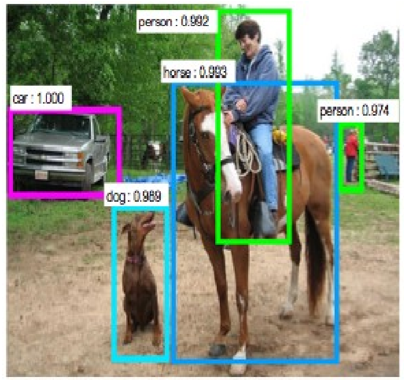
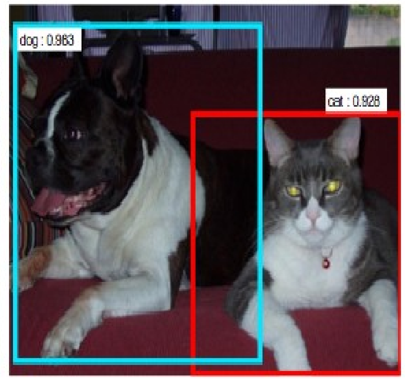
Neural Network Study on Lattice 1+1d Scalar Field Theory

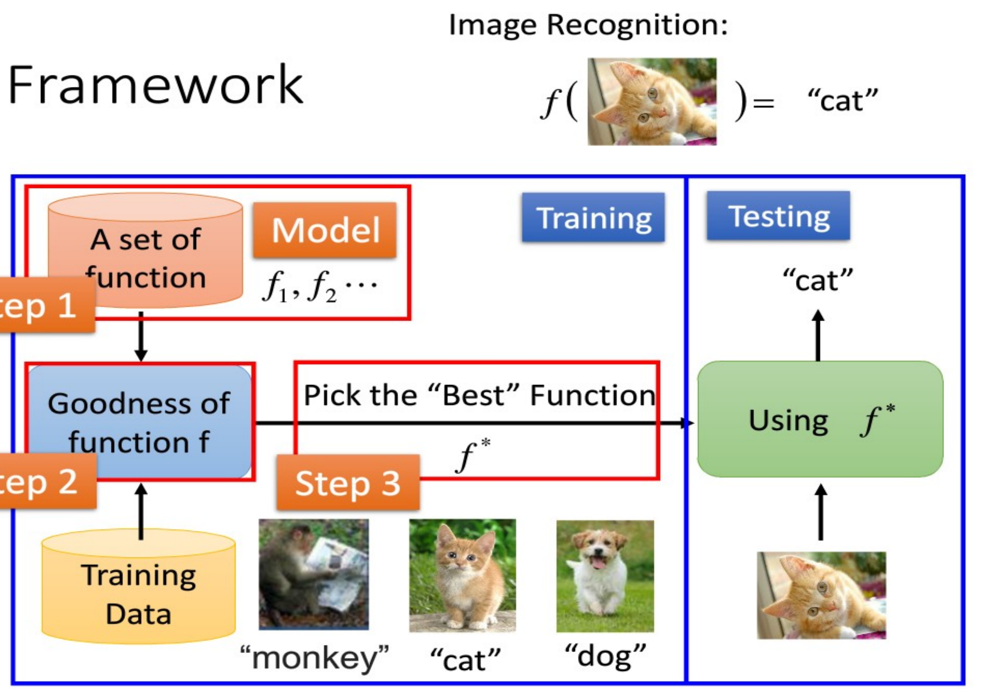
Kai Zhou (FIAS, Frankfurt, Germany)

In collaboration with :
Gergely Endrődi (ITP, Frankfurt, Germany)
Long-gang Pang (UC Berkeley, Berkeley, USA)
Horst Stoecker (FIAS, Frankfurt, Germany)

Introduction







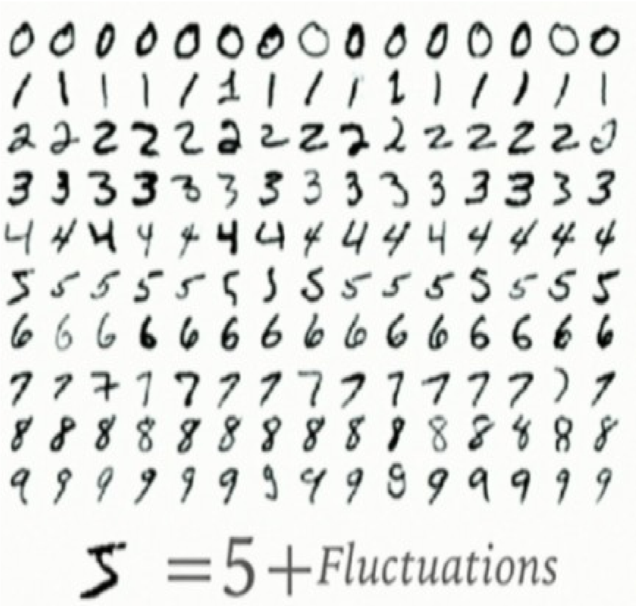
Find and **Decode** the mapping/representations into Deep Neural Network

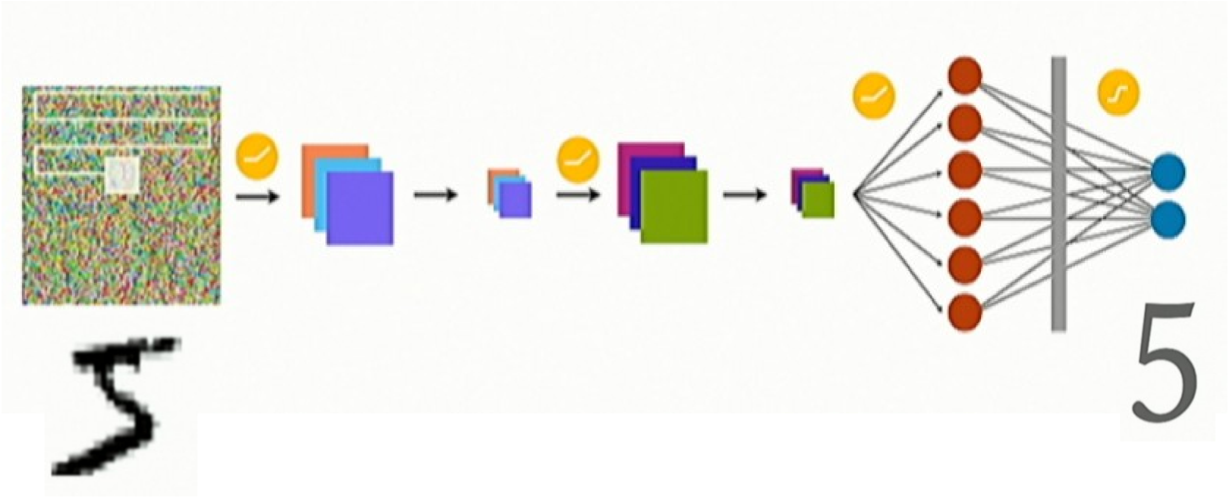
→ **function approximator**

Universal approximator (Hastad et al 86 & 91)

Introduction

Convolutional Neural Network has proved to be extremely powerful in Pattern Recognition, Image Classification

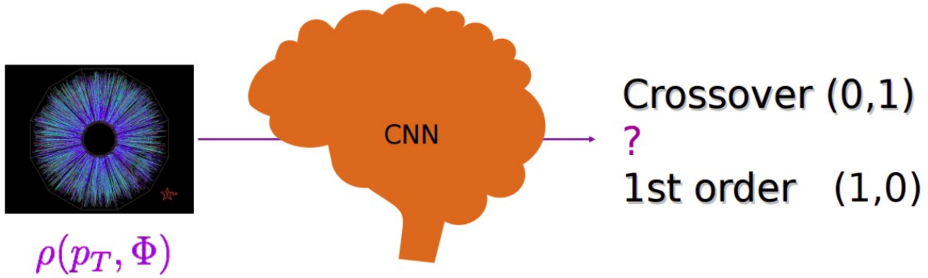




Deepthinkers Group at FIAS

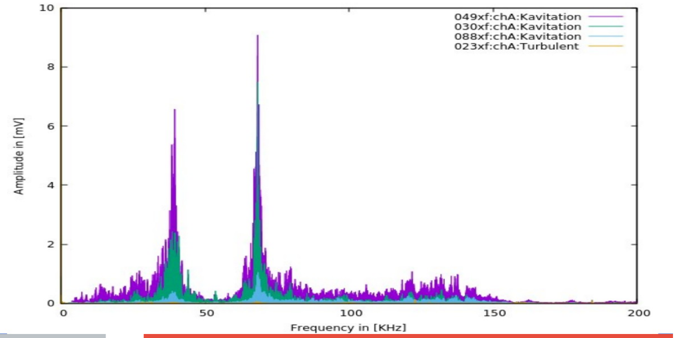
(1) Statistical physics/ lattice configuration analysis

(2) heavy-ion collisions : decode medium property (phase transition)



(3) hydrodynamic simulation : speed-up

(4) smart-valve: 'hear' the valve (leakage? Flow status?)



