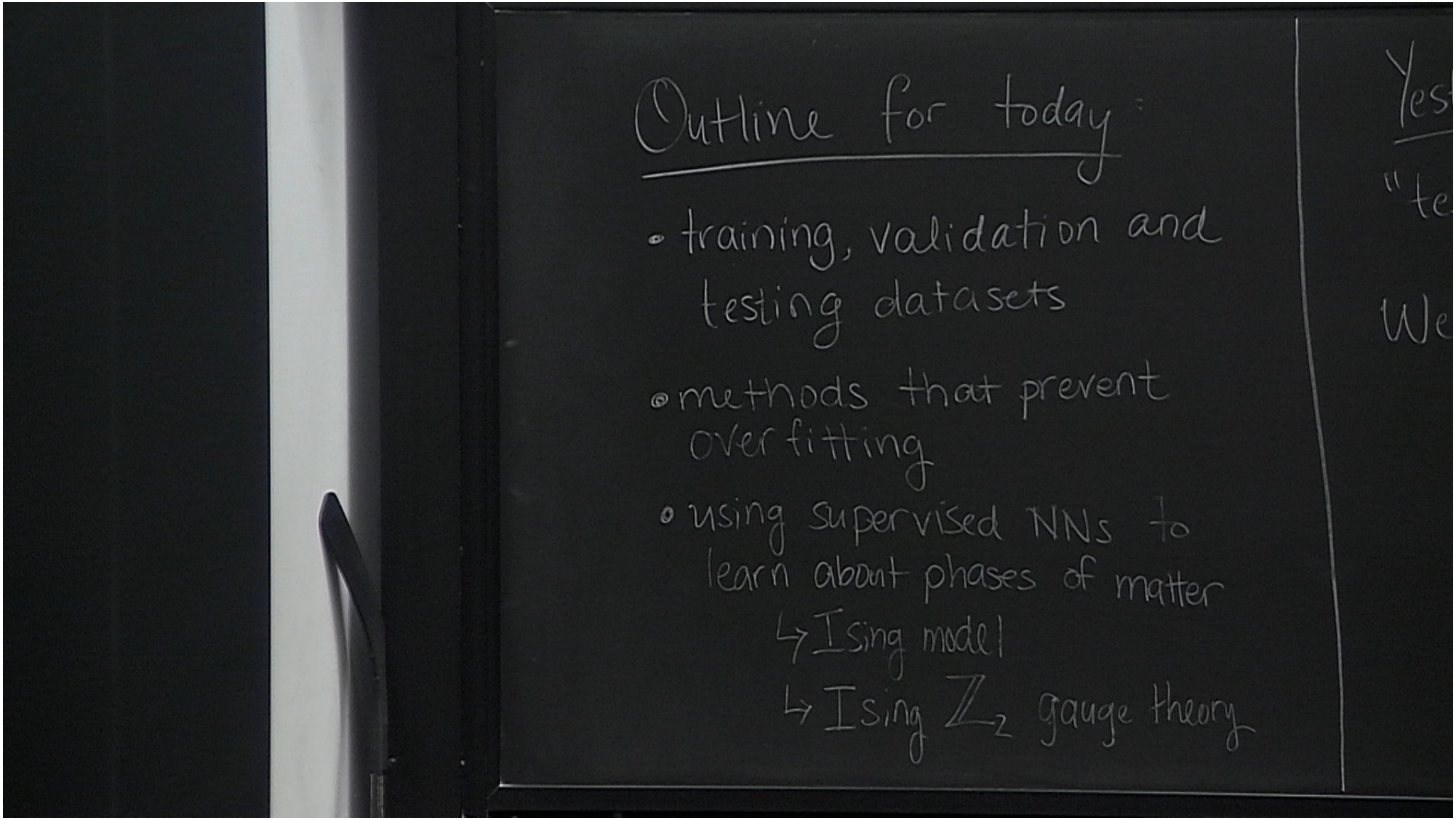
Title: PSI 2018/2019 - Machine Learning - Lecture 4

Speakers: Lauren Hayward Sierens

Collection: PSI 2018/2019 - Machine Learning (Hayward Sierens)

