Title: LECTURE: Generative Modelling

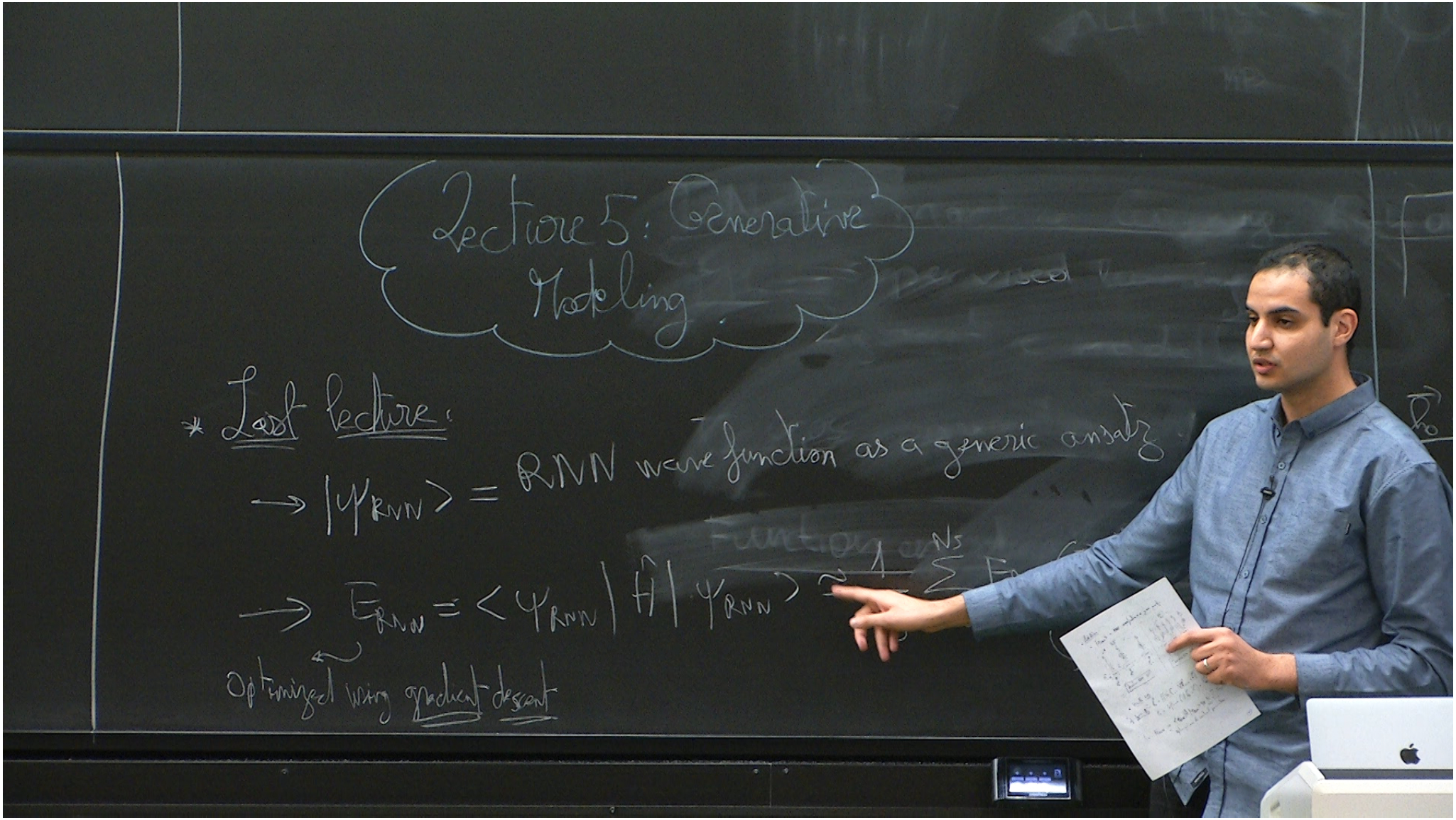
Speakers: Mohamed Hibat Allah

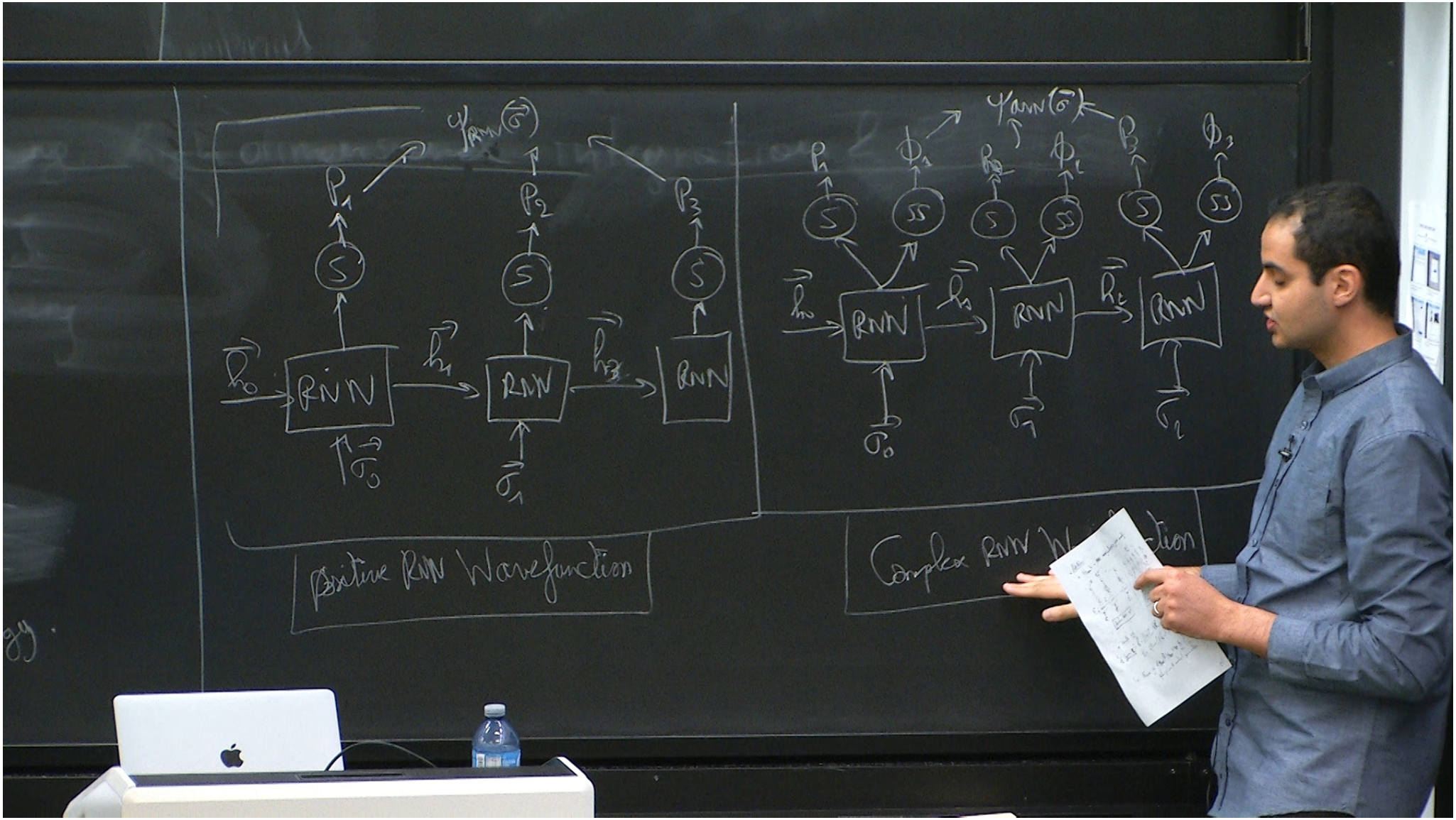
Collection: Quantum and AI Career Trajectories Mini-Course: Computational Methods and their Applications