Dualization approach for $\lambda\phi^4$

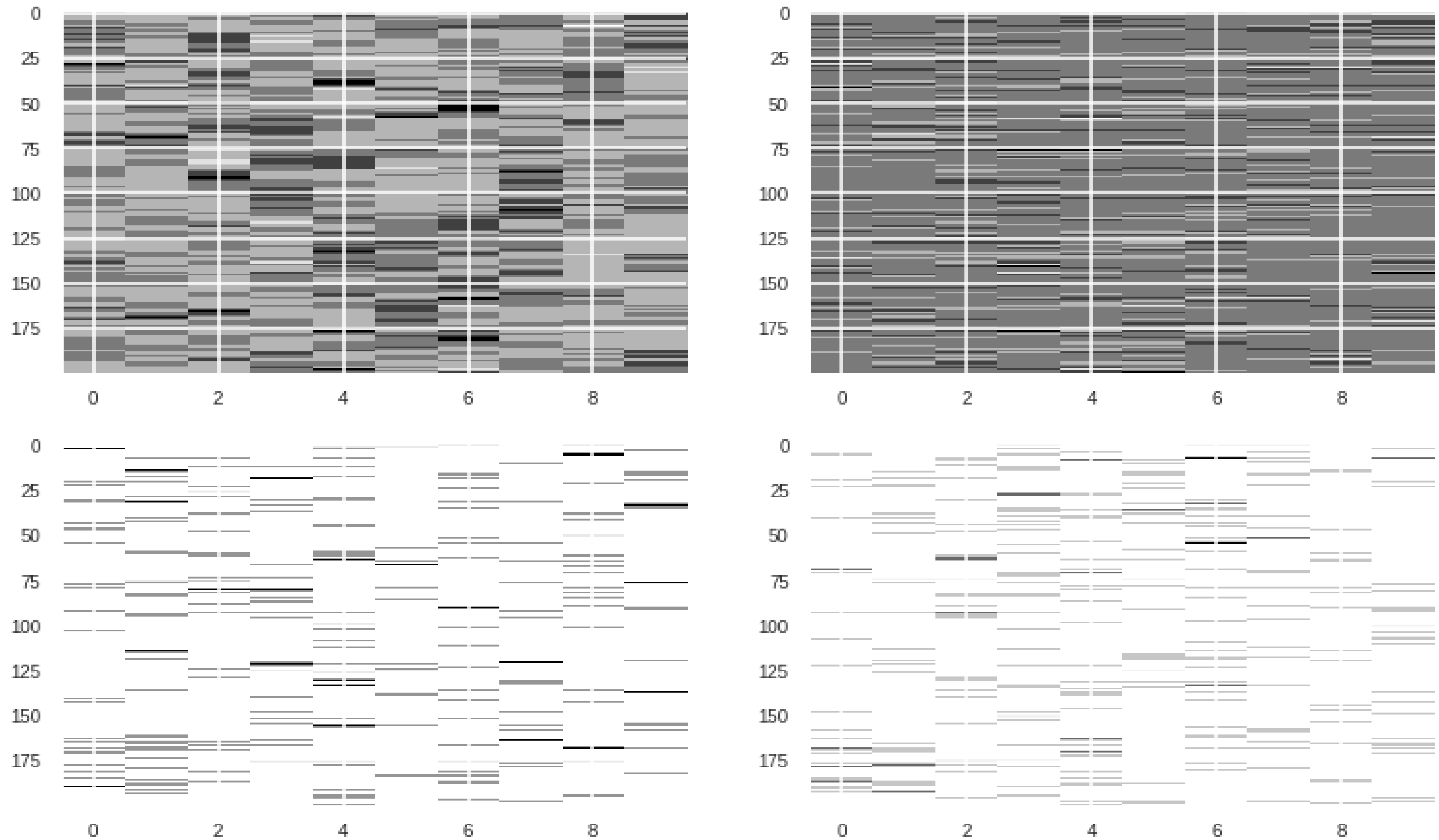
configurations – 4 integer-valued variables: $k_t, \quad k_x, \quad l_t, \quad l_x$

$$m = 0.1$$

$$\lambda = 1.0$$

$$N_t = 200$$

$$N_x = 10$$



Divergence constraint : $$\nabla \cdot k(n) = \sum_{\nu} [k_\nu(n) - k_\nu(n - \hat{\nu})] = 0$$

Observables : n and $|\phi|^2$

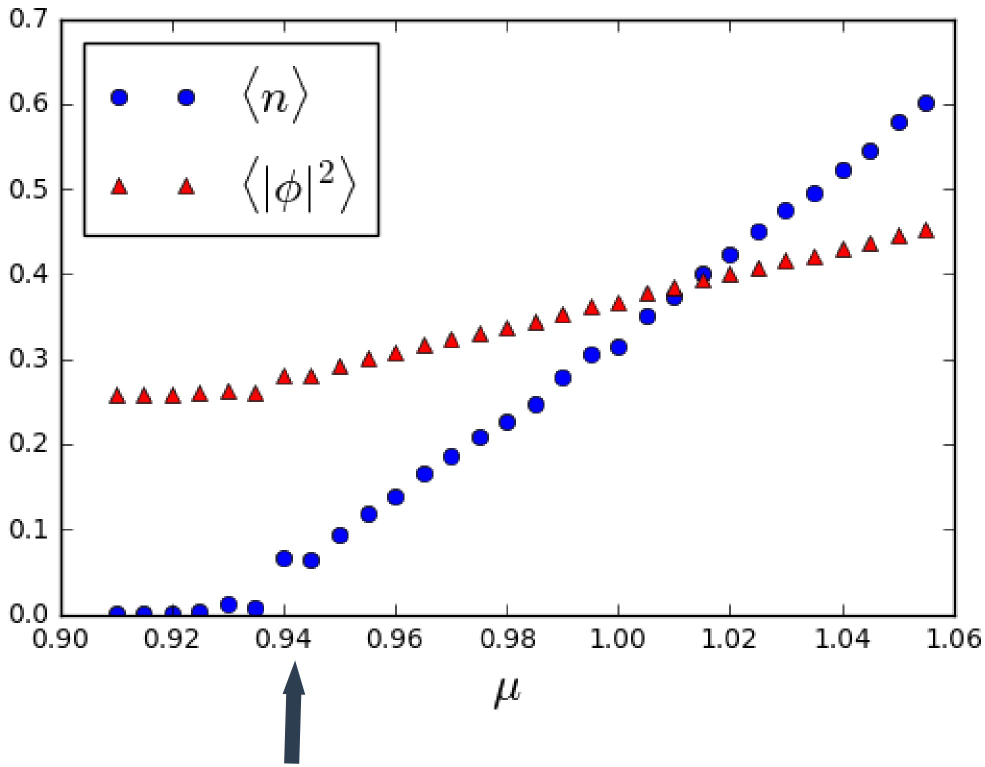
Grand canonical ensemble

$$m = 0.1$$

$$\lambda = 1.0$$

$$N_t = 200$$

$$N_x = 10$$



Condensation sets in at $\mu_{th} \sim m_{phys} \sim 0.94$

Observables : n and $|\phi|^2$

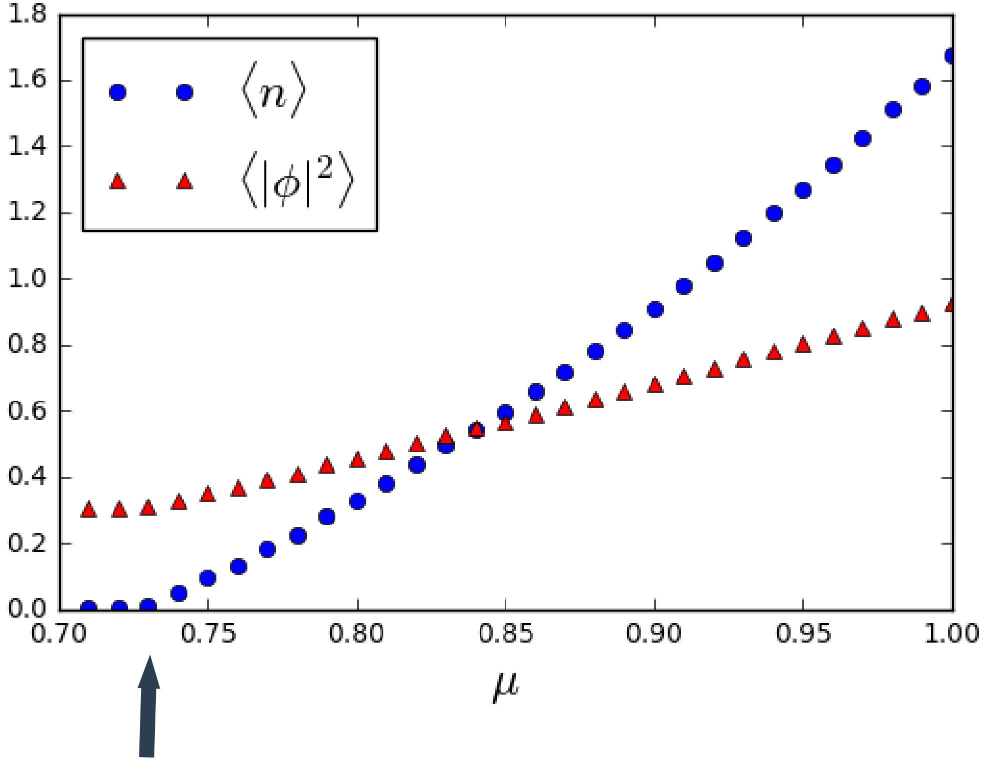
Grand canonical ensemble

$$m = 0.1$$

$$\lambda = 0.5$$

$$N_t = 200$$

$$N_x = 10$$



Condensation sets in at $\mu_{th} \sim m_{phys} \sim 0.73$

Exploring NN ability here

(1) Classification 1: differentiate configs with different interaction

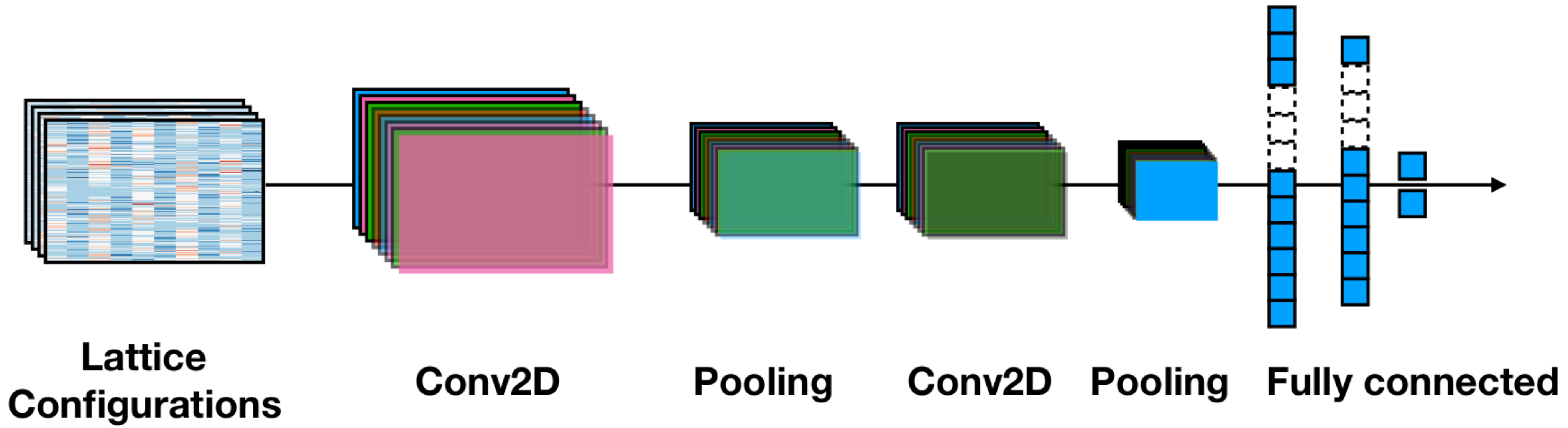
(2) Classification 2: detect phase transition based on config.

