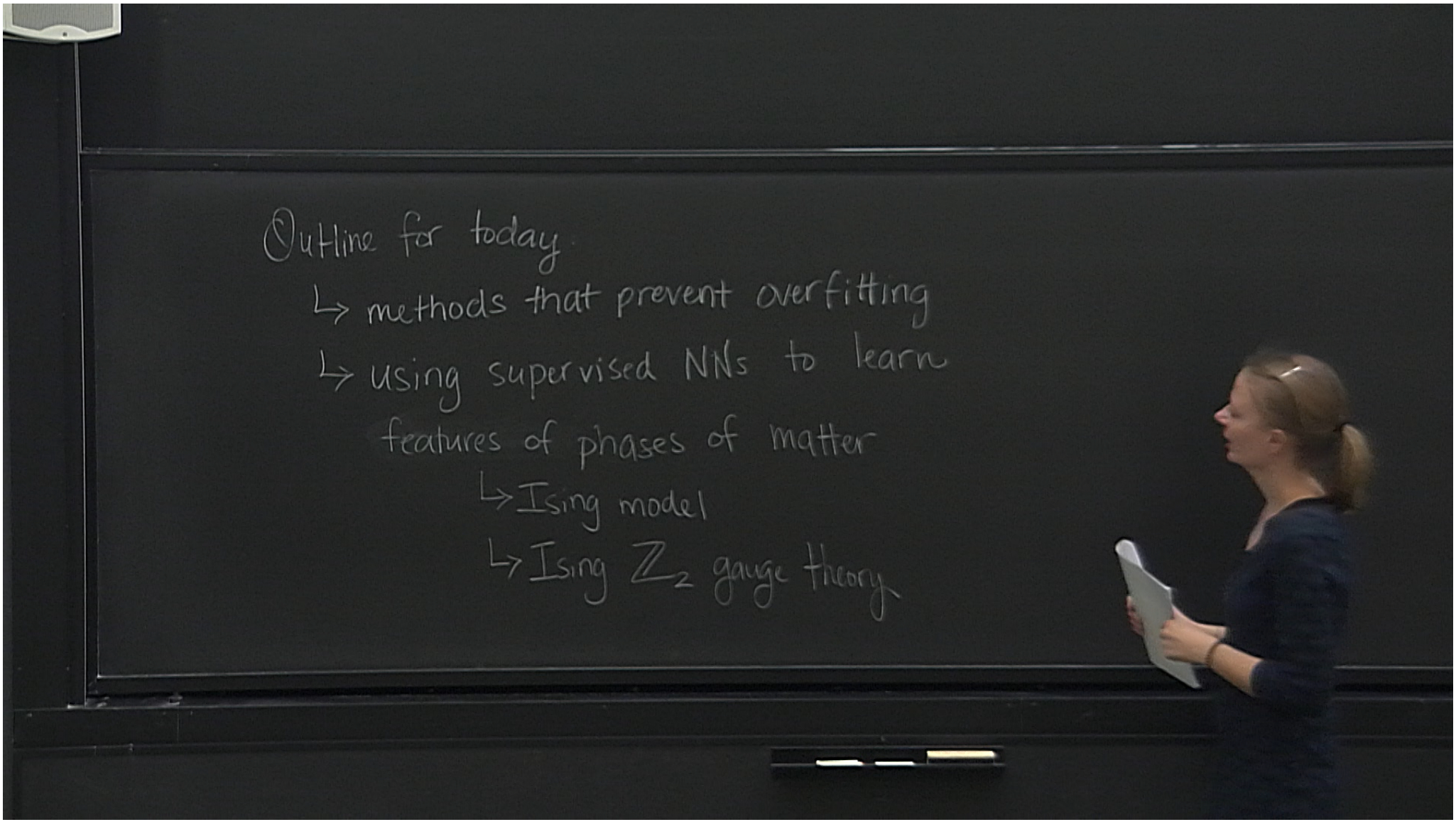
Title: PSI 2017/2018 - Machine Learning for Many Body Physics - Lecture 4

Date: Apr 12, 2018 09:00 AM

URL: http://pirsa.org/18040057

Abstract:


Outline for today:
↳ methods that prevent overfitting
↳ using supervised NNs to learn
features of phases of matter
↳ Ising model
↳ Ising $\mathbb{Z}_2$ gauge theory

Methods that prevent overfitting

1. L1 or L2 regularization

↳ adds an extra term to the cost function that depends on the magnitudes of the weights

