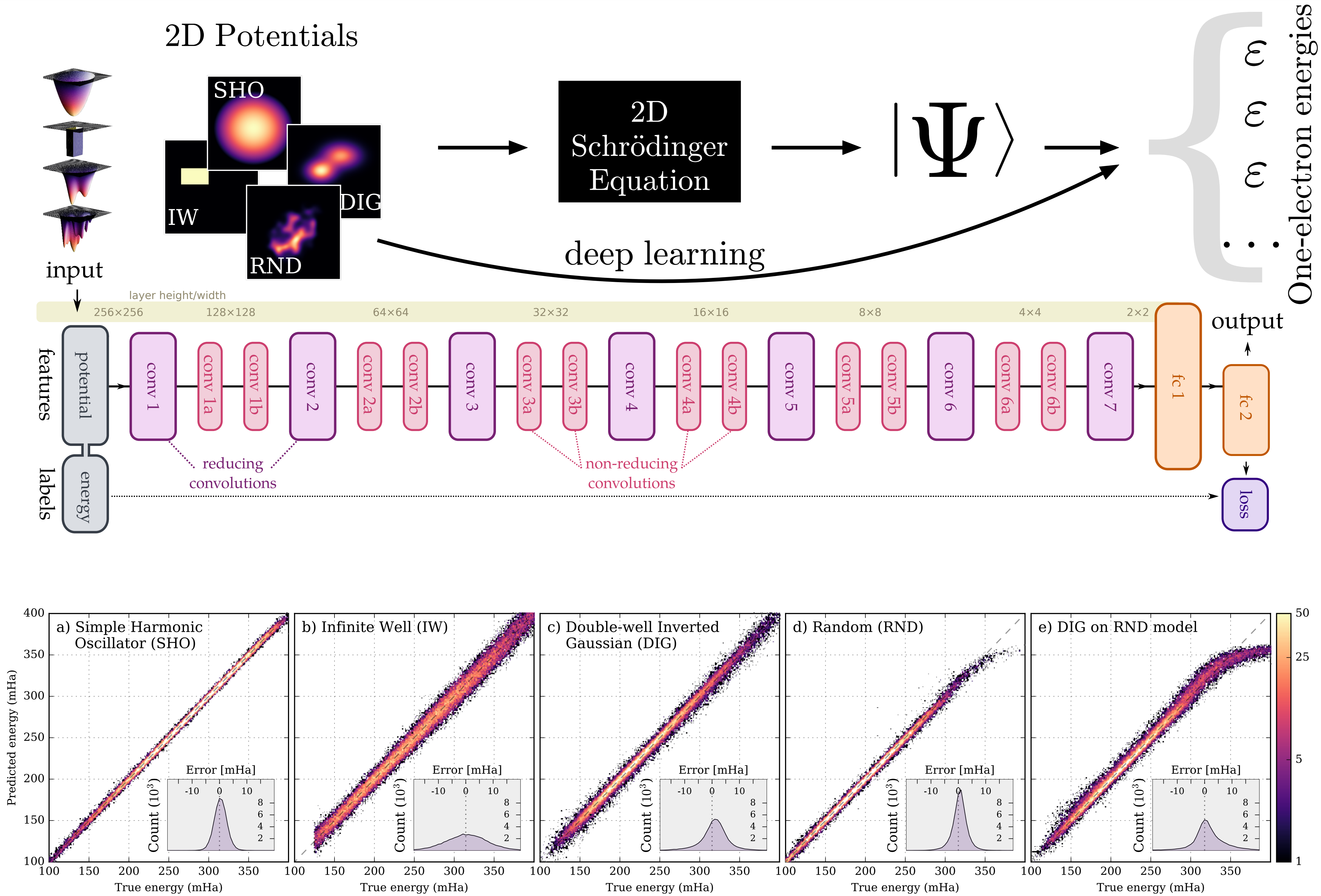
Machine learning phases of matter

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



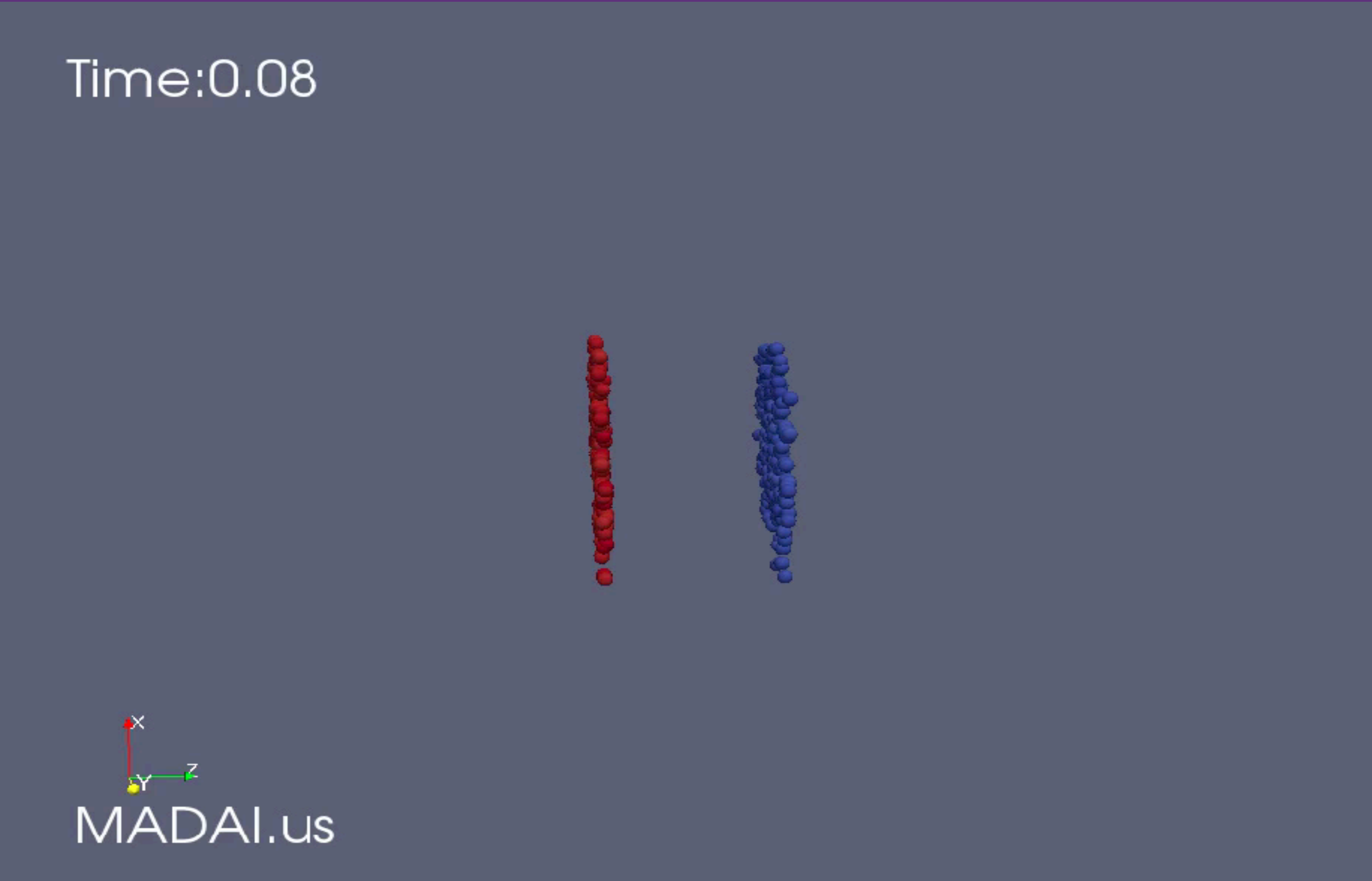
Determining phase from spin configurations

