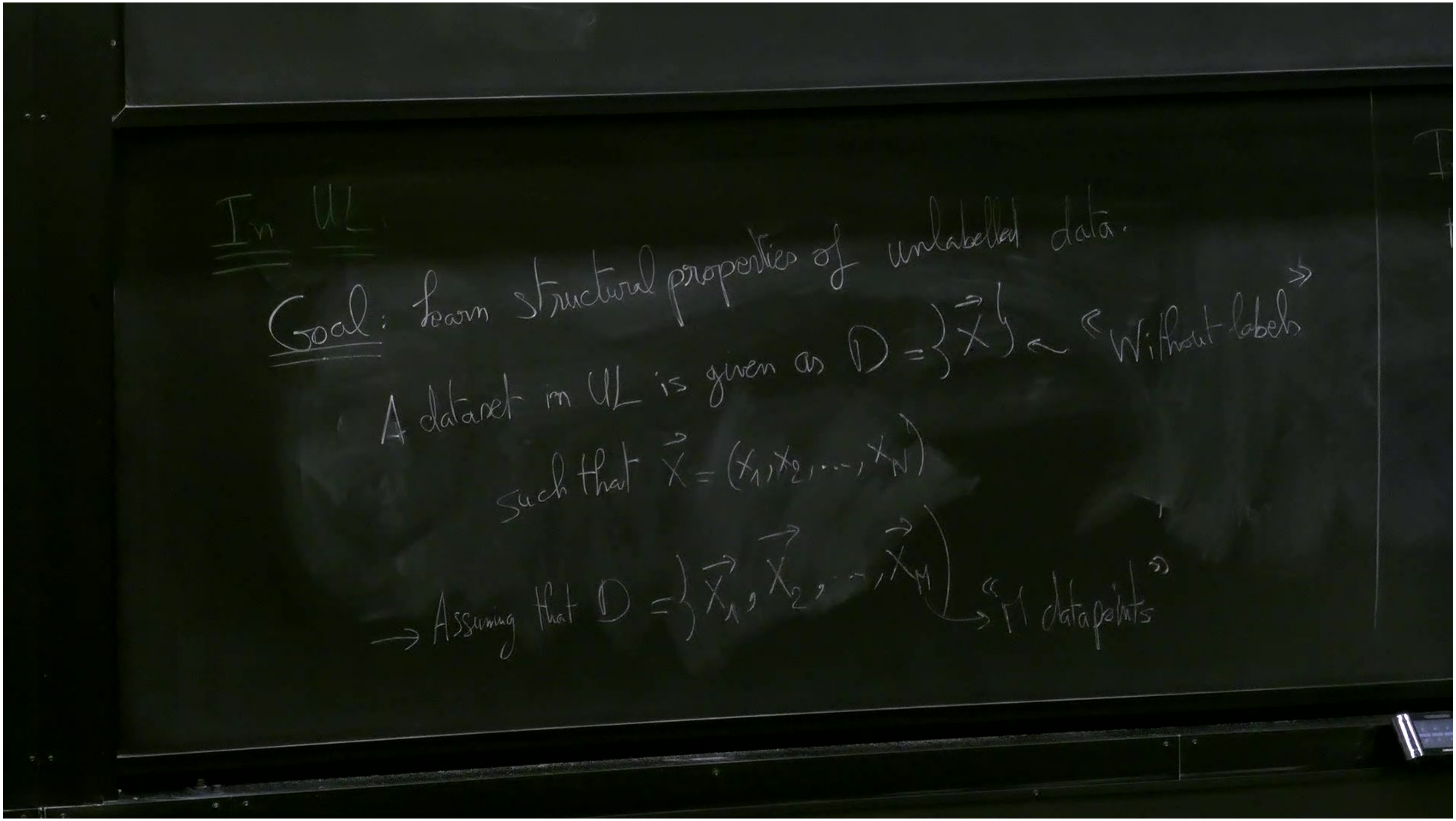
Title: Machine Learning Lecture

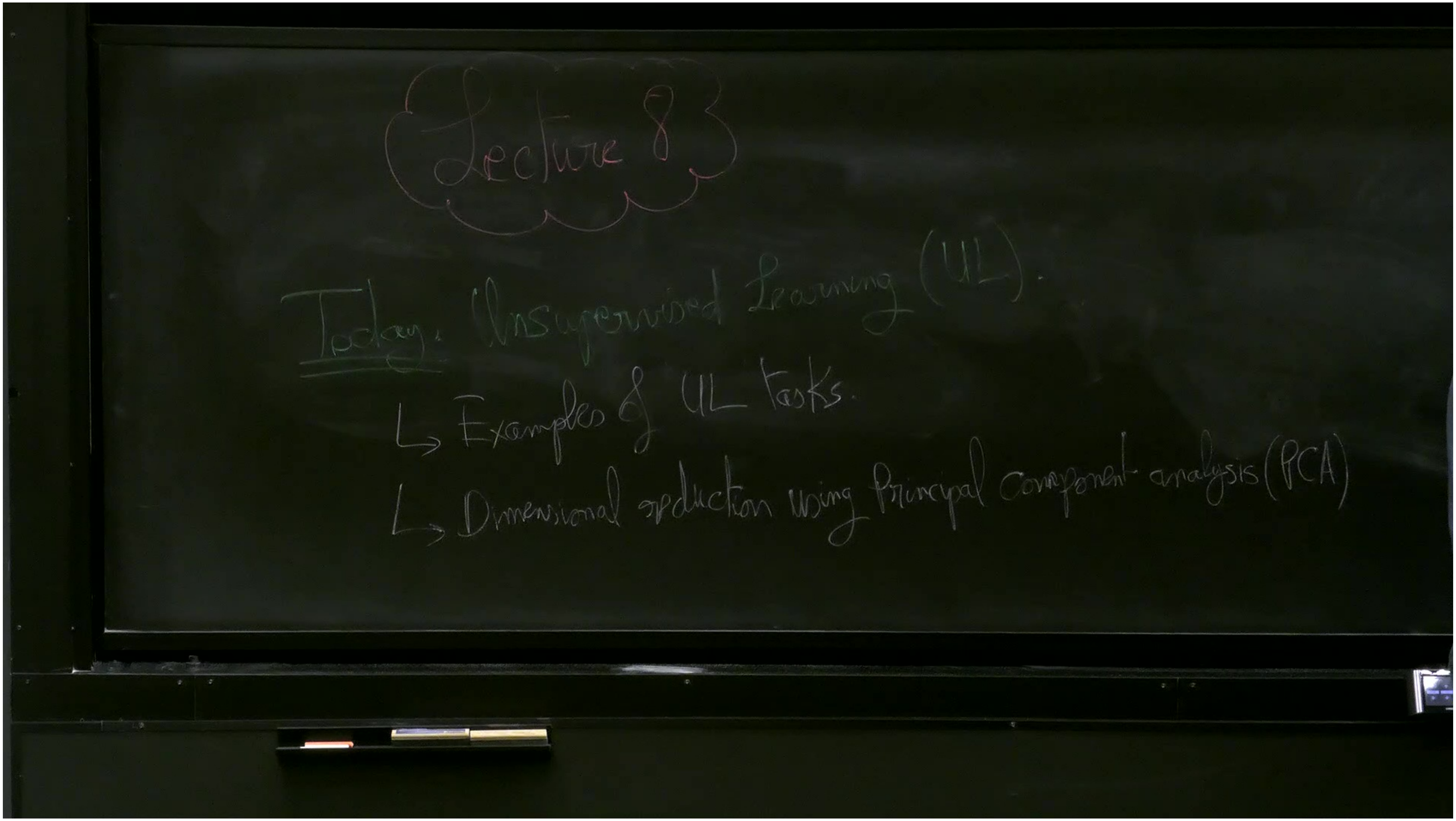
Speakers: Mohamed Hibat Allah

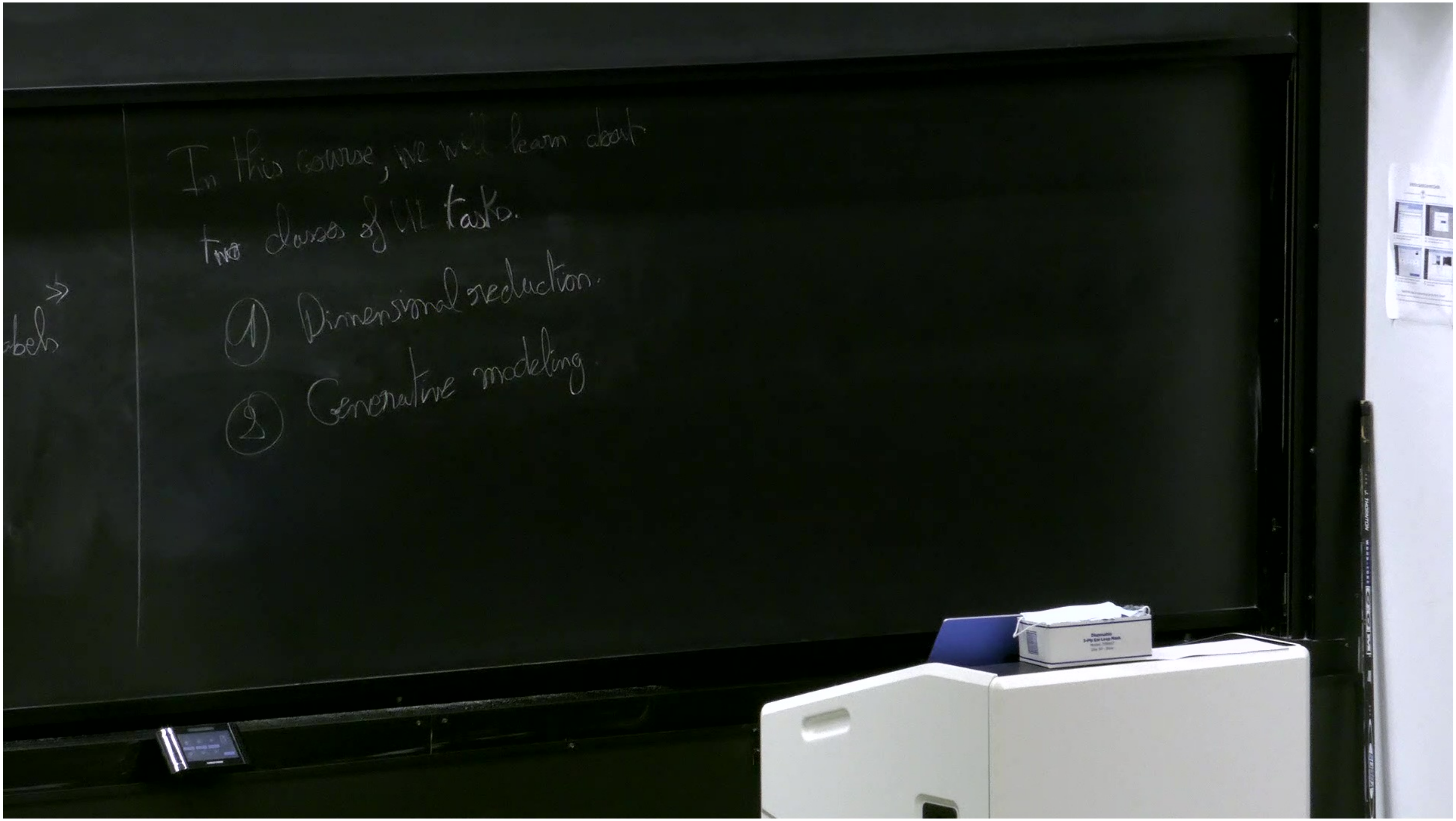
Collection: Machine Learning 2023/24

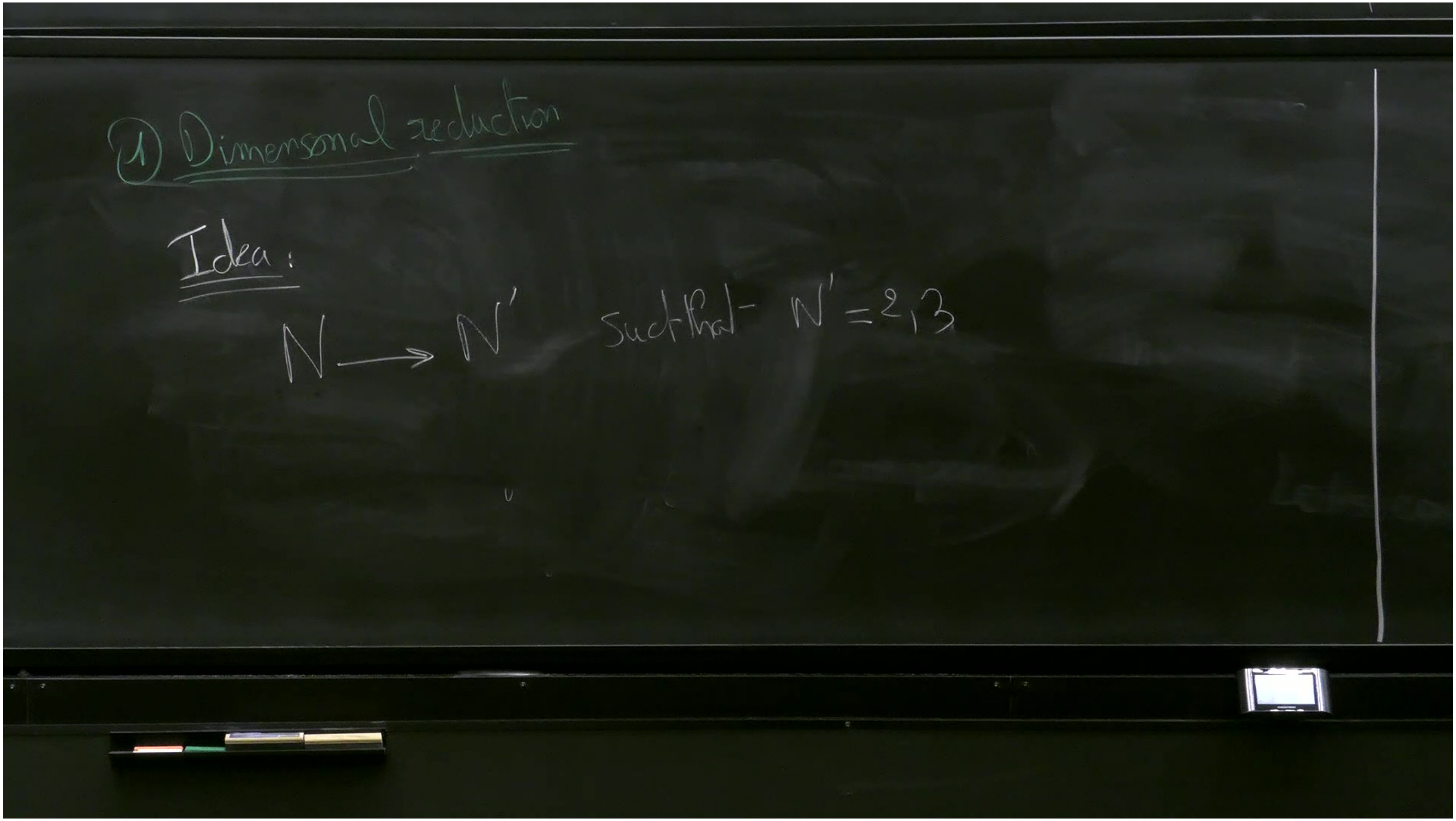
Date: April 19, 2024 - 3:45 PM