Date: March 28, 2019 - 9:00 AM

URL: http://pirsa.org/19030056

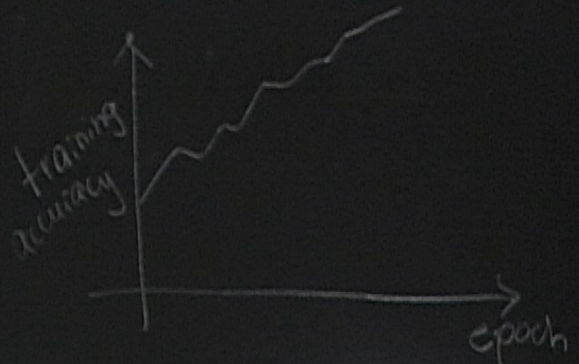

Outline for today:
• training, validation and testing datasets
• methods that prevent overfitting
• using supervised NNs to learn about phases of matter
↳ Ising model
↳ Ising $\mathbb{Z}_2$ gauge theory

Yesterday:

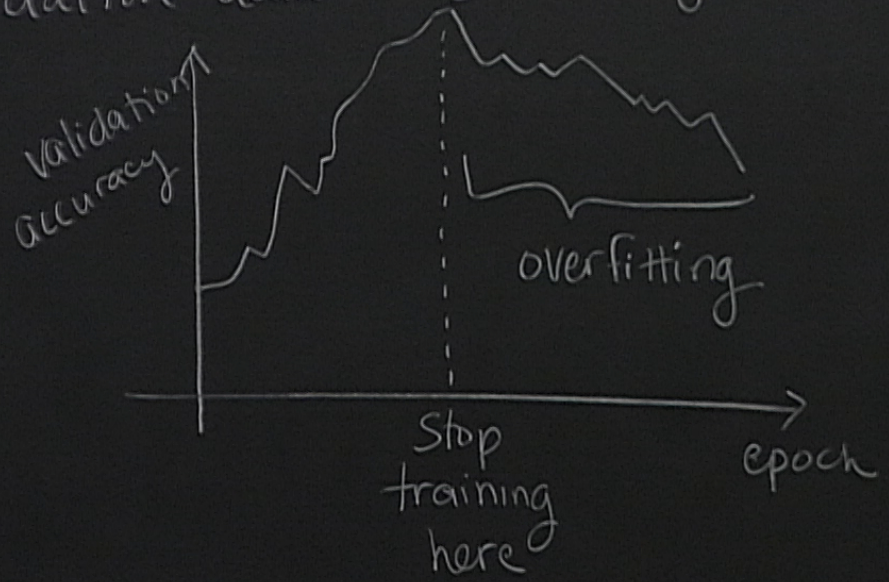
"testing data" should be "validation data".

We discussed dividing our data into two sets to recognize overfitting. We can do further analysis by dividing into three sets

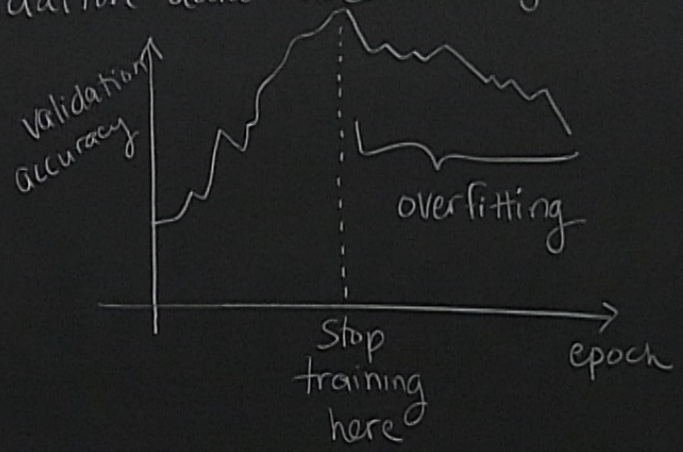
- training data: used to adjust the weights and biases