Deep learning in Physics (Solving Schrodinger Eqs)



Deep learning and the Schrodinger equation, by K. Mills, M. Spanner, Tamblyn (February 7, 2017)

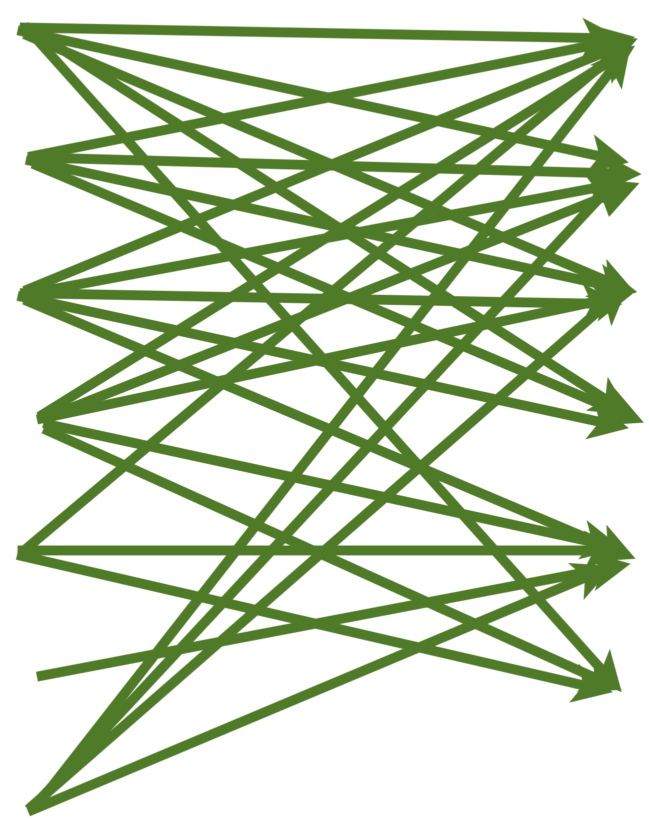
Relativistic high energy heavy ion collisions



Multiple parameters entangle with multiple observables

Model Parameter:

- eqn. of state
- shear viscosity
- initial state
- pre-equilibrium dynamics
- thermalization time
- quark/hadron chemistry
- particlization/freeze-out



experimental data:

- π/K/P spectra
- yields vs. centrality & beam
- elliptic flow
- HBT
- charge correlations & BFs
- density correlations

Fig from S. Bass QM2017 (Bayesian method)