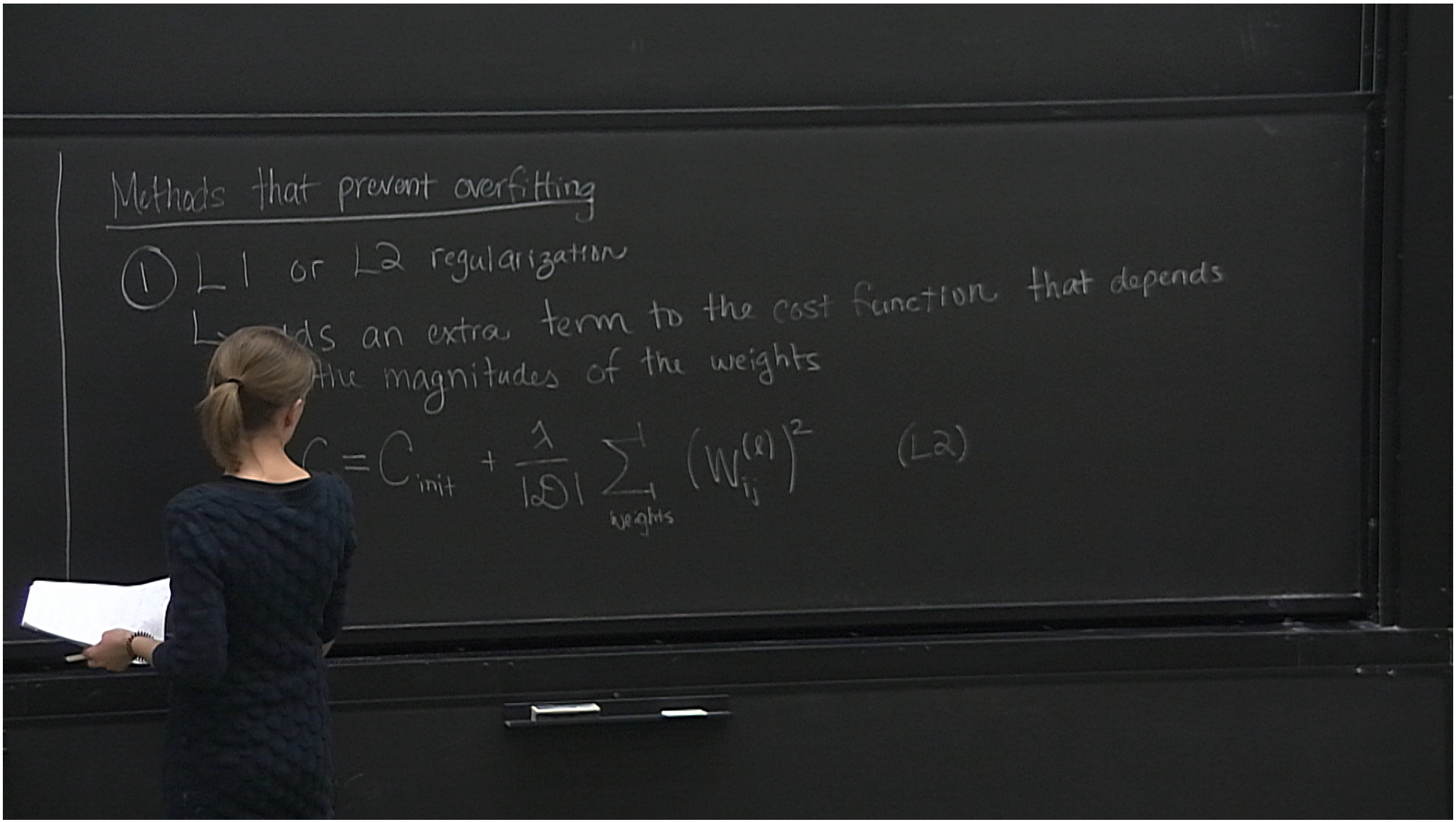
$$C = C_{init}$$

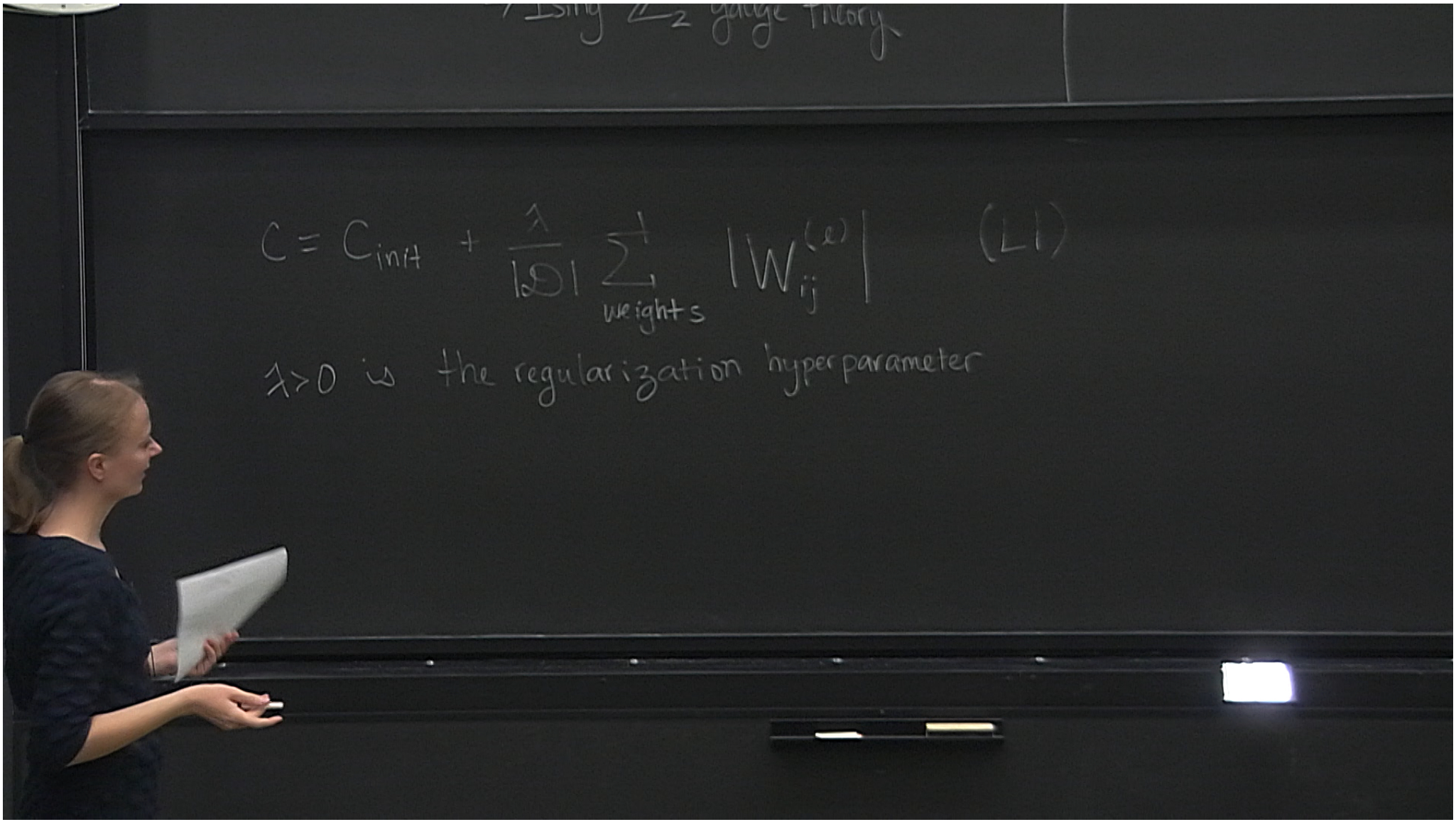
Methods that prevent overfitting

1. L1 or L2 regularization

↳ adds an extra term to the cost function that depends on the magnitudes of the weights

$$C = C_{init} + \frac{\lambda}{|\mathcal{D}|} \sum_{\text{weights}} \left(W_{ij}^{(\ell)}\right)^2$$


Methods that prevent overfitting
1 L1 or L2 regularization
L adds an extra term to the cost function that depends
the magnitudes of the weights
$C = C_{init} + \frac{\lambda}{|\mathcal{D}|} \sum_{\text{weights}} \left(W_{ij}^{(\ell)}\right)^2$ (L2)


Ising $\mathbb{Z}_2$ gauge theory
$C = C_{init} + \frac{\lambda}{|D|} \sum_{weights} |W_{ij}^{(\ell)}|$ (L1)
$\lambda > 0$ is the regularization hyperparameter

Ising $\mathbb{Z}_2$ gauge theory

$$C = C_{init} + \frac{\lambda}{|D|} \sum_{\text{weights}} |W_{ij}^{(\ell)}| \qquad (L1)$$

$\lambda > 0$ is the regularization hyperparameter

- λ too small

→ Ising $\mathbb{Z}_2$ gauge theory

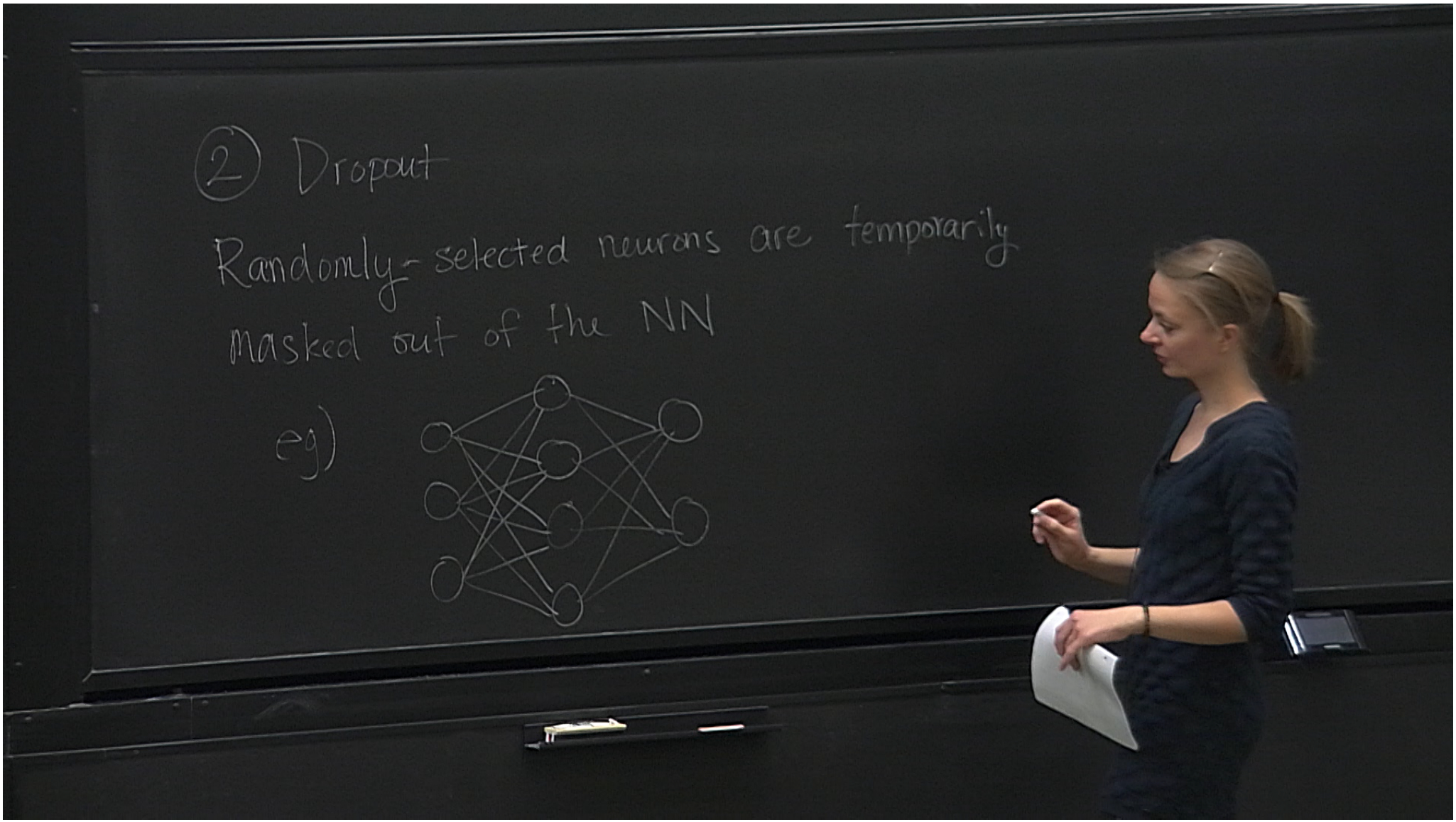
$$C = C_{init} + \frac{\lambda}{|\mathcal{D}|} \sum_{\text{weights}} |W_{ij}^{(\ell)}| \qquad (L1)$$

$\lambda > 0$ is the regularization hyperparameter

- λ too small → overfitting
- λ too large → underfitting

Effect: NN favours smaller weights

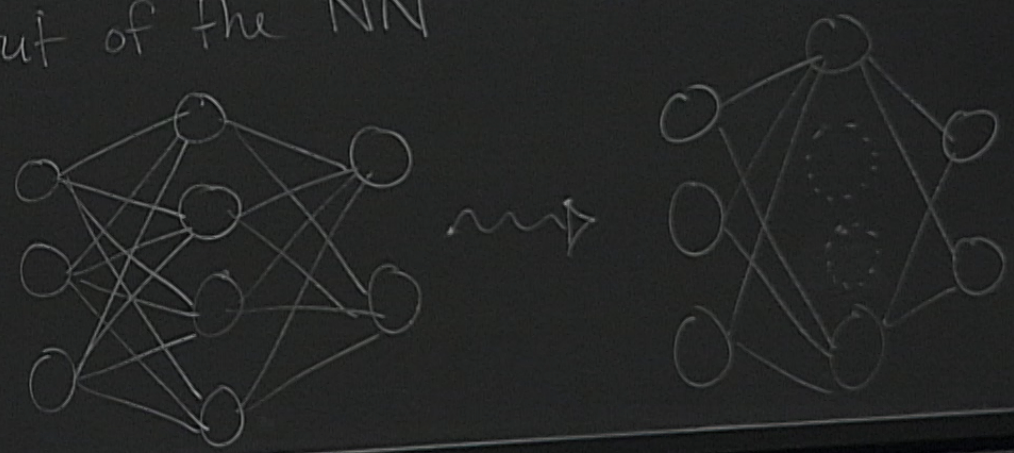
$\Rightarrow$ small change (added noise) to the training data will <u>not</u> have a large effect on the NN's behaviour