(3) Regression: learn physical observables

(4) Generative model : learn to generate new configs

Generate configs with proper distribution

DCNN Architechture – Classification



$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log \hat{y}_i + (1 - y_i)\log(1 - \hat{y}_i)] + \lambda \|\theta\|_2^2$$

DCNN Architechture – Classification 1

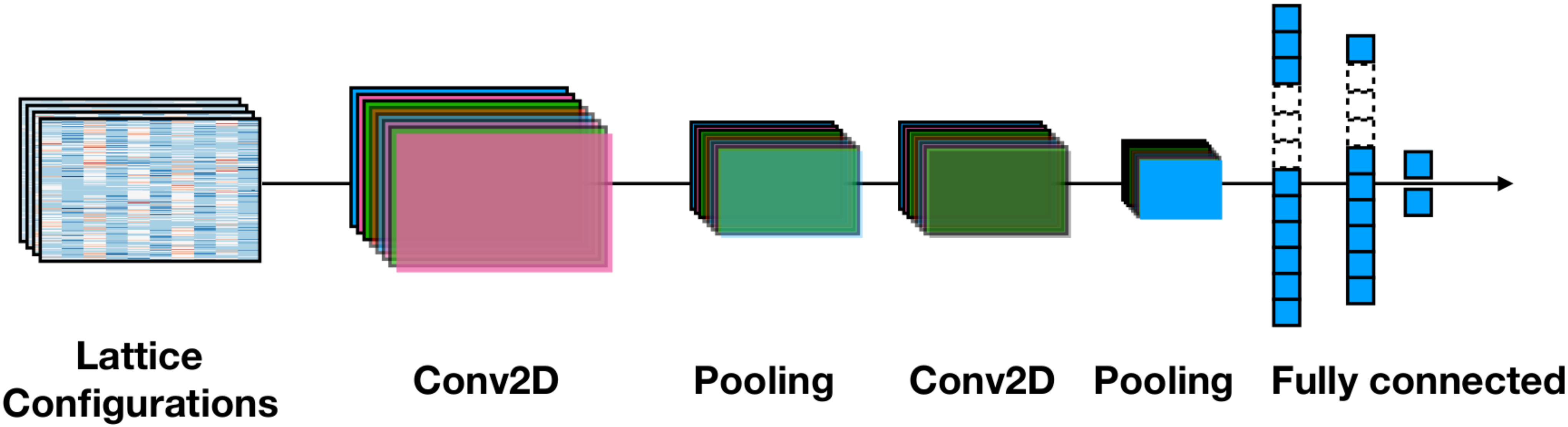
Can interaction information be decoded from microscopic configuration (same sized lattice)?

Train with two ensemble of configs (different λ) :

$$10k : \mu = 1.05 @ \lambda = 1.0 \quad \text{labeld} \quad y = (0,1)$$

$$10k : \mu = 0.85 @ \lambda = 0.5 \quad \text{labeld} \quad y = (1,0)$$

Test on other chemical potentials with the two couplings



DCNN Architechture – Classification 1

Can interaction information be decoded from microscopic configuration (same sized lattice)?

Train with two ensemble of configs (different λ) :

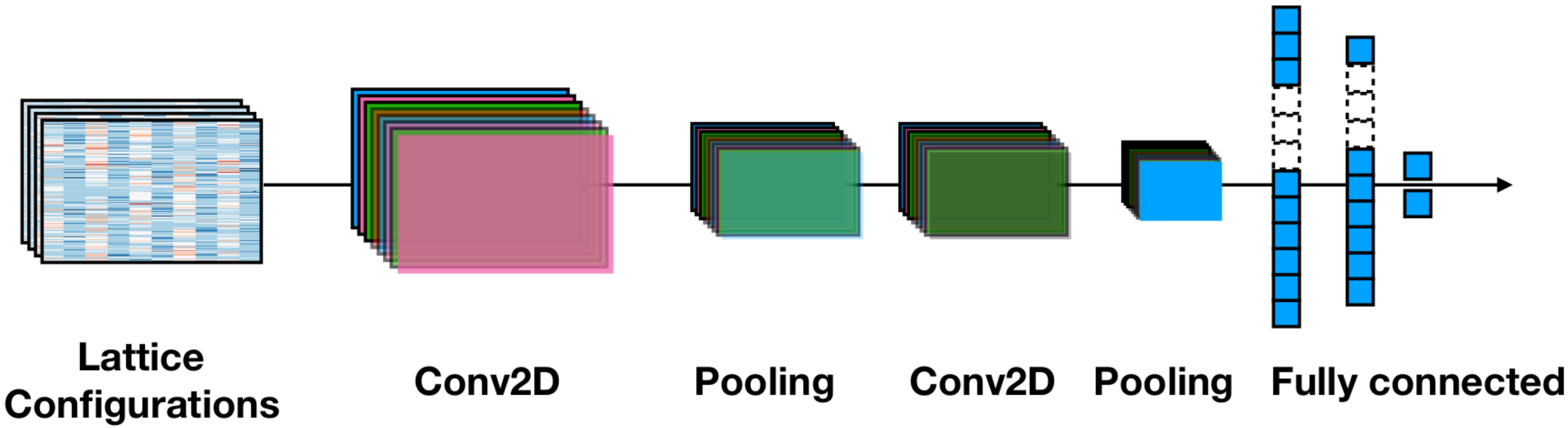
$$\mu = 1.05@\lambda = 1.0 \quad \text{labeld} \quad y = (0,1)$$

$$\mu = 0.85@\lambda = 0.5 \quad \text{labeld} \quad y = (1,0)$$

Test on other chemical potentials with the two couplings

→ $\text{perfetly classified with} \quad 99.8\% \quad \text{accuracy}$

DCNN Architechture – Classification 2

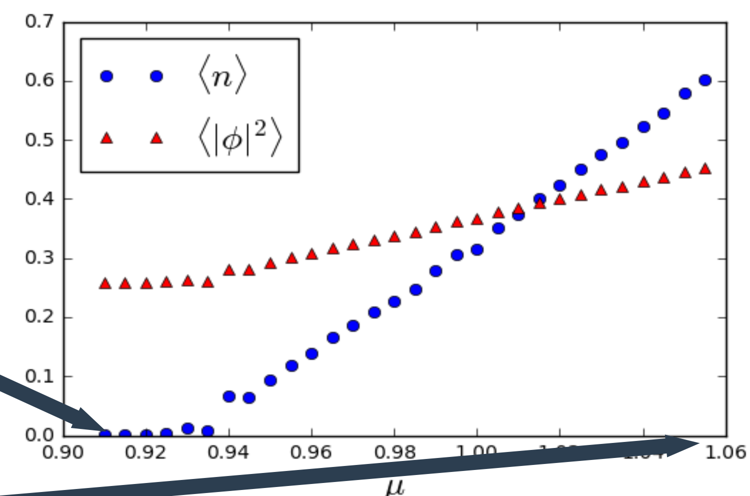


Training set consist only two ensembles of configuations at

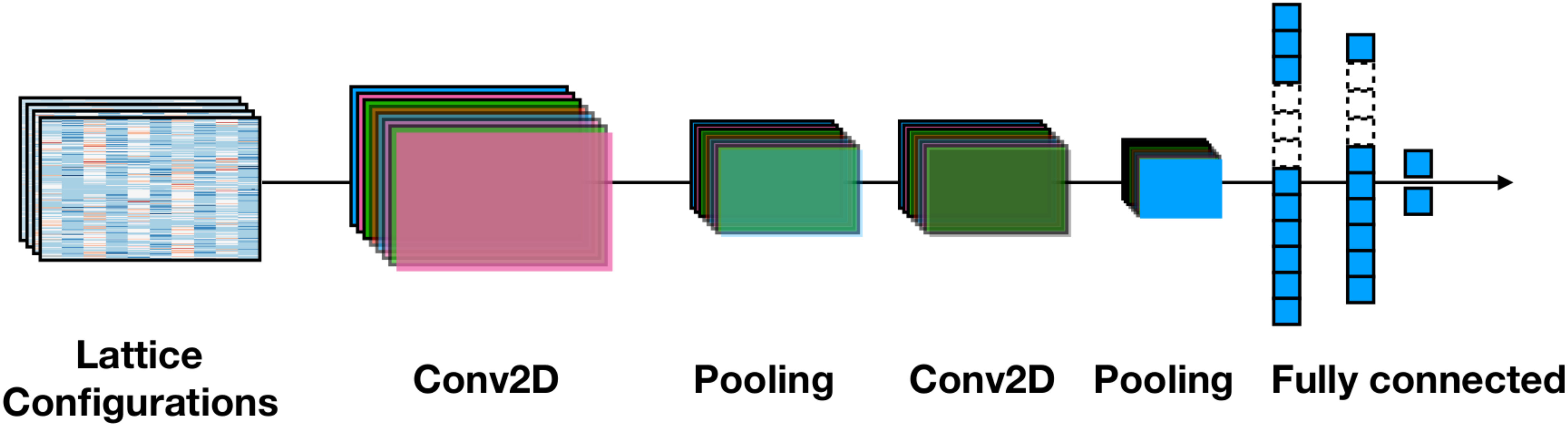
$$\mu = 0.91 \quad \text{with label} \quad y = (0,1)$$