- validation data: used to diagnose overfitting to the training data



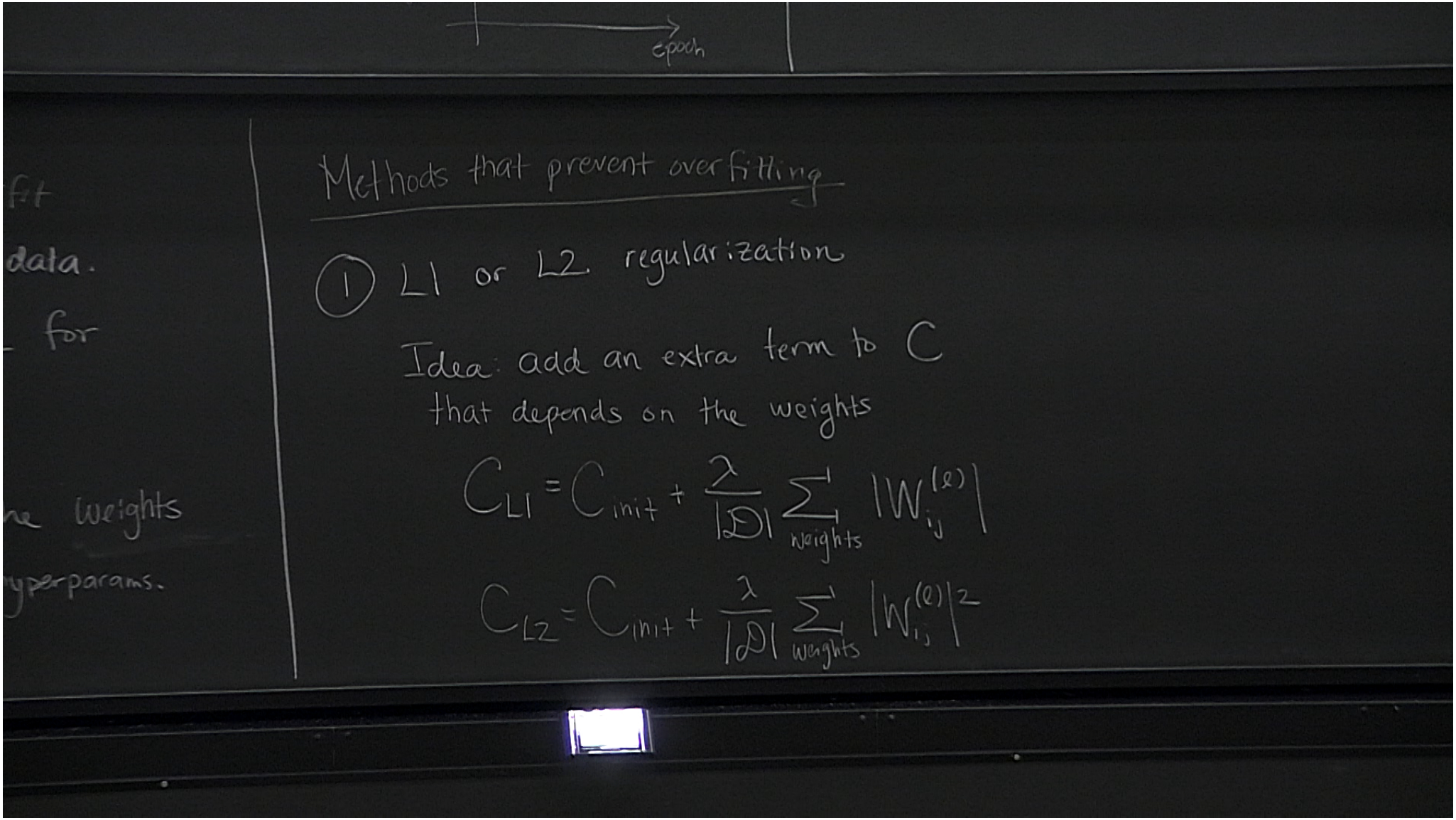
- Validation data: used to diagnose overfitting to the training data