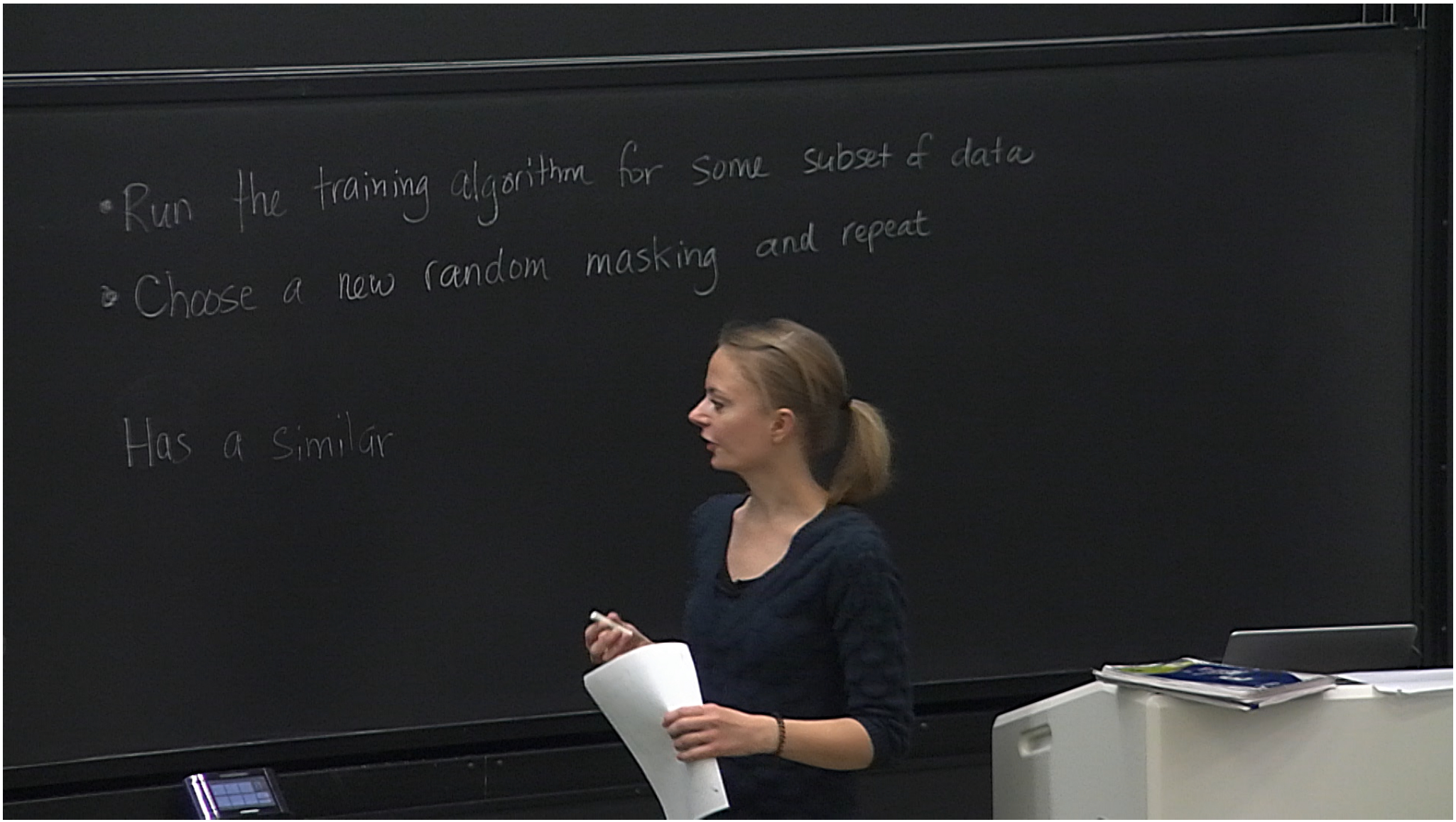

② Dropout
Randomly-selected neurons are temporarily
masked out of the NN
eg)

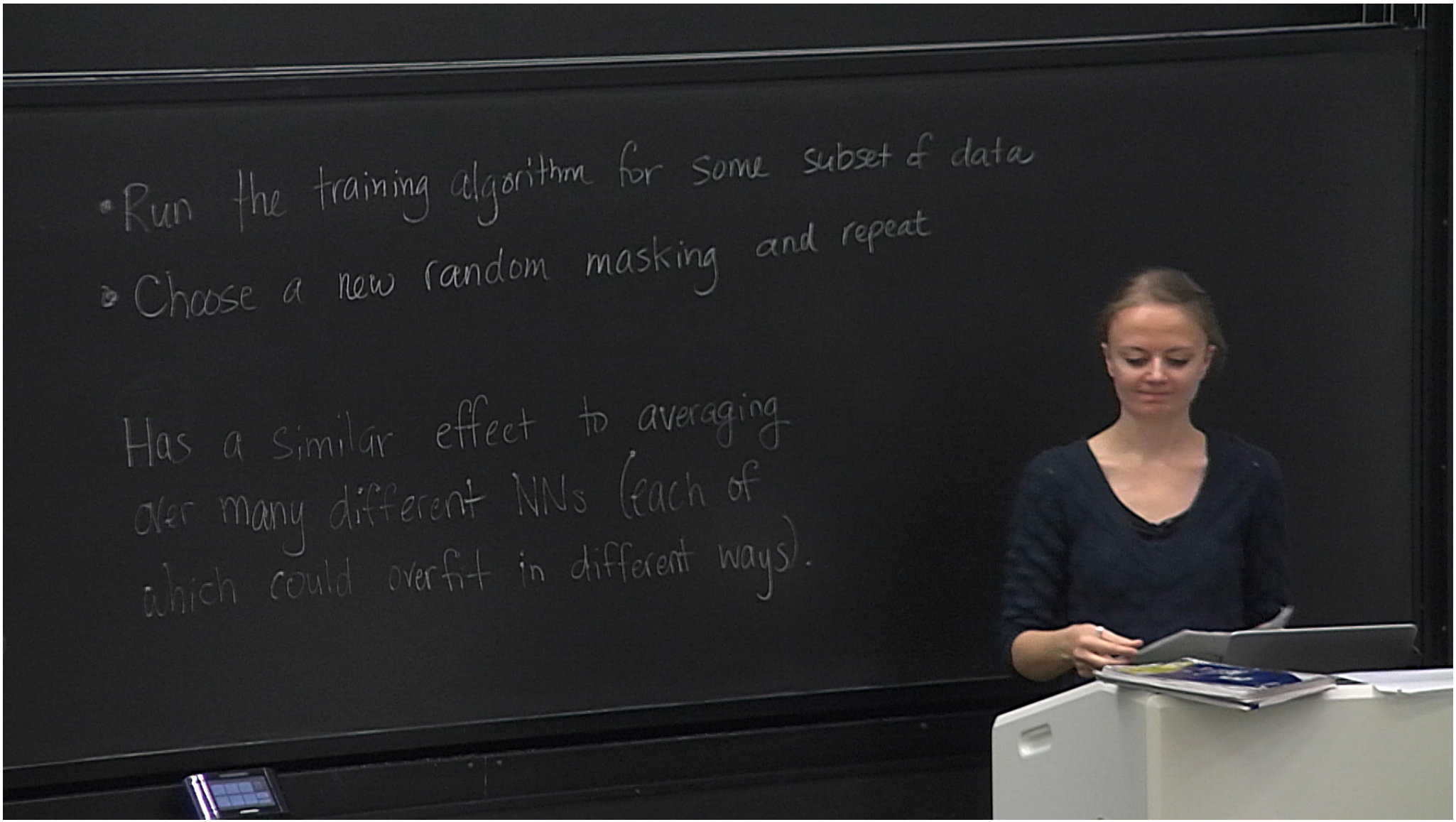
② Dropout

Randomly-selected neurons in the hidden layer(s) are temporarily masked out of the NN

eg)




• Run the training algorithm for some subset of data
• Choose a new random masking and repeat
Has a similar


• Run the training algorithm for some subset of data
• Choose a new random masking and repeat
Has a similar effect to averaging over many different NNs (each of which could overfit in different ways).

Distinguishing phases of matter using SL

↳ Carrasquilla & Melko, arXiv: 1605.01735

Distinguishing phases of matter using SL

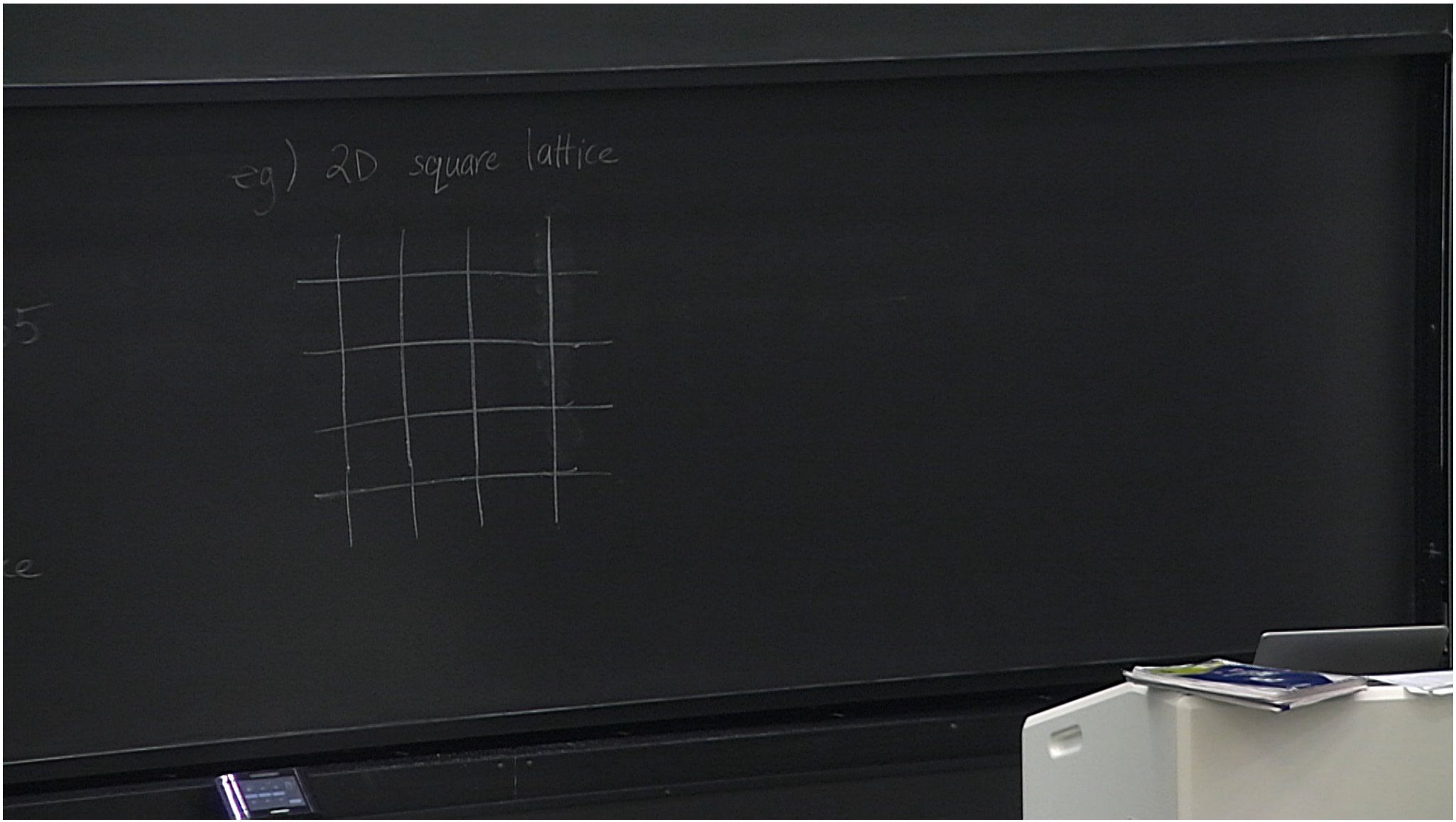
↳ Carrasquilla & Melko, arXiv: 1605.01735