state-of-the-art model-data comparison

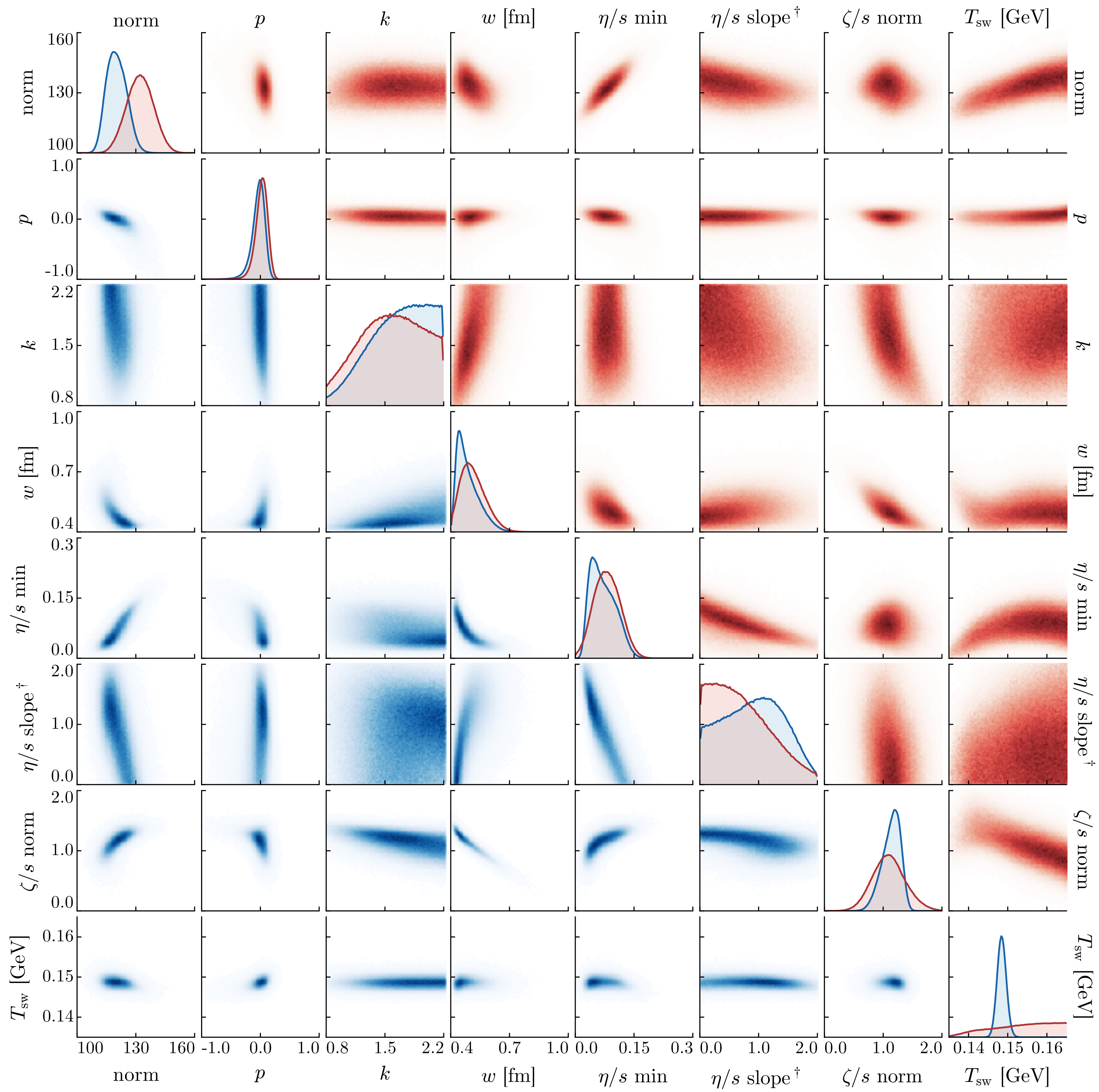


FIG. 7. Posterior distributions for the model parameters from calibrating to identified particles yields (blue, lower triangle) and charged particles yields (red, upper triangle). The diagonal has marginal distributions for each parameter, while the off-diagonal contains joint distributions showing correlations among pairs of parameters. †The units for η/s slope are [GeV^{-1}].

TABLE I. Input parameter ranges for the initial condition and hydrodynamic models.

Parameter	Description	Range
Norm	Overall normalization	100–250
p	Entropy deposition parameter	-1 to $+1$
k	Multiplicity fluct. shape	0.8–2.2
w	Gaussian nucleon width	0.4–1.0 fm
η/s hrg	Const. shear viscosity, $T < T_c$	0.3–1.0
η/s min	Shear viscosity at T_c	0–0.3
η/s slope	Slope above T_c	0–2 GeV^{-1}
ζ/s norm	Prefactor for $(\zeta/s)(T)$	0–2
T_{switch}	Particlization temperature	135–165 MeV

- Bayesian method

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)}$$

X: model —— Y: data

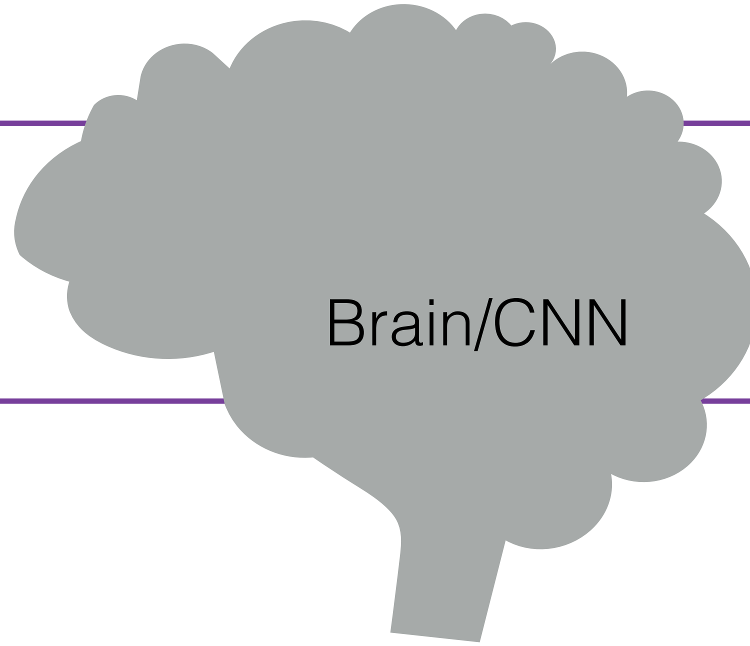
PRC 94.024907, J.E.Bernhard. et.el.

PRL. **114**, 202301, S. Pratt, et.el

Comments on the Bayesian method used in heavy ion physics

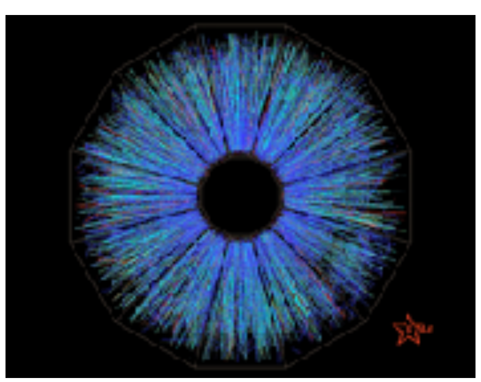
- Takes full use of the known features (expert-designed observables)
- The features are usually event-averaged for both model side and experimental side
- Can be improved by using more event-by-event information
- Relies on known features instead of learning new features from raw data (high dimensional data) or Monte Carlo simulations.

Brain/CNN neglects irrelevant features

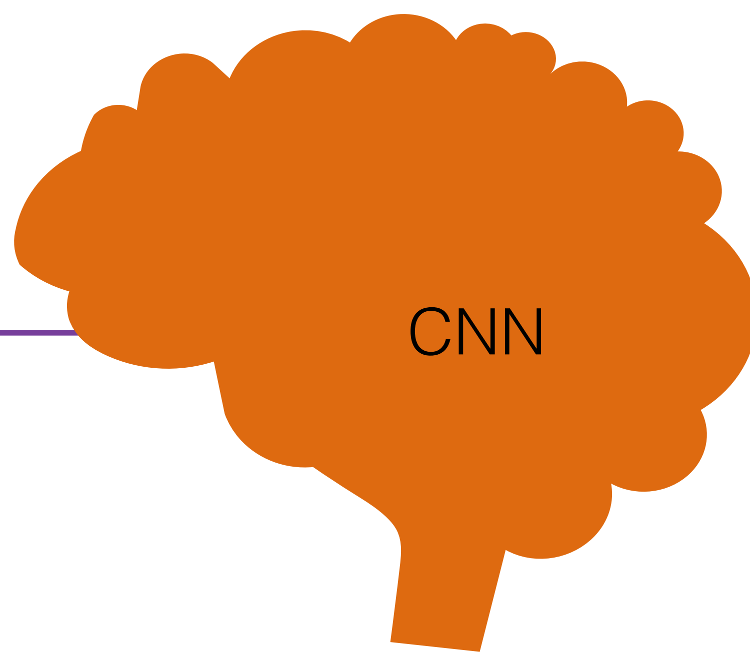


Dog

Cat



$\rho(p_T, \Phi)$



crossover or
1st order transition

Key idea for this proof-of-principle study

Supervised learning using deep convolution neural network with big amount of labeled training data (spectra, EoS type) from event-by-event relativistic hydrodynamics.

Open Source Libraries

Keras + TensorFlow in the present study

Keras is a high level neural network library, written in Python and capable of running on top of either TensorFlow or Theano.