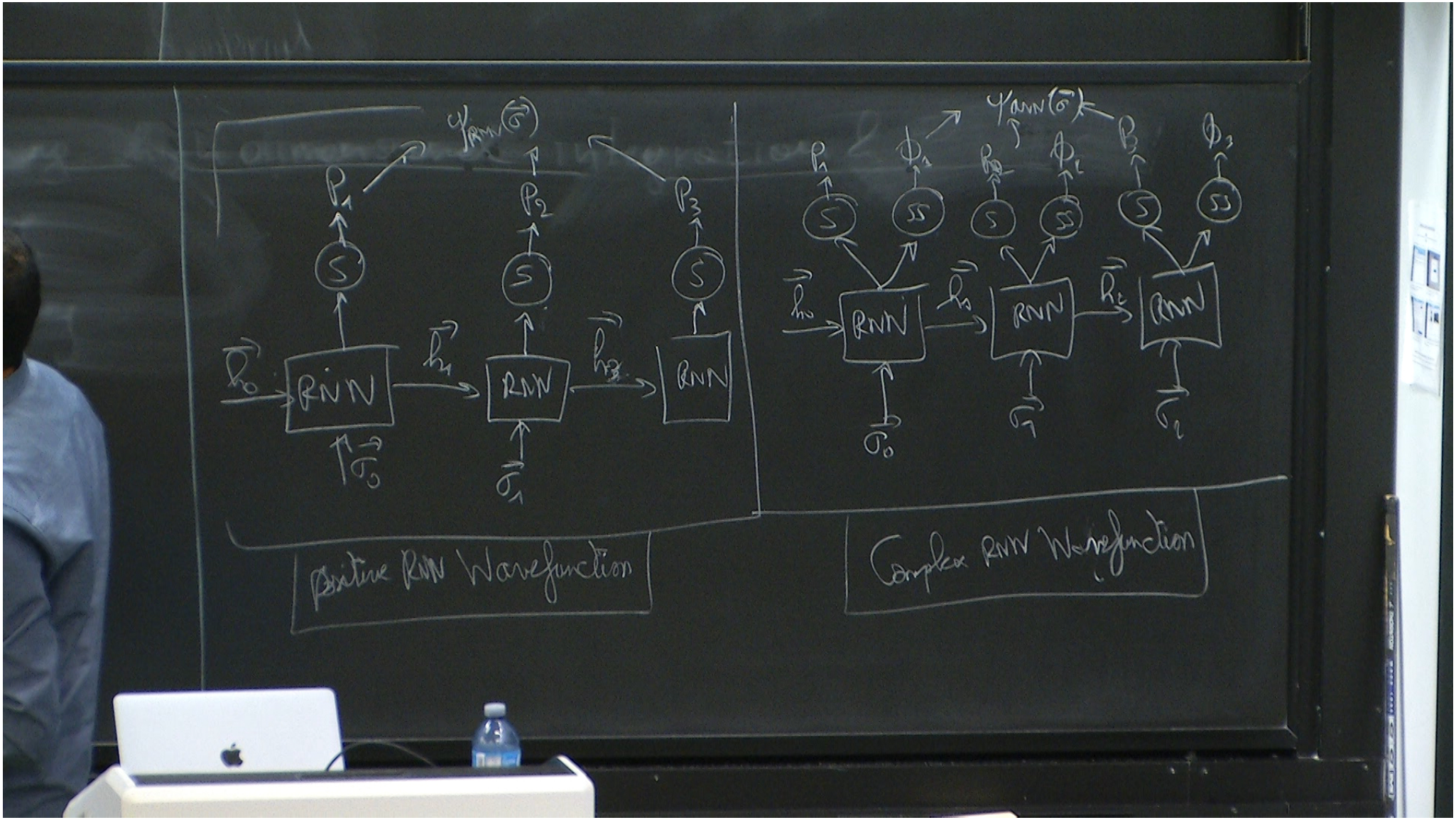
Date: May 12, 2023 - 11:00 AM

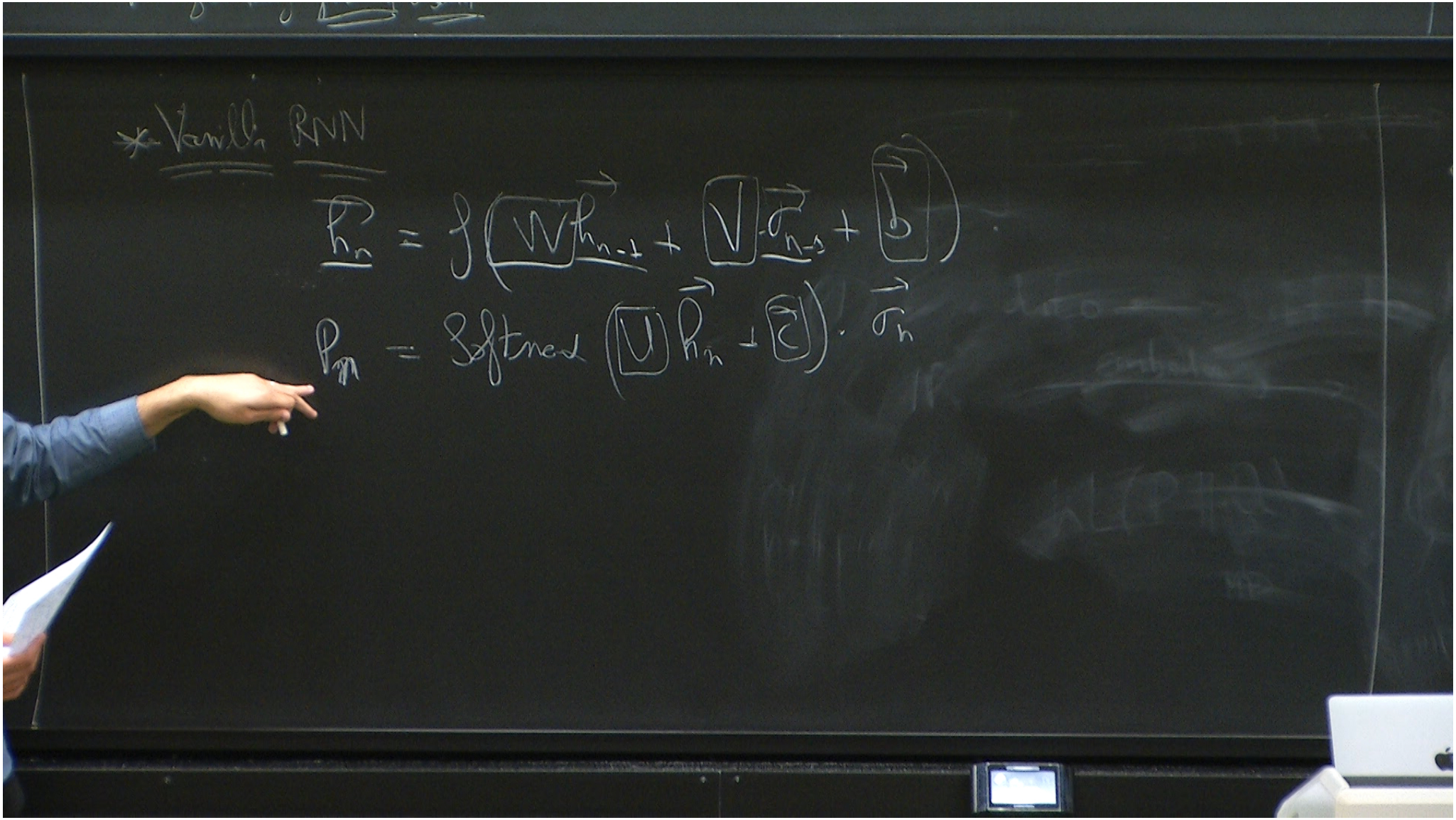
URL: https://pirsa.org/23050098

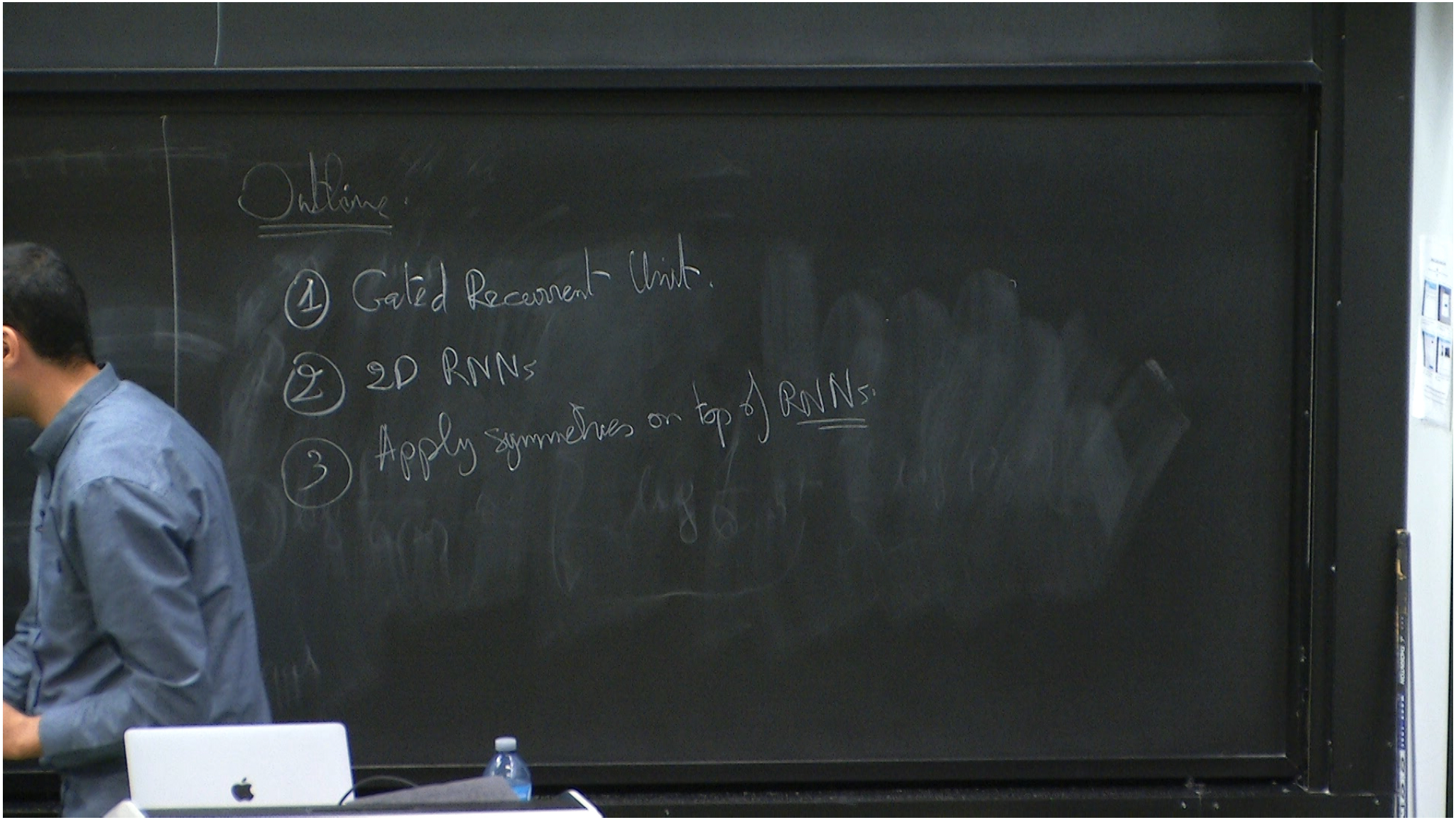
Abstract: TBC


Lecture 5: Generative Modeling
* Last lecture:
→ |ψ_RNN⟩ = RNN wave function as a generic ansatz
→ E_RNN = ⟨ψ_RNN| Ĥ |ψ_RNN⟩ ≈ 1/N_s Σ
Optimized using gradient descent


RNN
RNN
RNN
Positive RNN Wavefunction
RNN
RNN
RNN


RNN
RNN
RNN
S
S
S
Positive RNN Wavefunction
RNN
RNN
RNN
S
SS
Complex RNN Wavefunction


* Vanilla RNN
$\vec{h}_n = f\left(W\vec{h}_{n-1} + V\vec{\sigma}_{n-1} + \vec{b}\right)$
$p_m = \text{Softmax}\left(U\vec{h}_n + \vec{c}\right) \cdot \vec{\sigma}_n$


Outline:
1 Gated Recurrent Unit.
2 2D RNNs
3 Apply symmetries on top of RNNs.

(1) Gated Recurrent Units (GRU):

The city, which located in Ontario, and which very populated, is beautiful.

$$h_n = w h_{n-1} + v\sigma_{n-1} = w\left(w h_{n-2} + v\sigma_{n-2}\right) + v\sigma_{n-1}$$
$$\vdots$$
$$= w^{n-1} h_1 + w^{n-2} v\sigma_1 + \dots$$

(1) Gated Recurrent Units (GRU):

The city, which located in Ontario, and which very populated, is beautiful.

$$h_n = w h_{n-1} + v\sigma_{n-1} = w(w h_{n-2} + v\sigma_{n-2}) + v\sigma_{n-1}$$

$$\vdots$$

$$= \underbrace{w^{n-1}}\, h_1 + \underbrace{w^{n-2}}\, v\sigma_1 + \dots$$

$$|w| < 1$$

(1) Gated Recurrent Units (GRU):

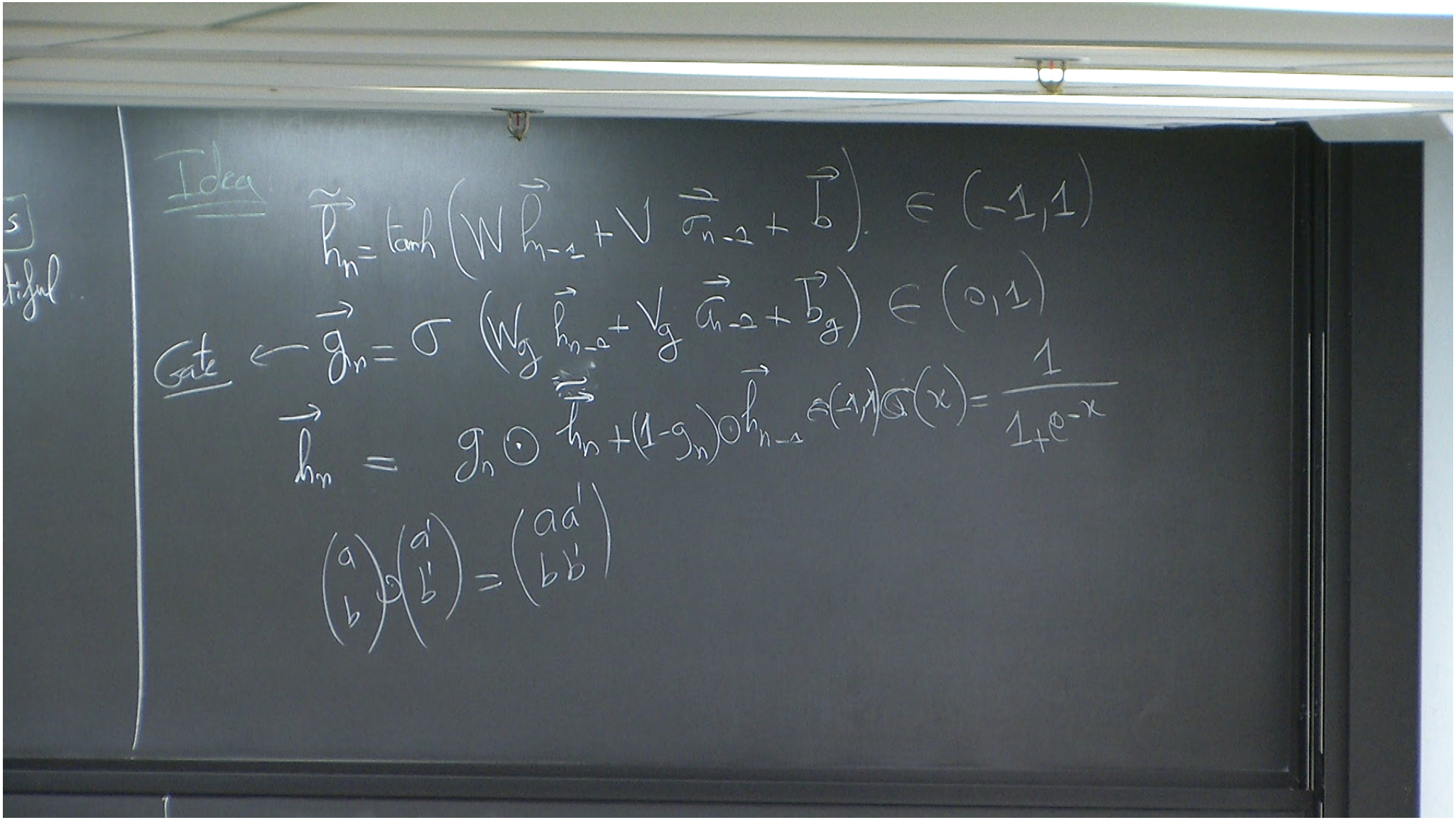
The city, which located in Ontario, and which very populated, is beautiful.

