Title: Machine Learning Lecture

Speakers: Mohamed Hibat Allah

Collection: Machine Learning 2023/24

Date: April 16, 2024 - 11:30 AM

URL: https://pirsa.org/24040050

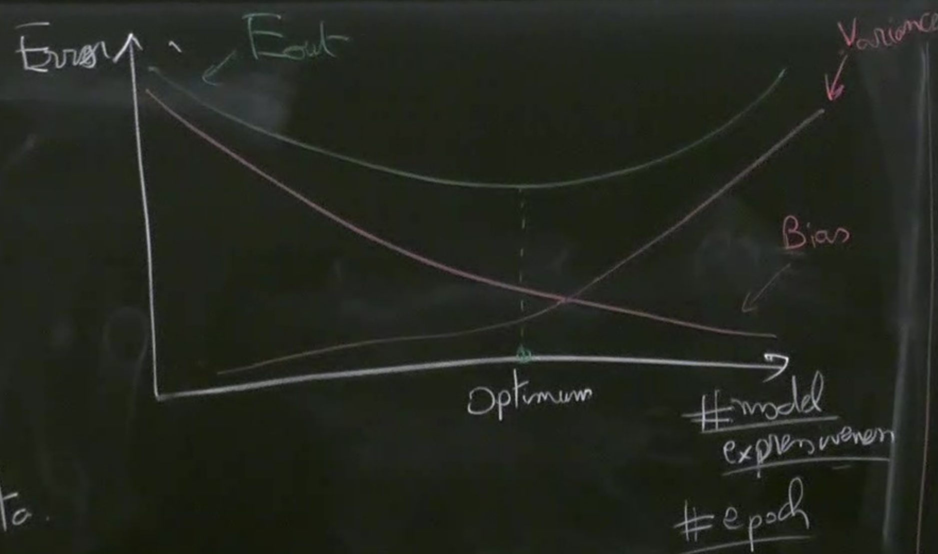
Lecture 6

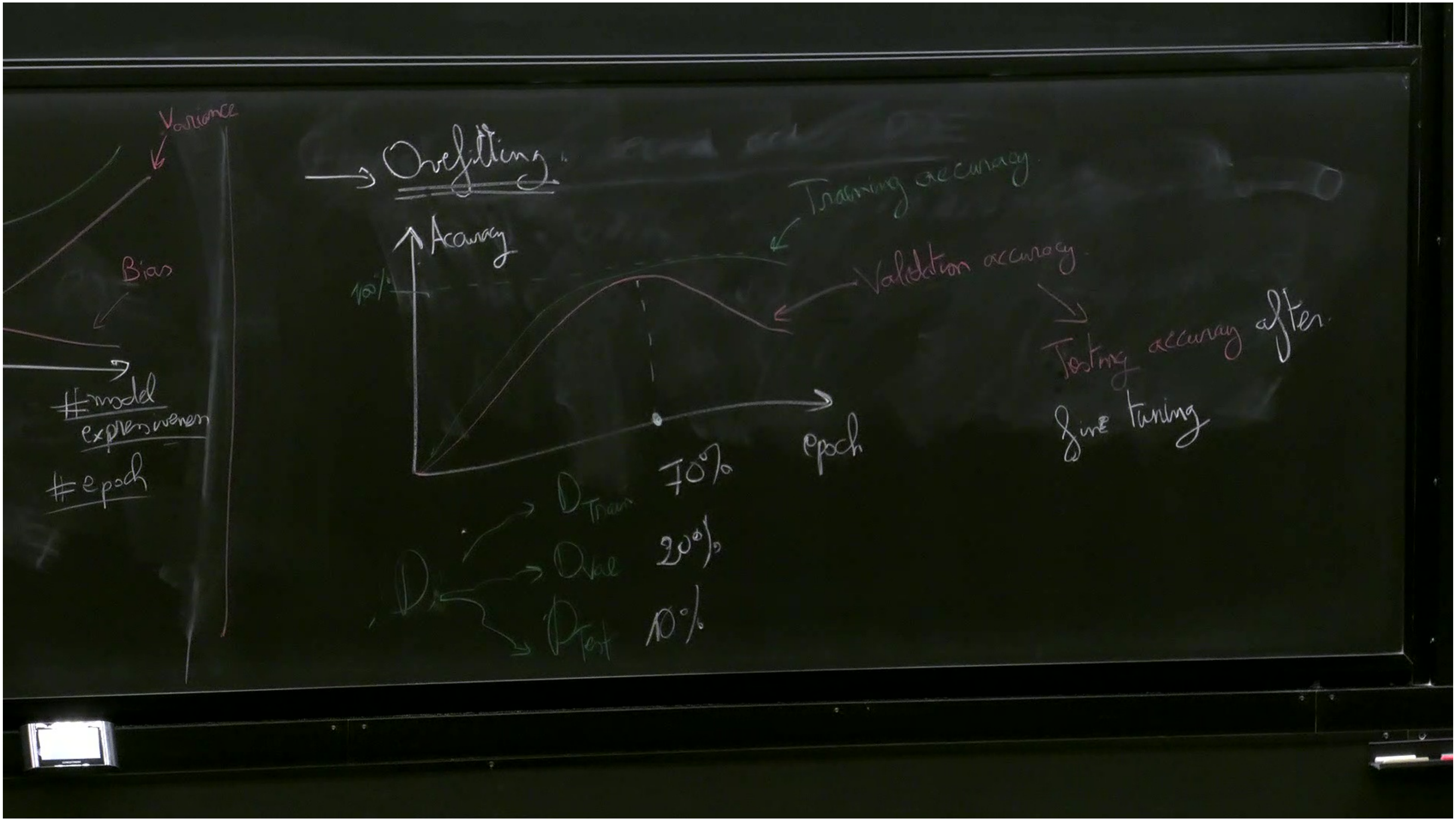
Last time:

↳ Bias-Variance tradeoff.

Bias: "Inability of the model to fit a data.

Variance: "Dependence of fit on the choice of the training data.





Variance
Bias
model expressiveness
epoch
Overfitting.
Accuracy
100%
Training accuracy
Validation accuracy
Testing accuracy after fine tuning
epoch
D
D_Train 70%
D_Val 20%
D_Test 10%

↳ Curing overfitting

↳ L1 + L2 regularization. MSE + L1/2

↳ Dropout

↳ Bagging + Ensemble methods.

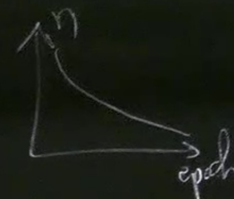
Outline for today

↳ Recap of Hyperparams for NN training.

↳ Learning phases of matter

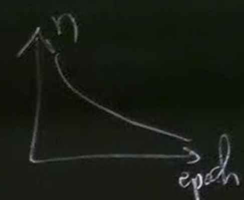
↳ Monte Carlo sampling.

Recap of Hyperparameters for NN training:



- ↳ Learning rate η , Adaptive learning rate
- ↳ # epochs
- ↳ # of neurons, # of layers
- ↳ Different GD schemes (plain GD, SGD, $\overbrace{\text{Mini-batch GD}}^{|B| \text{ batch size}}$, Momentum GD, Adam)

Recap of Hyperparameters for NN training:



- ↳ Learning rate η, Adaptive learning rate
- ↳ # epochs
- ↳ # of neurons, # of layers
- ↳ Different GD schemes (Plain GD, SGD, $\underbrace{\text{Mini-batch GD}}_{}$ |B| batch size, Momentum GD, Adam)
- ↳ Regularization "λ" in L1 + L2 + P_{drop} in dropout

→ Activation functions (Sigmoid, Tanh, ReLU,..)