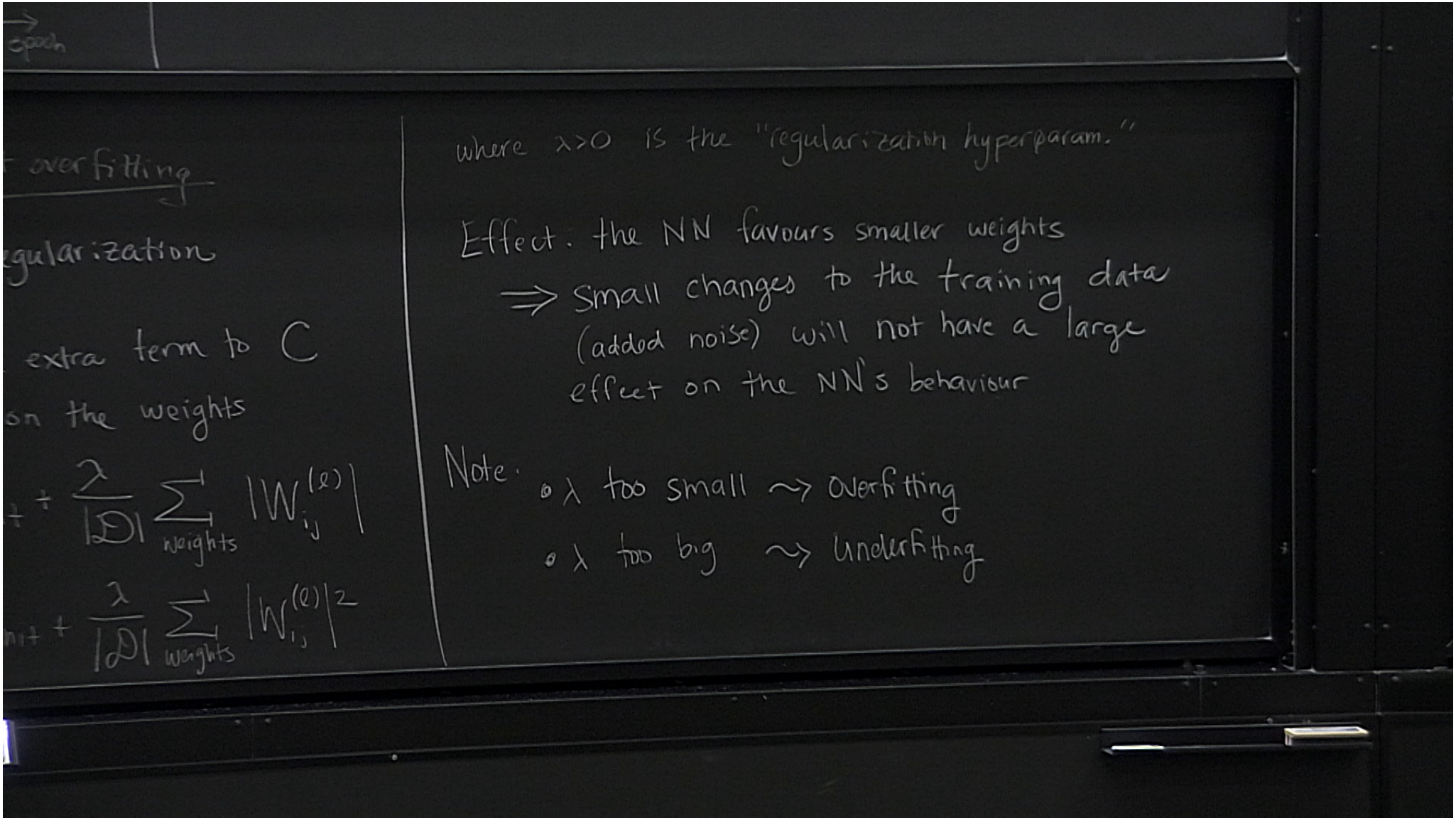


We could then adjust the hyperparams. and repeat the training with the goal of maximizing validation accuracy.

During this process we might overfit the hyperparams. to the validation data. So we report the final accuracy for a third "testing dataset".

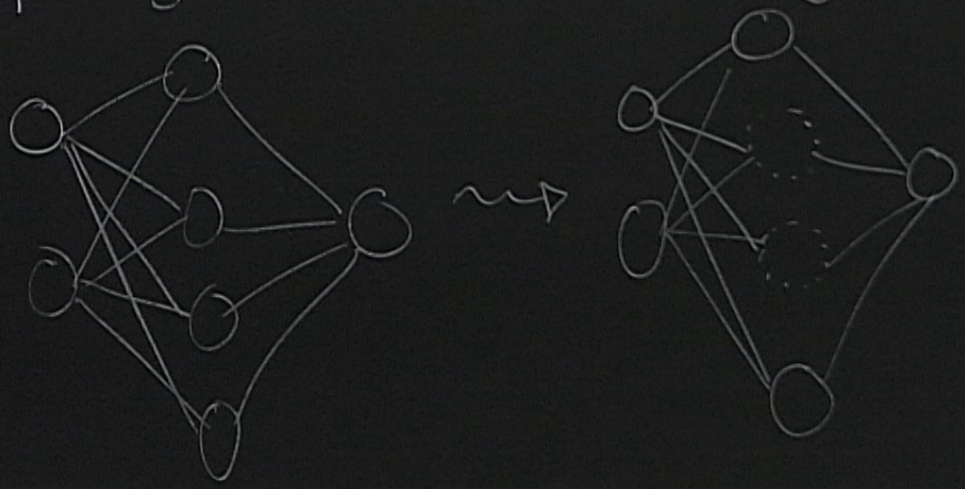
- testing data. never used to adjust the weights and biases AND never used to adjust the hyperparams.