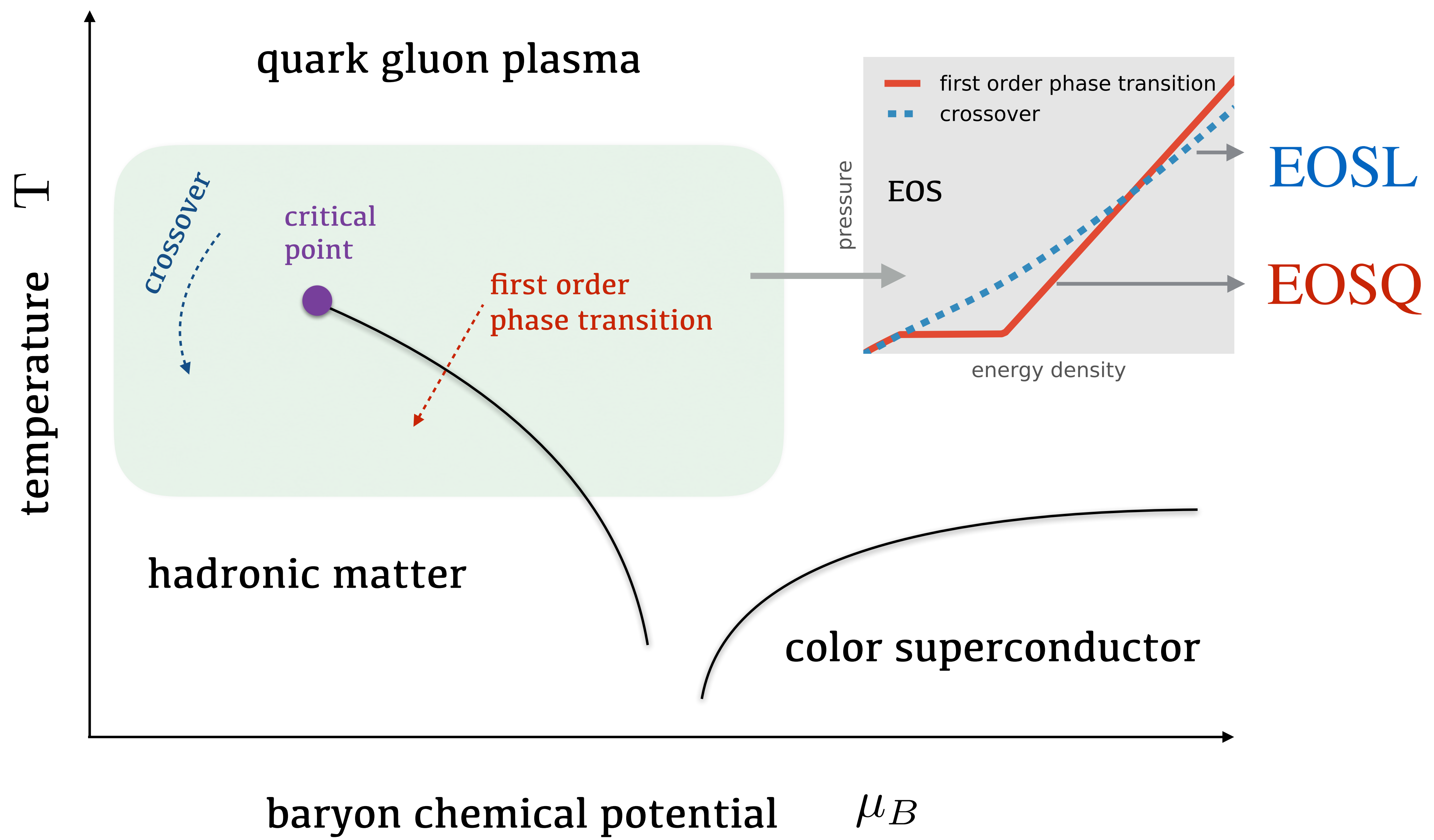
```
# Build one fully connected neural network (784->10->10 neurons) in Keras, for MNIST

from keras.models import Sequential
from keras.layers import Dense, Activation

model = Sequential()
model.add(Dense(output_dim=10, input_dim=784))
model.add(Activation("relu"))
model.add(Dense(output_dim=10))
model.add(Activation(“softmax"))
model.compile(loss='categorical_crossentropy', optimizer='sgd',
metrics=['accuracy'])
```

2017/01/15: Good news, Tensorflow chooses Keras!

EoS


quark gluon plasma
temperature T
crossover
critical point
first order phase transition
hadronic matter
color superconductor
baryon chemical potential μ_B
first order phase transition
crossover
EOS
pressure
energy density
EOSL
EOSQ

Model (3+1D viscous hydrodynamics)

CLVisc: a (3+1)D viscous hydrodynamics parallelized on GPU using OpenCL

$$\nabla_\mu T^{\mu\nu} = 0 \tag{1}$$

$$\Delta^{\mu\nu\alpha\beta} u^\lambda \nabla_\lambda \pi_{\alpha\beta} = -\frac{\pi^{\mu\nu} - \pi^{\mu\nu}_{\rm NS}}{\tau_\pi} - \frac{4}{3}\pi^{\mu\nu}\nabla_\lambda u^\lambda \tag{2}$$

where

$$T^{\mu\nu} = (\varepsilon + P)u^\mu u^\nu - P g^{\mu\nu} + \pi^{\mu\nu} \tag{3}$$

$$\Delta^{\mu\nu\alpha\beta} = \frac{1}{2}(\Delta^{\mu\alpha}\Delta^{\nu\beta} + \Delta^{\nu\alpha}\Delta^{\mu\beta}) - \frac{1}{3}\Delta^{\mu\nu}\Delta^{\alpha\beta} \tag{4}$$