and

$$\mu = 1.05 \quad \text{with label} \quad y = (1,0)$$



DCNN Architechture – Classification 2



Training set consist only two ensembles of configuations at

$$10k : \mu = 0.91 \quad \text{with label} \quad y = 0$$

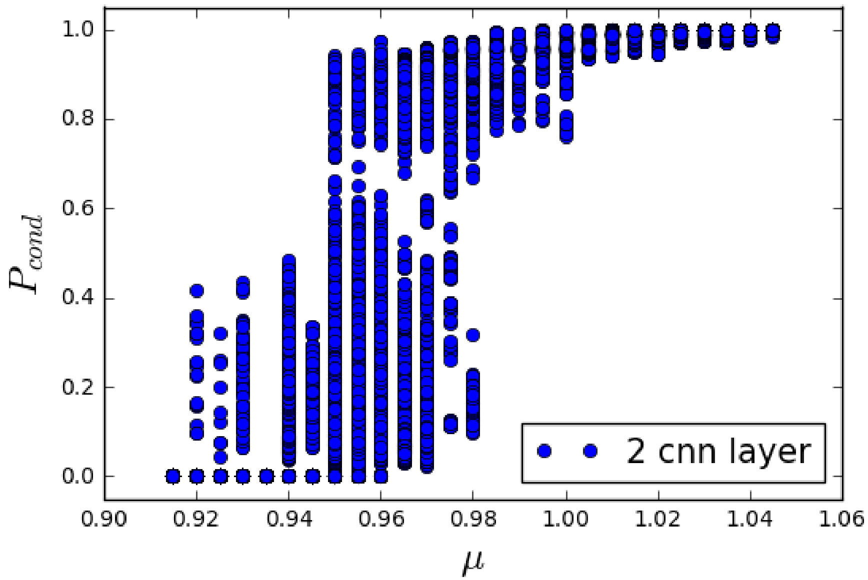
and

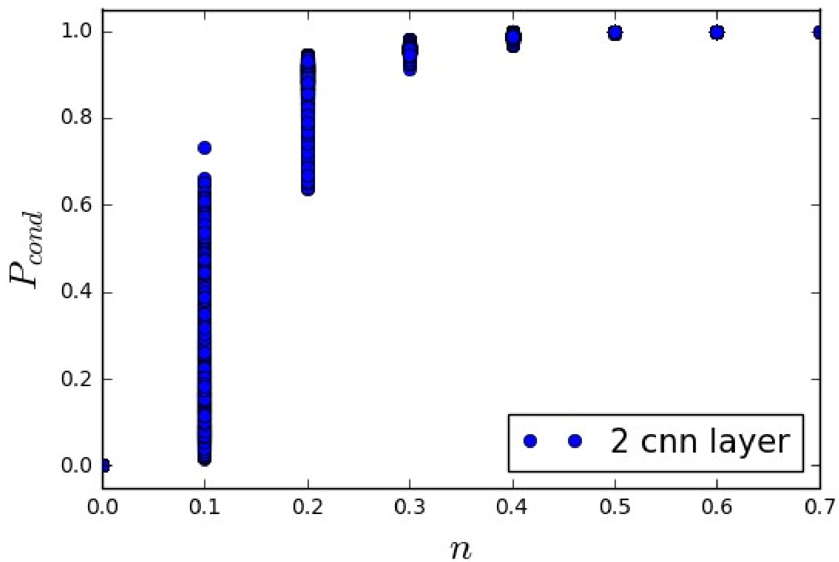
$$10k : \mu = 1.05 \quad \text{with label} \quad y = 1$$

Testing set consist of different ensembles of configurations at different chemical potential

$$0.91 < \mu < 1.05$$

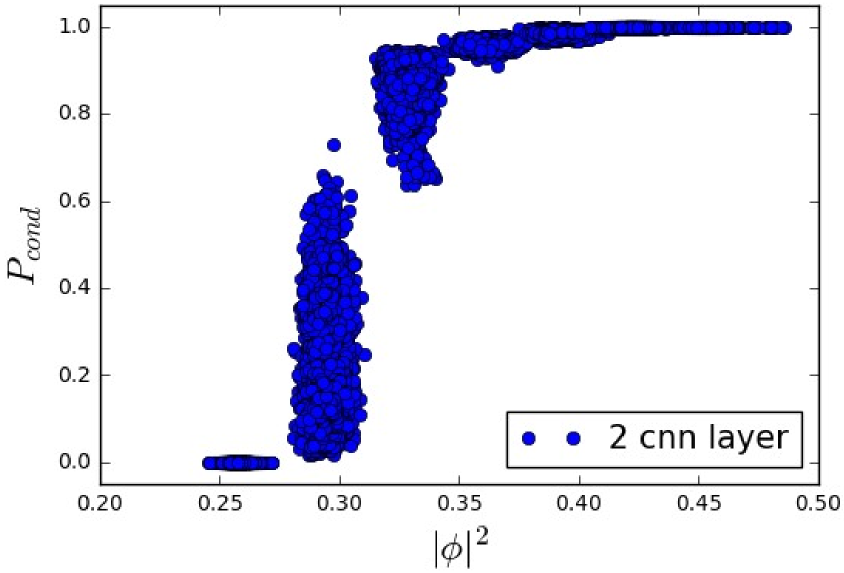
Condensation probability from DCNN



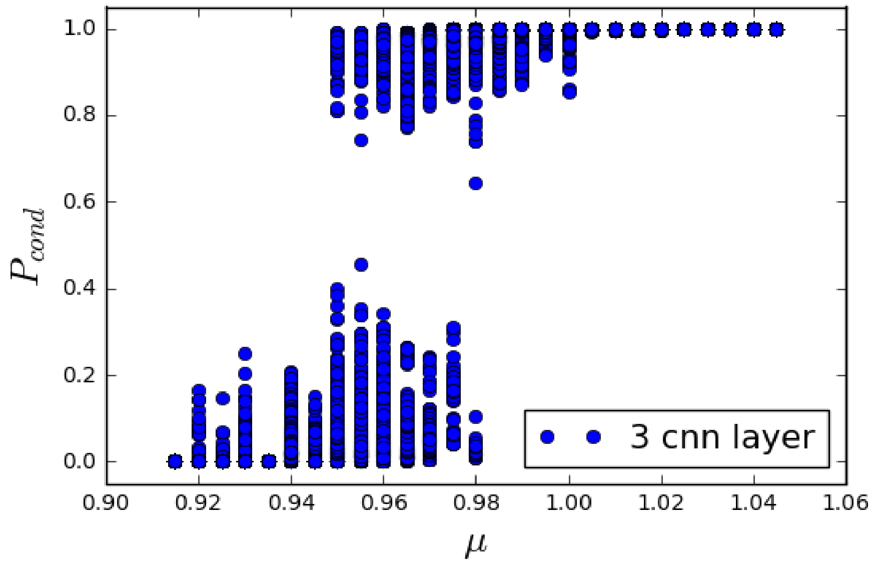


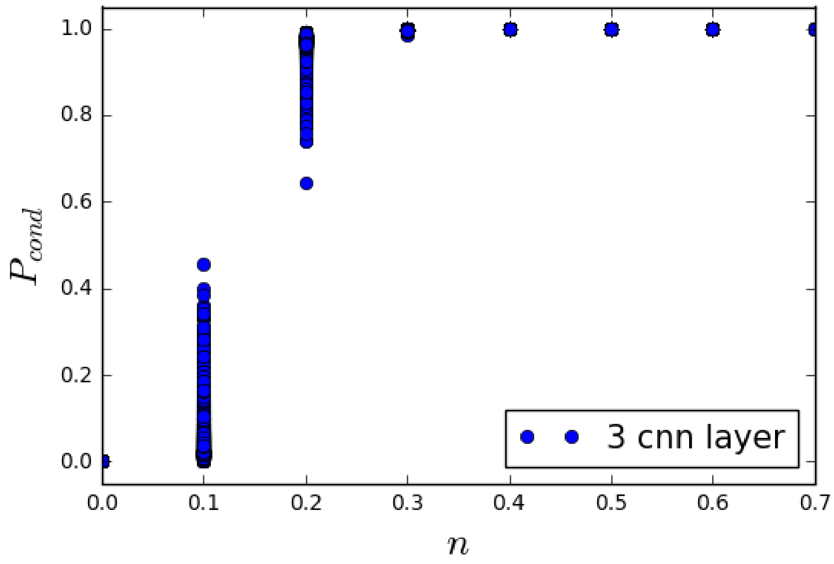
Non-linear correlation between P_cond and physical observables :