epoch
Methods that prevent overfitting
1 L1 or L2 regularization
Idea: add an extra term to C that depends on the weights
$C_{L1} = C_{init} + \frac{\lambda}{|D|} \sum_{weights} |W_{ij}^{(\ell)}|$
$C_{L2} = C_{init} + \frac{\lambda}{|D|} \sum_{weights} |W_{ij}^{(\ell)}|^2$


epoch
overfitting
egularization
extra term to C
on the weights
$+ \frac{\lambda}{|D|} \sum_{\text{weights}} |W_{ij}^{(\ell)}|$
$+ \frac{\lambda}{|D|} \sum_{\text{weights}} |W_{ij}^{(\ell)}|^2$
where $\lambda > 0$ is the "regularization hyperparam."
Effect: the NN favours smaller weights
⇒ small changes to the training data (added noise) will not have a large effect on the NN's behaviour
Note:
• λ too small ⇝ overfitting
• λ too big ⇝ underfitting

② Dropout

Idea: randomly-selected neurons are temporarily masked out during training



Run the training for some subset of the data. Then choose a new random mask and repeat

Effect: similar to averaging over many different NNs (each of which could overfit in different ways).

Distinguishing phases of matter using SL

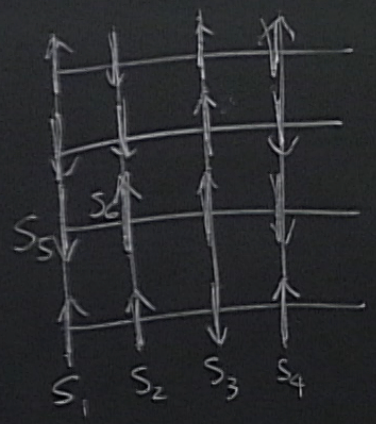
See Carrasquilla & Melko, arXiv:1605.01735

Classical Ising Model

$$H = -J \sum_{\langle ij \rangle} s_i s_j \quad , \quad s_i = +1 \text{ or } -1 \; (\uparrow \text{ or } \downarrow)$$

The d.o.f. are on sites of a lattice

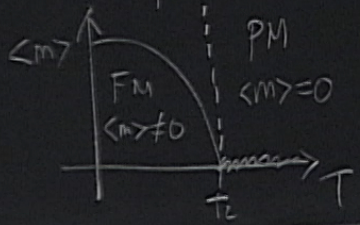
2D square lattice



phase transition at $T_c \approx 2.269J$ in the thermodynamic limit

- $T < T_c$: ferromagnetic (FM) phase
- $T > T_c$: paramagnetic (PM) phase

The order param. is $m = \frac{1}{N}\sum_i s_i$



Goal: train a NN to distinguish FM and PM phases <u>without</u> knowledge of the order param. nor the Hamiltonian

Input data : spin configs.

$$\vec{X} = [+1, -1, -1, +1, -1, \ldots, +1]$$
$$S_1 \quad S_2 \quad S_3$$

Labels: 0 or 1
(FM or PM)

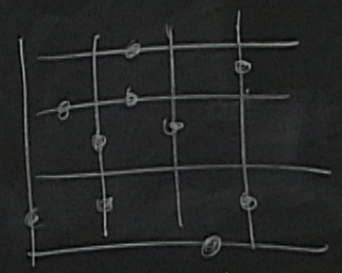
We will use MC simulation to generate spin config. samples (see tomorrow's lecture)

Classical Ising $\mathbb{Z}_2$ gauge theory

$$H = -J \sum_{\text{plaquettes}}{}' \left(\prod_{i \in p} s_i \right) \quad \text{with } s_i = +1 \text{ or } -1$$

d.o.f. are on the bonds of the lattice.