Classical Ising model

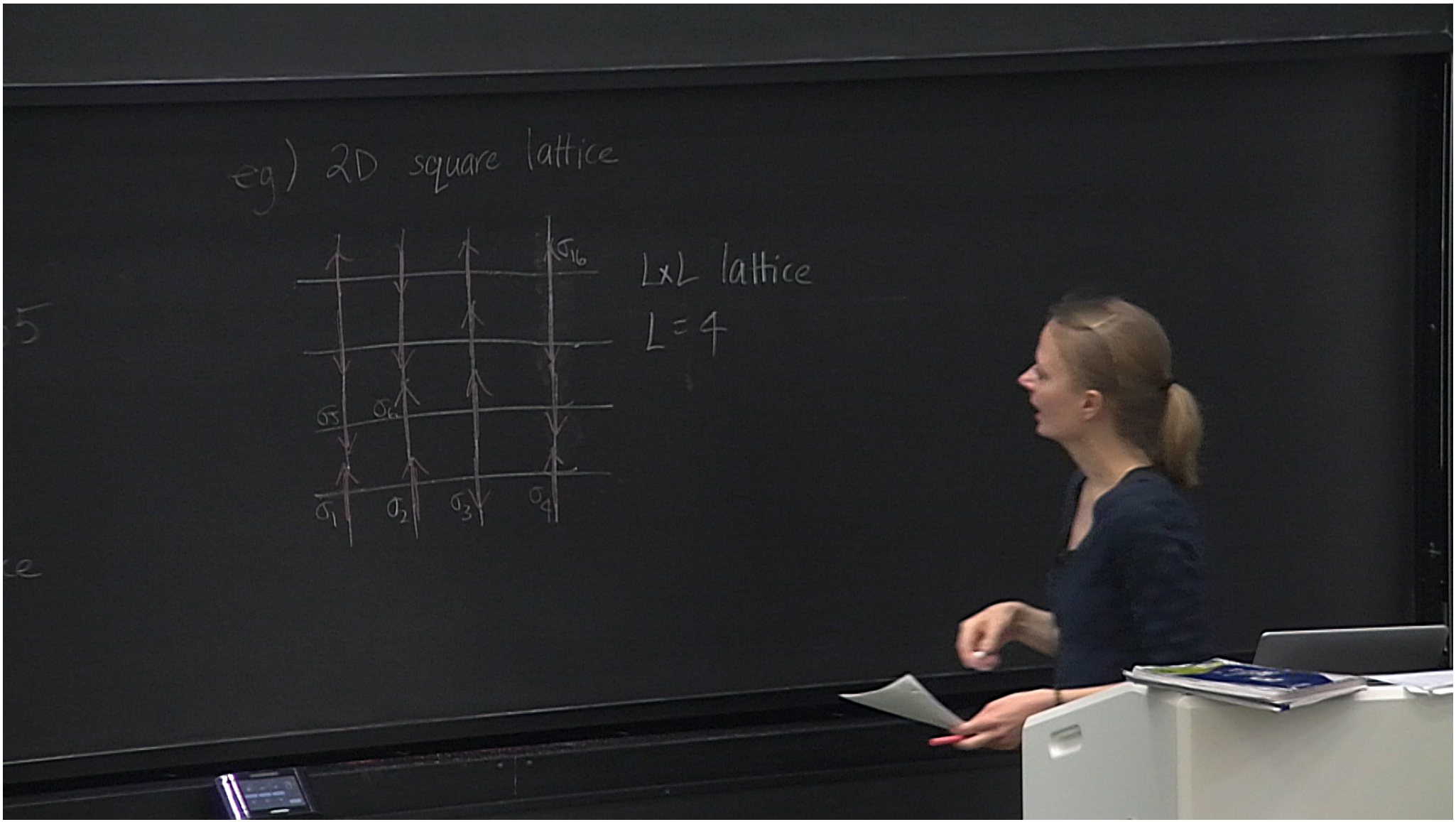
$$H = -J\sum_{\langle ij\rangle}\sigma_i\sigma_j - h\sum_i \sigma_i$$

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

$\sigma_i = +1$ or -1
$(\uparrow)$ $(\downarrow)$

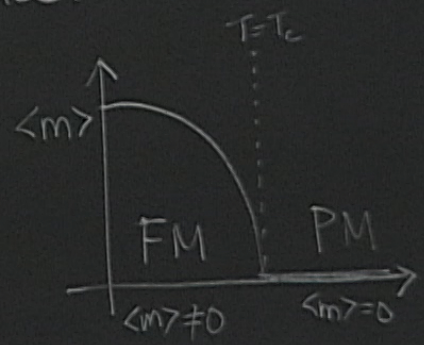

eg) 2D square lattice


eg) 2D square lattice


eg) 2D square lattice
σ_{16}
σ_5
σ_6
σ_1
σ_2
σ_3
σ_4
LxL lattice
L=4

On a 2D square lattice, we know that there is a phase transition from a ferromagnetic (FM) to a paramagnetic (PM) phase. When $h=0$: $T_c \cong 2.269J$

In the thermodynamic limit:



order parameter:

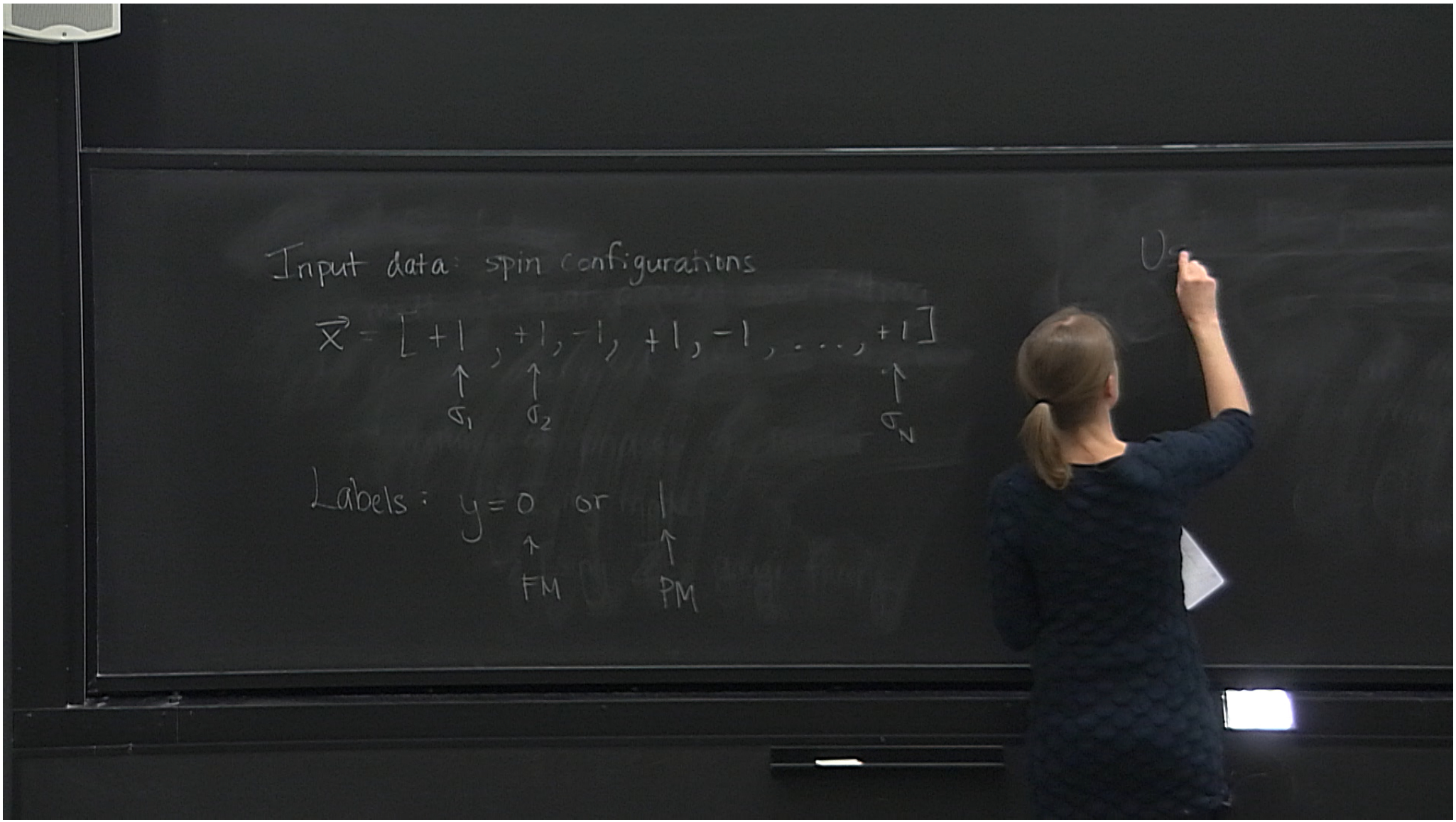
$$\langle m \rangle = \frac{1}{N} \sum_i \sigma_i$$

$\cong 2.269J$

rder parameter:

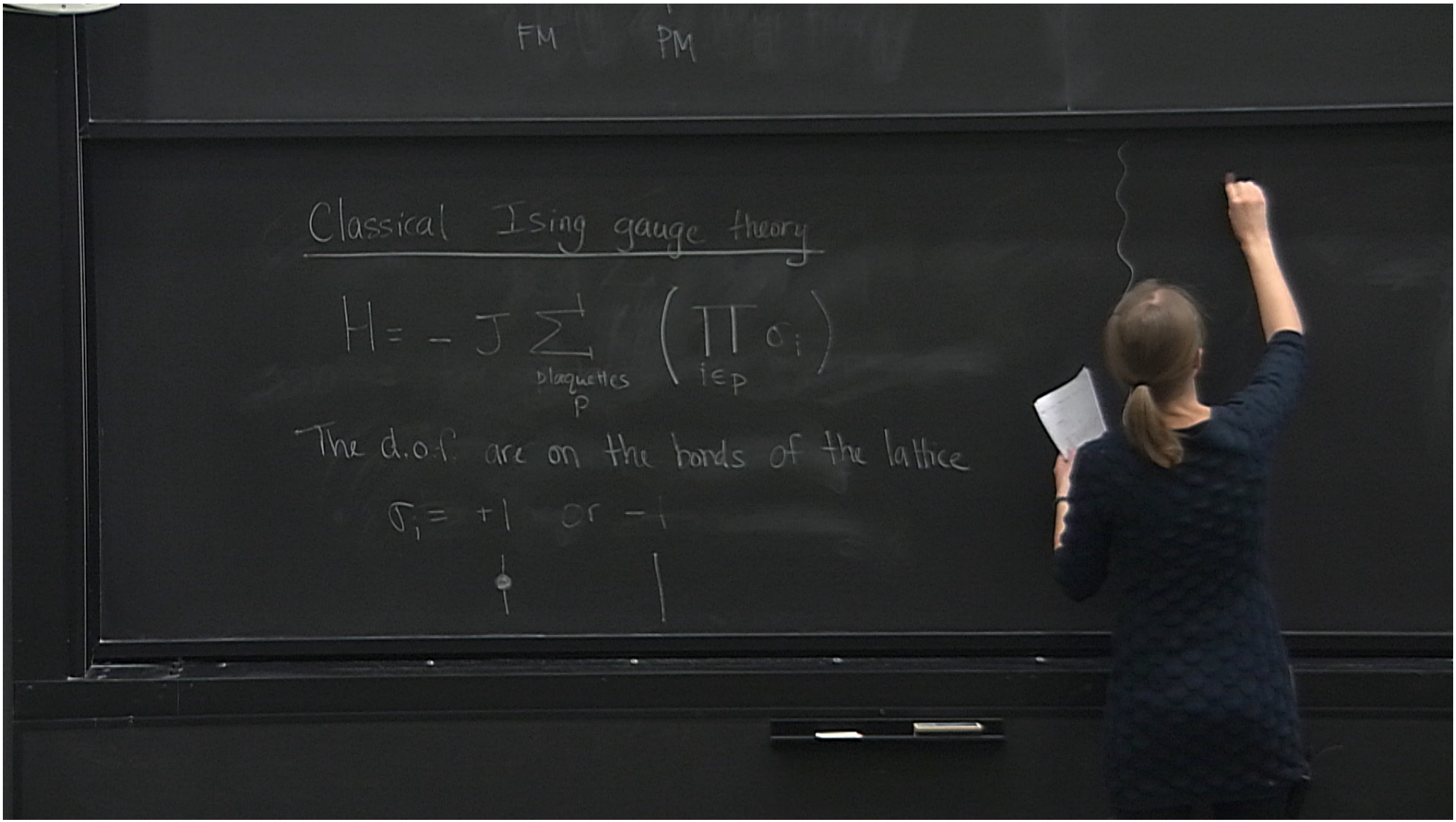
$$\langle m \rangle = \frac{1}{N}\sum_i \sigma_i$$

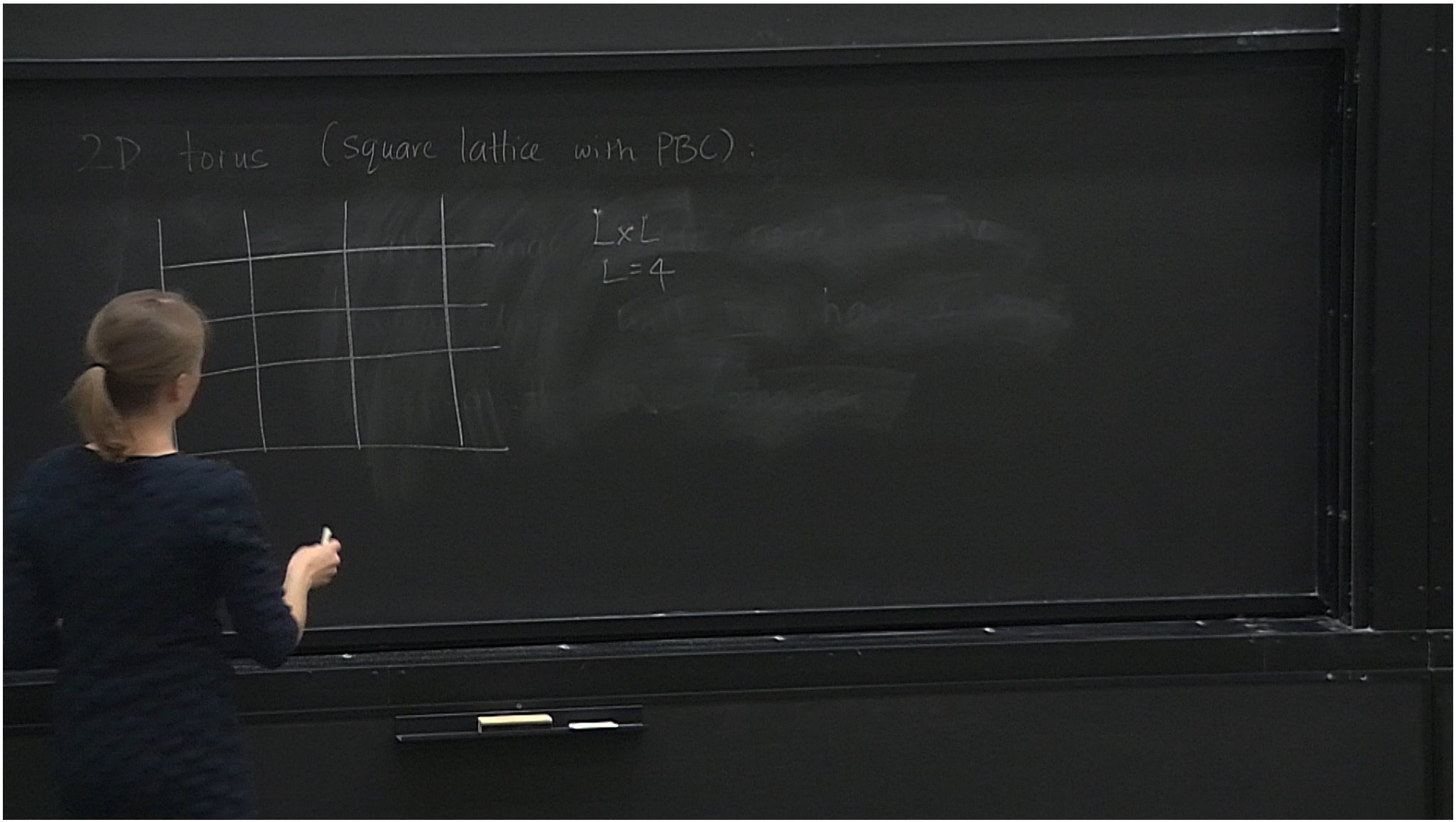
Goal: train a NN to learn how to classify the FM and PM labels (Homework #1)


Input data: spin configurations
$\vec{x} = [+1, +1, -1, +1, -1, \ldots, +1]$
σ_1 σ_2 σ_N
Labels: $y = 0$ or 1
FM PM

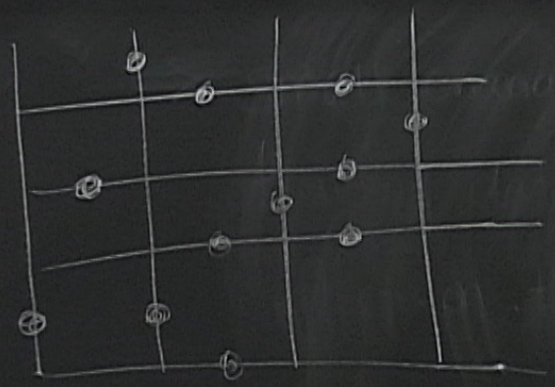
Use Monte Carlo (MC) to generate the spin configuration samples

(we will discuss MC methods tomorrow)


FM
PM
Classical Ising gauge theory
$H = -J \sum_{\substack{\text{plaquettes} \\ p}} \left(\prod_{i \in p} \sigma_i \right)$
The d.o.f. are on the bonds of the lattice
$\sigma_i = +1$ or -1