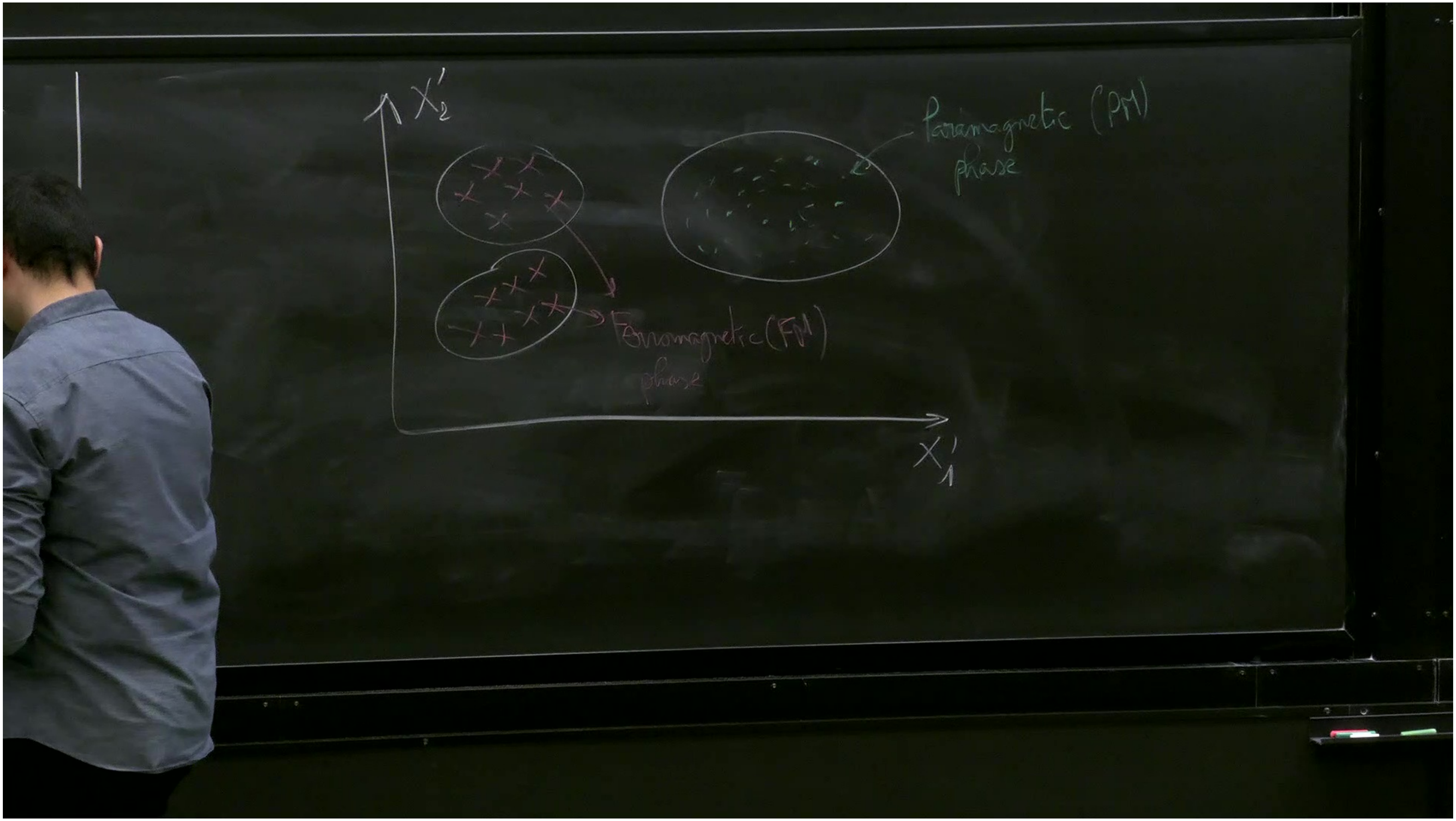
URL: https://pirsa.org/24040055


In UL:
Goal: learn structural properties of unlabelled data.
A dataset in UL is given as D = {X⃗} ← "Without labels"
such that X⃗ = (x1, x2, ..., xN)
→ Assuming that D = {X⃗1, X⃗2, ..., X⃗M} → "M datapoints"