n and squared field



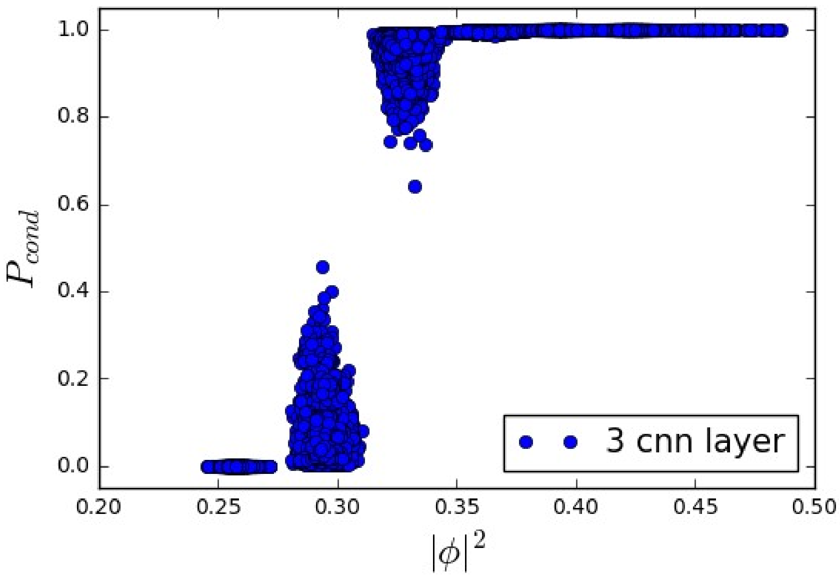
Condensation probability from DCNN





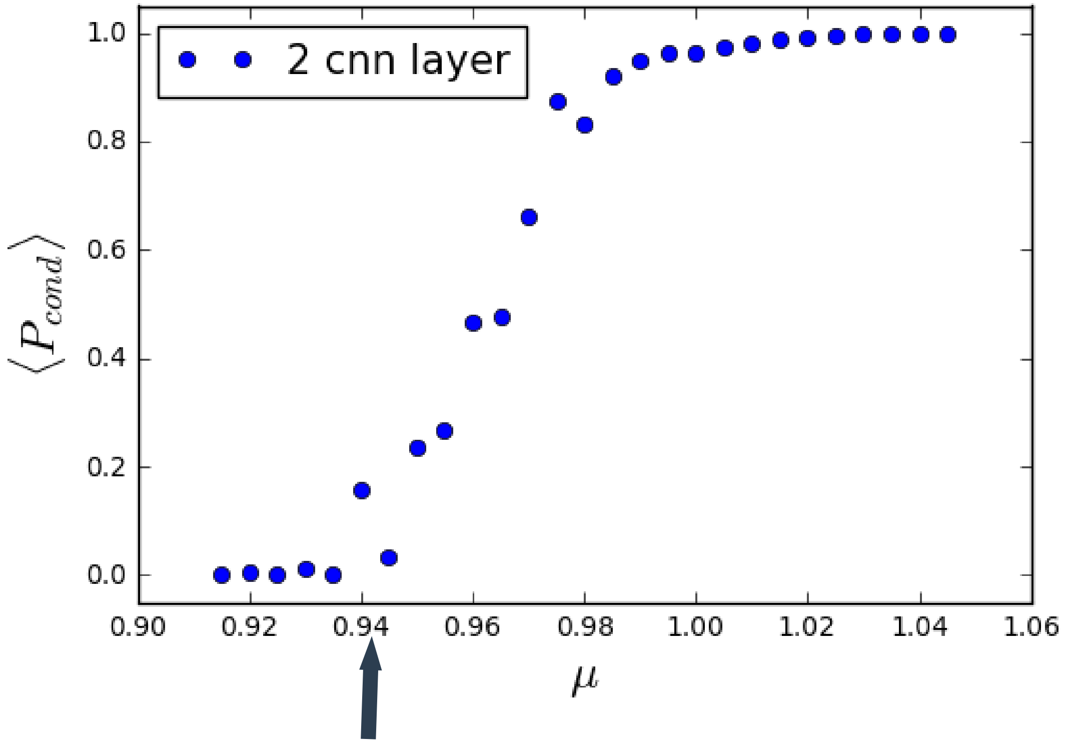
Adding one more CNN layer gives better expressive power to the network :

better distinguish ability



Ensemble average cond-probability

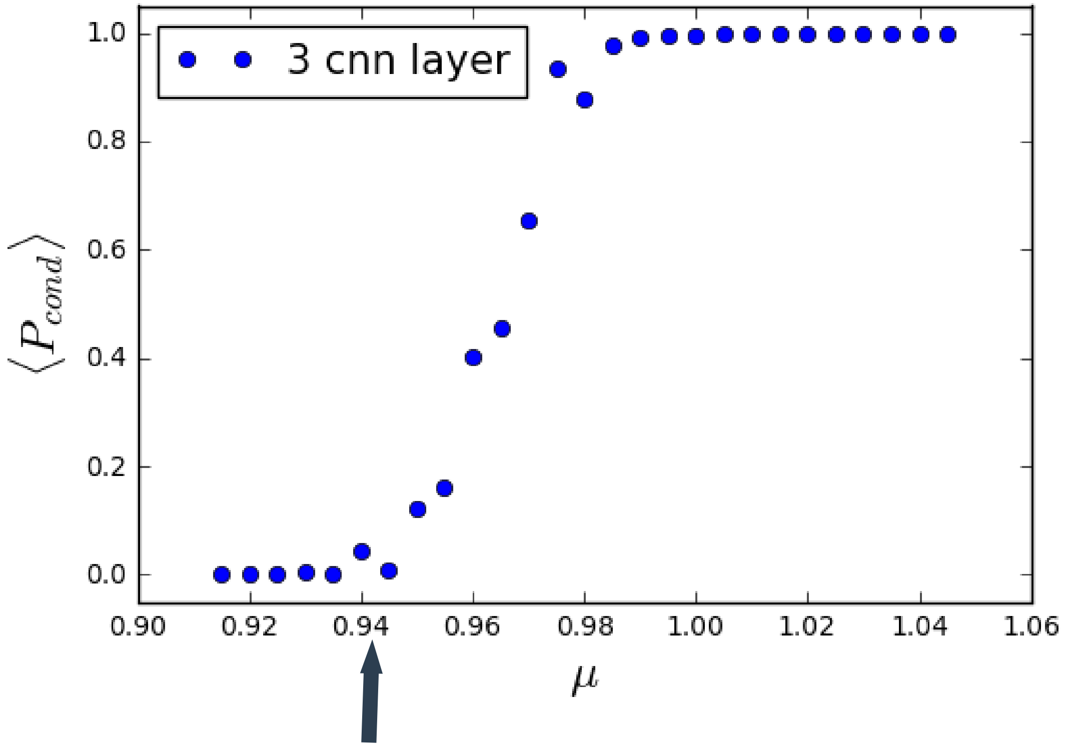
Classifier of the phases : $\langle n \rangle = 0 \quad \text{and} \quad \langle n \rangle \neq 0$



$$\mu_{th}(\langle P_{cond} \rangle > 0) \sim \mu_{th}(\langle n \rangle > 0)$$

Ensemble average cond-probability

Classifier of the phases : $\langle n \rangle = 0 \quad \text{and} \quad \langle n \rangle \neq 0$

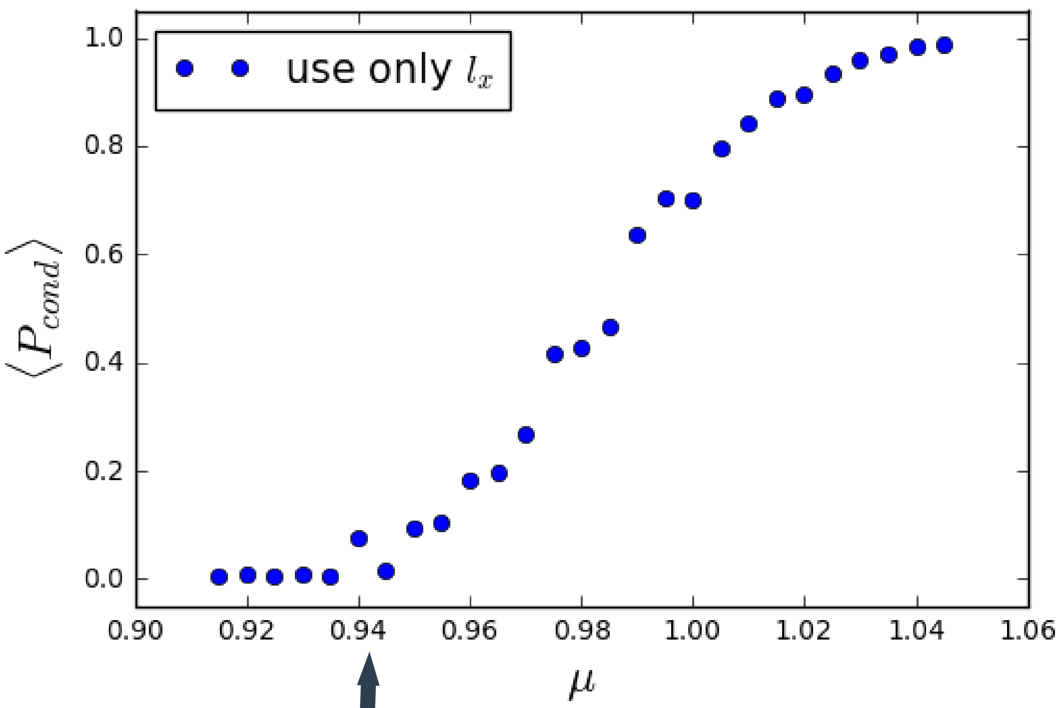


one more CNN layer sharpens the transition curve

$$\mu_{th}(\langle P_{cond} \rangle > 0) \sim \mu_{th}(\langle n \rangle > 0)$$

Discard kt information

Is phase transition information also encoded in l variable ? Yes!