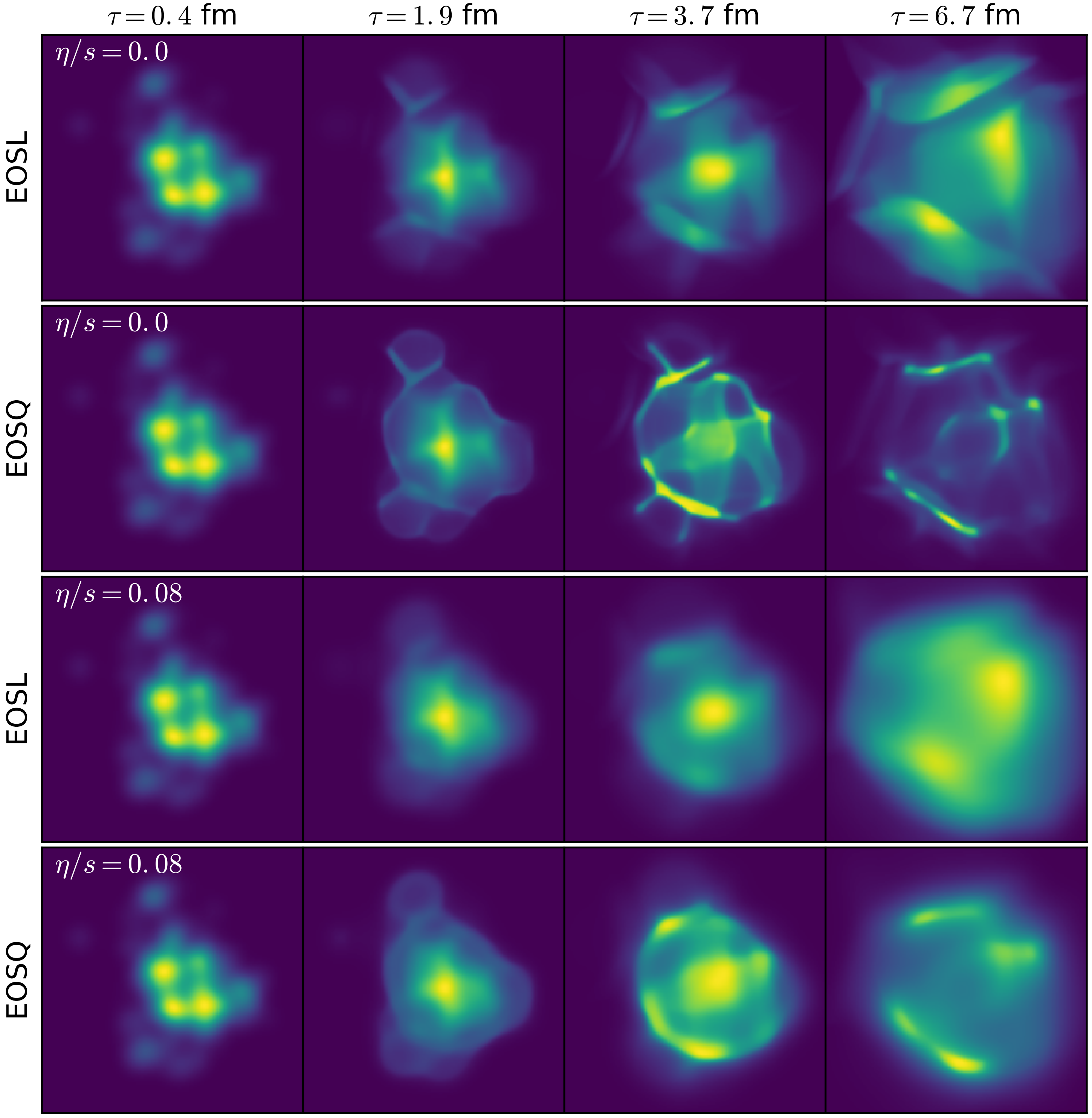
$$\Delta^{\mu\nu} = g^{\mu\nu} - u^\mu u^\nu,\ g^{\mu\nu} = diag(1, -1, -1, -\tau^{-2}) \tag{5}$$

ε and P are the energy density and pressure, u^μ is the fluid velocity vector. ∇_μ is the covariant derivative.

- Constraints: $P = P(\varepsilon)$, $u_\mu u^\mu = 1$, $u_\mu \pi^{\mu\nu} = 0$, $\pi^\mu_\mu = 0$.

CLVisc, L.G. Pang, B.W. Xiao, Y. Hatta, X.N.Wang, PRD 2015

Initial state fluctuation to final state correlation



Cooper-Frye Particalization

$$E\frac{dN_i}{dp^3} = \frac{dN_i}{dY p_T dp_T d\phi} = \frac{g_i}{(2\pi)^3}\int p^\mu d\Sigma_\mu f_{eq}(1+\delta f)$$

where
$$f_{eq} = \frac{1}{\exp\left((p\cdot u - \mu_i)/T_f\right) \pm 1}$$
$$\delta f = (1 \mp f_{eq})\frac{p_\mu p_\nu \pi^{\mu\nu}}{2T_f^2(\varepsilon + P)}$$

Training dataset

$\rho(p_T, \Phi)$ for charged pions at mid-rapidity

TRAINING DATASET	$\eta/s = 0$		$\eta/s = 0.08$	
	EOSL	EOSQ	EOSL	EOSQ
Au-Au $\sqrt{s_{NN}} = 200\,\mathrm{GeV}$	7935	5828	500	500
Pb-Pb $\sqrt{s_{NN}} = 2.76\,\mathrm{TeV}$	5467	3328	500	500

- CLVisc + AMPT initial condition + GPUs on GSI-GreenCube = (~22000 events, doubled by left-right flipping, 10% for validation).
- τ_0 is 0.4 fm for Au+Au and 0.2 fm for Pb+Pb collisions
- Tfrz=0.137 GeV

Testing dataset

TESTING DATASET GROUP 1 : iEBE-VISHNU + MC-Glauber						
Centrality:	$\eta/s \in [0, 0.05]$		$\eta/s \in (0.05, 0.10]$		$\eta/s = (0.10, 0.16]$	
10-60%	EOSL	EOSQ	EOSL	EOSQ	EOSL	EOSQ
Au-Au $\sqrt{s_{NN}} = 200$ GeV	650	850	900	750	200	950
Pb-Pb $\sqrt{s_{NN}} = 2.76$ TeV	500	650	600	644	499	150
TESTING DATASET GROUP 2 : CLVisc + IP-Glasma						
Au-Au $\sqrt{s_{NN}} = 200$ GeV	EOSL			EOSQ		
b$\lesssim$8 fm & $\eta/s = 0$	4165			4752		