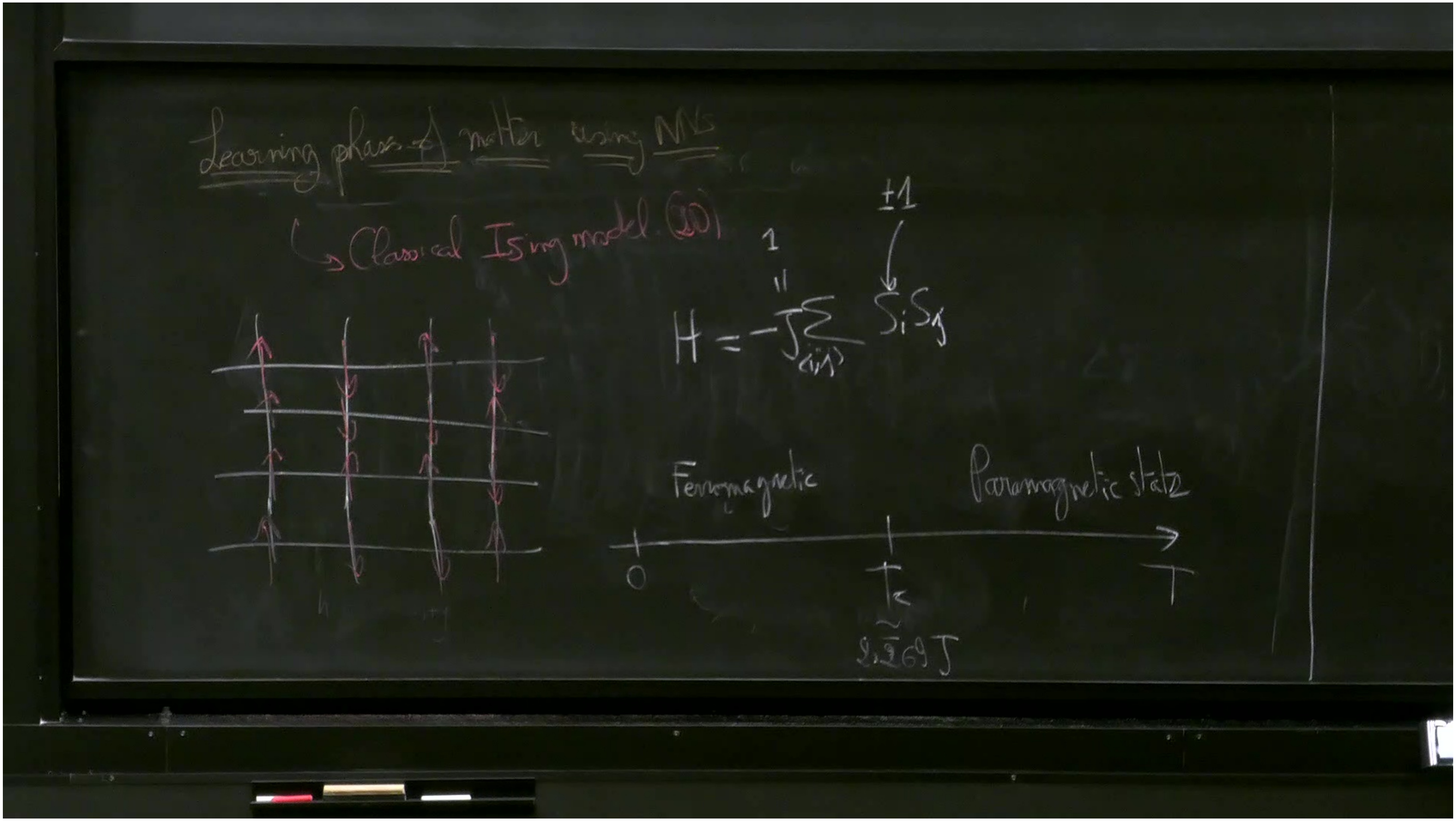
→ % of D_{Train}, D_{Val}, D_{Test}

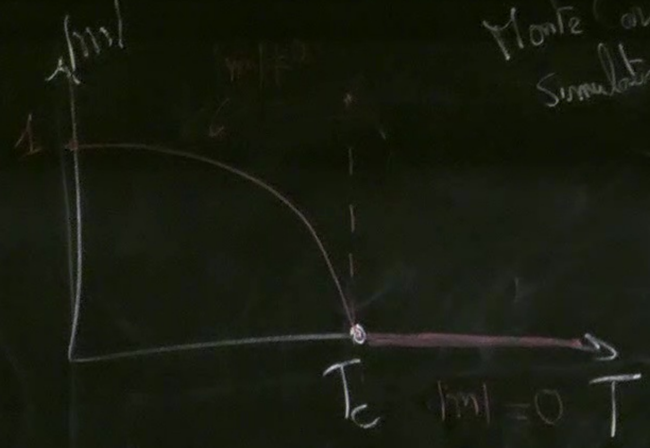
→ Data preprocessing

→ Initilization of parameters of NNs

→ Choice of cost function

→


Learning phases of matter using NNs