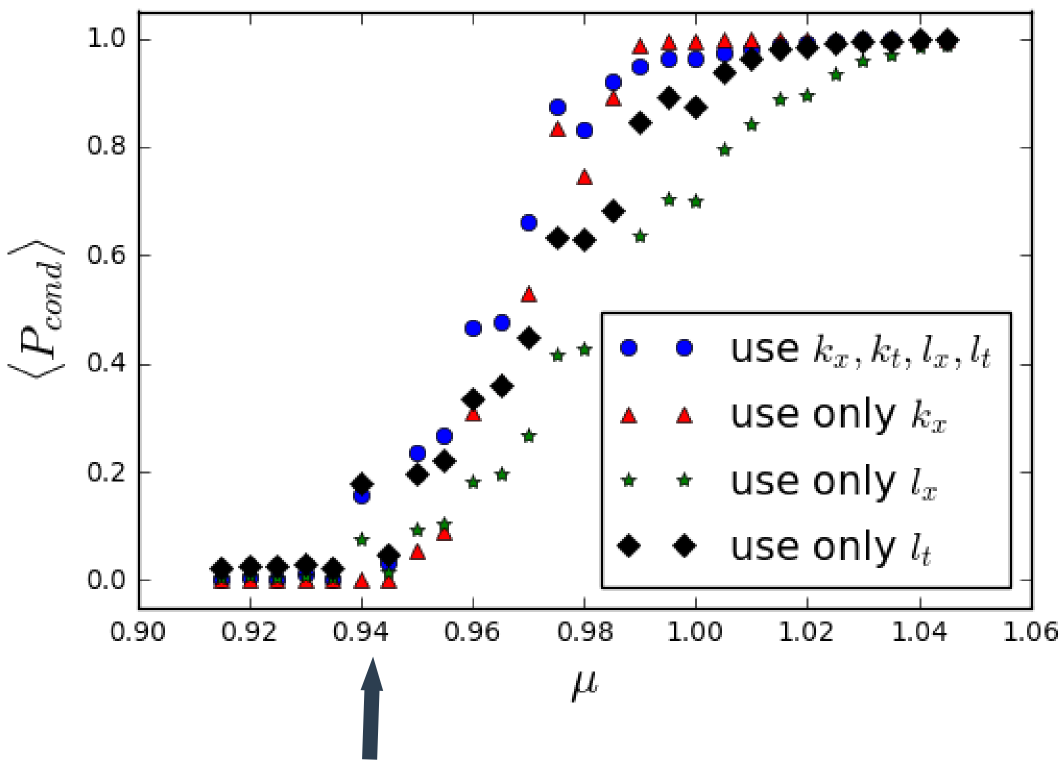


beyond conventional knowledge, indicating hidden structures in the l variables and not only in the k_t variables.

$$\mu_{th}(\langle P_{cond} \rangle > 0) \sim \mu_{th}(\langle n \rangle > 0)$$

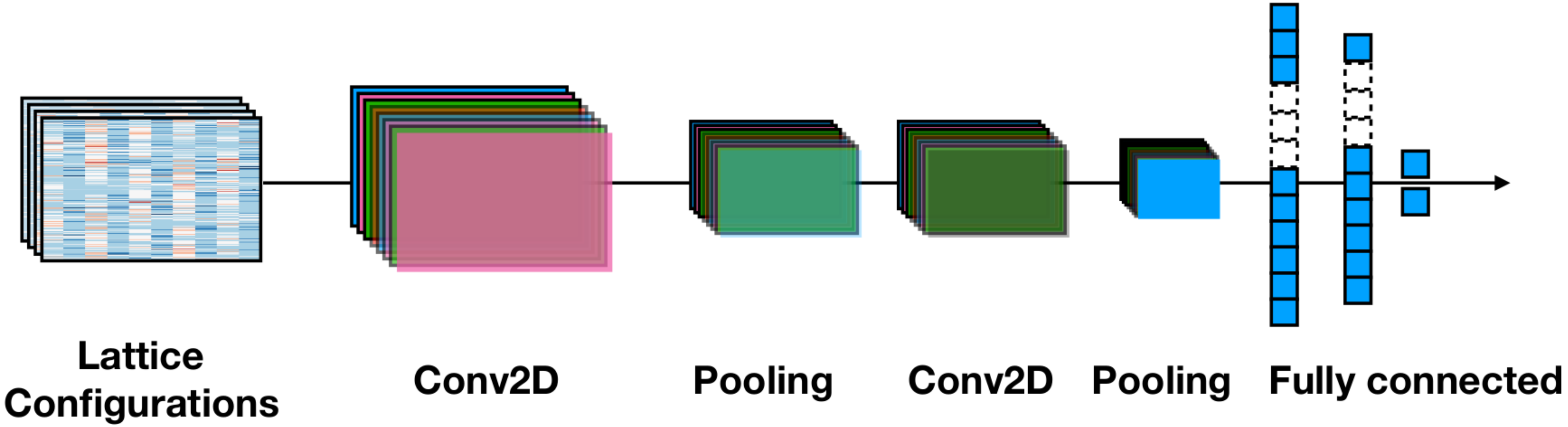
Try different field component variables

The same transition point



$$\mu_{th}(\langle P_{cond} \rangle > 0) \sim \mu_{th}(\langle n \rangle > 0)$$

DCNN Architechture – Regression



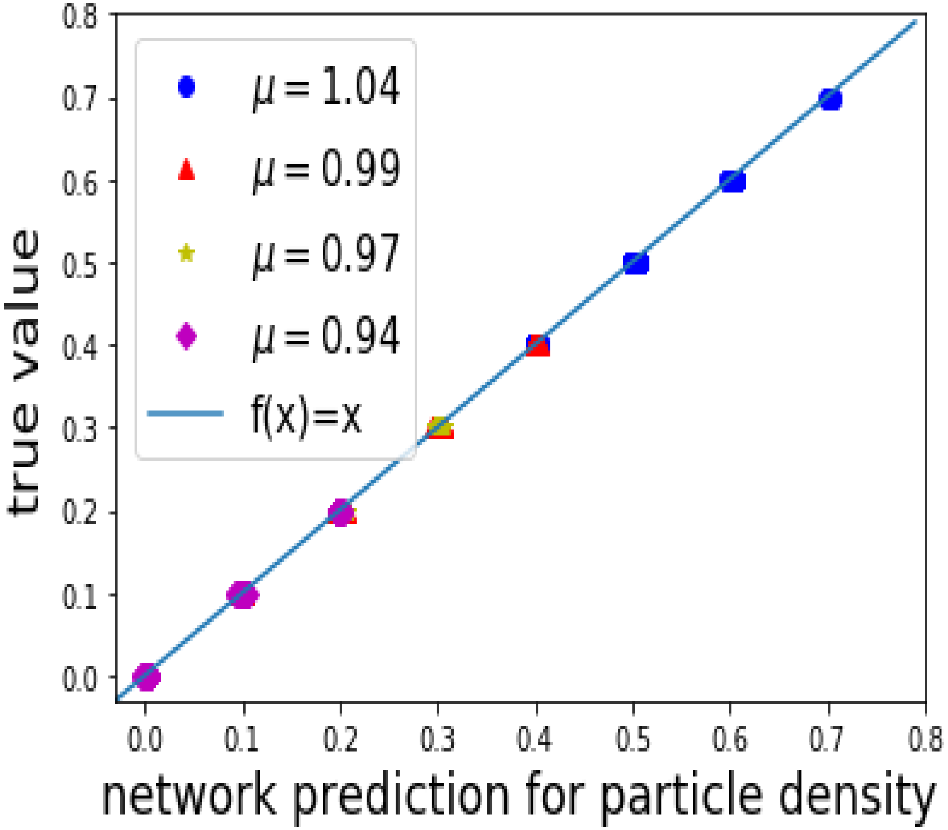
$$\mathcal{L} = -\frac{1}{2N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 + \lambda\|\theta\|_2^2$$