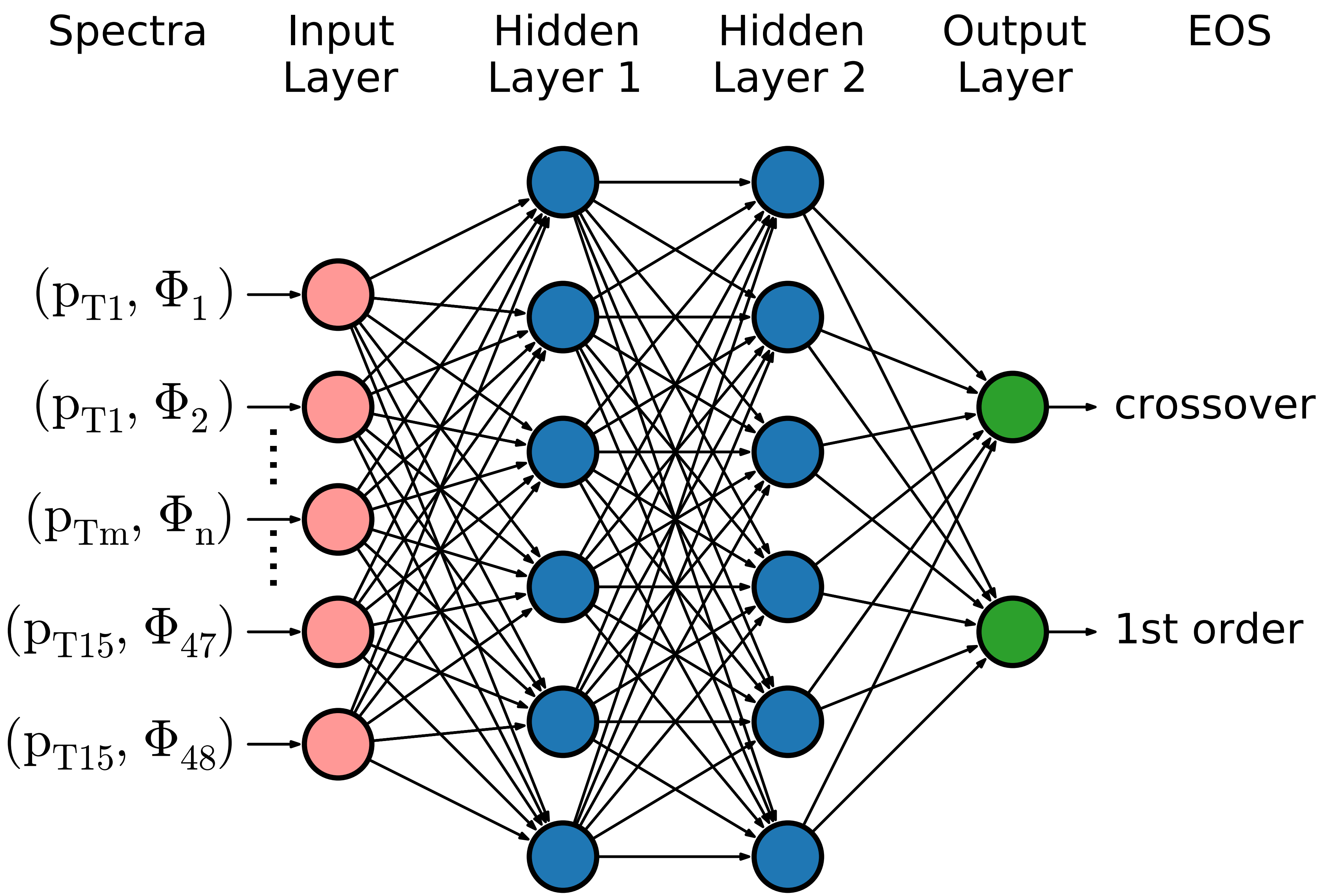
- iEBE-VISHNU: another viscous hydro with different numerical solver for partial differential equations and different initial condition

- τ_0 is 0.6 fm for all the testing dataset.

- Tfrz in [0.11GeV, 0.14 GeV] for iEBE-VISHNU

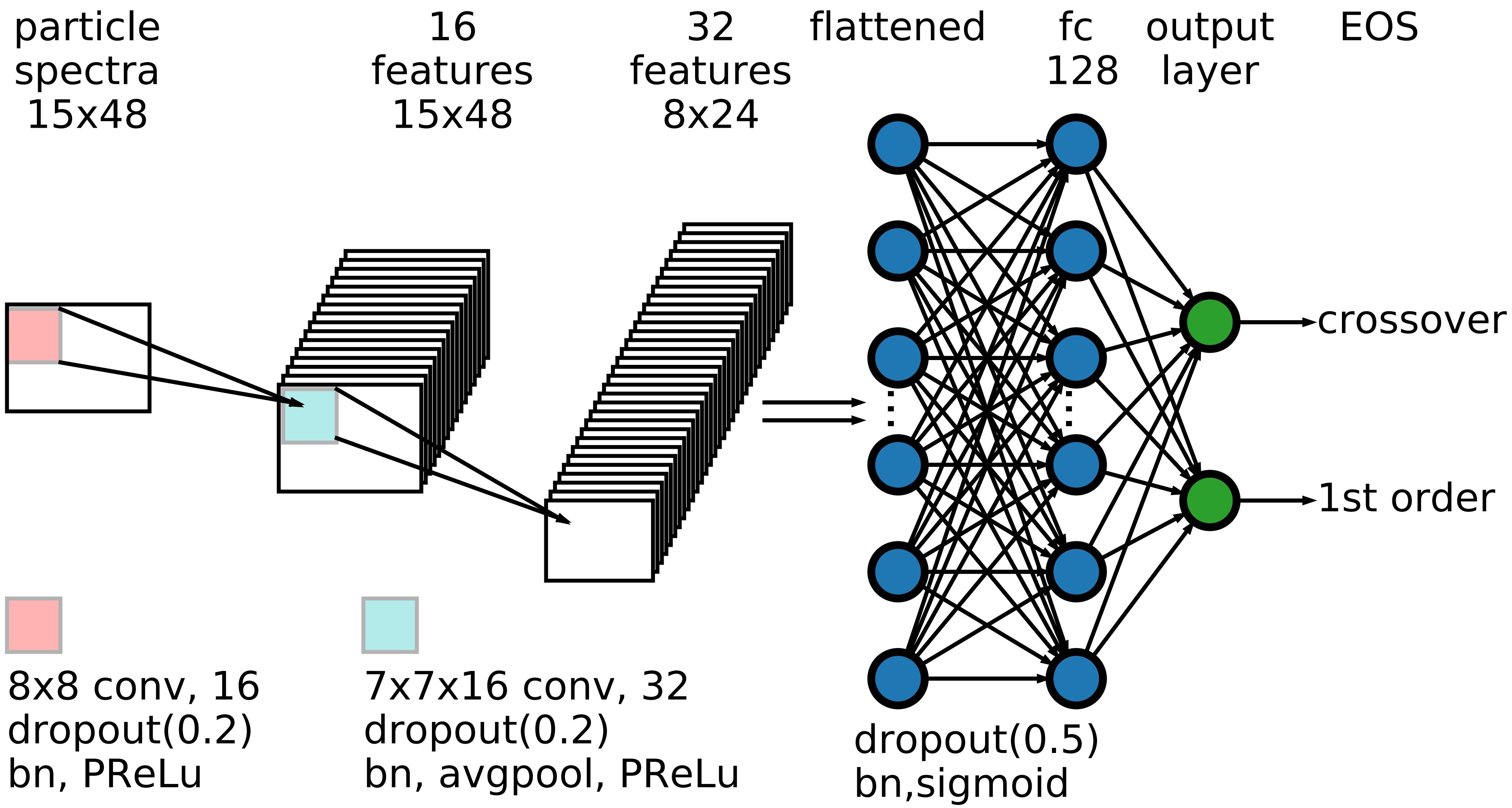
iEBE-VISHNU: C. Shen, Z. Qiu, H. Song, J. Bernhard, S. Bass, and U. Heinz, Comput. Phys. Commun. 199, 61 (2016)