Classical Ising model (2D)
±1
1
H = −J Σ⟨i,j⟩ S_i S_j
Ferromagnetic
Paramagnetic state
0
T_c
T
2.269 J



Monte Carlo Simulation $\longleftarrow$ $D = \{(\vec{x}, \vec{y})\}$

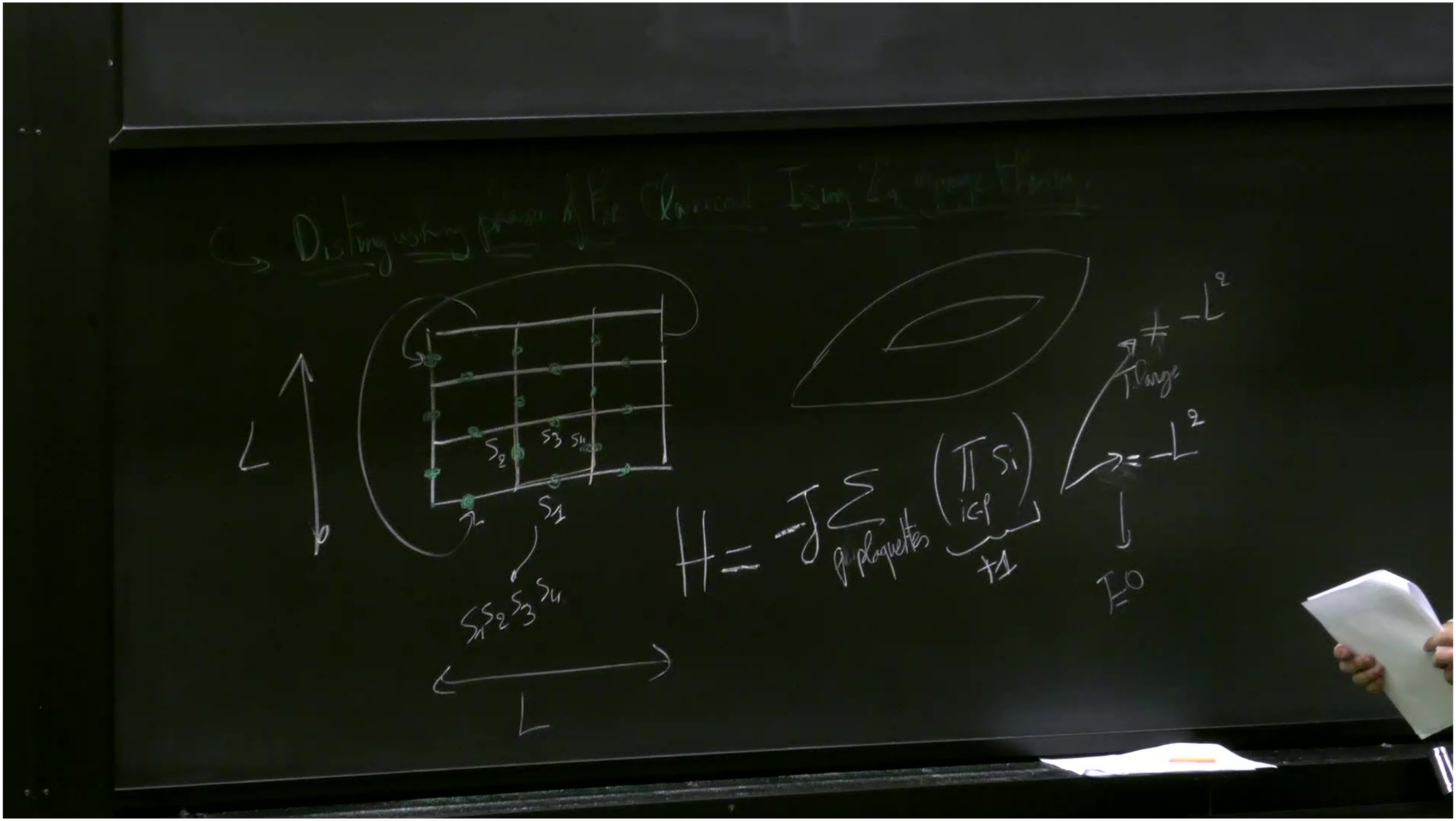
$$\vec{x} = [S_1, S_2, \dots\dots, S_{L^2}]$$