$$\text{Training set}: \quad \mu = 0.91 \quad \text{and} \quad \mu = 1.05$$

regression for particle density n

Note, for training, only used $\mu = 0.91 \quad \text{and} \quad \mu = 1.05$

$$n = \frac{1}{N_x N_t a} \sum_n k_t(n)$$



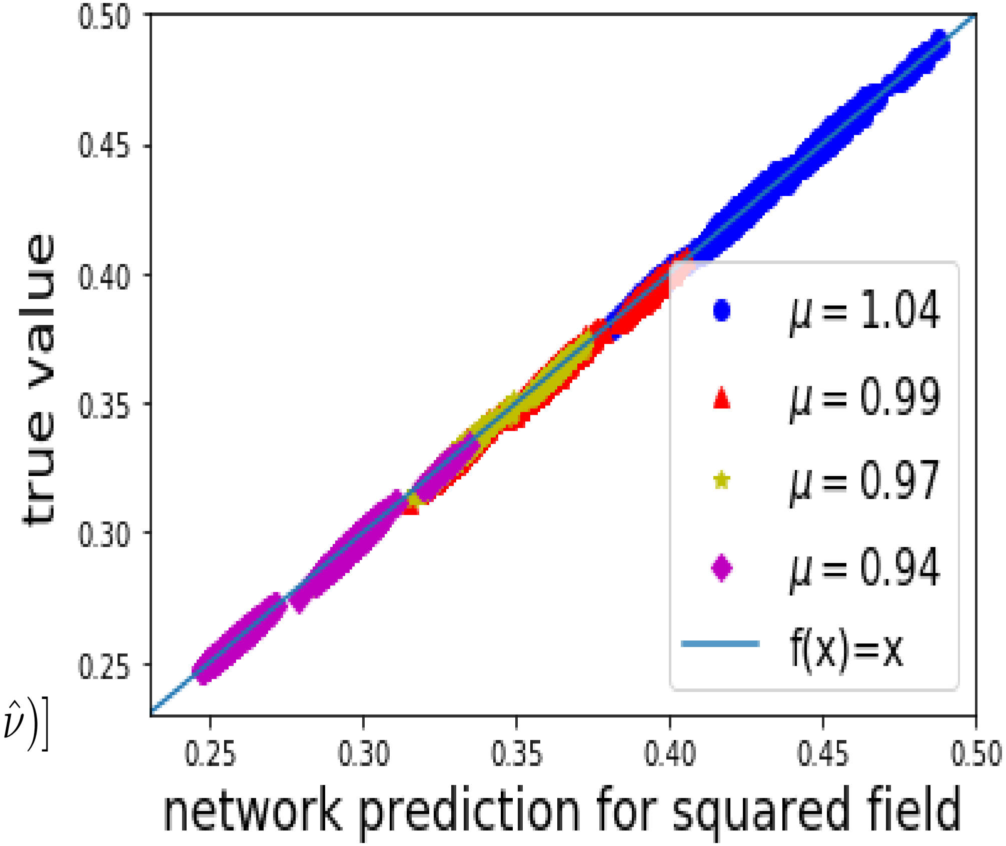
regression for squared field $|\phi|^2$

Note, for training, only used $\mu = 0.91$ and $\mu = 1.05$

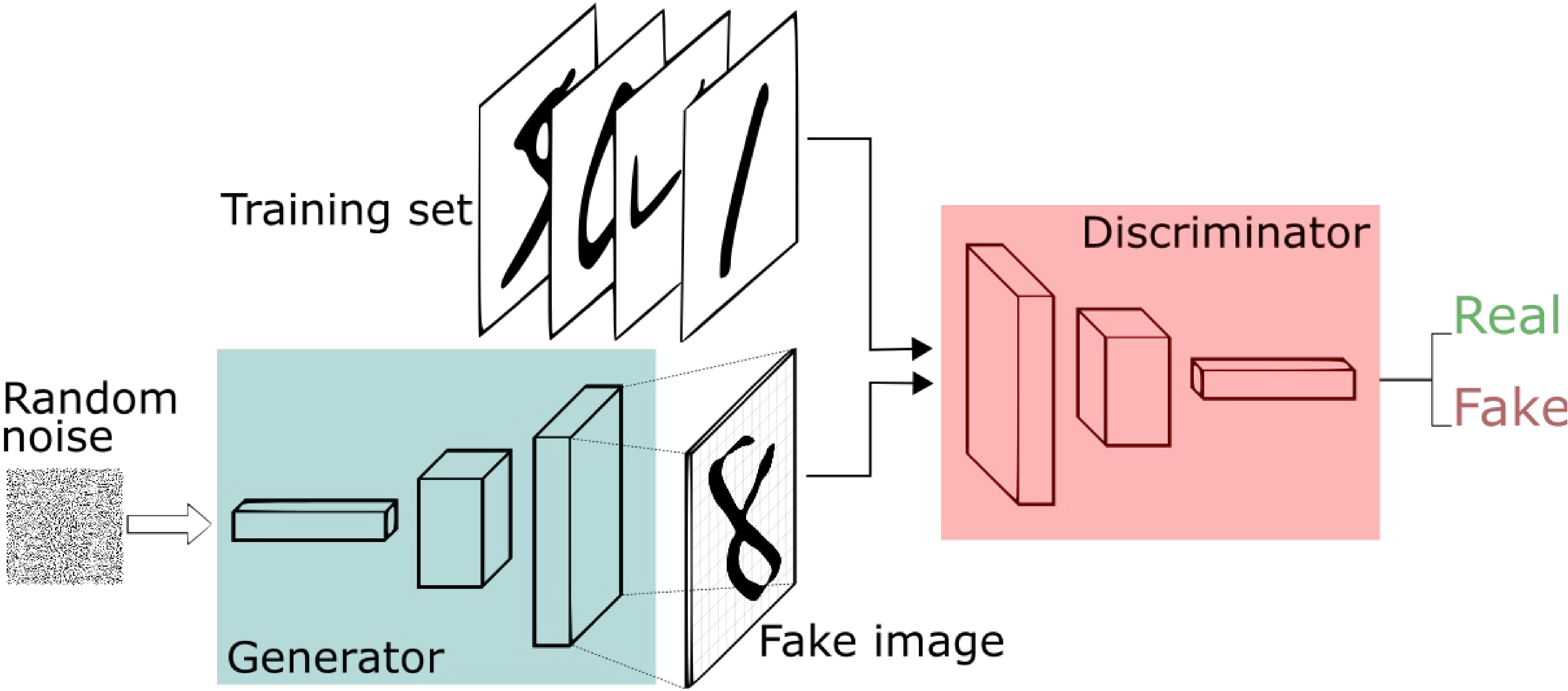
$$|\phi|^2 = \frac{1}{N_x N_t} \sum_n \frac{W[s(n)+2]}{W[s(n)]}$$

$$W[s(n)] = \int_0^\infty \mathrm{d}r\, r^{s(n)+1}\, e^{-(4+m^2)r^2 - \lambda r^4}$$

$$s(n) = \sum_\nu [|k_\nu(n)| + |k_\nu(n-\hat{\nu})| + 2(\ell_\nu(n) + \ell_\nu(n-\hat{\nu})]$$



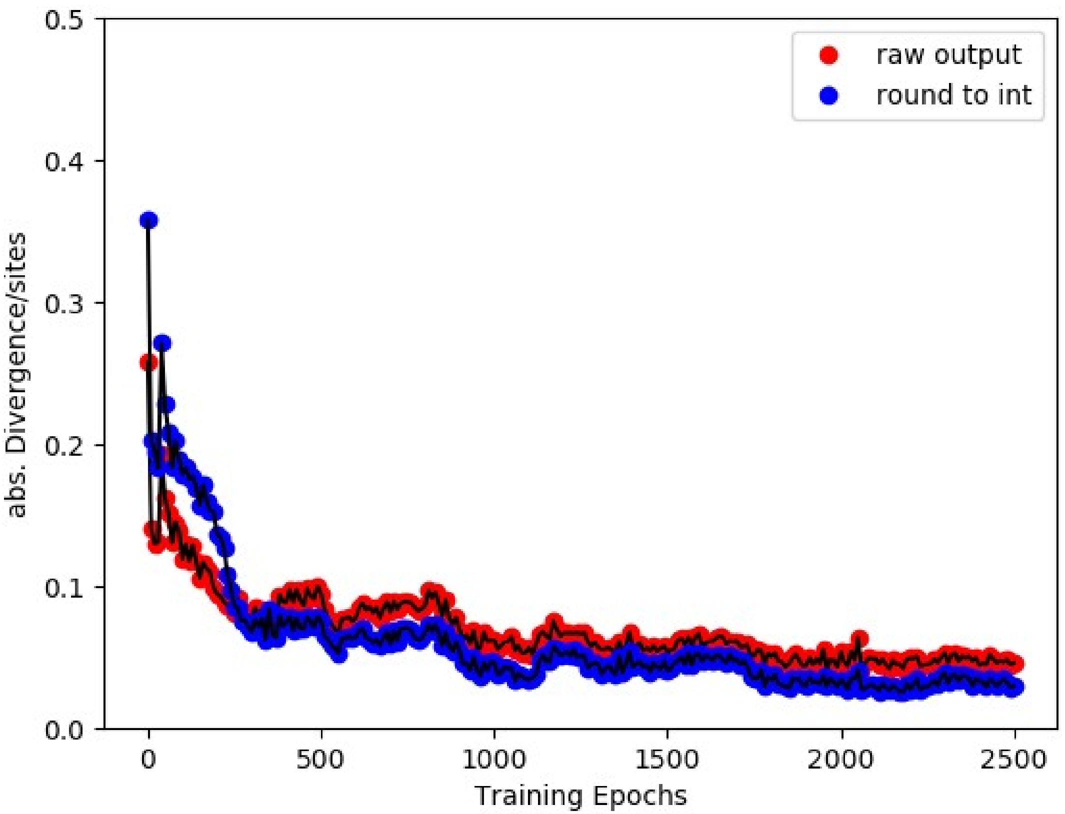
Gnerative Adversarial Network



GAN – generate proper configurations

The divergence condition automatically get learned :

Physical configs can be generated



$$\nabla \cdot k(n) = \sum_{\nu} [k_{\nu}(n) - k_{\nu}(n - \hat{\nu})] = 0$$

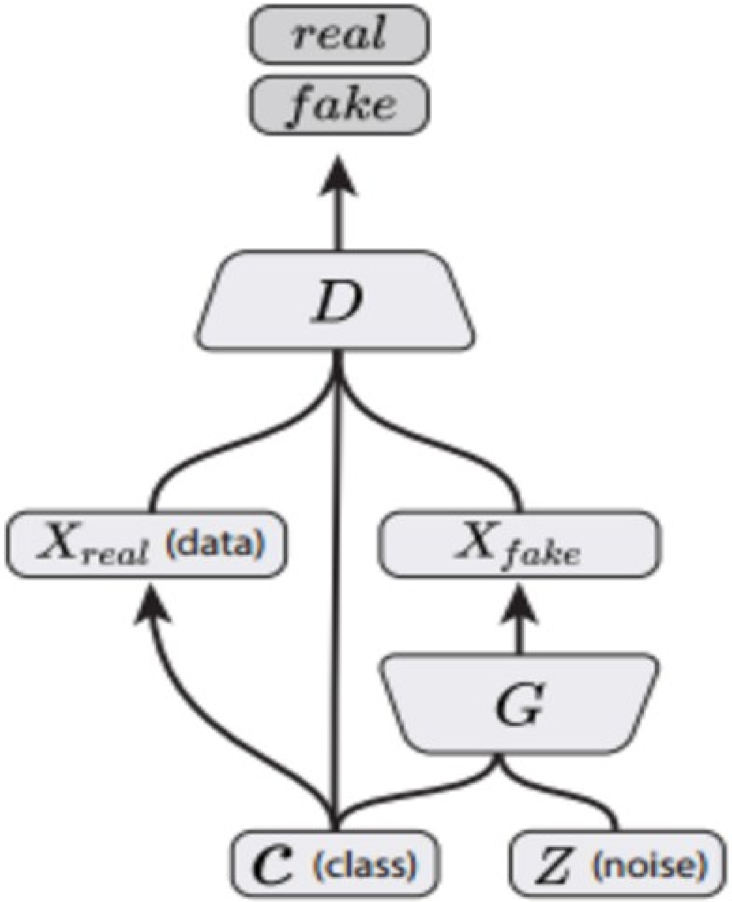
Conditional GAN

make GAN conditional on particle density n,

We train GAN using one esemble with $\mu = 1.05$ labled as well by n (including <u>n=0.4, 0.5, 0.6, 0.7</u>),