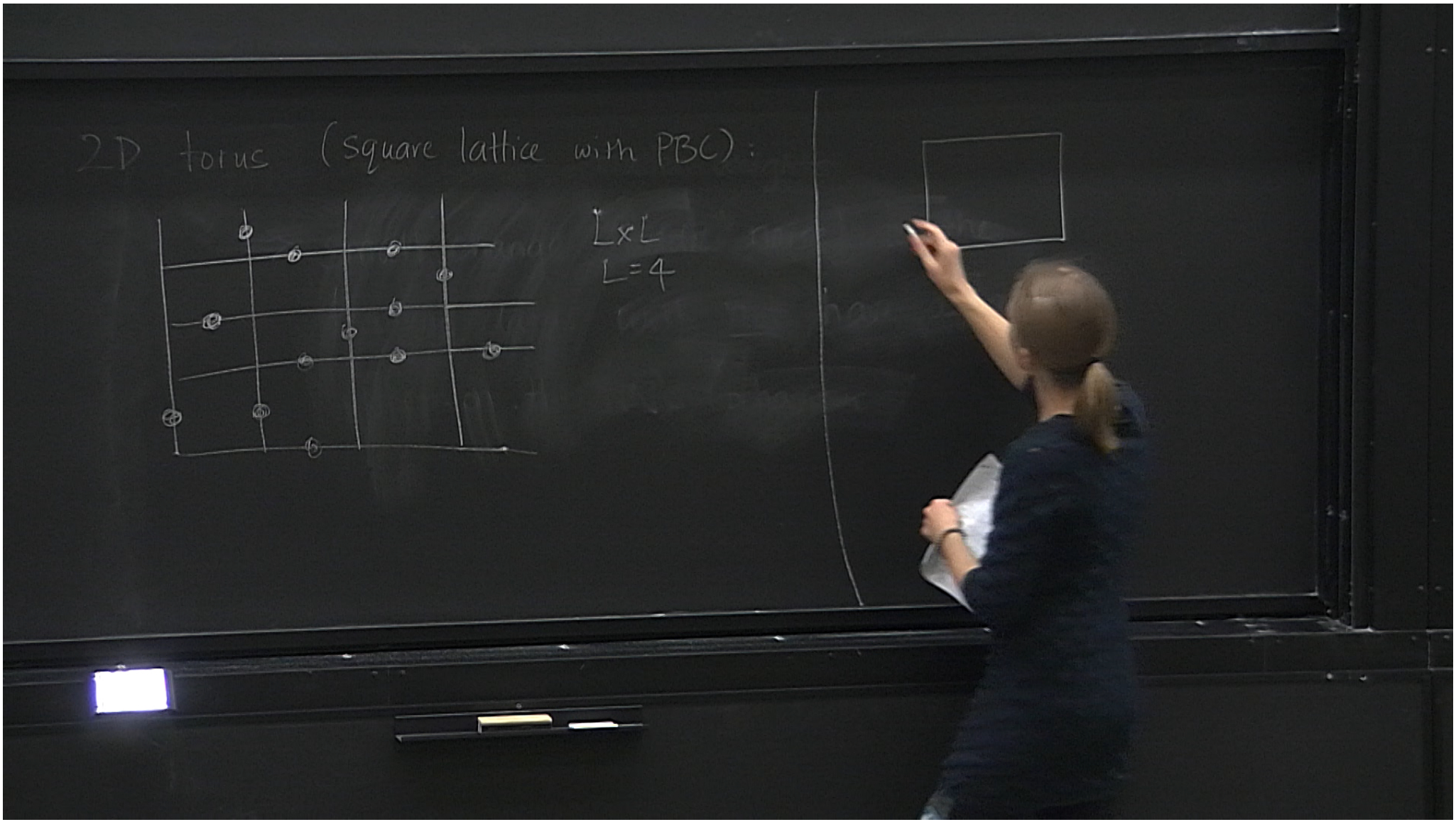

2D torus (square lattice with PBC):
L×L
L=4

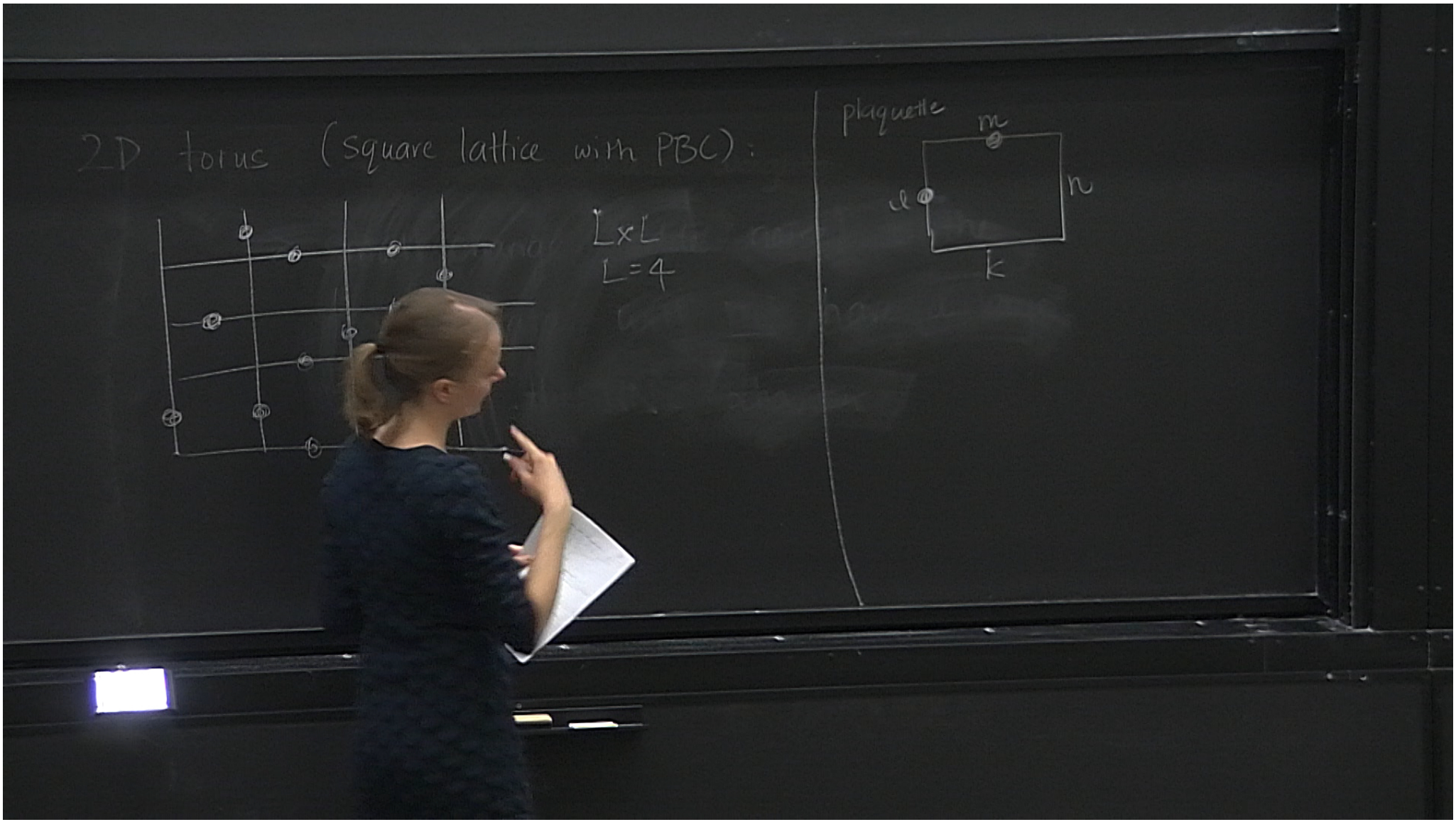
2D torus (square lattice with PBC):

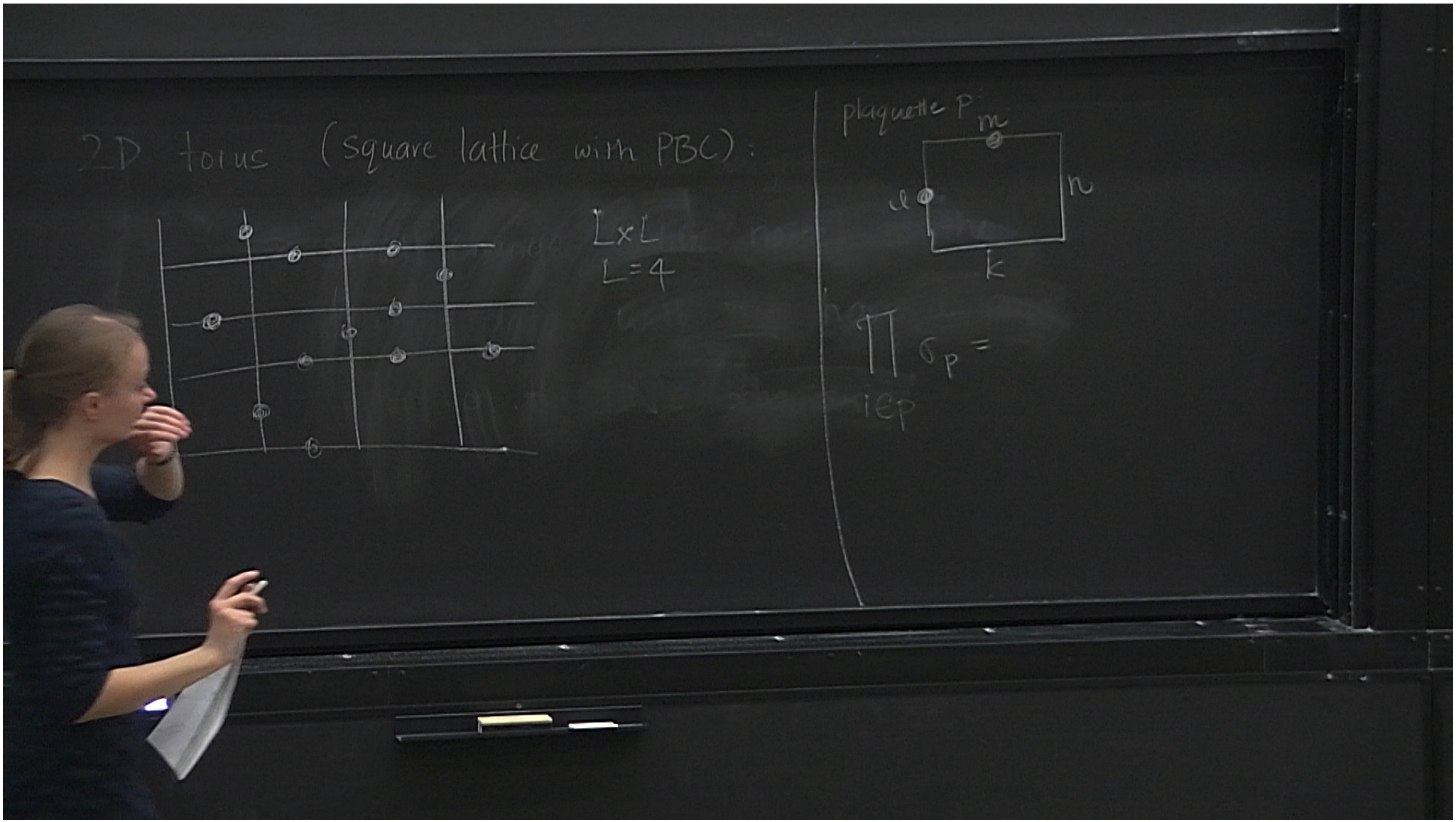


$L \times L$

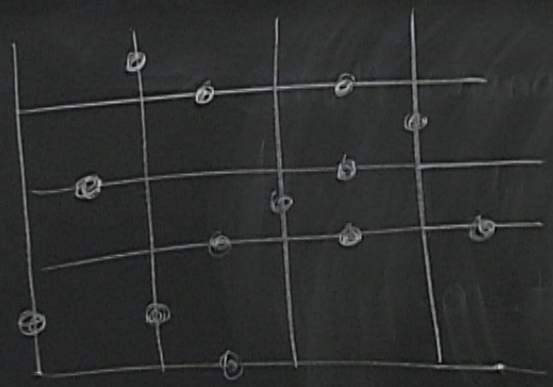
$L = 4$


2D torus (square lattice with PBC):
L×L
L=4


2D torus (square lattice with PBC):
L×L
L=4
plaquette
m
n
l
k


2D torus (square lattice with PBC):
L×L
L=4
plaquette p:
m
n
l
k
$\prod_{i \in p} \sigma_p =$

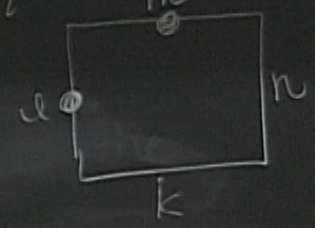
2D torus (square lattice with PBC):