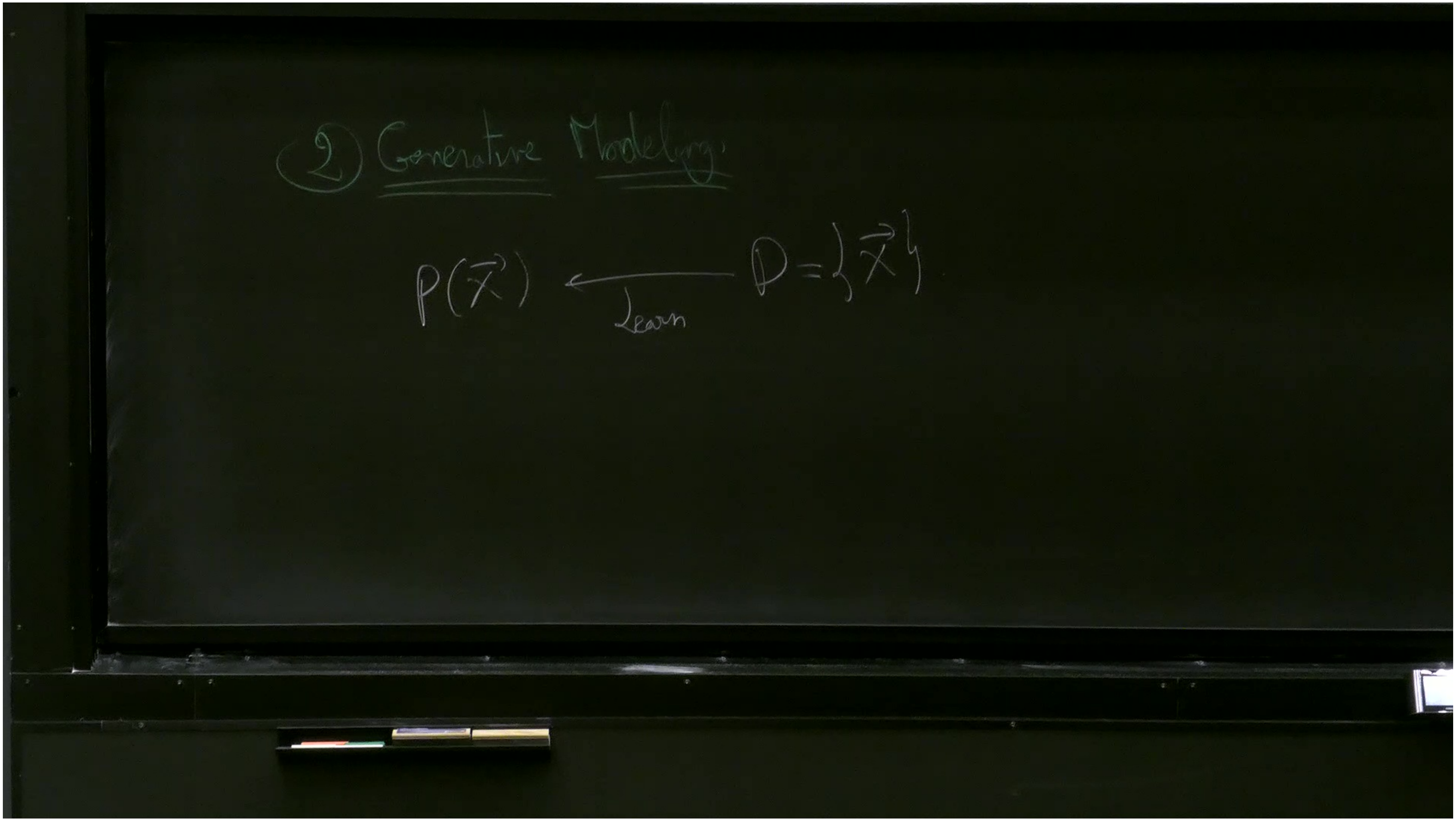

Lecture 8
Today: Unsupervised Learning (UL).
↳ Examples of UL tasks.
↳ Dimensional reduction using Principal component analysis (PCA)

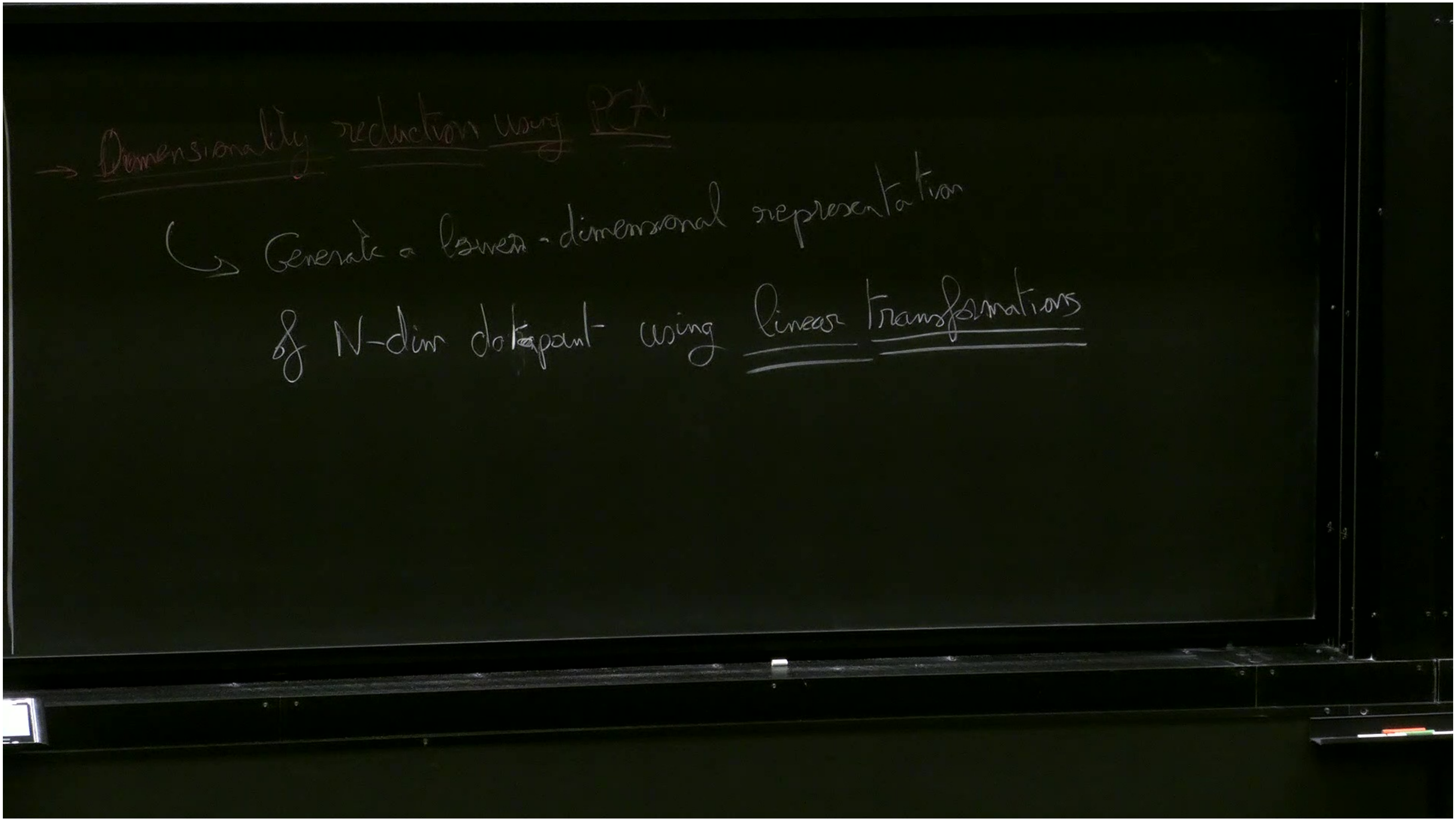

In this course, we will learn about
two classes of UL tasks.
(1) Dimensional reduction.
(2) Generative modeling.