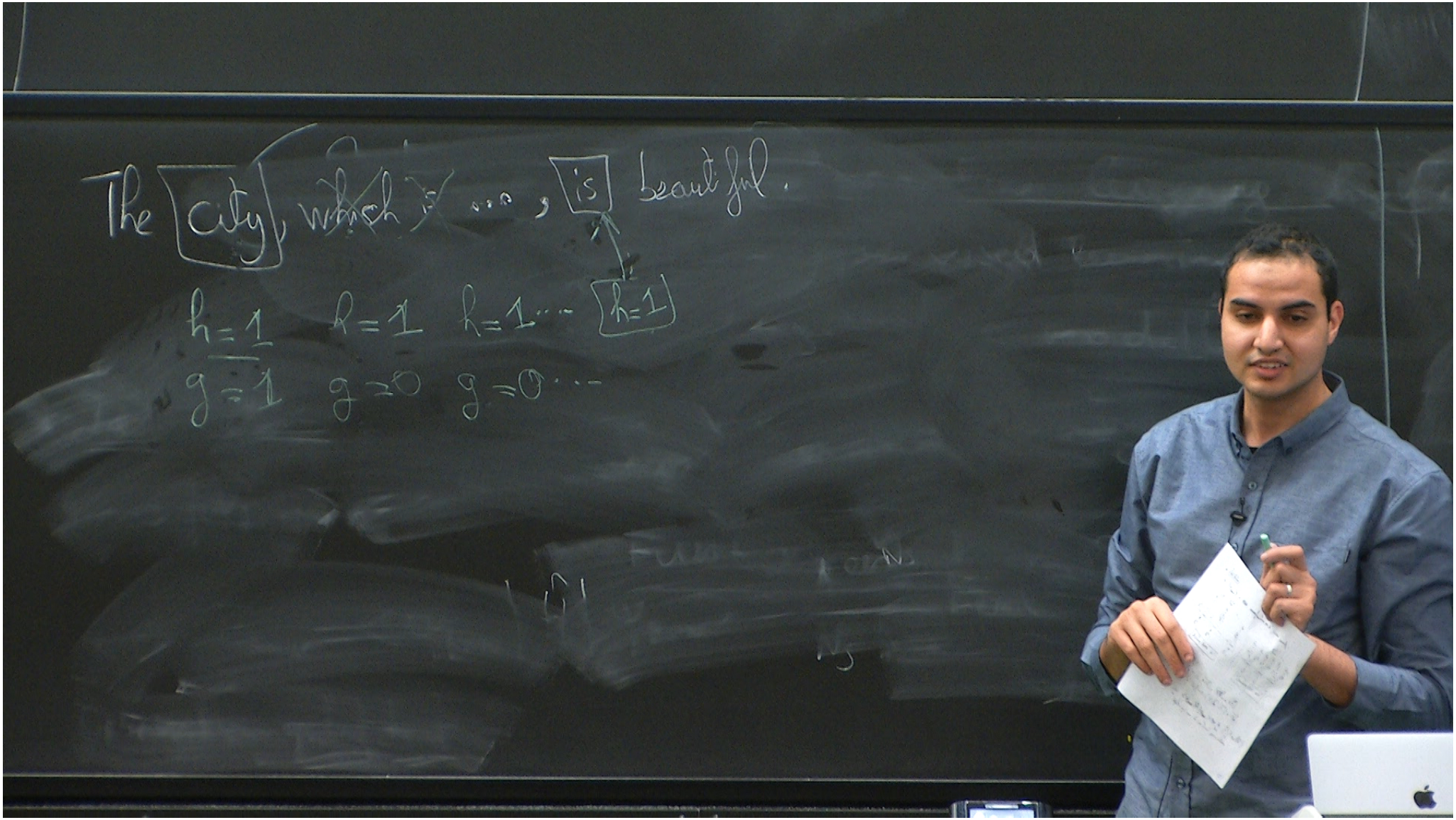
$$h_n = w h_{n-1} + v\sigma_{n-1} = w(w h_{n-2} + v\sigma_{n-2}) + v\sigma_{n-1}$$