First attempt with fully connected neural network

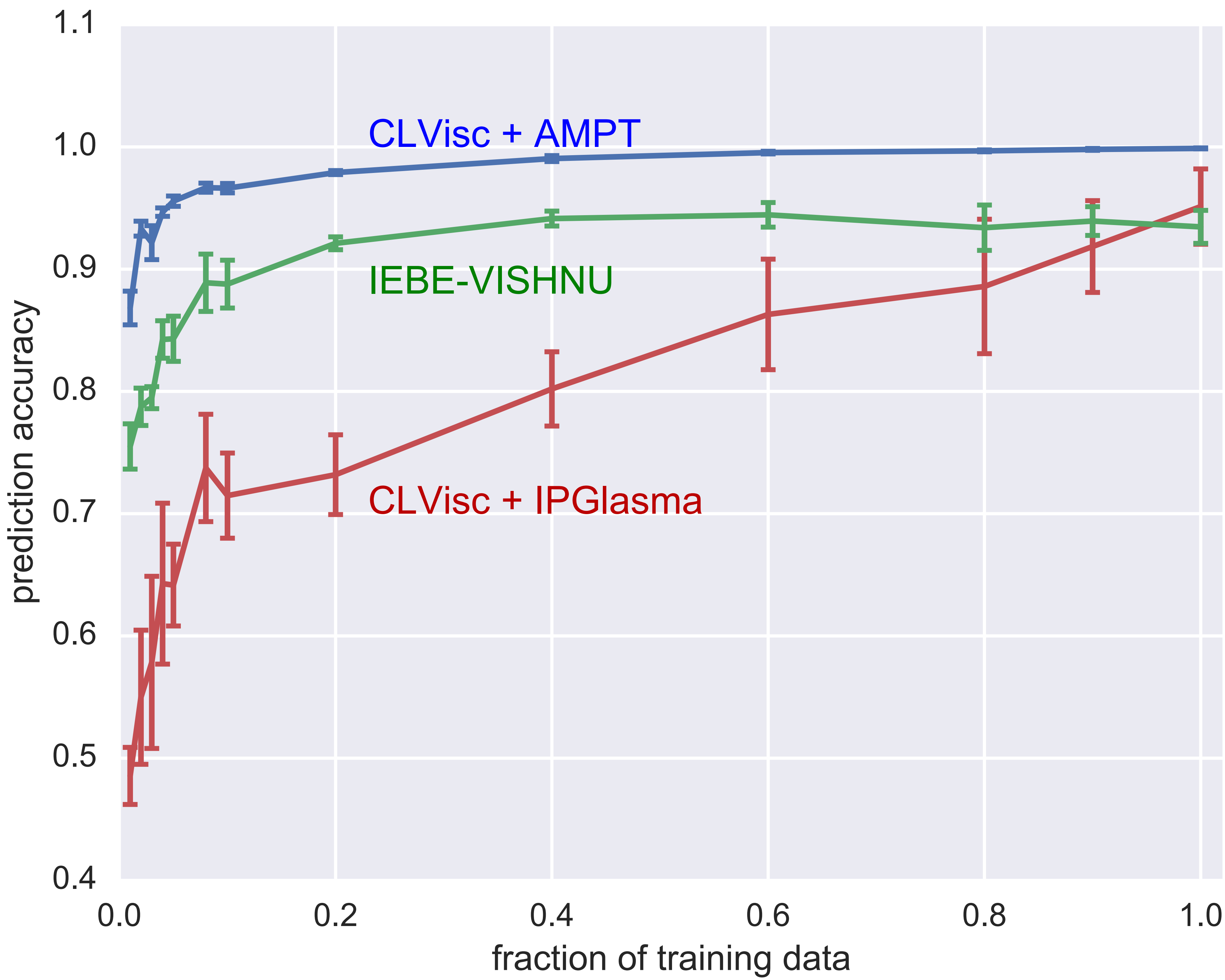


Overfit to the training dataset! Does not work for testing dataset.

CNN architecture for EoS-meter



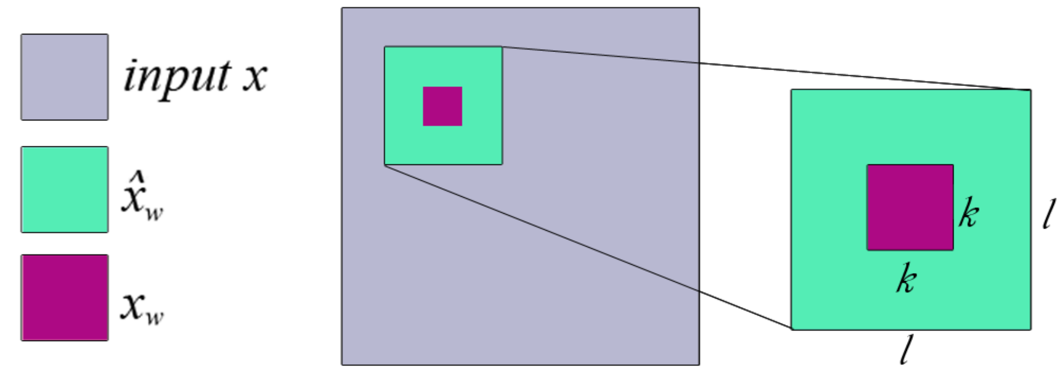
Prediction Accuracy & Uncertainty in 10-fold cross validation