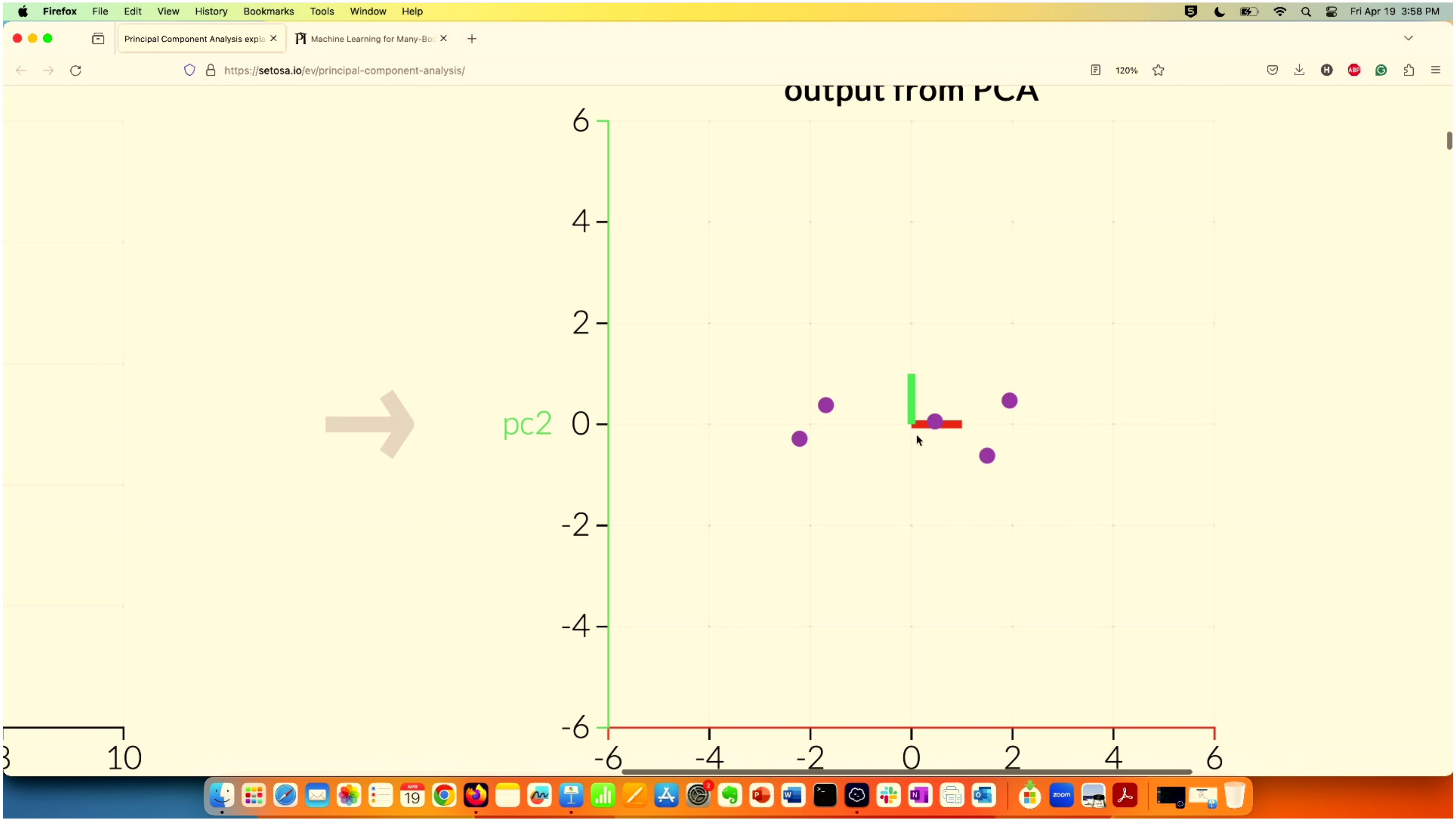

① Dimensional reduction
Idea:
N ⟶ N' such that N' = 2, 3

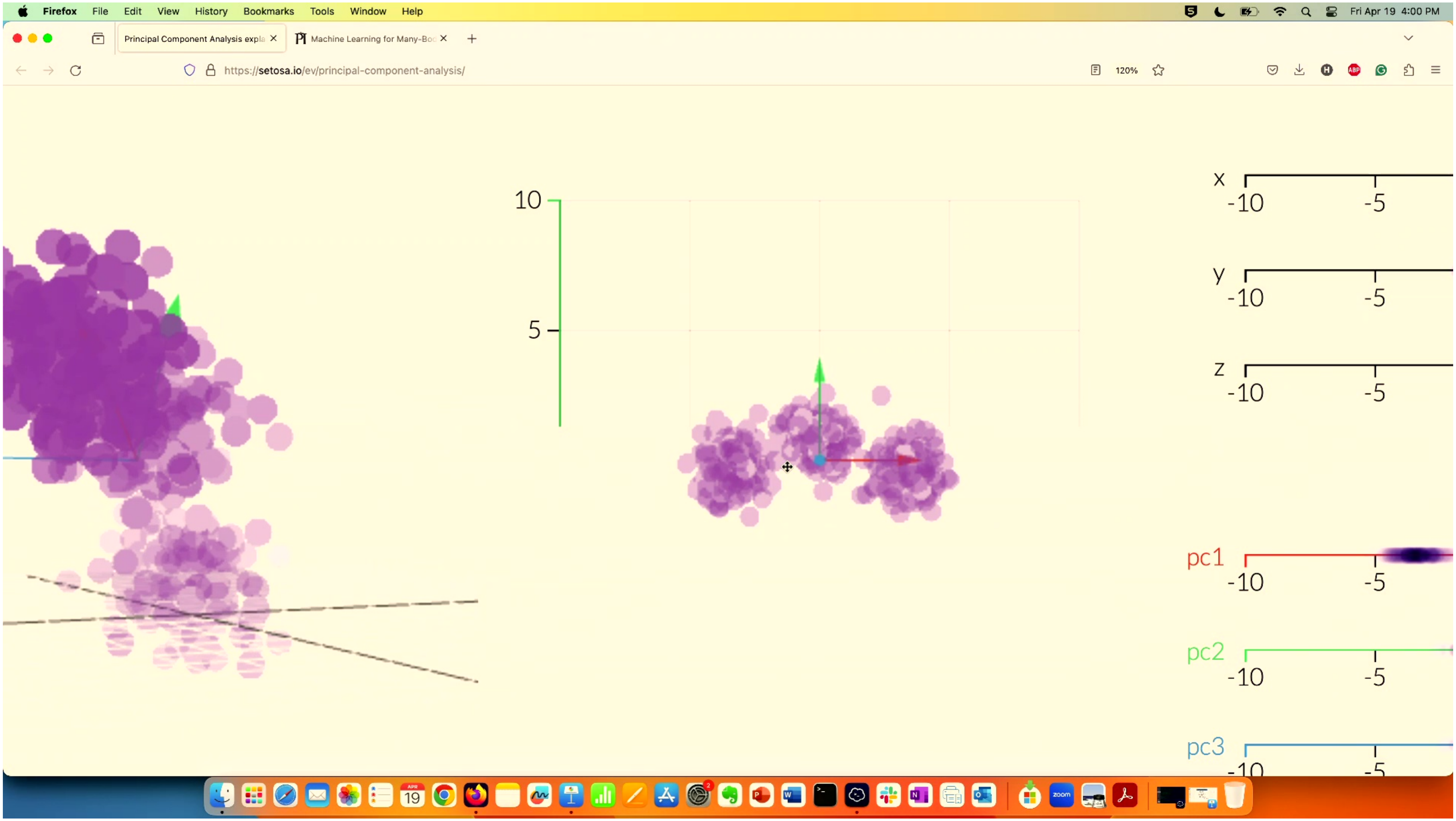


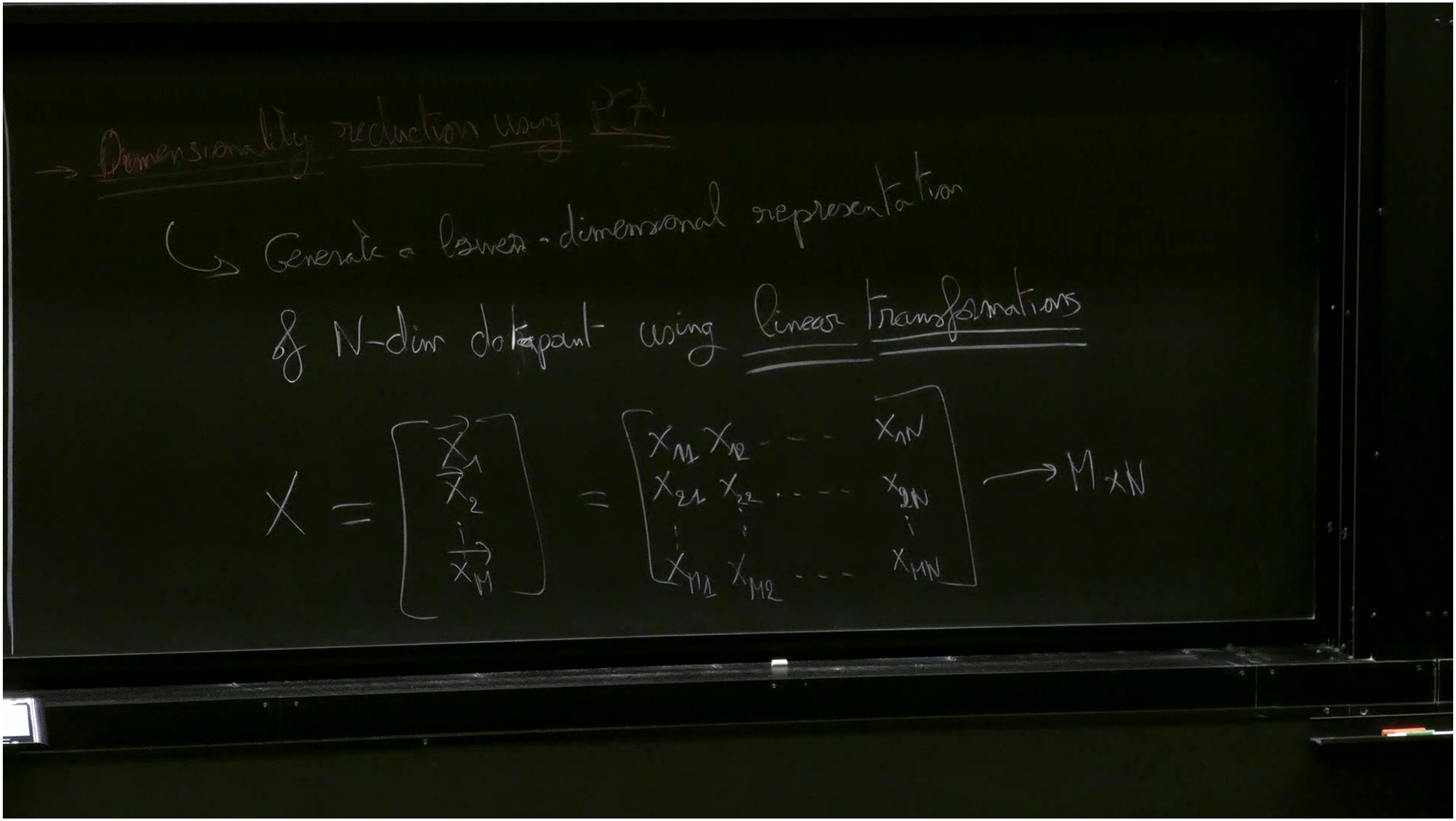
X_2'
Paramagnetic (PM) phase
Ferromagnetic (FM) phase
X_1'

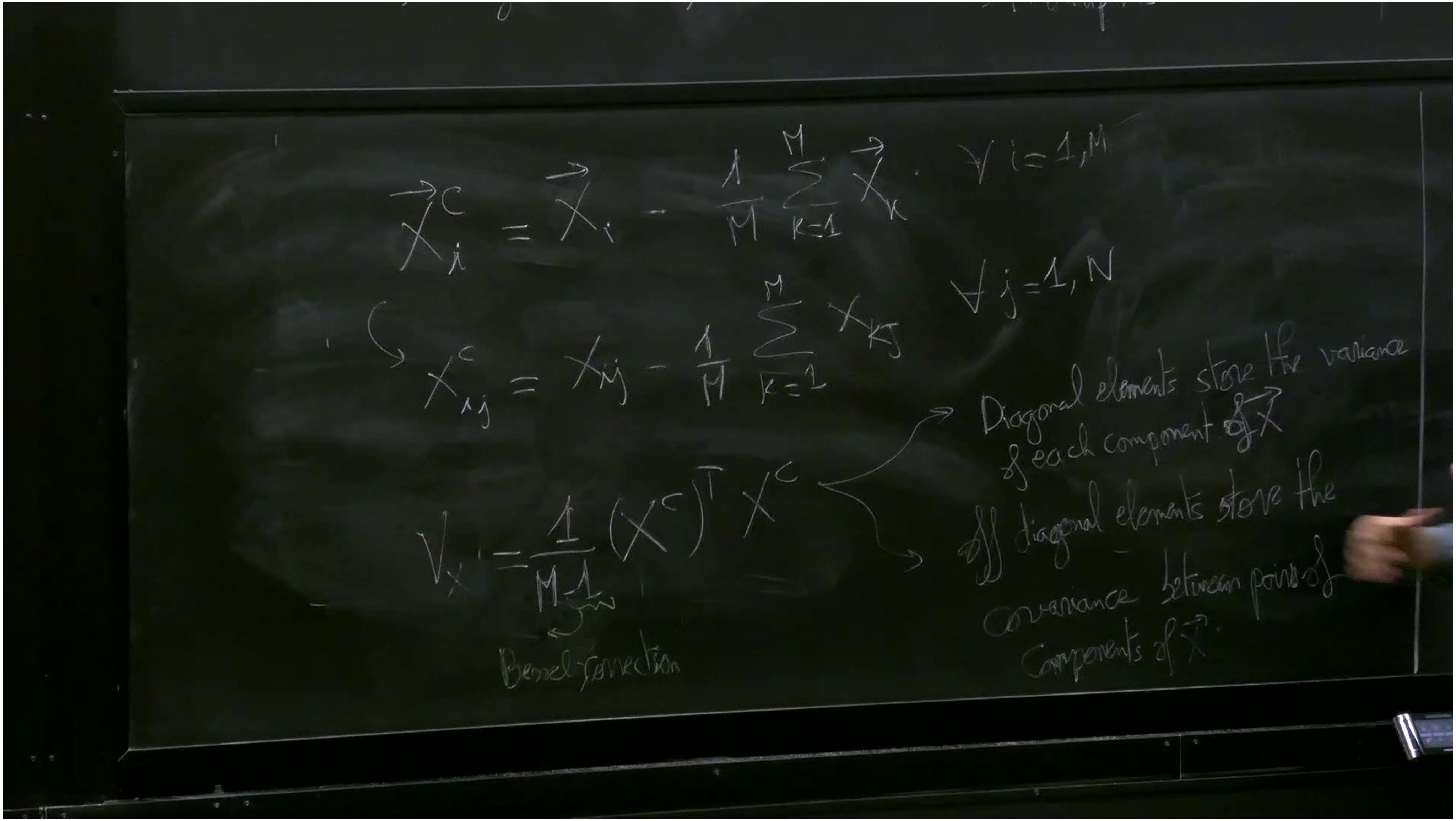

2 Generative Modeling:
P(x) ← Learn D = {x}


→ Dimensionality reduction using PCA:
↳ Generate a lower-dimensional representation
of N-dim datapoint using linear transformations


output from PCA
pc2
6
4
2
0
-2
-4
-6
-6
-4
-2
0
2
4
6
10


Firefox File Edit View History Bookmarks Tools Window Help
Fri Apr 19 4:00 PM
Principal Component Analysis expla
Machine Learning for Many-Bo
https://setosa.io/ev/principal-component-analysis/
120%
10
5
x
-10 -5
y
-10 -5
z
-10 -5
pc1
-10 -5
pc2
-10 -5
pc3
-10 -5