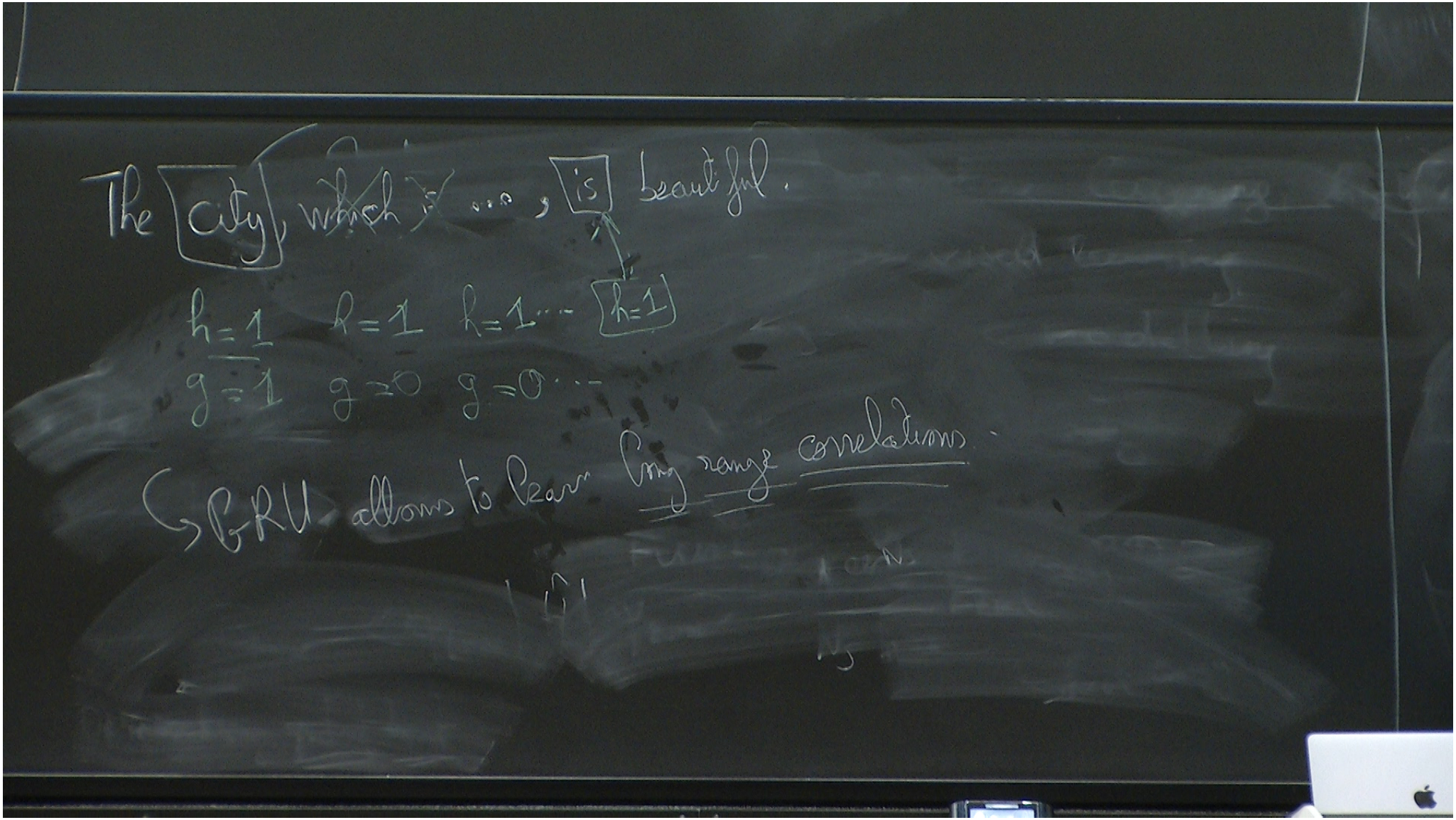
$$\vdots$$

$$= \underbrace{w^{n-1}}\, h_1 + \underbrace{w^{n-2}}\, v\sigma_1 + \dots$$

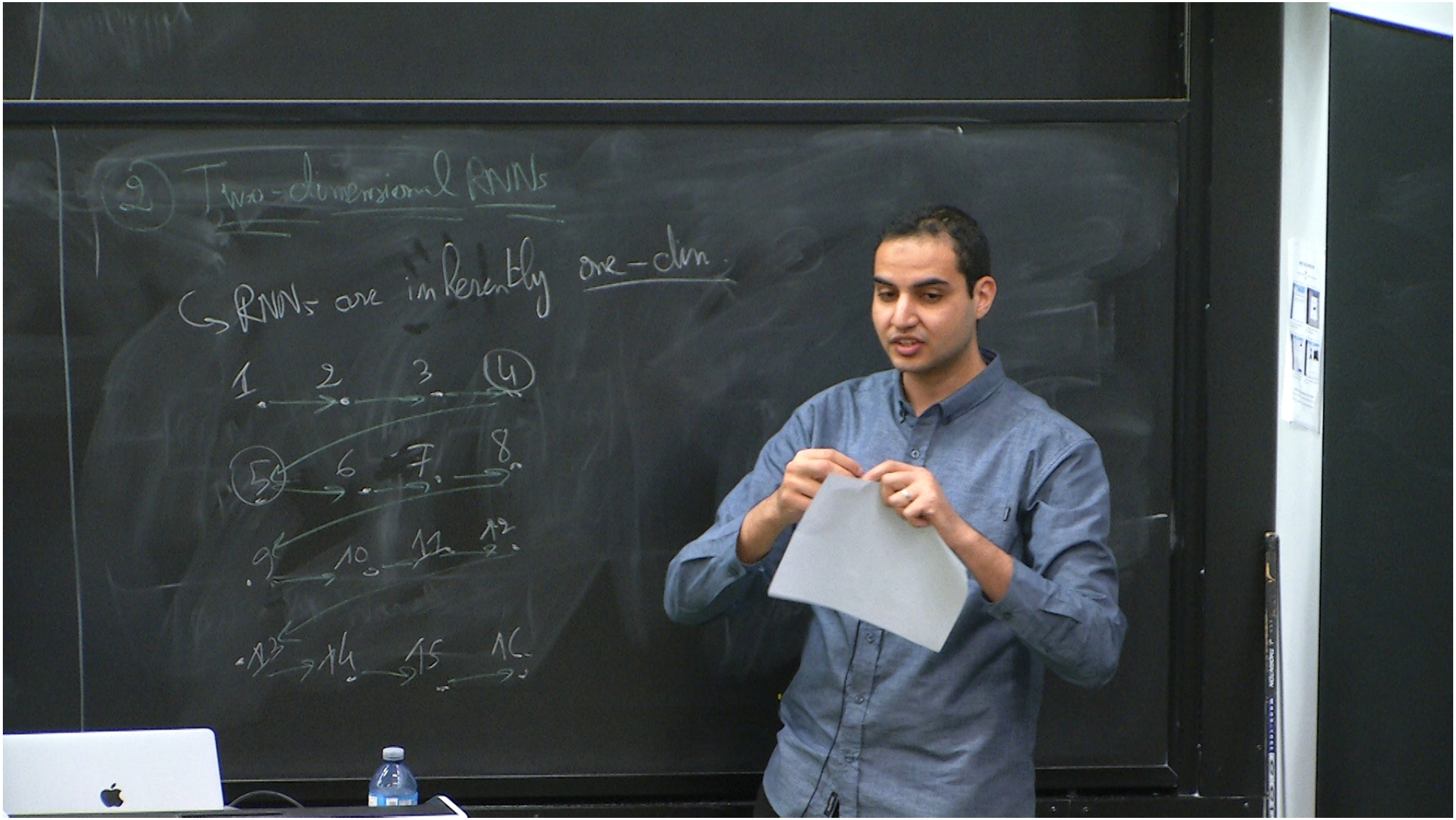
$|w| < 1 \longrightarrow$ Exponential decay


Idea
$\tilde{\vec{h}}_n = \tanh\left(W \vec{h}_{n-1} + V \vec{a}_{n-2} + \vec{b}\right) \in (-1,1)$
Gate
$\vec{g}_n = \sigma\left(W_g \vec{h}_{n-2} + V_g \vec{a}_{n-2} + \vec{b}_g\right) \in (0,1)$
$\vec{h}_n = g_n \odot \tilde{\vec{h}}_n + (1-g_n) \odot \vec{h}_{n-1}$
$\sigma(x) = \frac{1}{1+e^{-x}}$
$\begin{pmatrix} a \\ b \end{pmatrix} \odot \begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} aa' \\ bb' \end{pmatrix}$

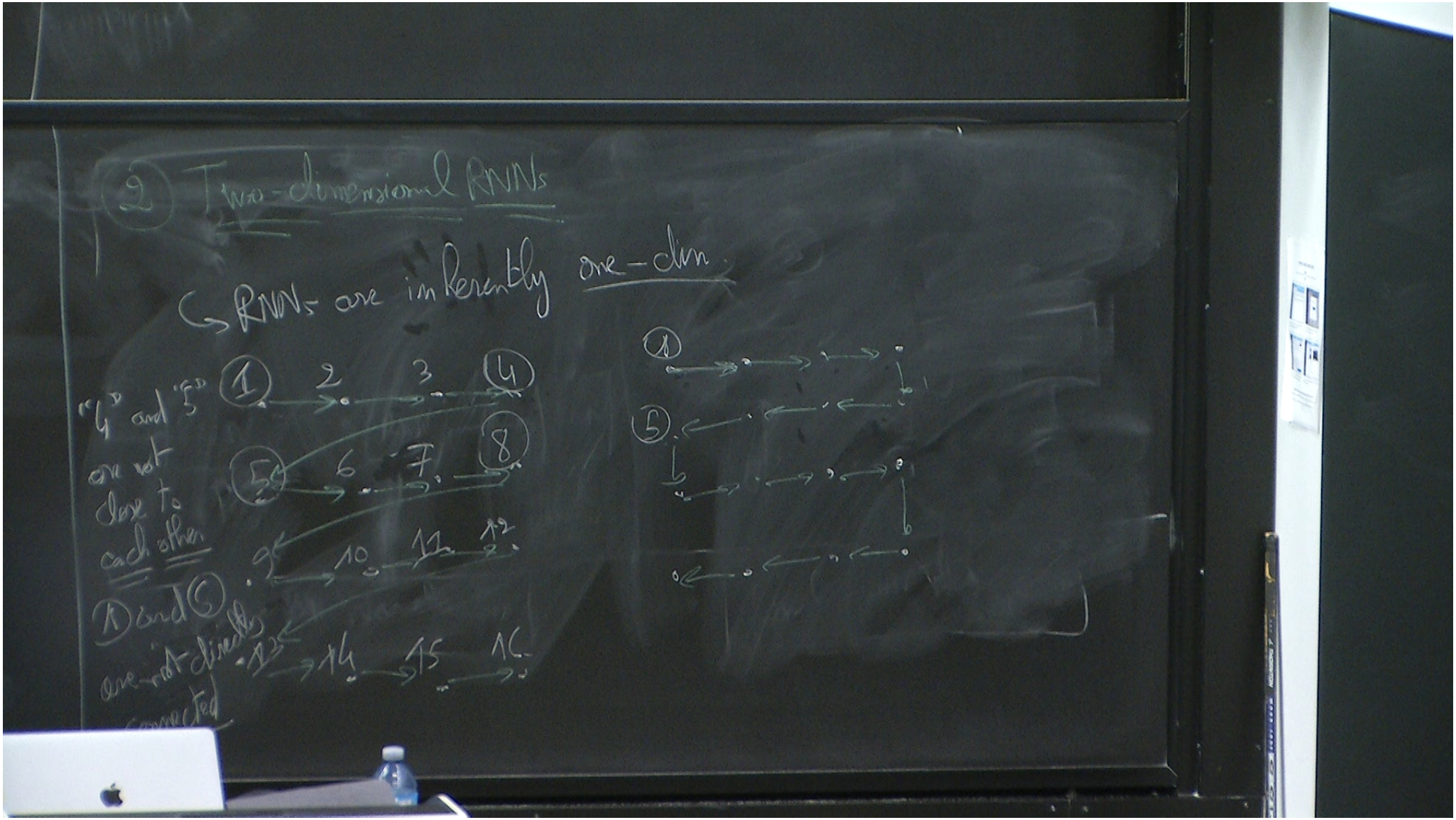

The city, which ... , is beautiful.
h=1 h=1 h=1 ... h=1
g=1 g=0 g=0 ...

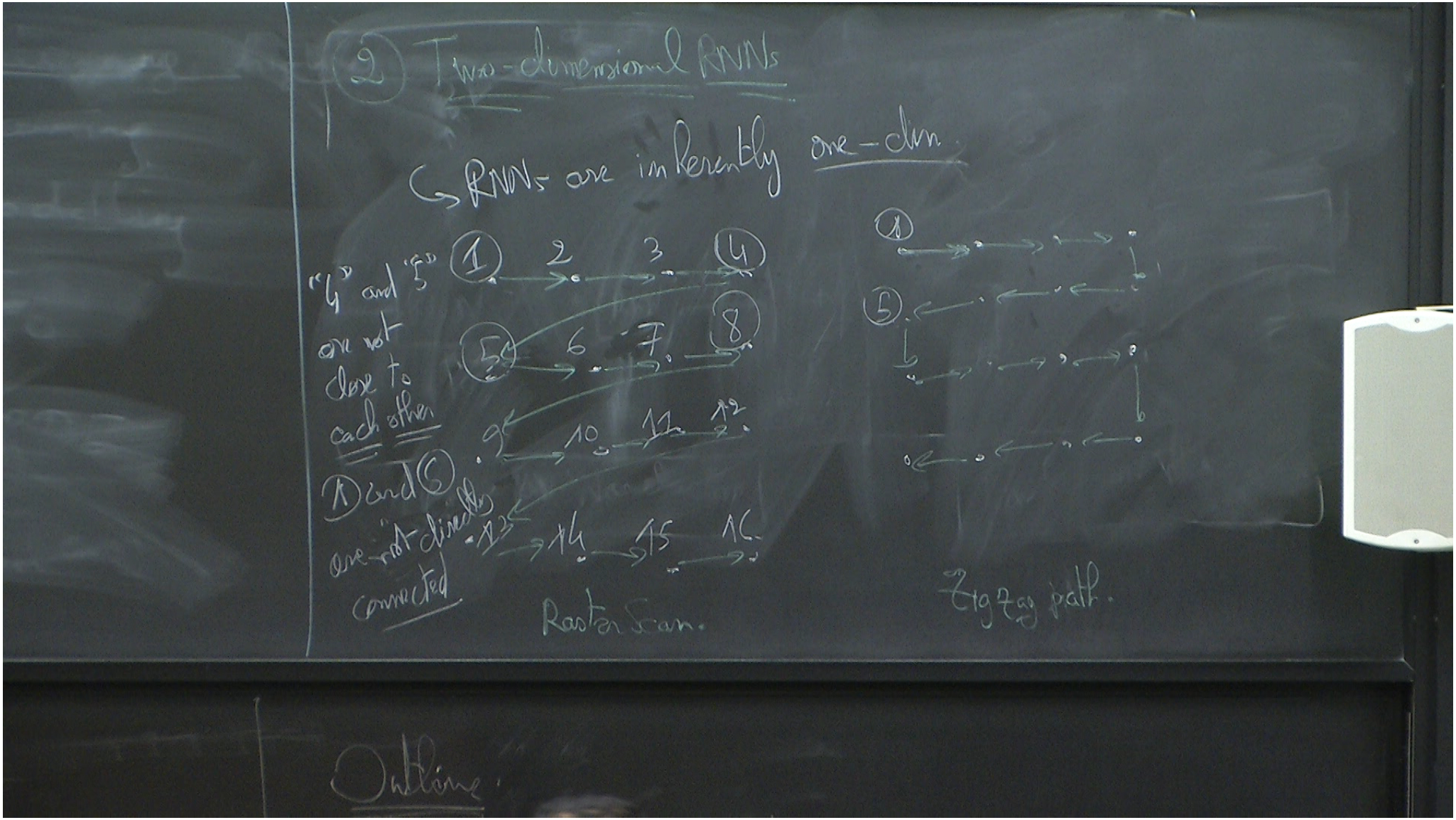

The city, which ... , is beautiful.
h=1 h=1 h=1 ... h=1
g=1 g=0 g=0 ...
↳ GRU allows to learn long range correlations.