$L \times L$

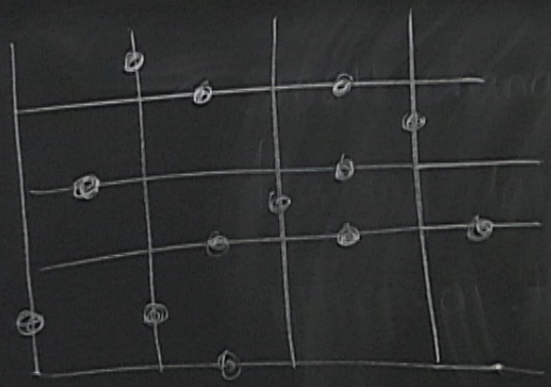
$L = 4$

plaquette P:



$$\prod_{i \in P} \sigma_i = \sigma_k \sigma_l \sigma_m \sigma_n$$

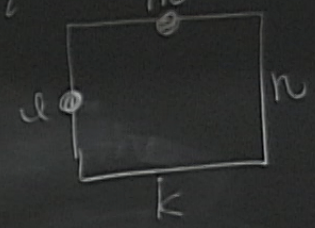
2D torus (square lattice with PBC):



$L \times L$

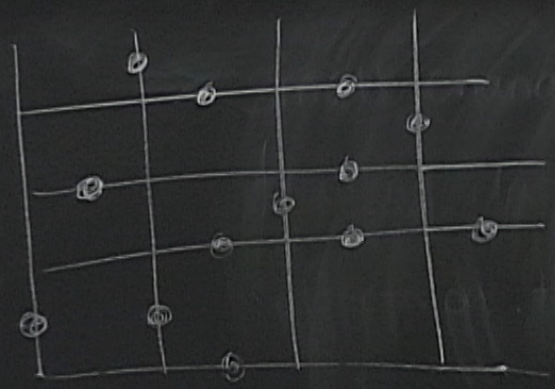
$L = 4$

plaquette P:



$$\prod_{i \in P} \sigma_i = \sigma_k \sigma_l \sigma_m \sigma_n$$

2D torus (square lattice with PBC):



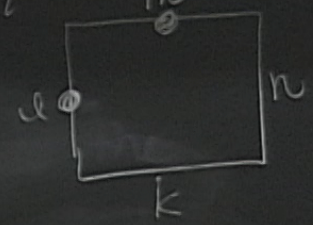
$L \times L$

$L = 4$

# of spins

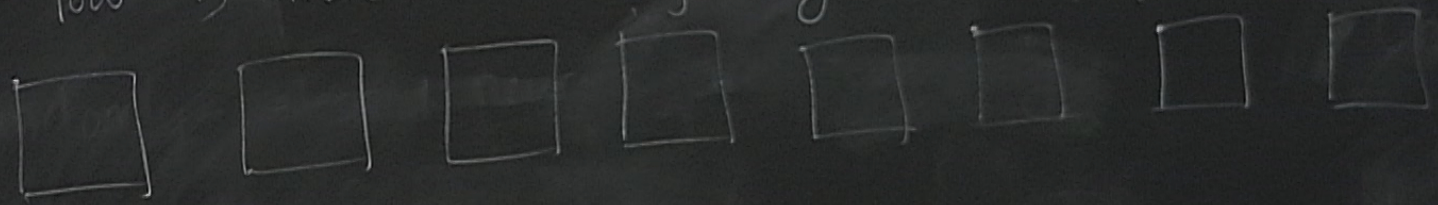
$= 2L^2$

plaquette P

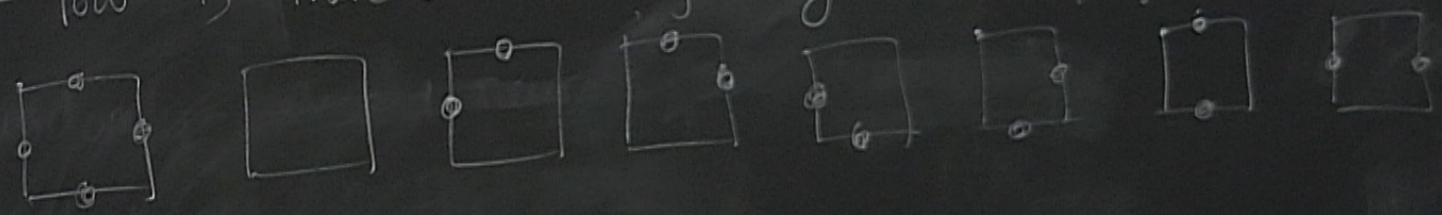


$$\prod_{i \in P} \sigma_i = \sigma_k \sigma_l \sigma_m \sigma_n$$

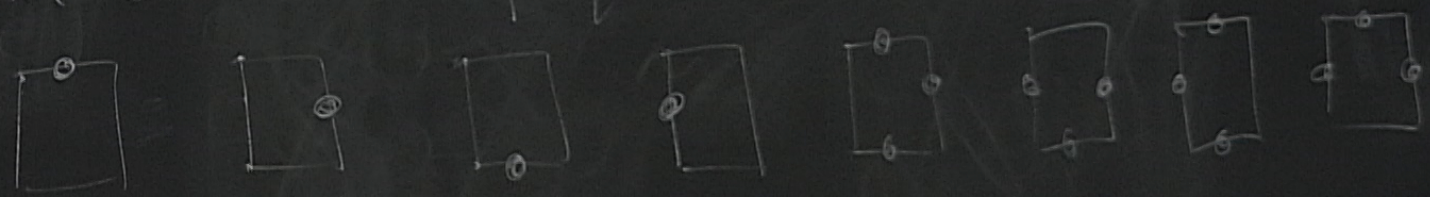
At low T, there are 8 energetically-favoured plaquettes:

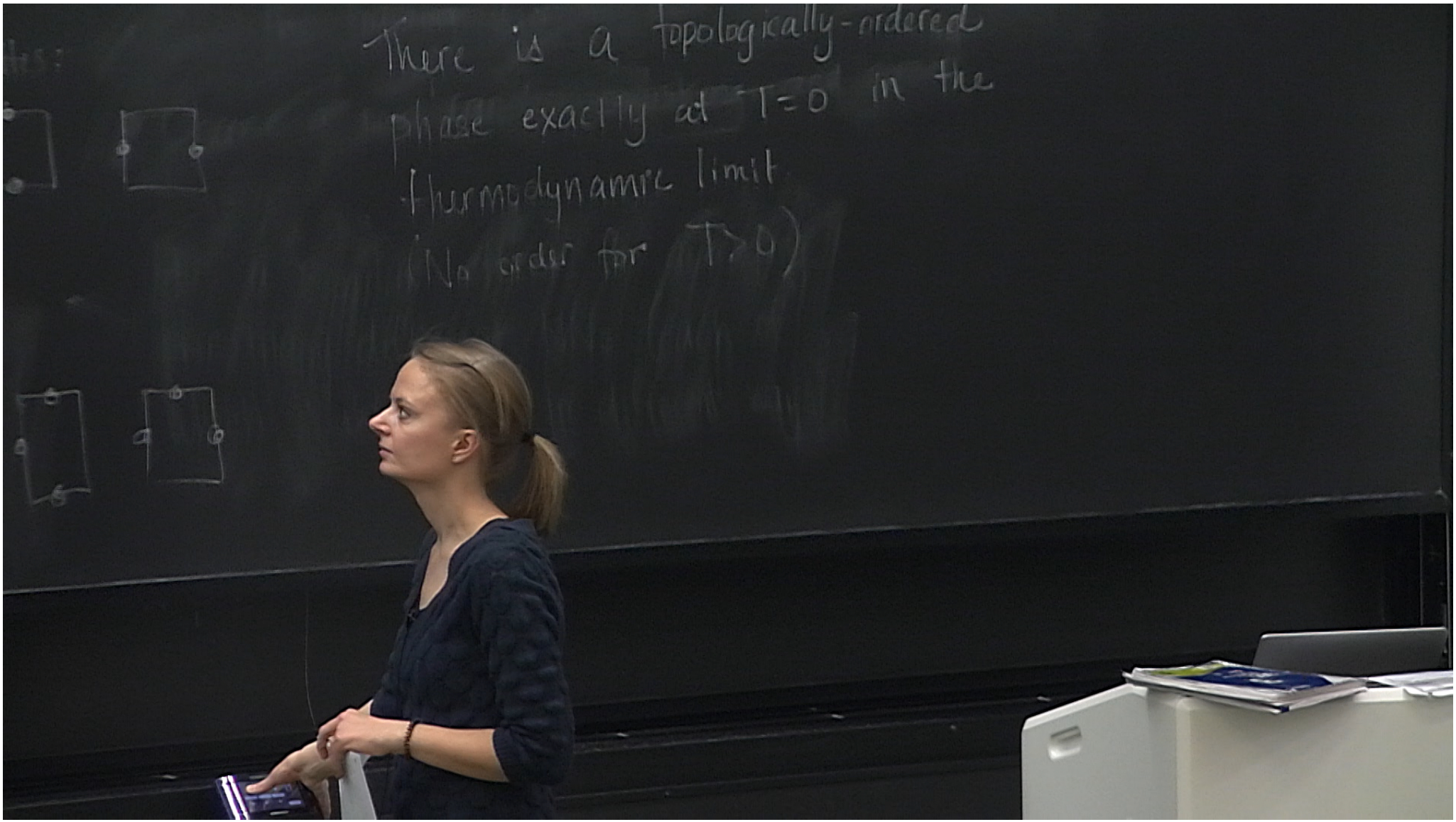


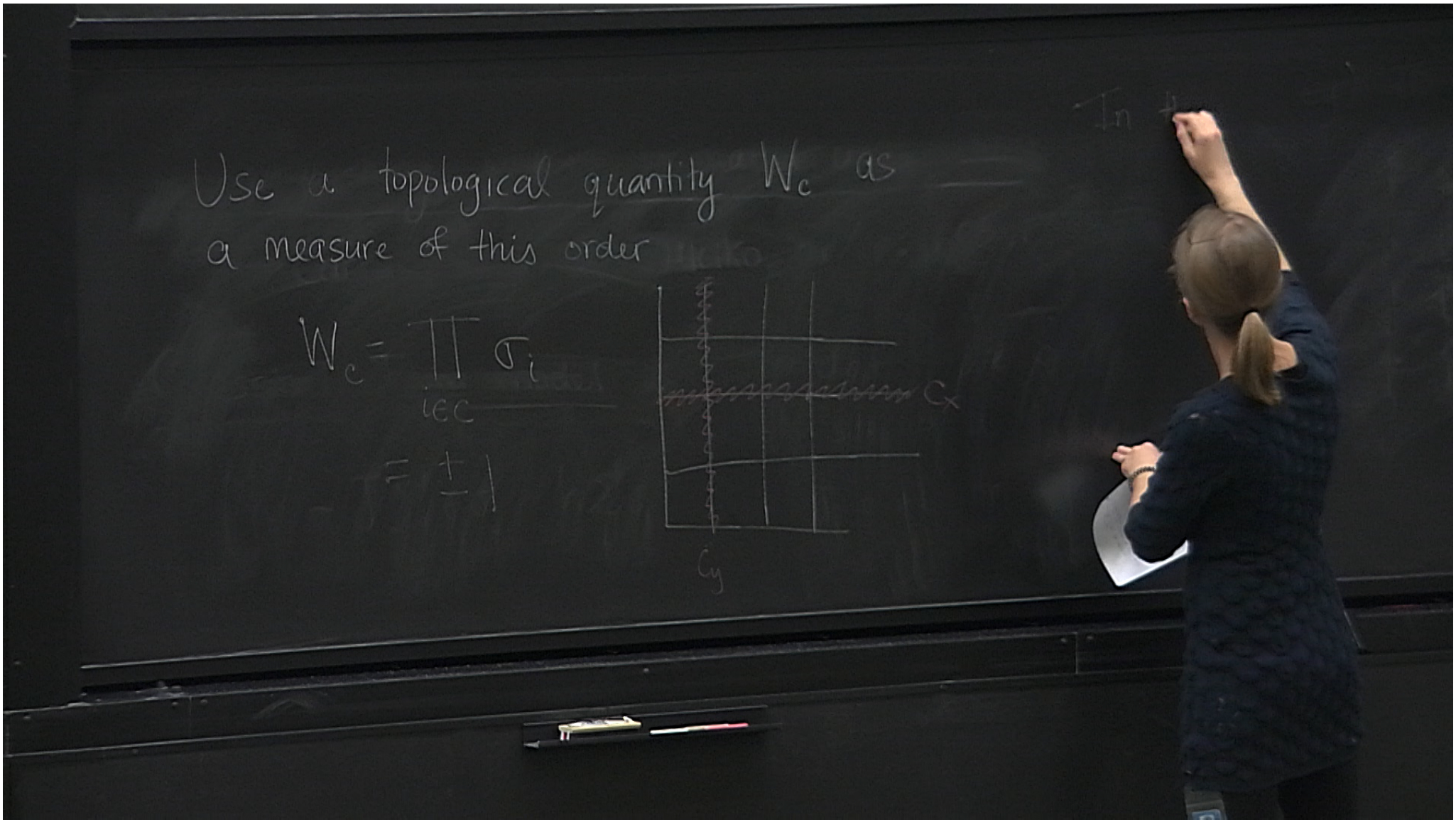
At low T, there are 8 energetically-favoured plaquettes:



And 8 "excited" plaquettes

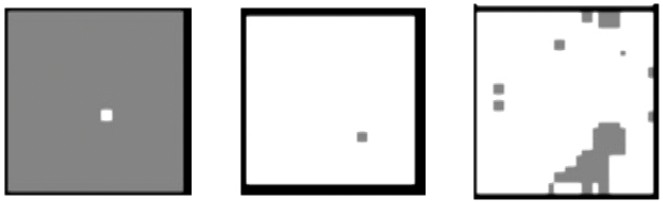



There is a topologically-ordered
phase exactly at T=0 in the
thermodynamic limit
(No order for T>0)


Use a topological quantity W_c as
a measure of this order
$W_c = \prod_{i \in C} \sigma_i$
$= \pm 1$
C_x
C_y

Ising model classification

Ferromagnetic phase:



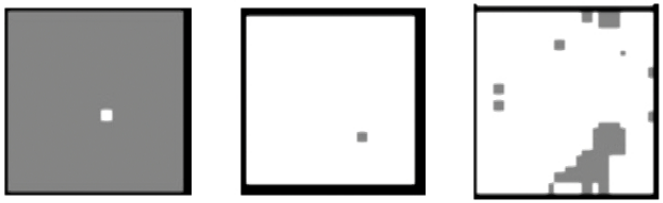
Paramagnetic phase:

Carrasquilla and Melko, arXiv:1605.01735

PERIMETER SCHOLARS INTERNATIONAL

Ising model classification

Ferromagnetic phase:



?

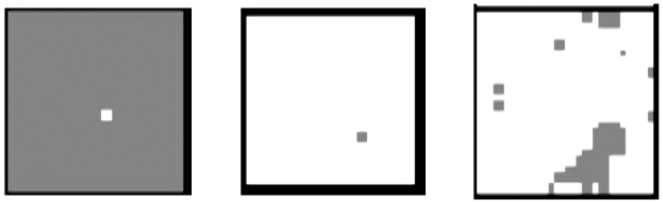
Paramagnetic phase:

Carrasquilla and Melko, arXiv:1605.01735

PERIMETER SCHOLARS INTERNATIONAL

Ising model classification

Ferromagnetic phase:



?

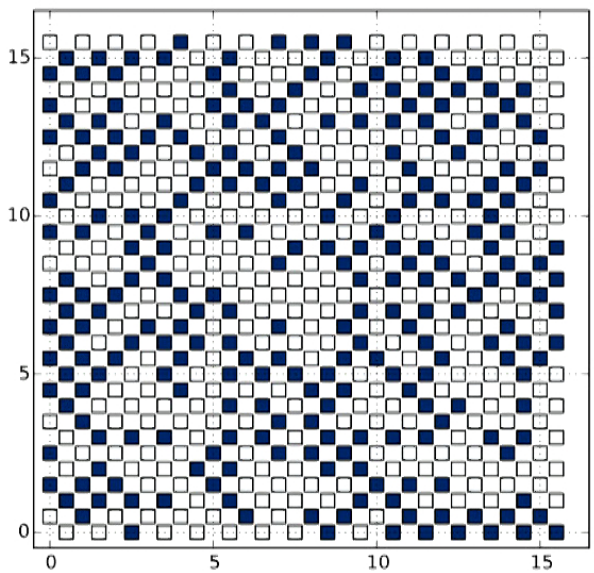
Paramagnetic phase:

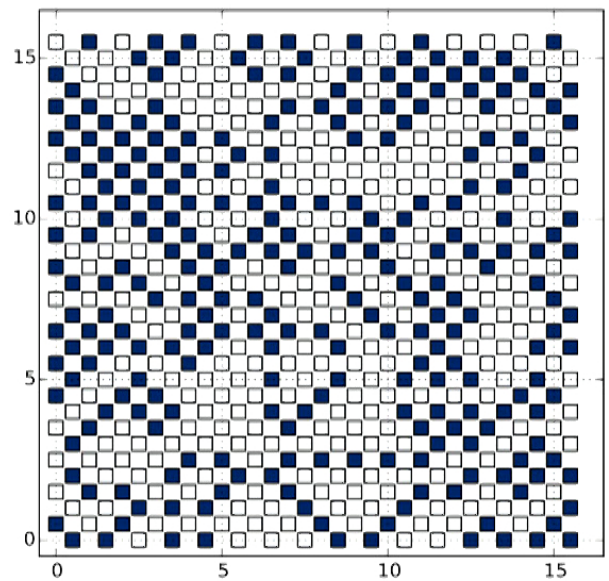
?

Carrasquilla and Melko, arXiv:1605.01735

PERIMETER SCHOLARS INTERNATIONAL

Ising gauge theory classification

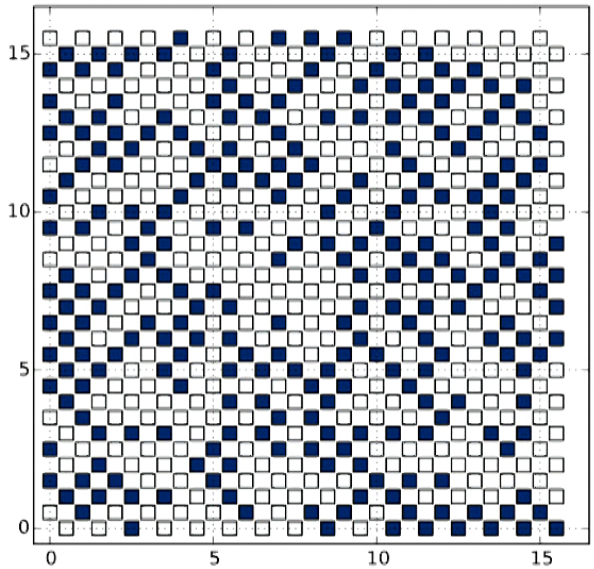




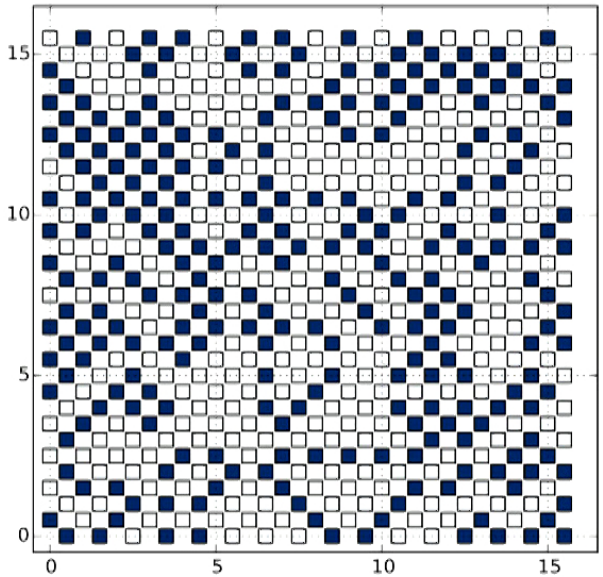
Carrasquilla and Melko, arXiv:1605.01735

PERIMETER SCHOLARS INTERNATIONAL

Ising gauge theory classification



High-temperature state



Ground state

Carrasquilla and Melko, arXiv:1605.01735

PERIMETER SCHOLARS INTERNATIONAL