→ Dimensionality reduction using PCA
↳ Generate a lower-dimensional representation
of N-dim datapoint using linear transformations
$X = \begin{bmatrix} \vec{x}_1 \\ \vec{x}_2 \\ \vdots \\ \vec{x}_M \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1N} \\ x_{21} & x_{22} & \cdots & x_{2N} \\ \vdots & \vdots & & \vdots \\ x_{M1} & x_{M2} & \cdots & x_{MN} \end{bmatrix} \longrightarrow M \times N$


$\vec{X}_i^c = \vec{X}_i - \frac{1}{M}\sum_{k=1}^{M} \vec{X}_k$ $\forall i = 1, M$
$x_{ij}^c = x_{ij} - \frac{1}{M}\sum_{k=1}^{M} x_{kj}$ $\forall j = 1, N$
$V_X = \frac{1}{M-1} (X^c)^T X^c$
Bessel correction
Diagonal elements store the variance of each component of $\vec{X}$
Off diagonal elements store the covariance between pairs of components of $\vec{X}$.

→ Let's find P such that $V_{X'}$ diagonal

$$X' = X^c P$$

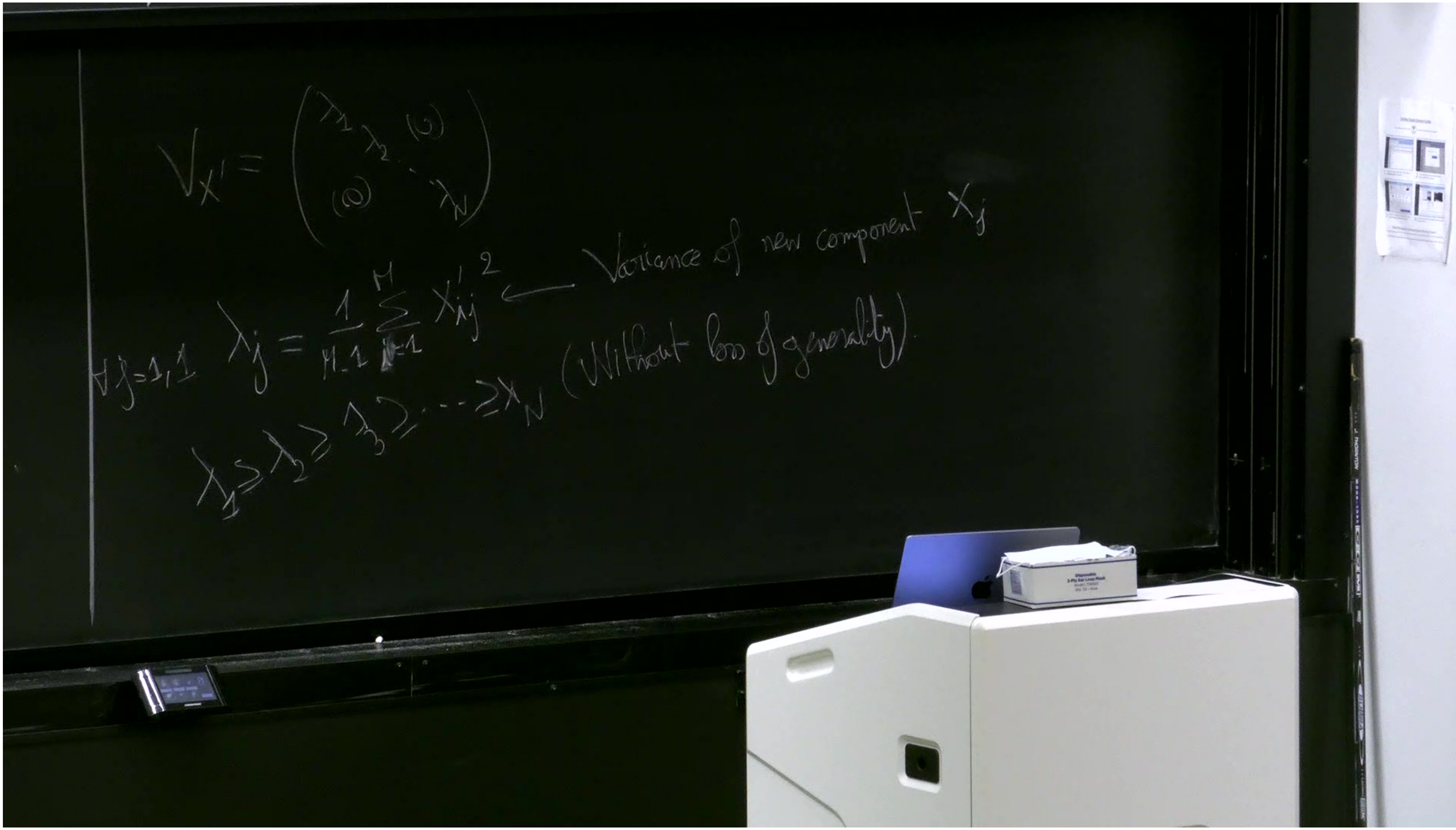
X': $M \times N$, X^c: $M \times N$, P: $N \times N$

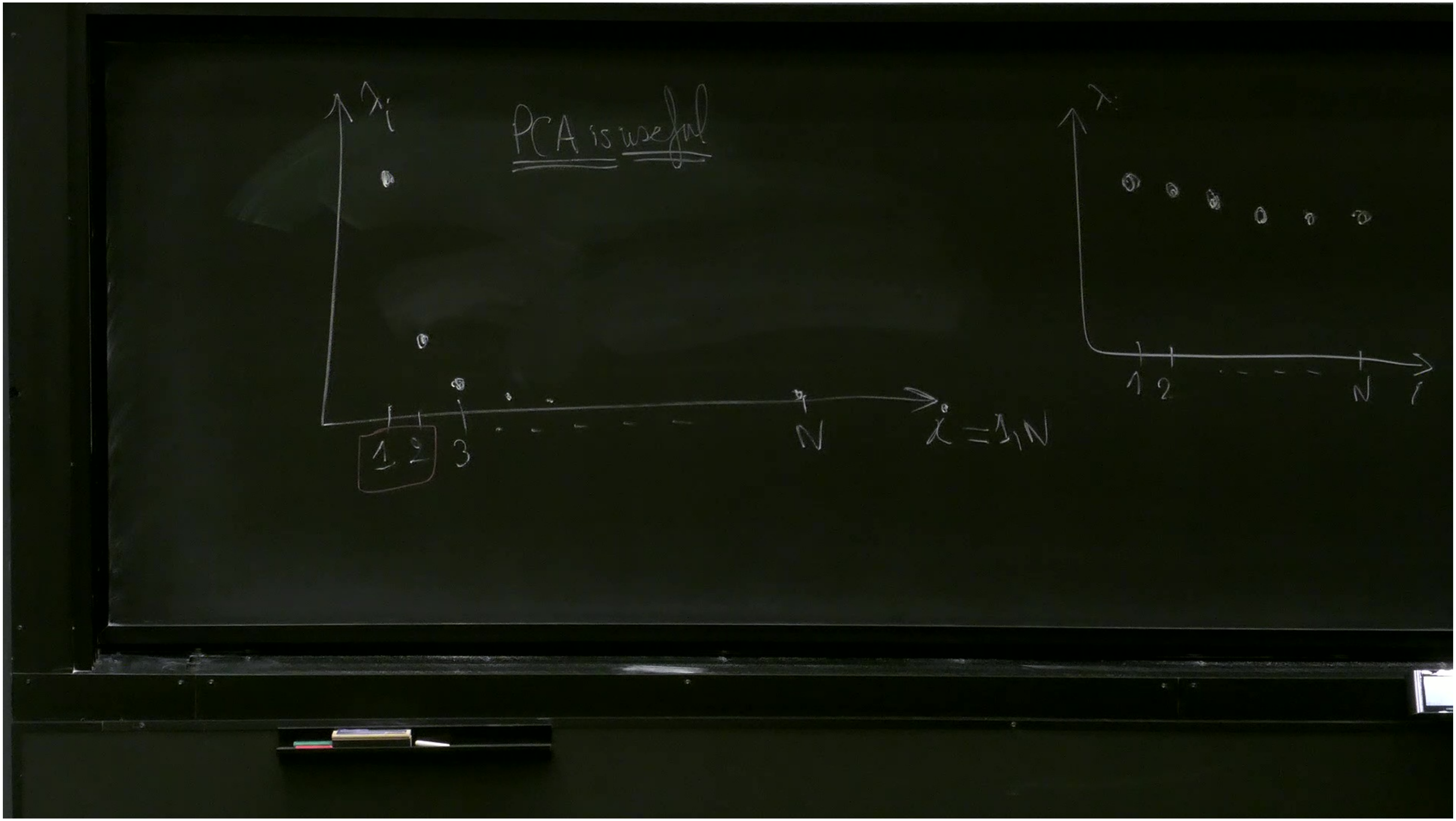
$$V_{X'} = \frac{1}{M-1}(X')^T X' = \frac{1}{M-1}(X^c P)^T (X^c P)$$