Once trained, in generating stage,

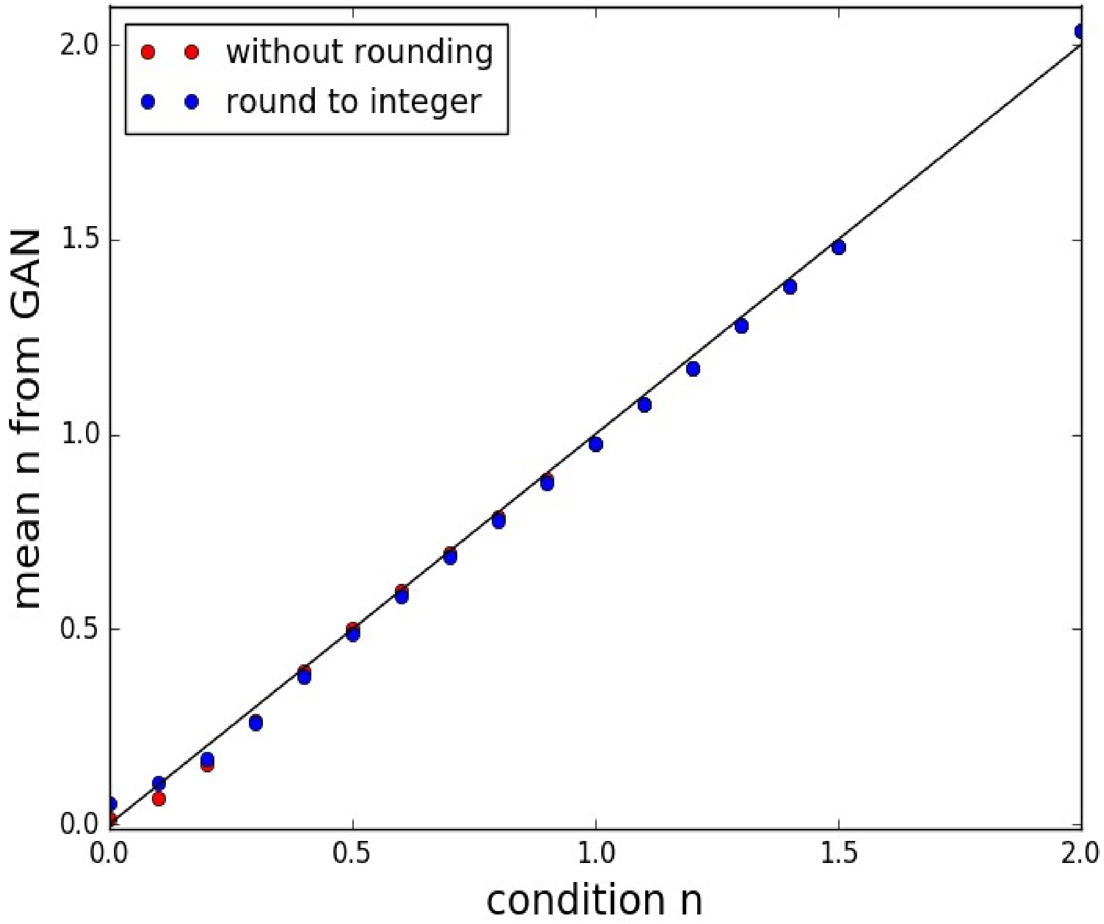
We specify different n values.



Conditional GAN
(Mirza & Osindero, 2014)

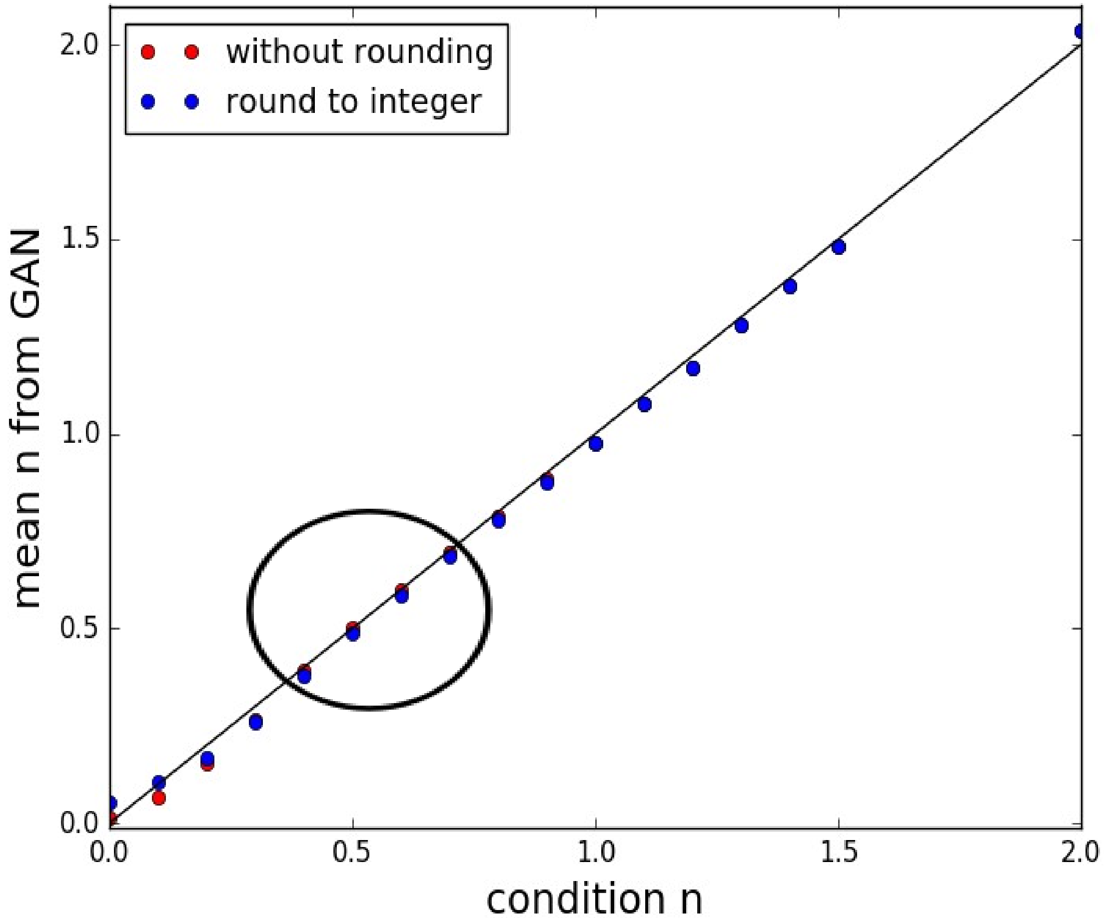
Conditional GAN

mean value for n and squared field of generated ensemble is controlled by condition in c-GAN.



Conditional GAN

mean value for n and squared field of generated ensemble is controlled by condition in c-GAN.



Results

(1) Classification 1: NN can identify interaction information

(2) Classification 2: NN can pin down phase transition point

(3) Regression: NN can learn physical observable (non-linear interpolation)

(4) Generative model : GAN can generate physical configs

For canonical ensemble, GAN can generate beyond training examples

Thanks!

Dualization approach for $\lambda\phi^4$

Euclidean continuum action for complex 1+1d scalar field

$$S^{\text{cont}} = \int_0^L \mathrm{d}x_1 \int_0^{1/T} \mathrm{d}x_2 \left[(D_\nu\phi)^*(D_\nu\phi) + m^2\phi^*\phi + \lambda(\phi^*\phi)^2\right] , \qquad D_\nu = \partial_\nu + i\mu\delta_{\nu,2}$$

On a lattice with n labels the lattice sites:

$$S^{\text{lat}} = \sum_n \left\{(4+m^2)\phi^*(n)\phi(n) + \lambda[\phi^*(n)\phi(n)]^2 - \sum_{\nu=1,2}\left[e^{\mu\delta_{\nu,2}}\phi^*(n)\phi(n+\hat{\nu}) + e^{-\mu\delta_{\nu,2}}\phi^*(n)\phi(n-\hat{\nu})\right]\right\}$$

Partition function is defined from path integral:

$$\mathcal{Z} = \int \phi \, \exp\left(-S^{\text{lat}}[\phi]\right)$$

Dualization approach for $\lambda\phi^4$

Flux representation for partition function :

$$\mathcal{Z} = \sum_{\{k,\ell\}} \prod_n \left\{ e^{\mu k_t(n)} \cdot W[s(n)] \cdot \delta[\nabla \cdot k(n)] \cdot \prod_\nu A[k_\nu(x), \ell_\nu(x)] \right\}$$

$$W[s(n)] = \int_0^\infty \mathrm{d}r \, r^{s(n)+1} \, e^{-(4+m^2)r^2 - \lambda r^4}$$

$$s(n) = \sum_\nu \left[|k_\nu(n)| + |k_\nu(n-\hat{\nu})| + 2(\ell_\nu(n) + \ell_\nu(n-\hat{\nu}) \right]$$

$$A[k_\nu(x), \ell_\nu(x)] = \frac{1}{(\ell_\nu(n) + |k_\nu(n)|)! \, \ell_\nu(n)!}$$

Divergence constraint : $\nabla \cdot k(n) = \sum_\nu [k_\nu(n) - k_\nu(n-\hat{\nu})] = 0$

Observables : n and $|\phi|^2$

Net particle density and squared field expectation

$$\langle n \rangle = \frac{T}{L} \frac{\partial \log \mathcal{Z}}{\partial \mu}$$

$$\langle |\phi|^2 \rangle = \frac{T}{L} \frac{\partial \log \mathcal{Z}}{\partial (m^2)}$$

Flux representation for above :

$$n = \frac{1}{N_x N_t a} \sum_n k_t(n)$$

$$|\phi|^2 = \frac{1}{N_x N_t} \sum_n \frac{W[s(n)+2]}{W[s(n)]}$$

GAN – distribution

Zero-sum game – Nash equilibrium

$$G^* = \arg\min_G \max_D (-\mathcal{L}_D(G, D))$$

$$\mathcal{L}_D = -\mathbb{E}_{\hat{x}\sim p_r(\hat{x})}[\log(D(\hat{x}))] - \mathbb{E}_{z\sim p(z)}[\log(1 - D(G(z)))]$$

$$\mathcal{L}_G = \mathbb{E}_{z\sim p(z)}[\log(1 - D(G(z)))]$$

$$D^*(\hat{x}) = \frac{p_r(\hat{x})}{p_r(\hat{x}) + p_g(\hat{x})}$$