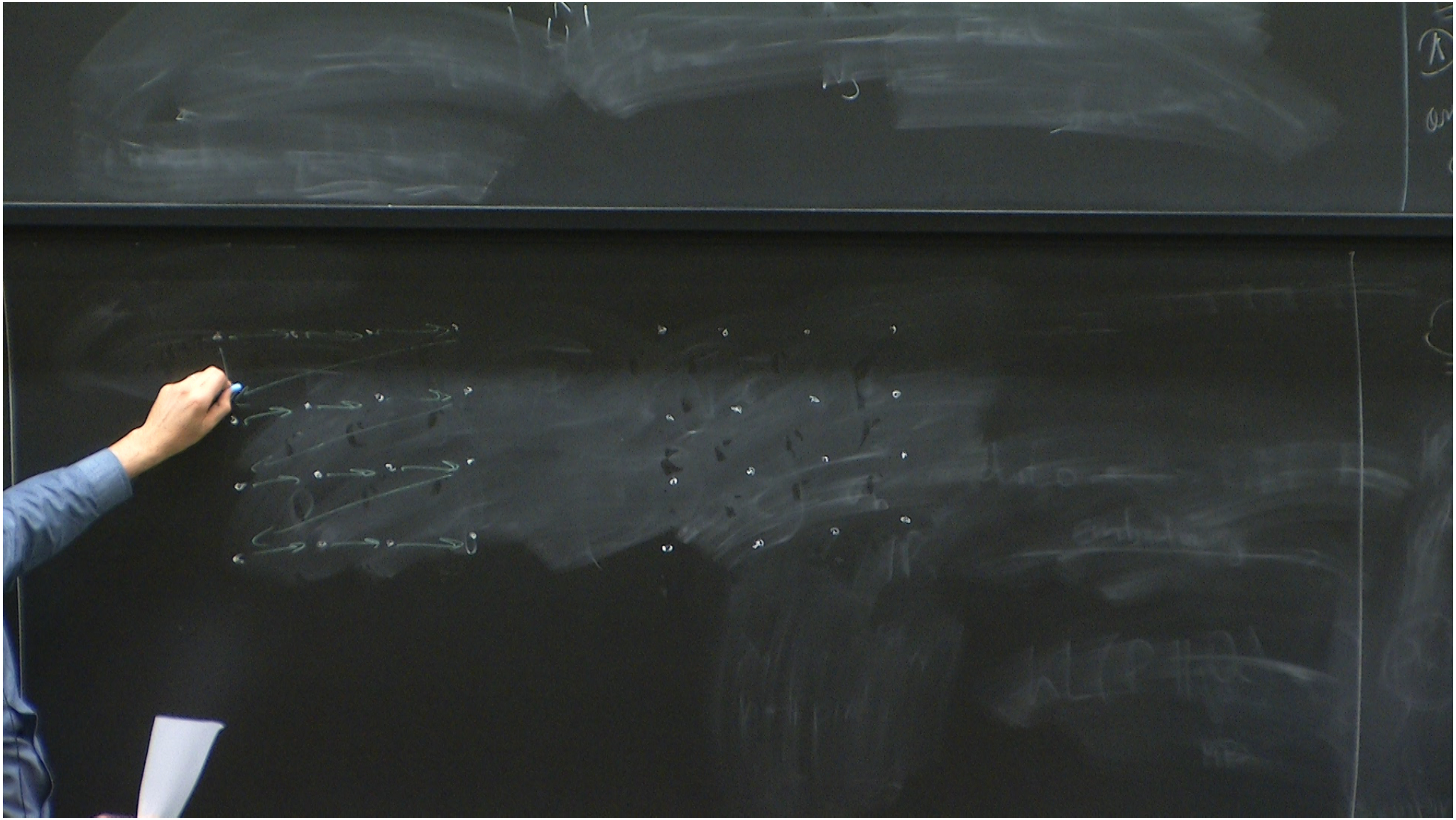

2 Two-dimensional RNNs
RNNs are inherently one-dim.
1. 2. 3. 4.
5. 6. 7. 8.
9. 10. 11. 12.
13. 14. 15. 16.

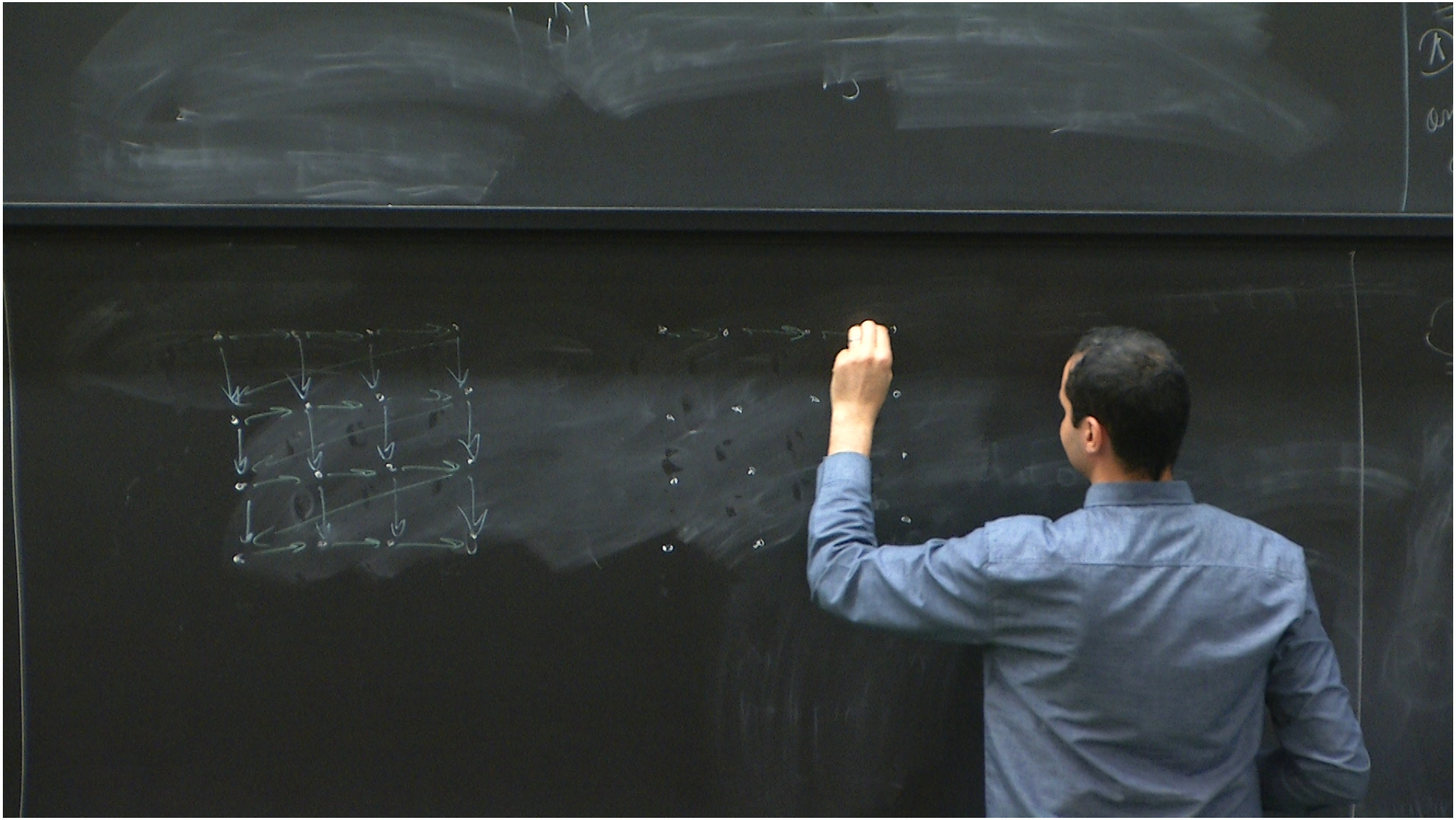

2 Two-dimensional RNNs
↳ RNNs are inherently one-dim
1 2 3 4
5 6 7 8
9 10 11 12
13 14 15 16

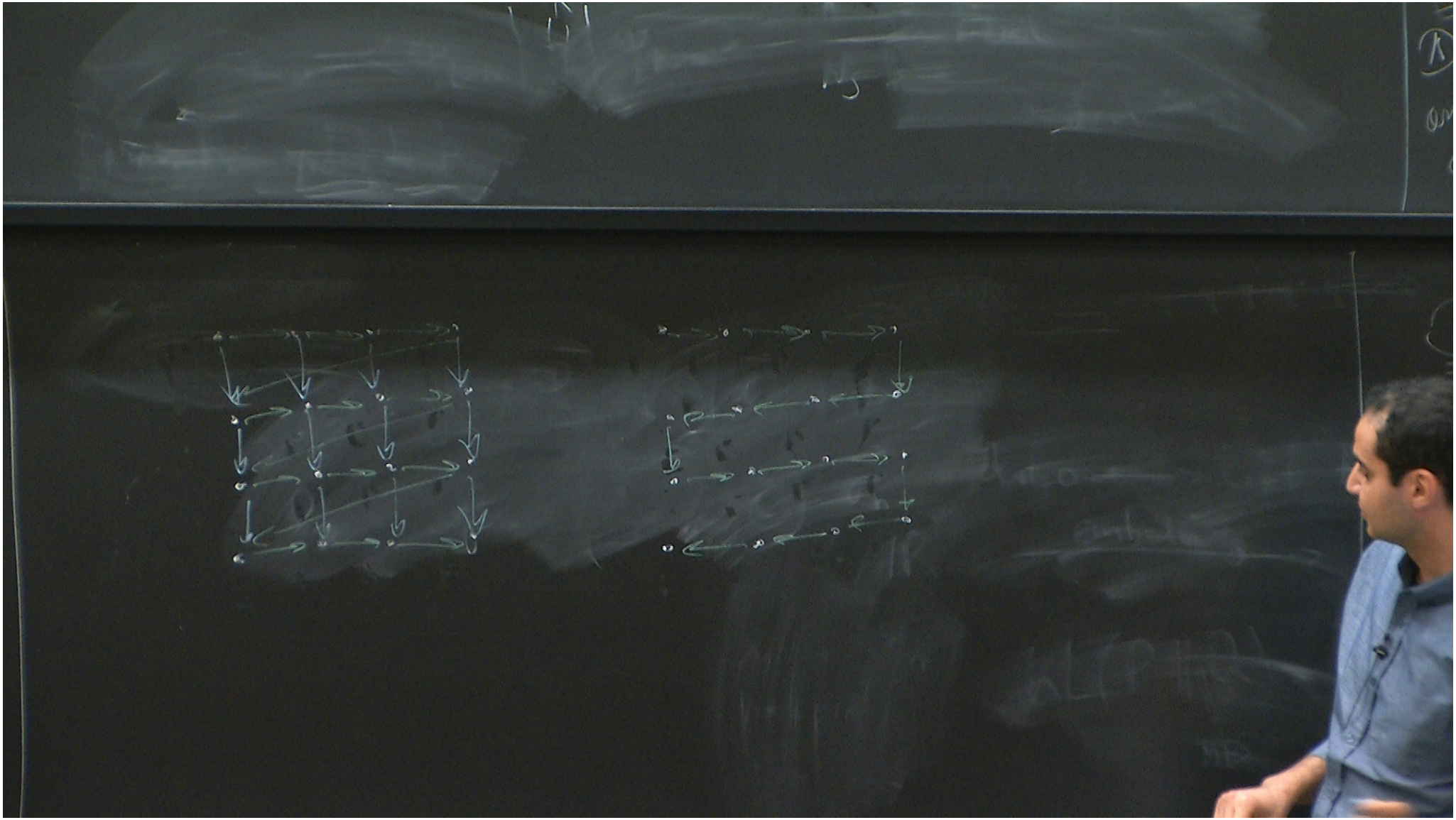

2 Two-dimensional RNNs
↳ RNNs are inherently one-dim.
"4" and "5" are not close to each other
1 and 16 are not directly connected