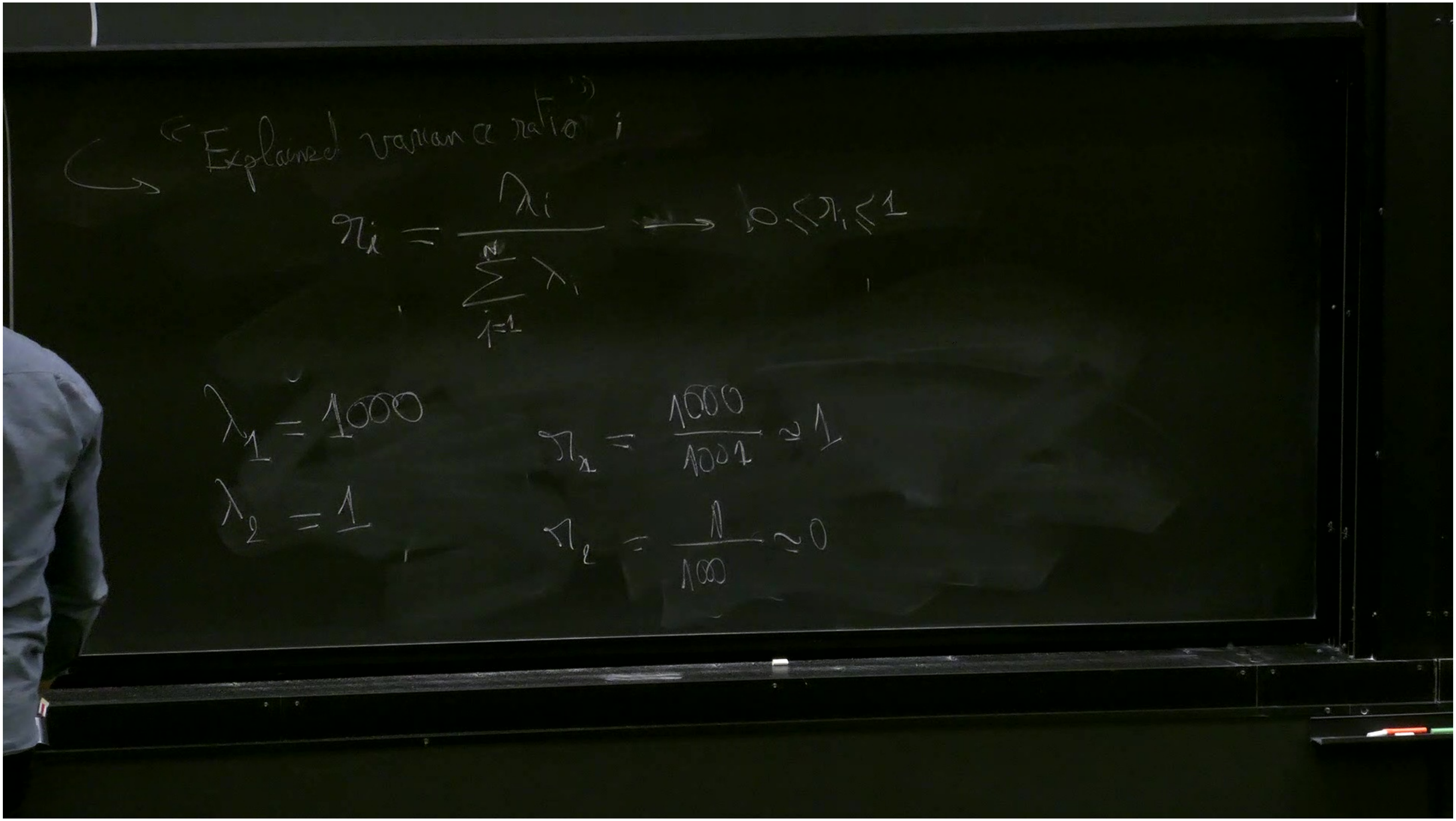
$$= P^T\left(\frac{1}{M-1}(X^c)^T X^c\right)P = P^T V_X P$$

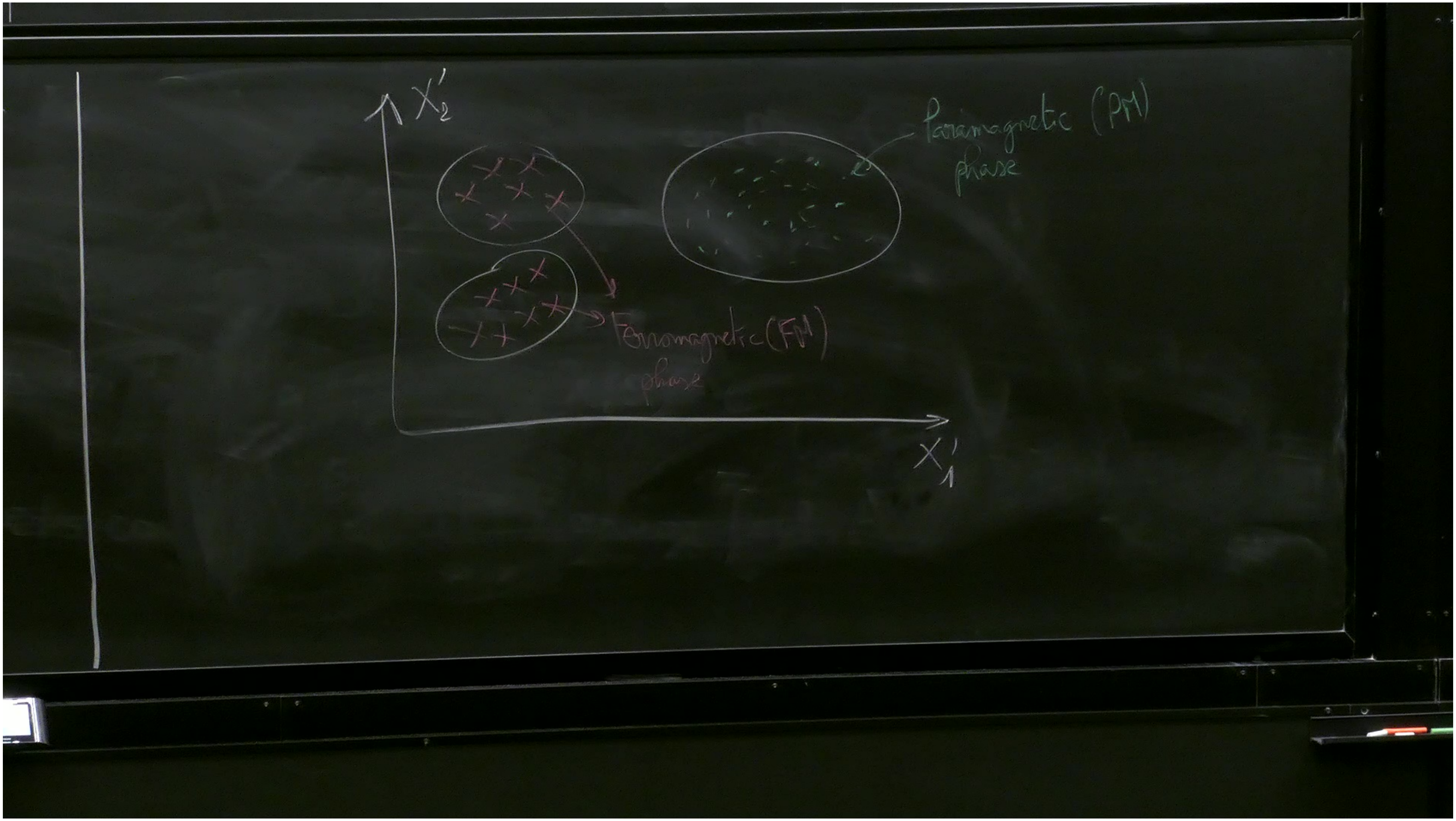
↙ Diagonal matrix

$P^{-1} \leftrightarrow P^T$

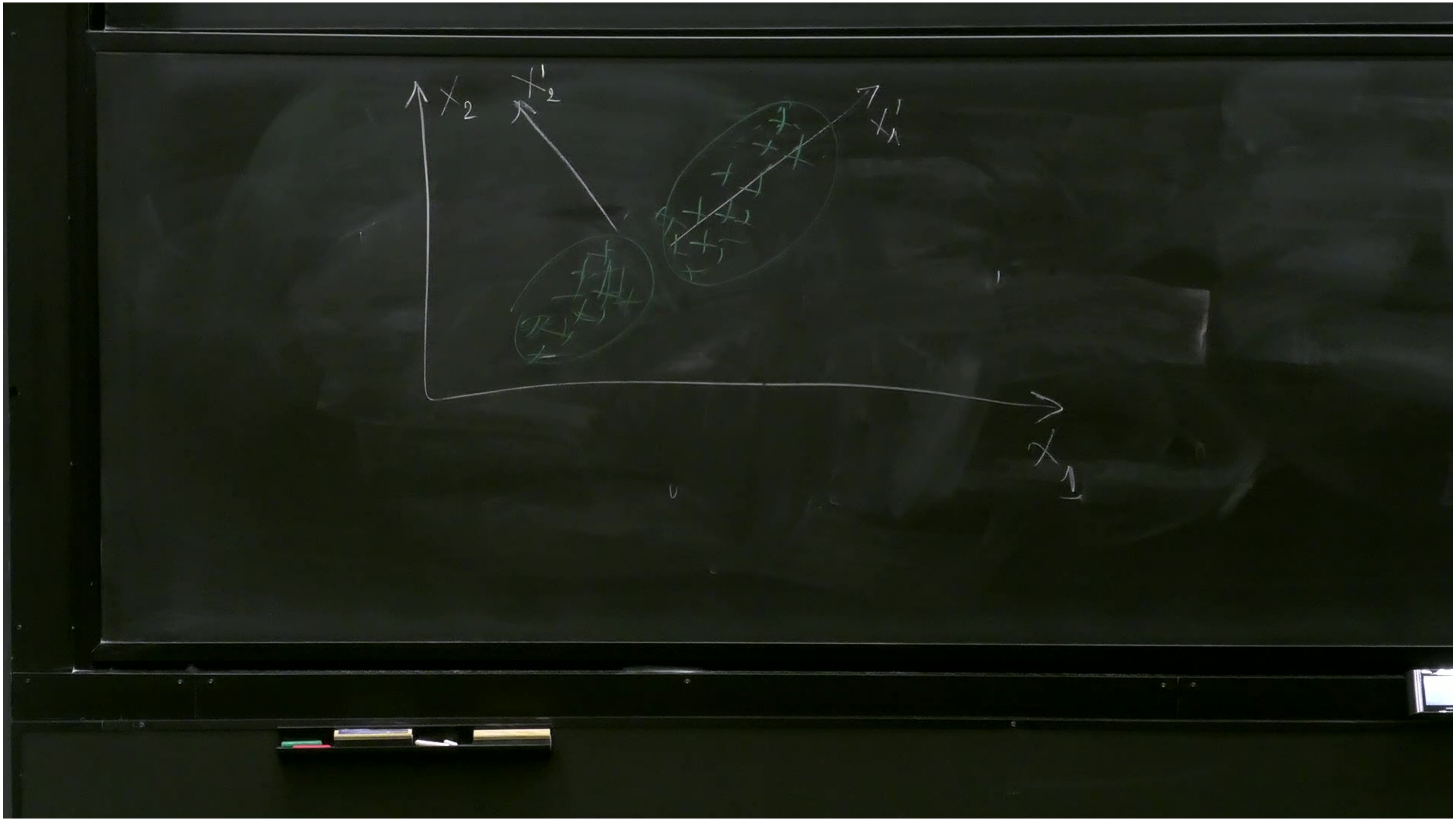

$V_{X'} = \begin{pmatrix} \lambda_1 & & (0) \\ & \lambda_2 & \\ (0) & & \ddots \; \lambda_N \end{pmatrix}$
$\lambda_j = \frac{1}{M-1} \sum_{i=1}^{M} {X'_{ij}}^2$ ← Variance of new component X_j
$\forall j = 1, 1$
$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_N$ (Without loss of generality).