On a 2D torus



$\Rightarrow s_i = +1$

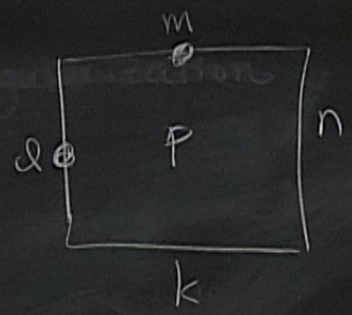
$\Rightarrow s_i = -1$

(periodic boundary conditions)

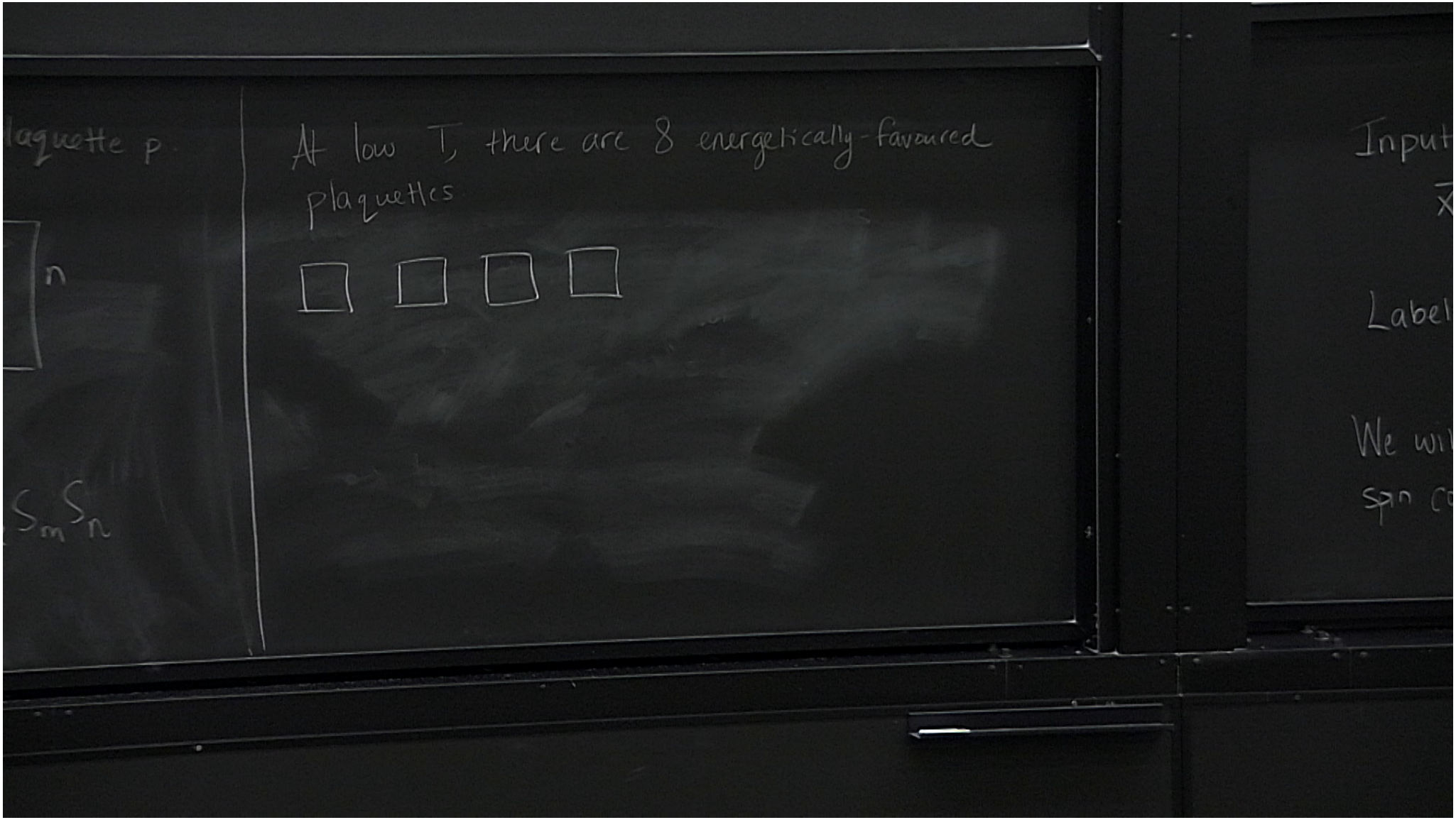
epoch

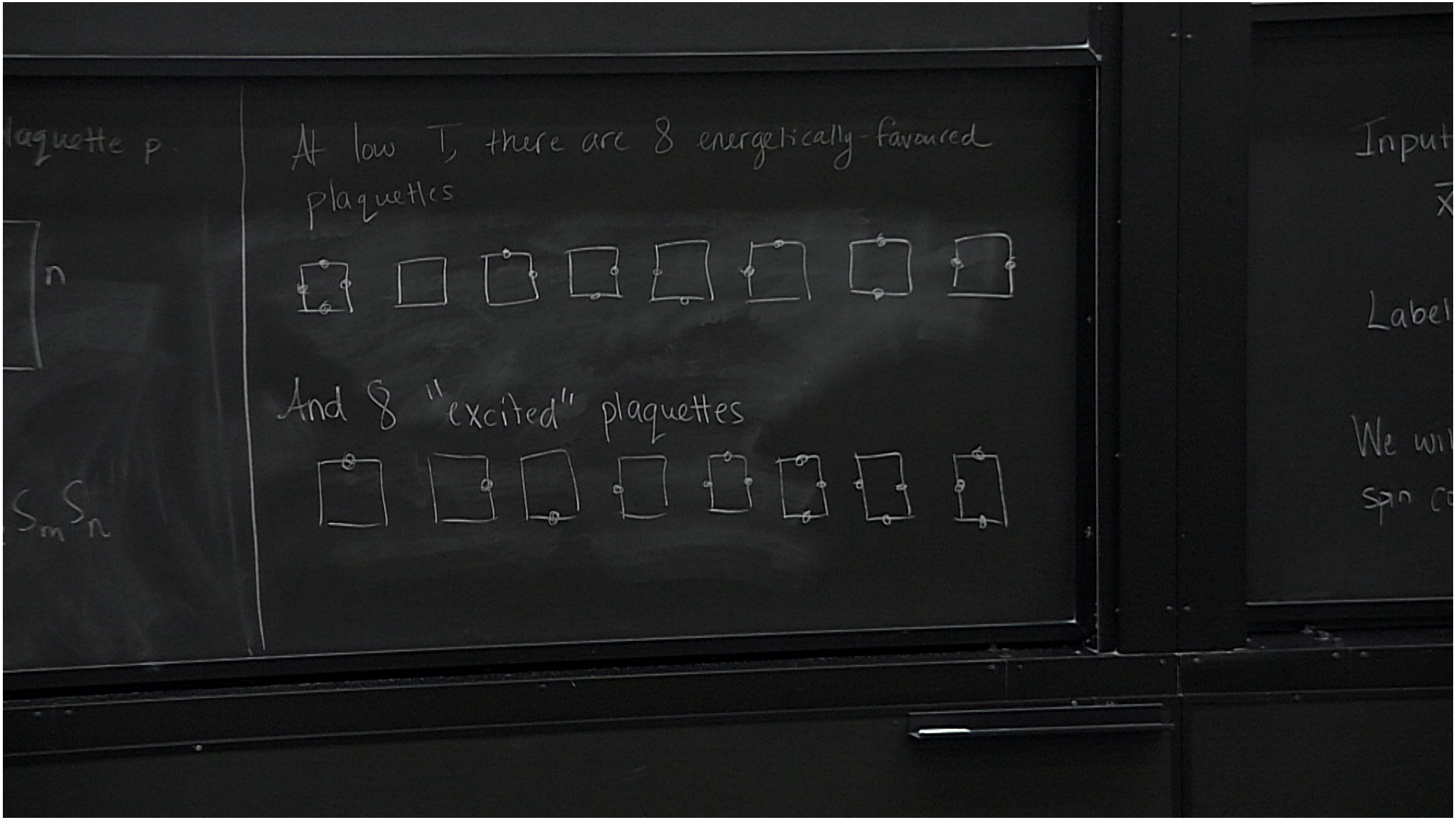
$= +1$ or -1

Consider one plaquette p.