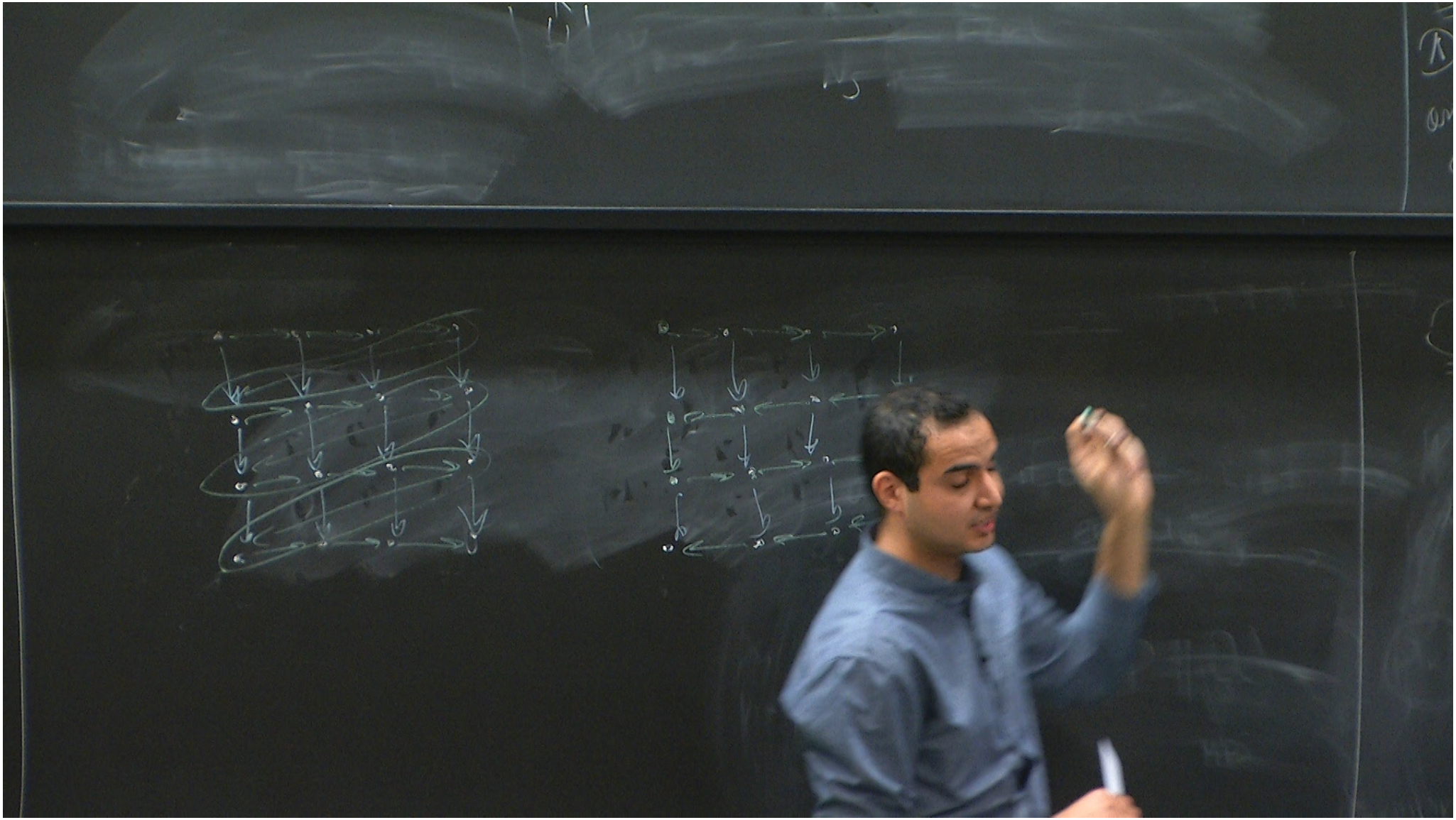

2 Two-dimensional RNNs
↳ RNNs are inherently one-dim
"4" and "5" are not close to each other
1 and 6 are not directly connected

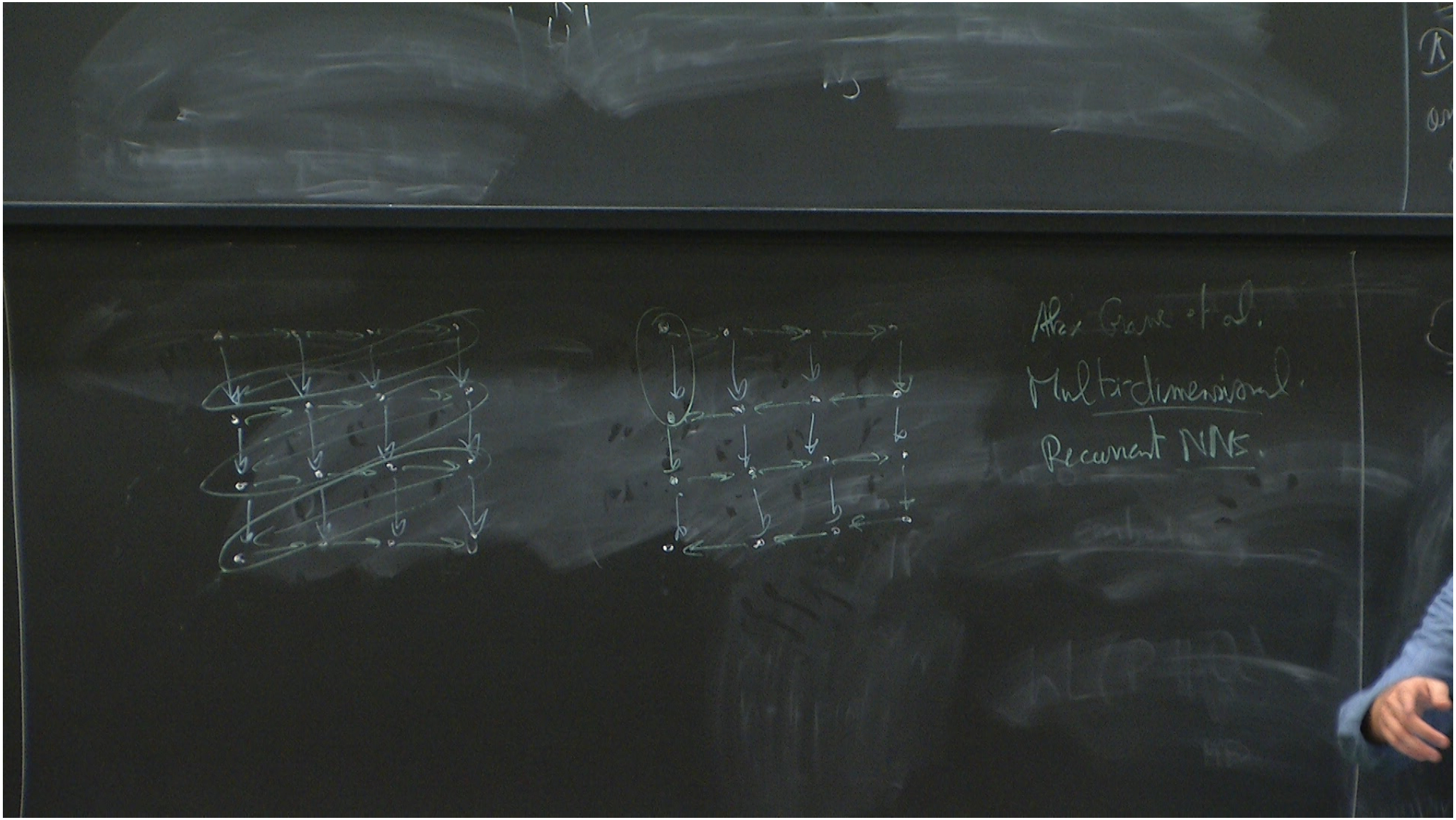

② Two-dimensional RNNs
↳ RNNs are inherently one-dim.
"4" and "5" are not close to each other
① and ⑯ are not directly connected
Raster Scan.
Zig-zag path.
Outline:

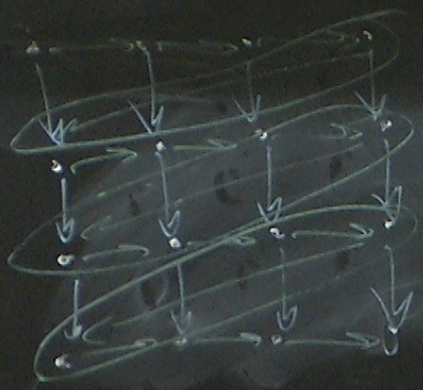


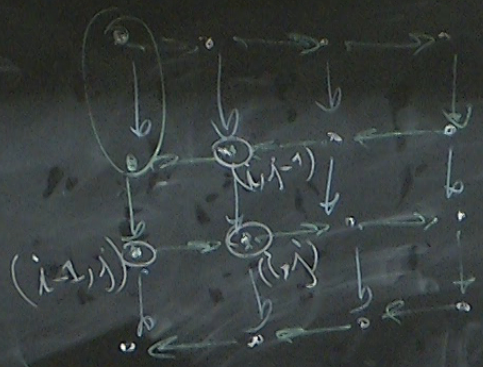







Alex Graves et al.
Multi-dimensional
Recurrent NNs.





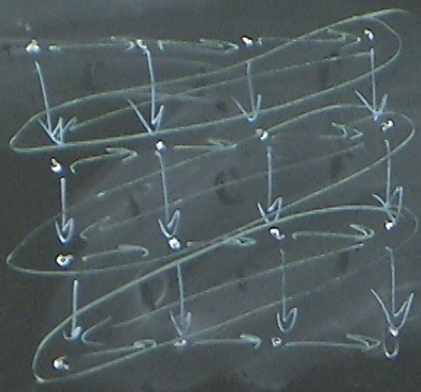
After Graves et al.

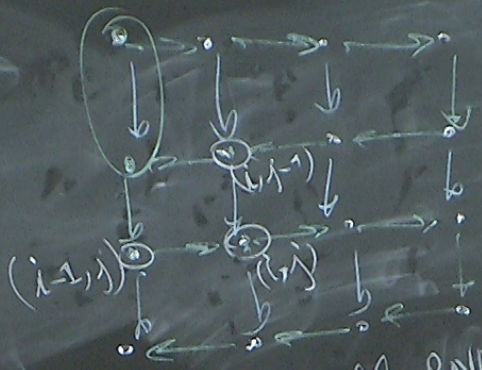
Multi-dimensional Recurrent NNs.