PCA is useful
λi
1 2 3
N
i = 1, N
λi
1 2
N i


"Explained variance ratio":
$r_i = \frac{\lambda_i}{\sum_{i=1}^{N} \lambda_i} \longrightarrow 0 \leq r_i \leq 1$
$\lambda_1 = 1000$
$\lambda_2 = 1$
$r_1 = \frac{1000}{1001} \simeq 1$
$r_2 = \frac{1}{100} \simeq 0$



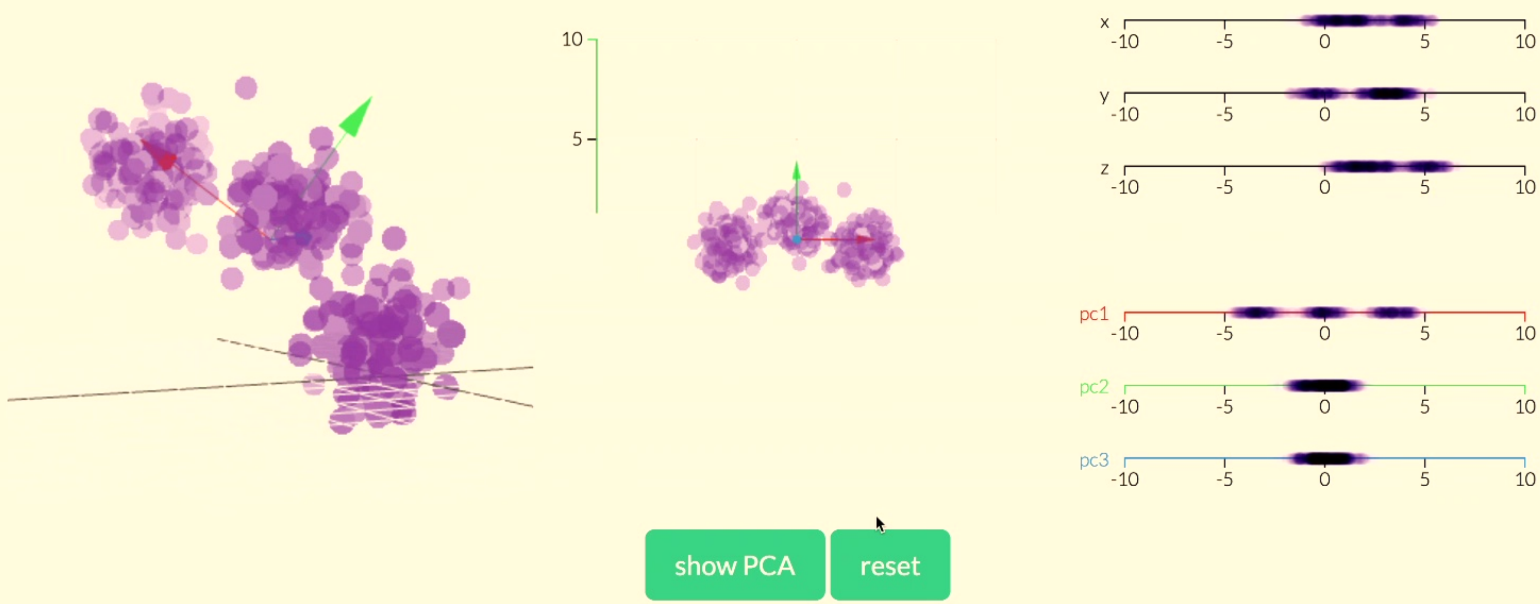
X_2'
Paramagnetic (PM) phase
Ferromagnetic (FM) phase
X_1'



x_2
x'_2
x'_1
x_1

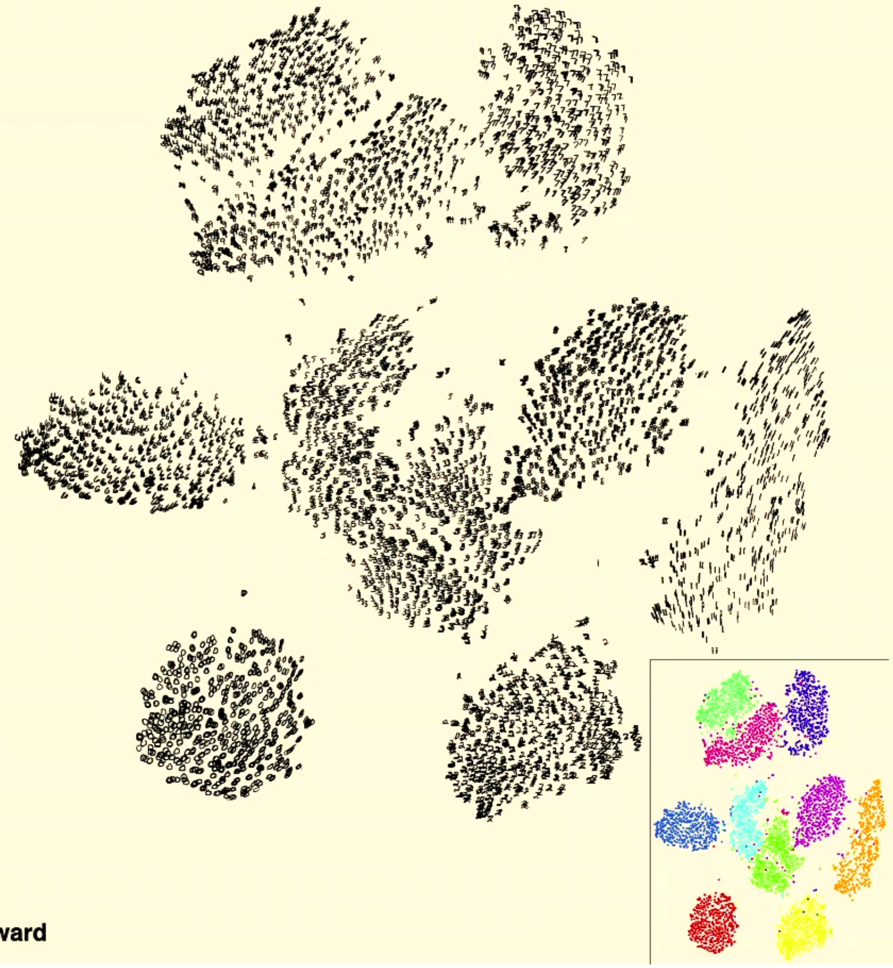
3D example

With three dimensions, PCA is more useful, because it's hard to see through a cloud of data. In the example below, the original data are plotted in 3D, but you can project the data into 2D through a transformation no different than finding a camera angle: rotate the axes to find the best angle. To see the "official" PCA transformation, click the "Show PCA" button. The PCA transformation ensures that the horizontal axis PC1 has the most variation, the vertical axis PC2 the second-most, and a third axis PC3 the least. Obviously, PC3 is the one we drop.