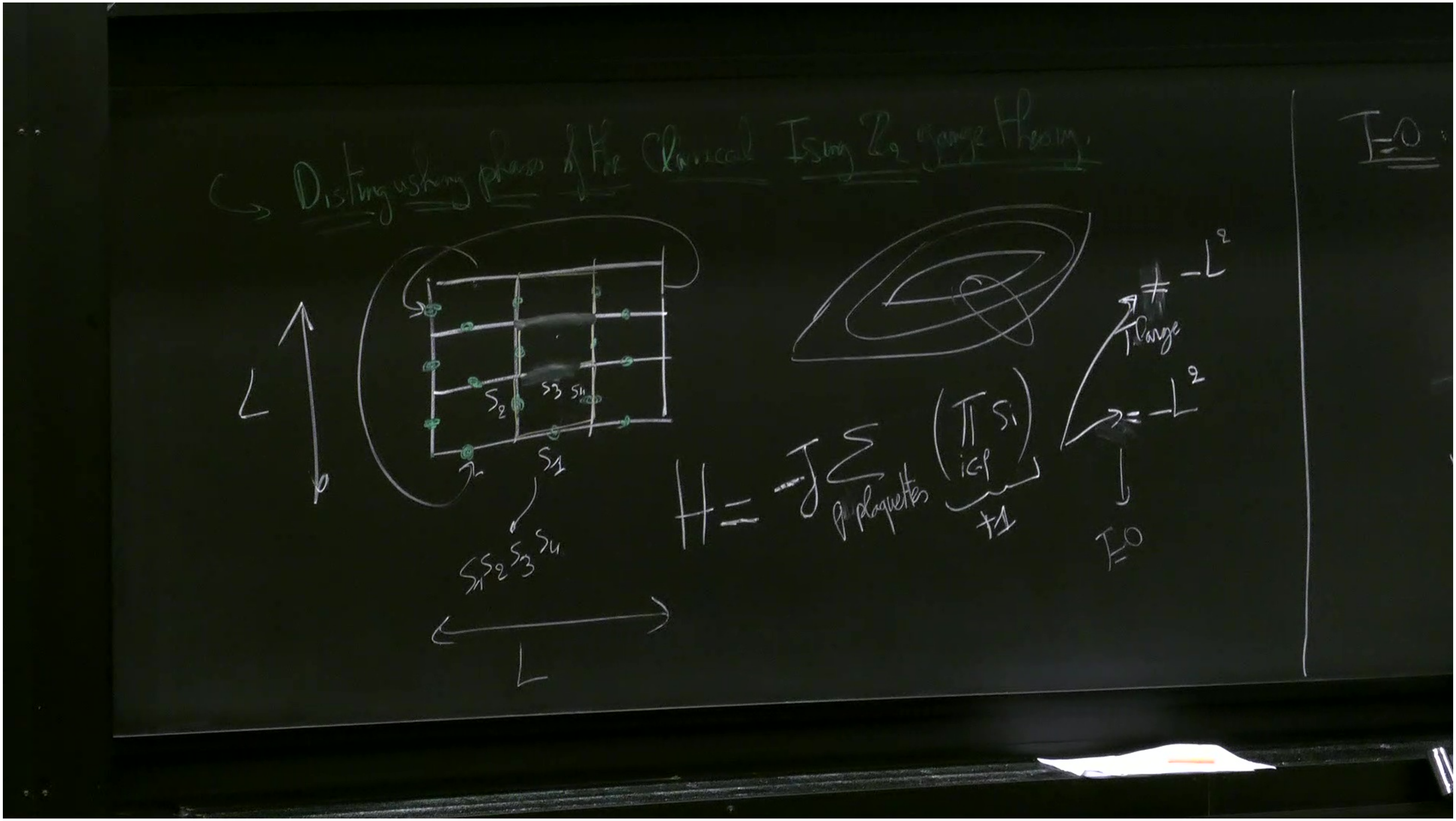
± 1

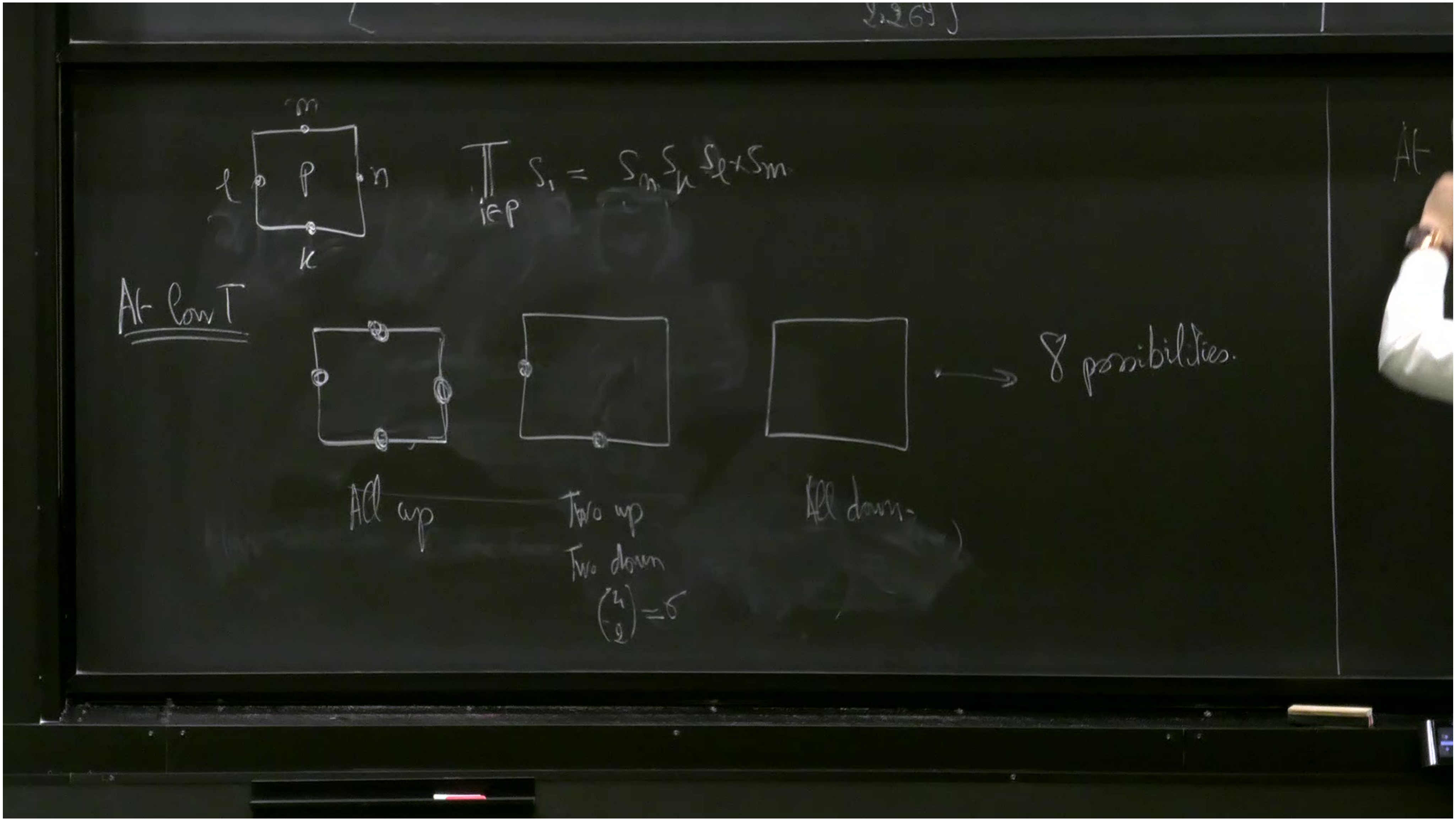
$$\vec{y} \nearrow 0 \quad (FM)$$
$$\vec{y} \searrow 1 \quad (PM)$$