1D

$$\vec{h}_n = f(W\vec{h}_{n-1} + U\vec{x}_{n-1} + \vec{b})$$

$$\vec{h}_{i,j} = f(W_v\vec{h}_{i,j-1} + V_v\vec{x}_{i,j-1} + W_h\vec{h}_{i-1,j} + V_h\vec{x}_{i-1,j} + \vec{b})$$





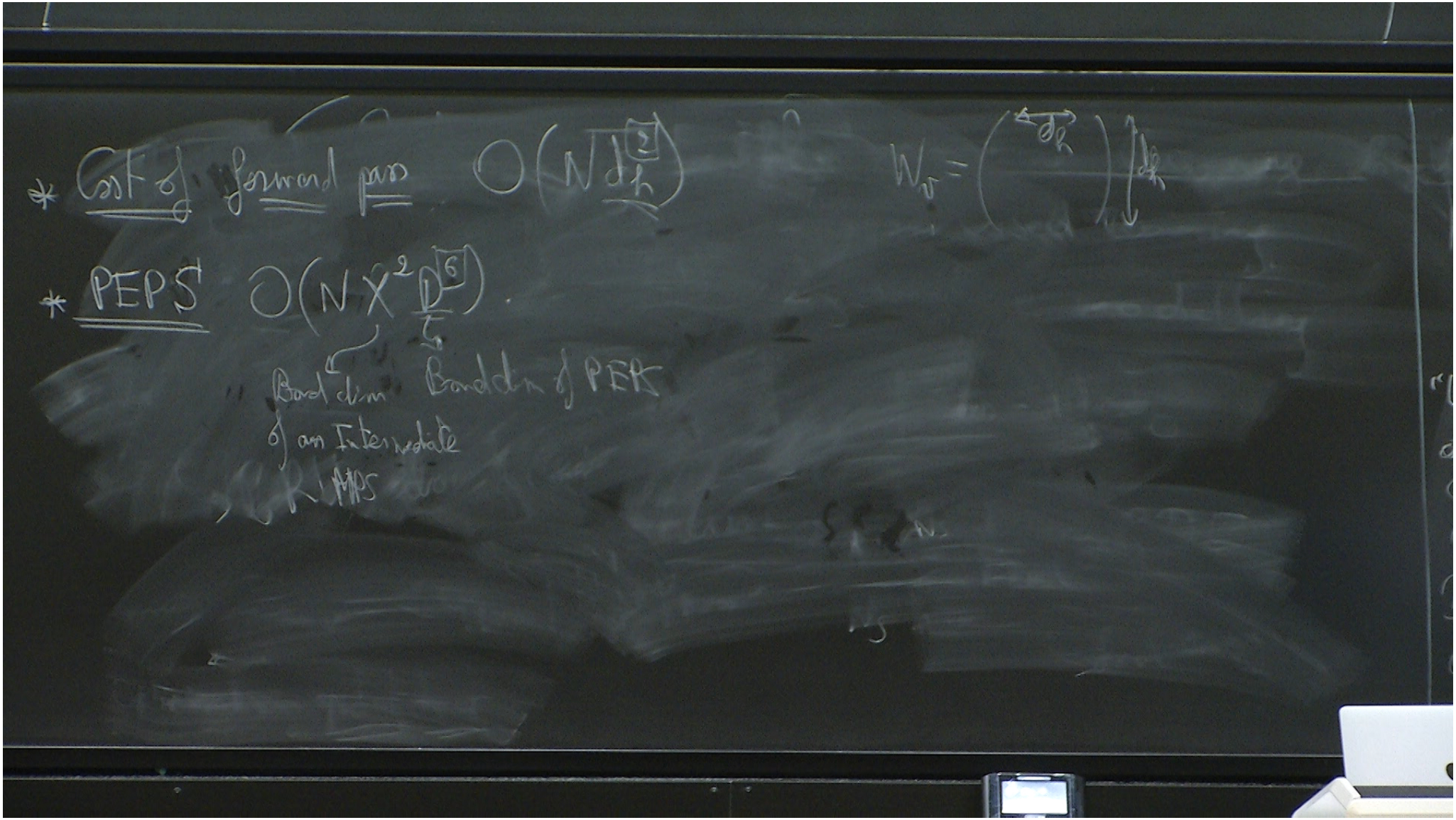
Alex Graves et al.
Multi-dimensional
Recurrent NNs.

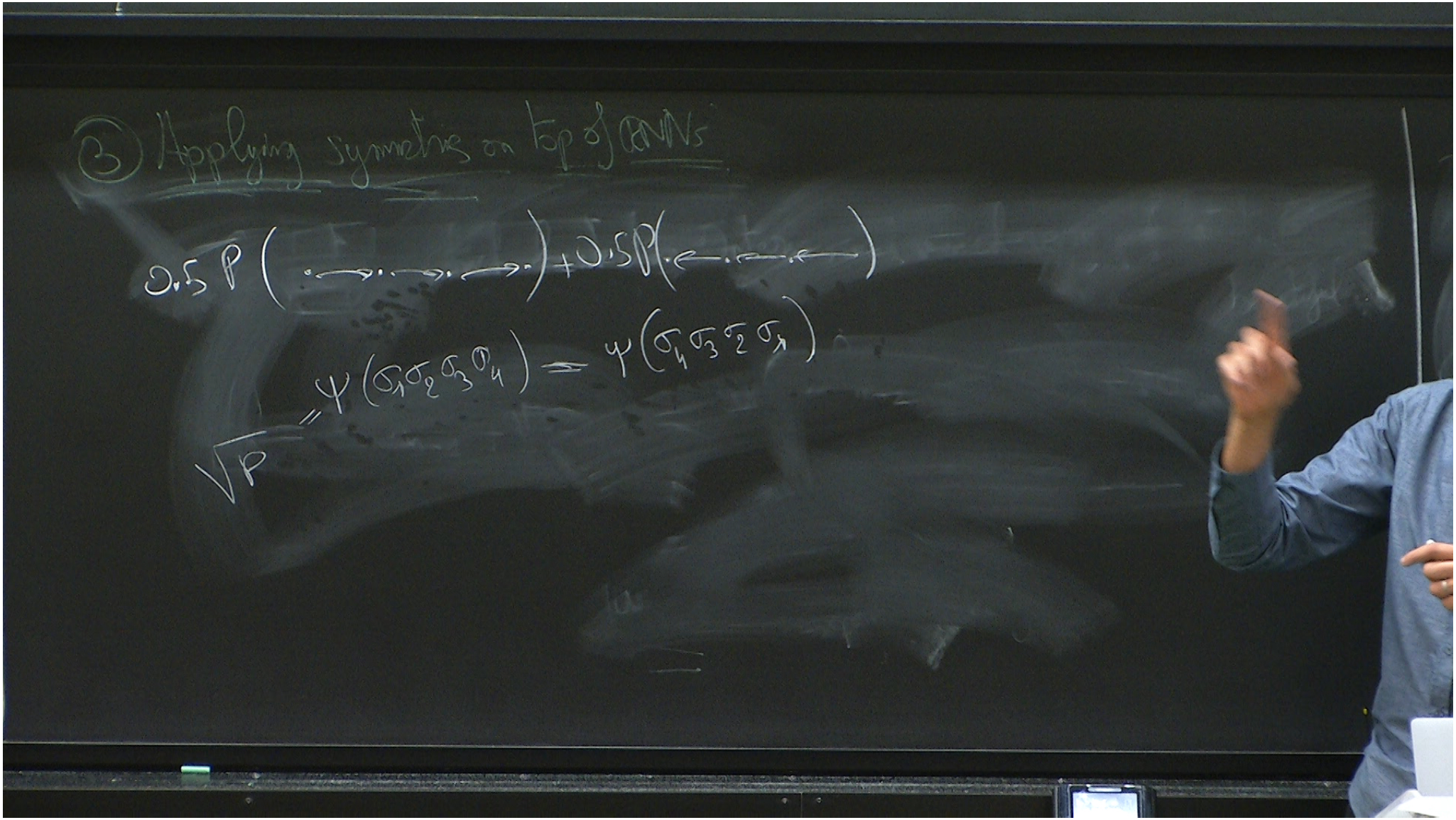
1D

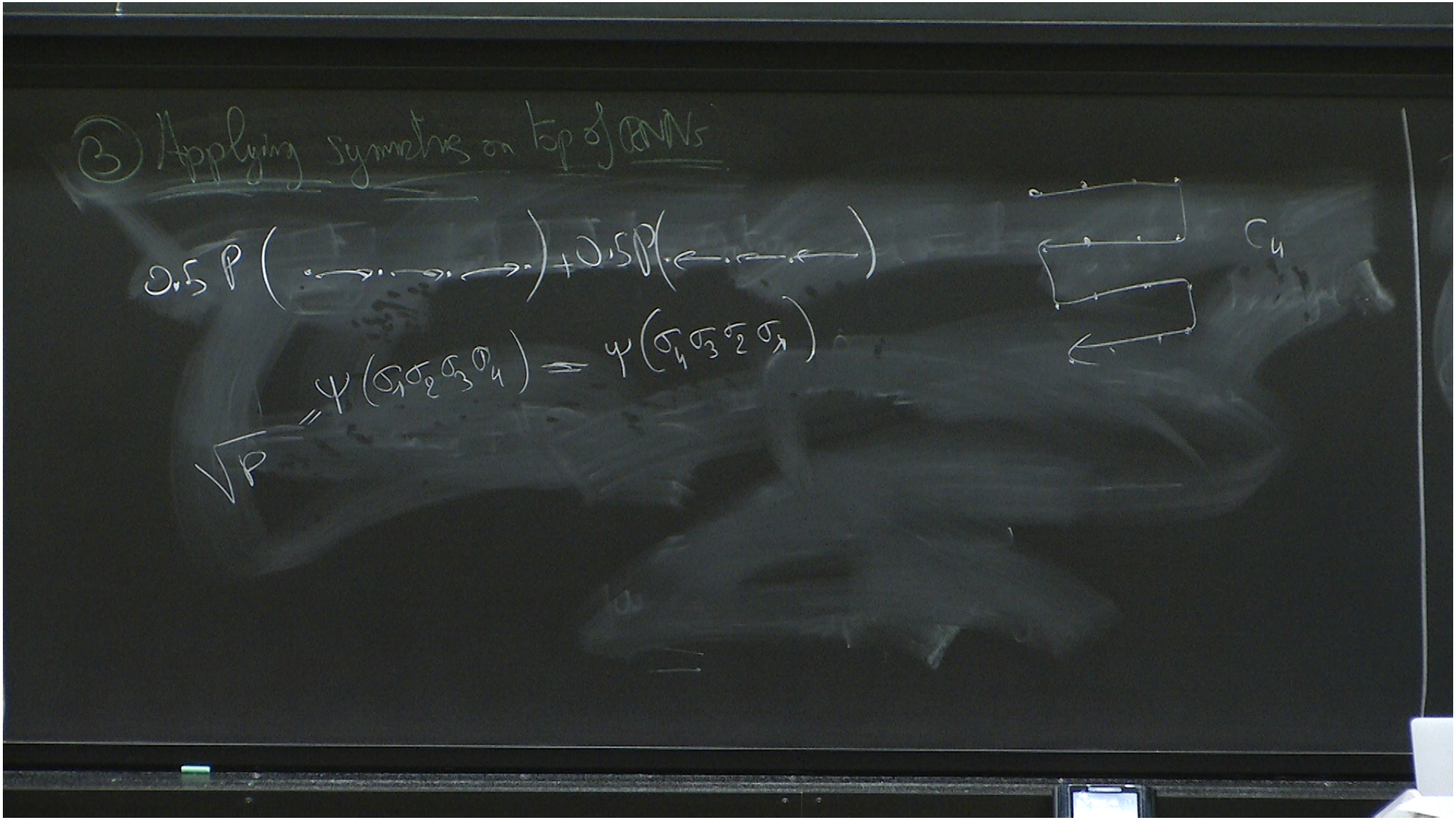
$$\vec{h}_n = f\left(W\vec{h}_{n-1} + U\vec{x}_{n-1} + \vec{b}\right).$$