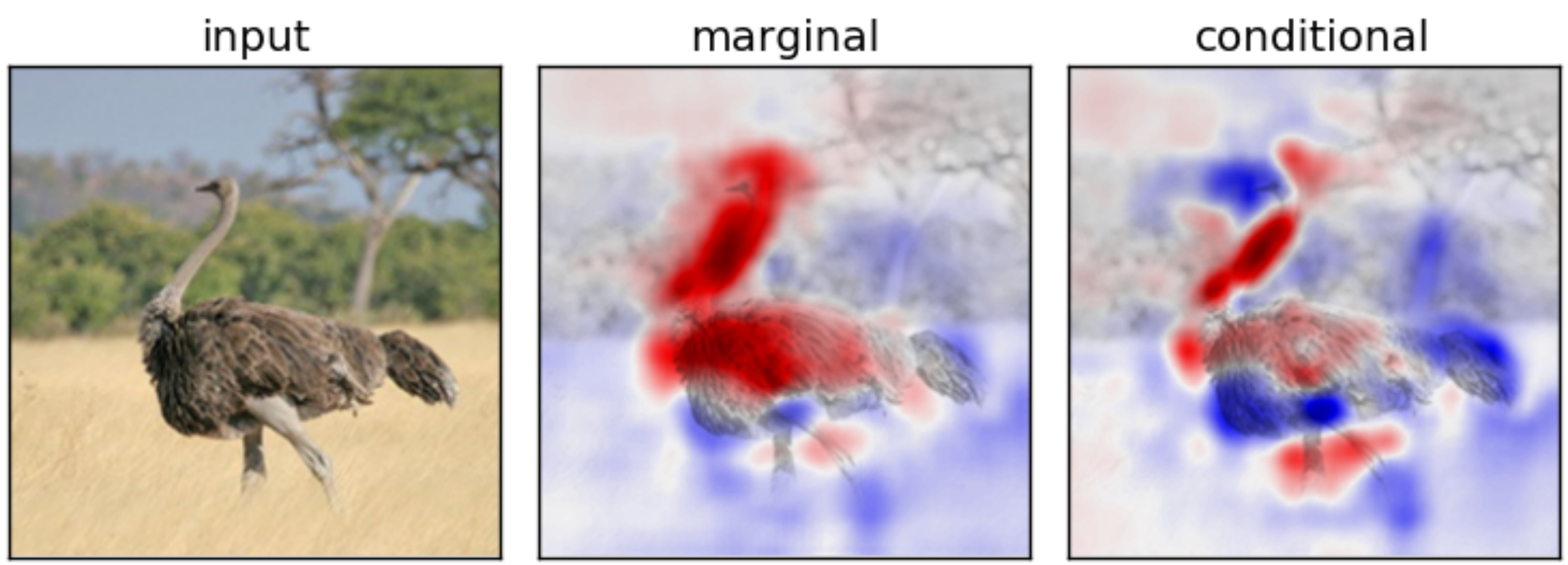


Prediction Difference Analysis

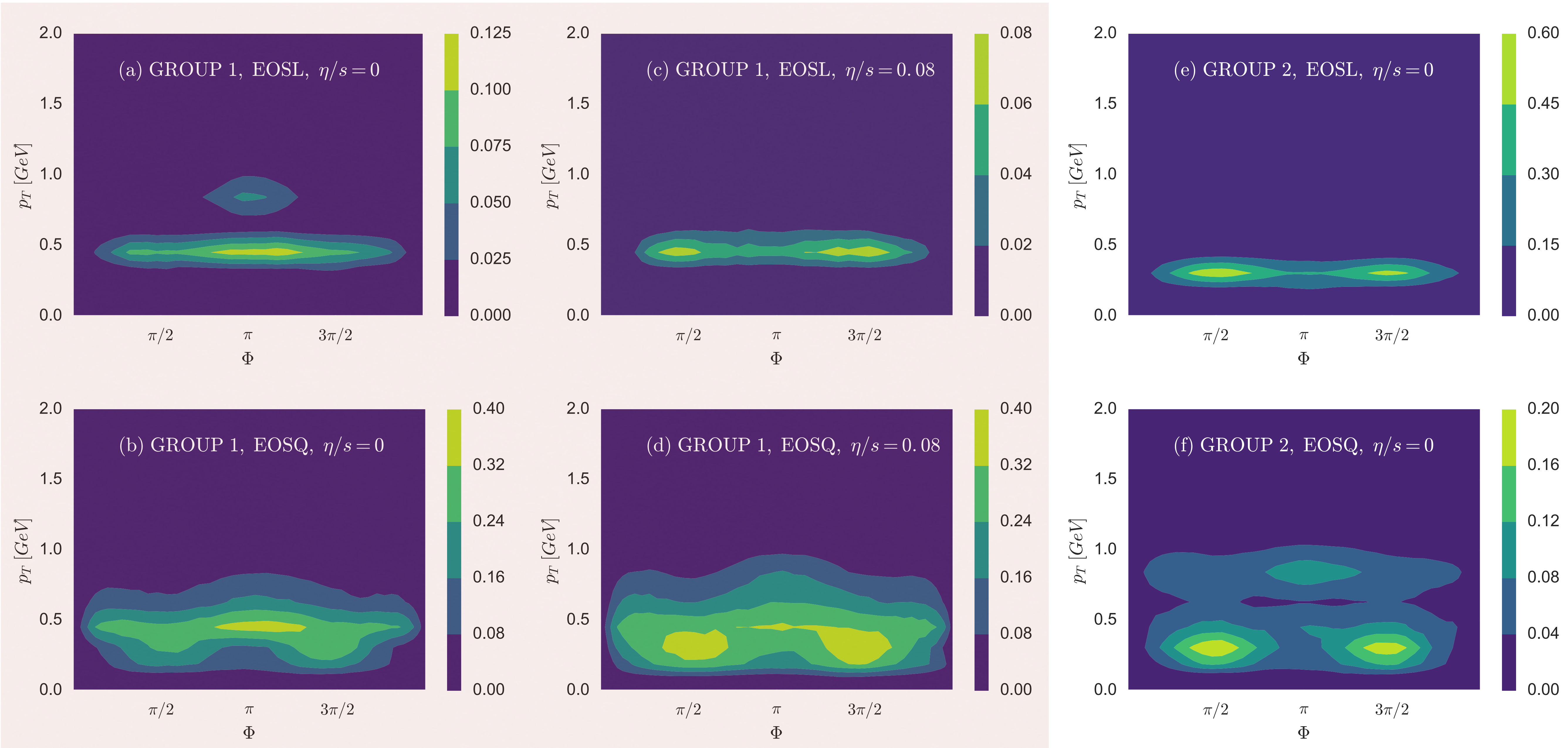
VISUALIZING DEEP NEURAL NETWORK DECISIONS: PREDICTION DIFFERENCE ANALYSIS

Luisa M Zintgraf[1,3], **Taco S Cohen**[1], **Tameem Adel**[1], **Max Welling**[1,2]
[1]University of Amsterdam, [2]Canadian Institute of Advanced Research, [3]Vrije Universiteit Brussel
`{lmzintgraf,tameem.hesham}@gmail.com`, `{t.s.cohen, m.welling}@uva.nl`





Prediction difference by marginally or conditionally sampling the value of one feature from mixed events.



GROUP 1 GROUP 2

- Importance regions are different for different testing datasets
- eta/s introduces a small difference

Summary and Outlook

- We firmly demonstrate that the "encoders" from QCD transition onto the spectra do exist.
- Deep CNN provides a powerful "decoder" to extract the QCD transition from final spectra (regardless the initial fluctuations).
- Prediction difference analysis highlights the most relevant features for classification.

OutLook

- Extend the model to work with exp. data
- Extract other parameters like temperature dependent shear viscosity or other physical properties.