$$m = \frac{\sum_i s_i}{N}$$

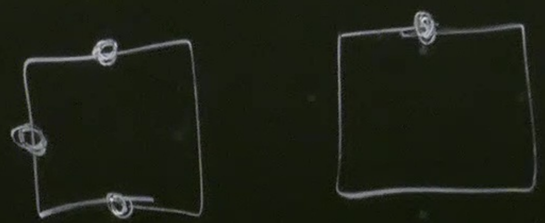
Carrasquilla, Melko.
« Machine learning phases of Matter », Nature 2017


L
s_1 s_2 s_3 s_4
$s_1 s_2 s_3 s_4$
$H = -J \sum_{\text{plaquettes}} \underbrace{\left(\prod_{i \in p} s_i\right)}_{\pm 1}$
$-L^2$
T large
T=0


Distinguishing phases of the Classical Ising Z2 gauge theory.
L
L
S1
s1 s2 s3 s4
H = -J Σ plaquettes (Π i∈p si)
±1
-L^2
T large
-L^2
T=0
T=0


m
l
p
n
k
$\prod_{i \in p} S_i = S_n S_k S_l \times S_m$
At low T
All up
Two up
Two down
$\binom{4}{2} = 6$
All down
8 possibilities.

At high temperature:



4 possibilities   4 possibilities

8 possibilities

L

Monte Carlo (MC) sampling:

$\beta = \frac{1}{T}$

$$P_\beta(\vec{x}) = \frac{e^{-\beta E(x)}}{Z} \longrightarrow \text{Approximate sampling}$$

$$\langle m \rangle = \sum_{x} P(\vec{x})\, m(\vec{x})$$

$$\simeq \frac{1}{M} \sum_{i=1}^{M} m(\vec{x}^{(i)})$$

$$Z = \sum_{x} e^{-\beta E}$$

$\#$ of possibilities $= 2^N$

$$\{x^{(i)}\}_{i=1}^{M}$$

L

→ Monte Carlo (MC) sampling:

$\beta = \frac{1}{T}$

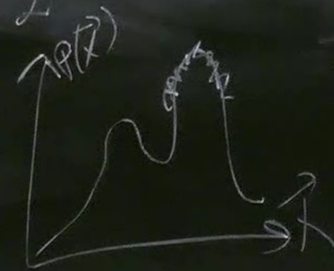
$$P_\beta(\vec{x}) = \frac{e^{-\beta E(x)}}{Z} \longrightarrow \text{Approximate sampling}$$

$$\langle m \rangle = \sum_{\vec{x}} P(\vec{x})\, m(\vec{x})$$

$$\simeq \frac{1}{M} \sum_{i=1}^{M} m(\vec{x}^{(i)})$$

$$Z = \sum_{\vec{x}} e^{-\beta E}$$

# of possibilities $= 2^N$



$\{\vec{x}^{(i)}\}_{i=1}^{M} \longrightarrow$ System at temperature T