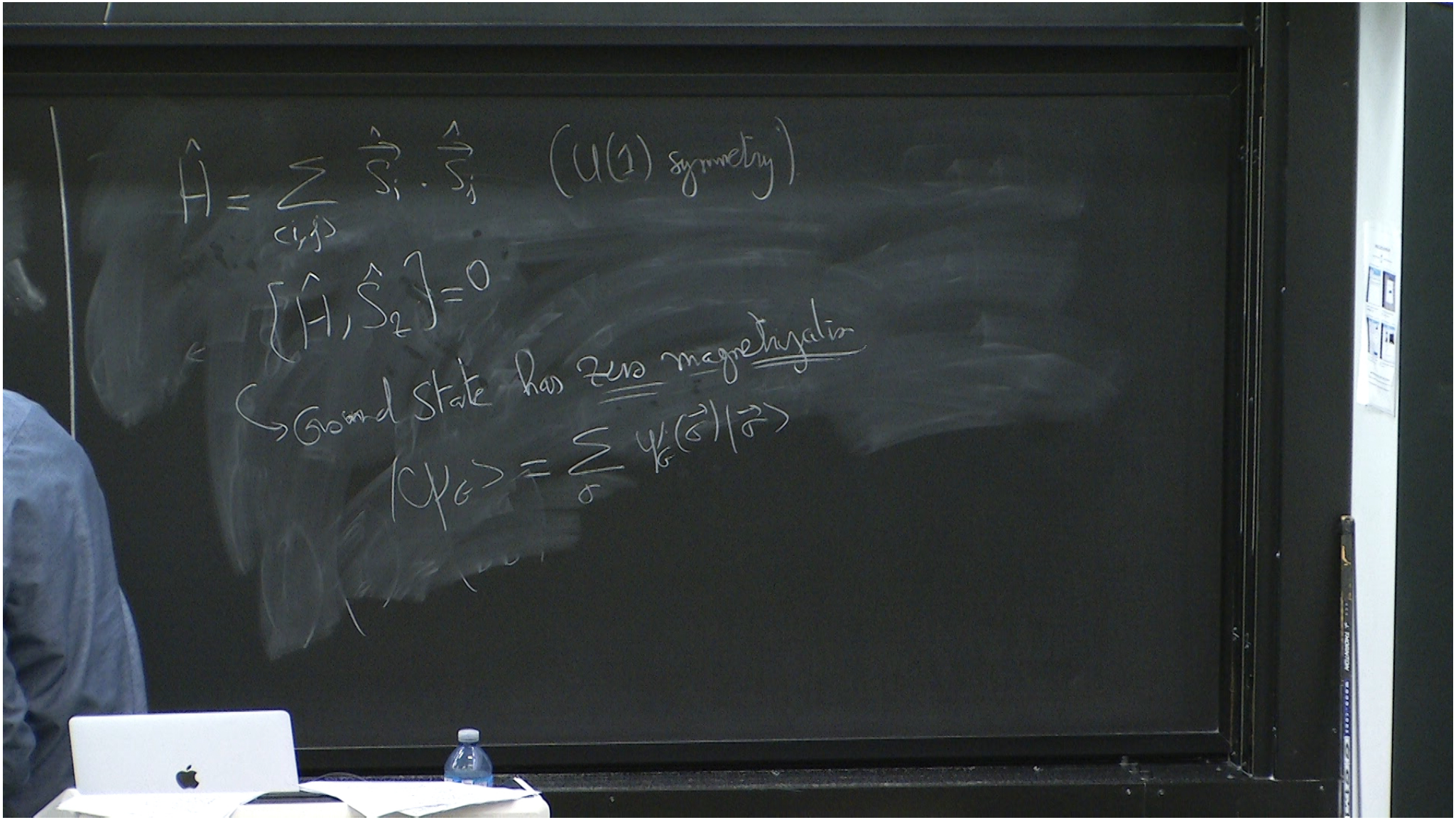
2D Vanilla RNN

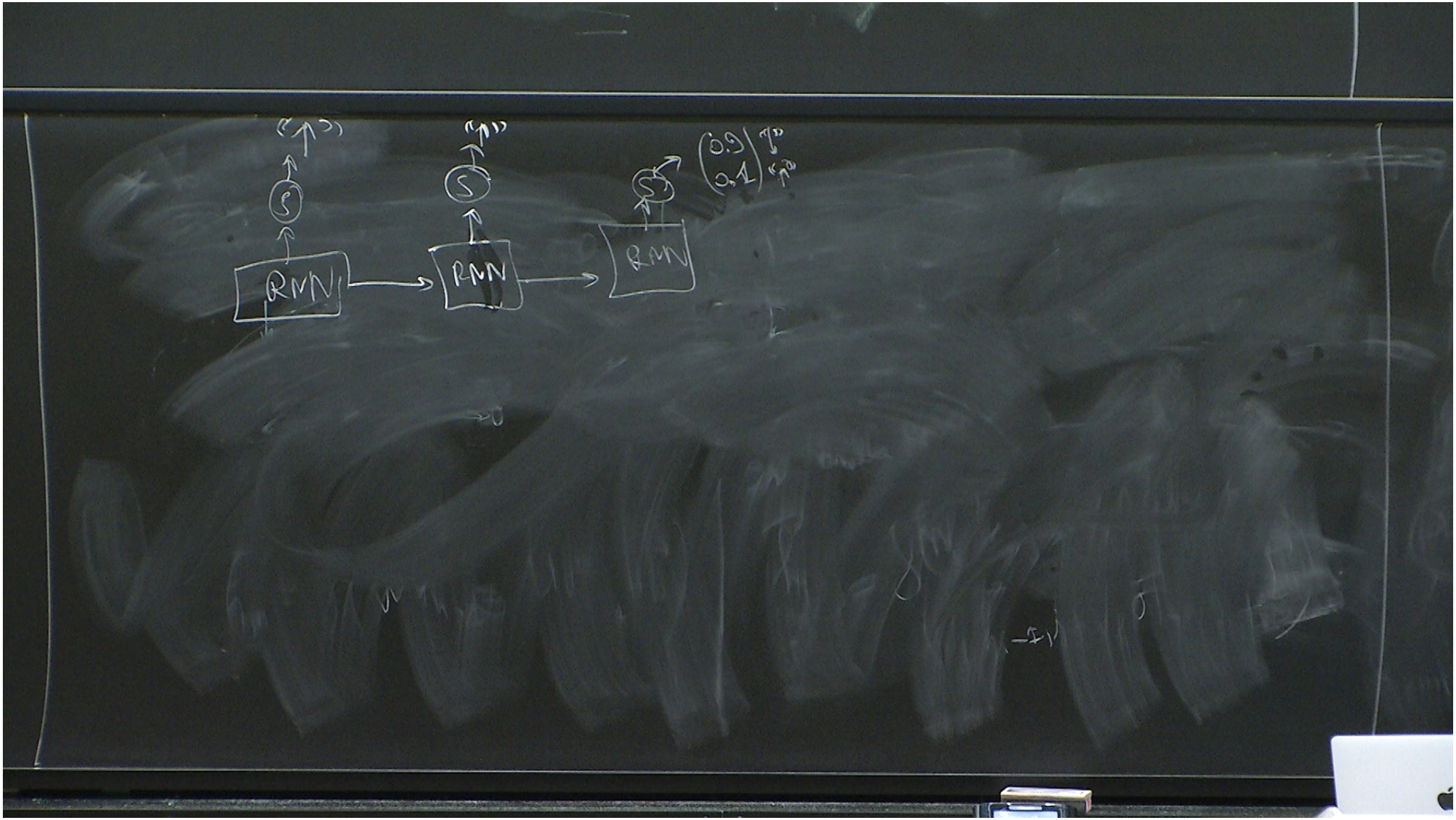
$$\vec{h}_{i,j} = f\left(W_x\vec{h}_{i,j-1} + V_x\vec{x}_{i,j-1} + W_y\vec{h}_{i-1,j} + V_y\vec{x}_{i-1,j} + \vec{b}\right)$$

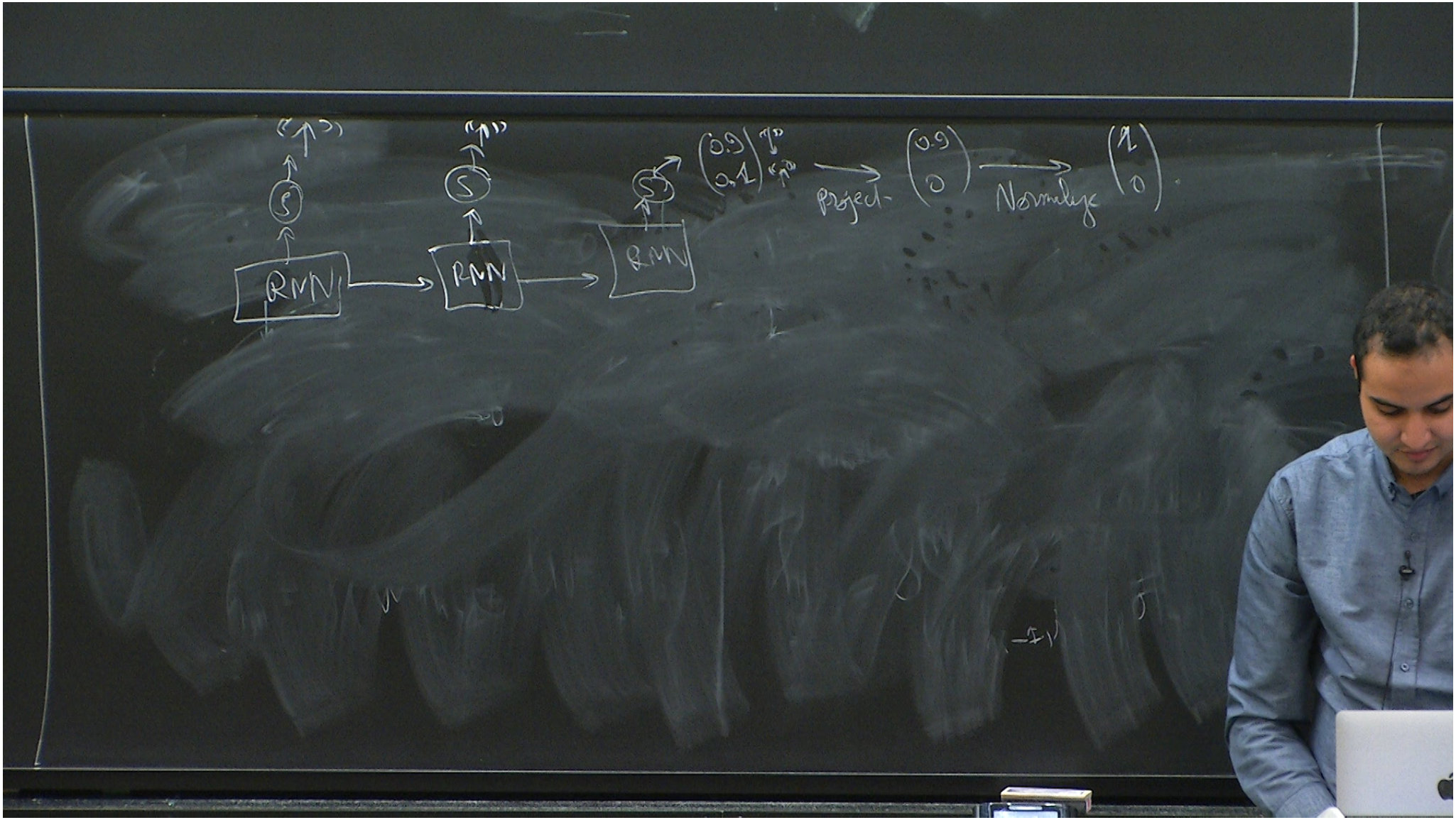

* Cost of forward pass $O(N d_h^2)$
$W_v =$
* PEPS $O(N X^2 D^6)$.
Bond dim of an Intermediate
Bond dim of PEPS


3 Applying symmetries on top of CNNs
0.5 P(→ → →) + 0.5 P(← ← ←)
√P = ψ(σ1σ2σ3σ4) = ψ(σ4σ3σ2σ1)


3 Applying symmetries on top of cNNs
0.5 P( ) + 0.5 P( )
$\sqrt{P} = \psi(\sigma_1\sigma_2\sigma_3\sigma_4) = \psi(\sigma_4\sigma_3\sigma_2\sigma_1)$
C_4


$\hat{H} = \sum_{\langle i,j \rangle} \hat{\vec{S}}_i \cdot \hat{\vec{S}}_j$ (U(1) symmetry)
$[\hat{H}, \hat{S}_z] = 0$
↳ Ground state has zero magnetization
$|\psi_G\rangle = \sum_{\sigma} \psi_G(\vec{\sigma}) |\vec{\sigma}\rangle$


RNN
RNN
RNN
0.9
0.1


RNN
RNN
RNN
Project
Normalize

RNN wave functions optimization