$$\prod_{i \in p} S_i = S_k S_l S_m S_n$$

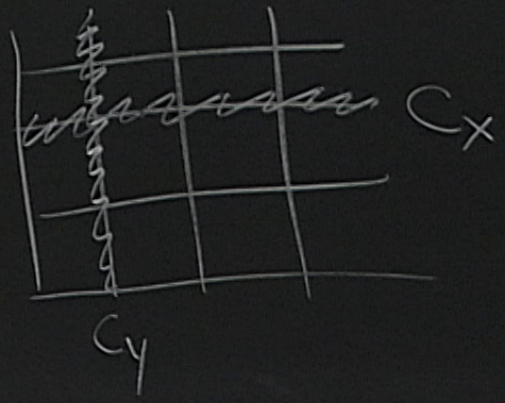

laquette p.
n
$S_m S_n$
At low T, there are 8 energetically-favoured plaquettes
Input
$\bar{x}$
Label
We wi
spin c


laquette p.
n
$S_m S_n$
At low T, there are 8 energetically-favoured plaquettes
And 8 "excited" plaquettes
Input
$\bar{x}$
Label
We wi
spin c

There is a topologically-ordered state exactly at $T=0$ in the TD limit.
(no order for $T>0$)

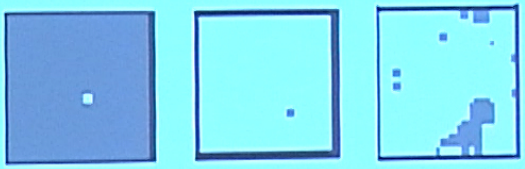
We can use the Wilson loop W_c to measure topological order

$$W_c = \prod_{i \in C} s_i = \pm 1$$



Ising model classification

Ferromagnetic phase:



Paramagnetic phase:

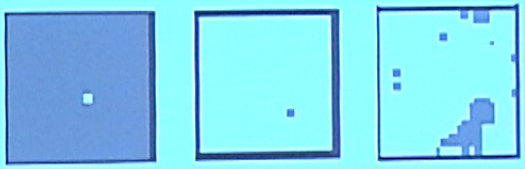


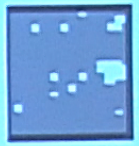
Carrasquilla and Melko, arXiv:1605.01735

PERIMETER SCHOLARS INTERNATIONAL

Ising model classification

Ferromagnetic phase:





?

Paramagnetic phase:



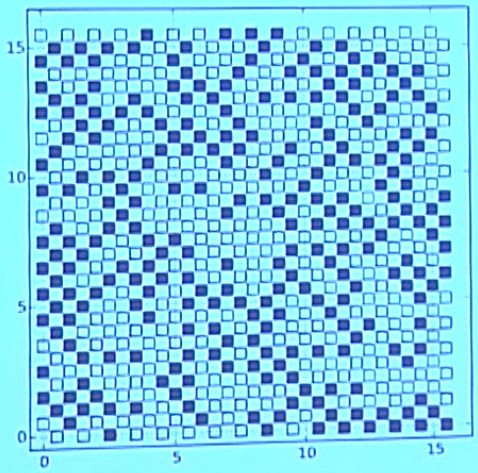


?

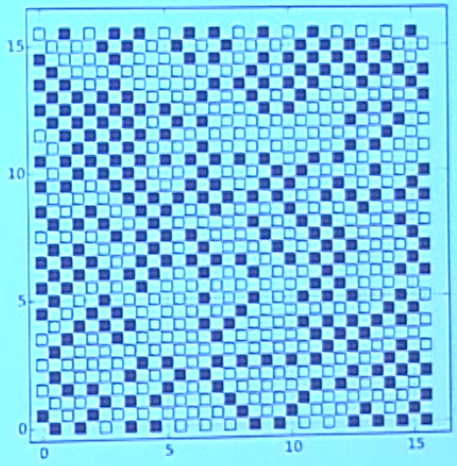
Carrasquilla and Melko, arXiv:1605.01735

PERIMETER SCHOLARS INTERNATIONAL

Ising gauge theory classification



High-temperature state



Ground state

Carrasquilla and Melko, arXiv:1605.01735

PERIMETER SCHOLARS INTERNATIONAL