Dimensional reduction in 2D

van der Maaten, https://lvdmaaten.github.io/tsne



Credit: Lauren Hayward

PERIMETER SCHOLARS INTERNATIONAL

Dimensional reduction in 3D

van der Maaten, https://lvdmaaten.github.io/tsne

Credit: Lauren Hayward

PERIMETER SCHOLARS INTERNATIONAL

Dimensional reduction in 2D

Mehta et al, arXiv:1803.08823

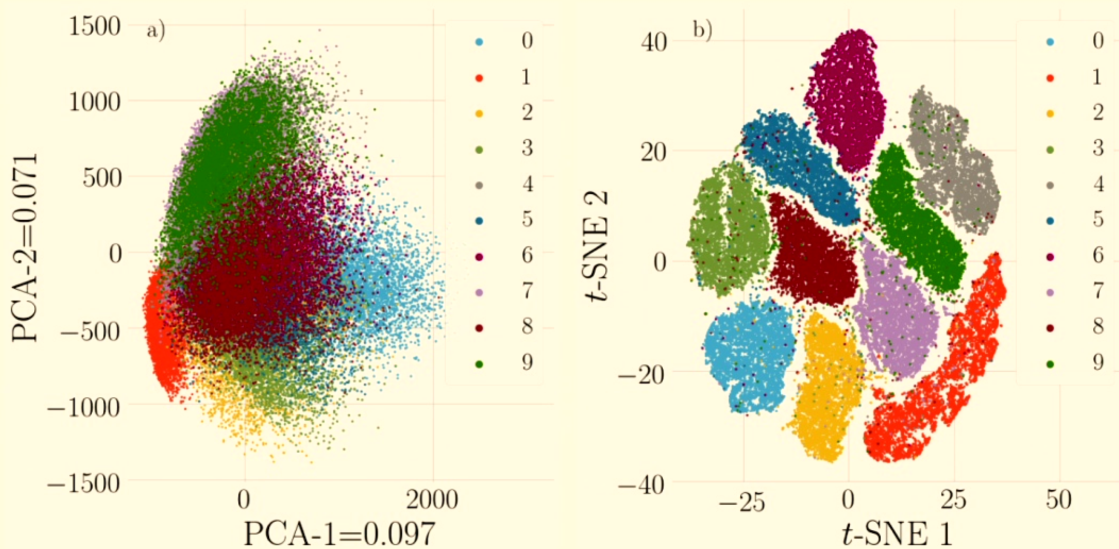


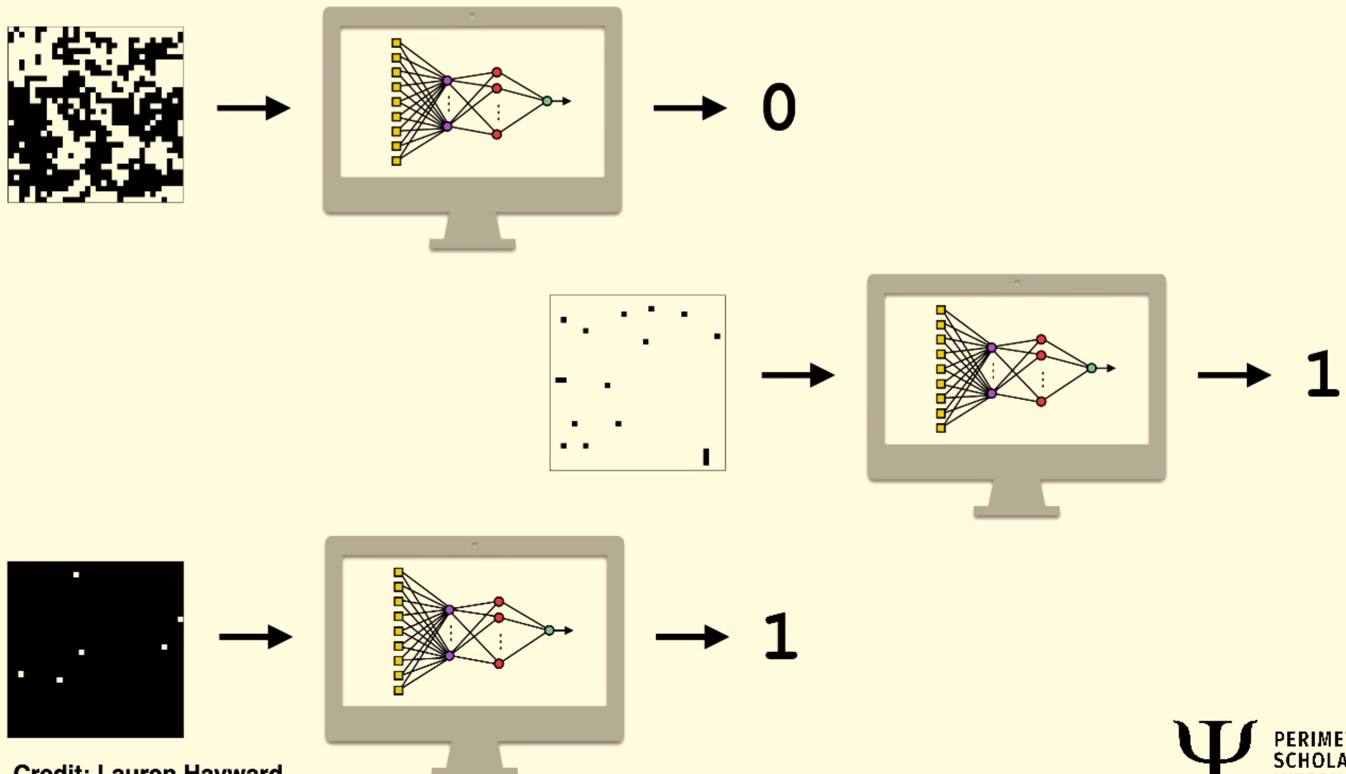
FIG. 54 Visualization of the MNIST handwritten digits training dataset (here $N = 60000$). (a) First two principal components. (b) t-SNE

Credit: Lauren Hayward

PERIMETER SCHOLARS INTERNATIONAL

Classifying phases of matter with supervised learning

Two-dimensional Ising model:

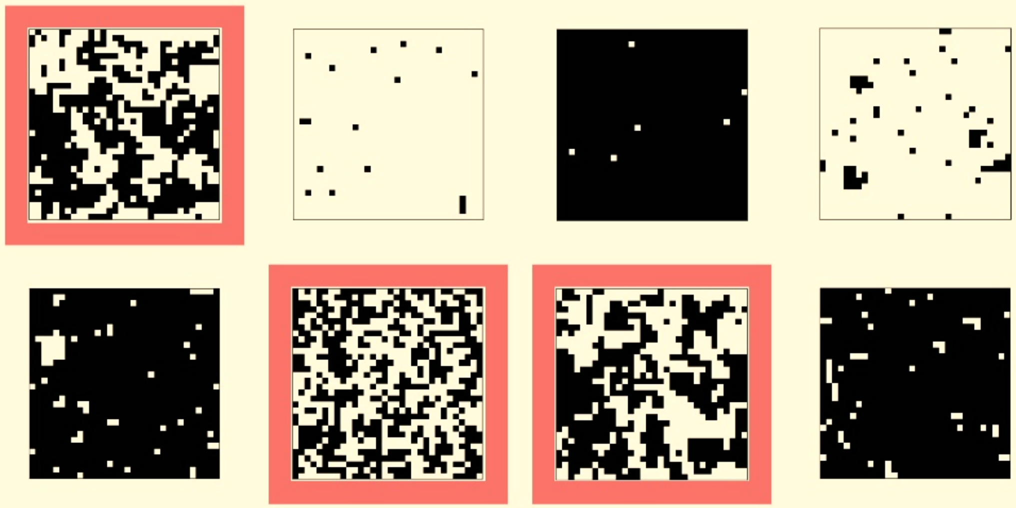


Credit: Lauren Hayward

PERIMETER SCHOLARS INTERNATIONAL

Identifying phases of matter without labels

Two-dimensional Ising model:



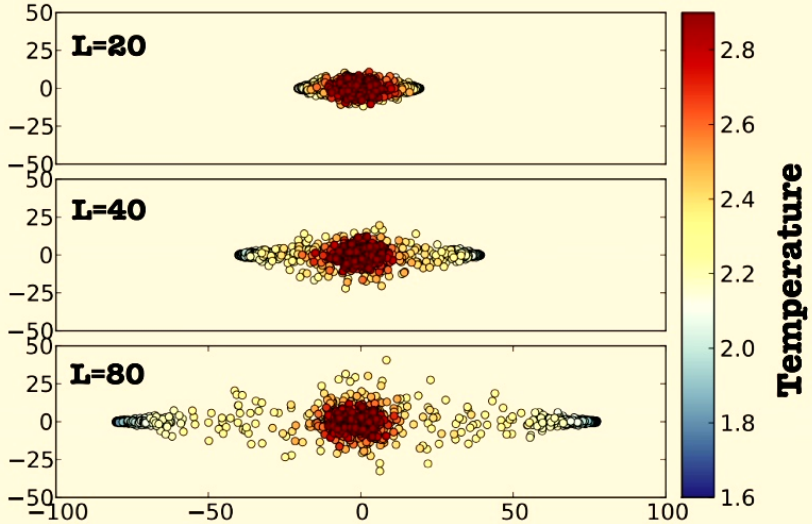
Group #1

Credit: Lauren Hayward

PERIMETER SCHOLARS INTERNATIONAL

Identifying phases of matter with PCA

Two-dimensional Ising model:



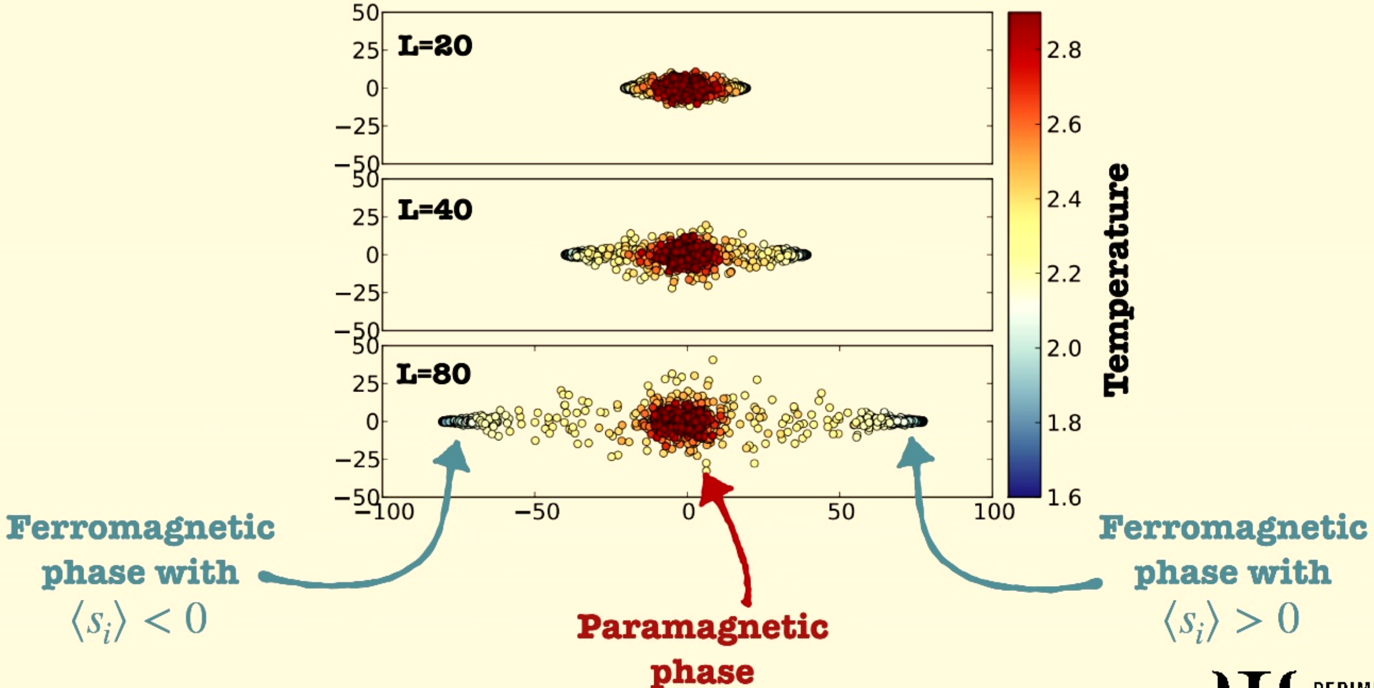
L. Wang, Phys. Rev. B **94**, 195105 (2016)

Credit: Lauren Hayward

PERIMETER SCHOLARS INTERNATIONAL

Identifying phases of matter with PCA

Two-dimensional Ising model:



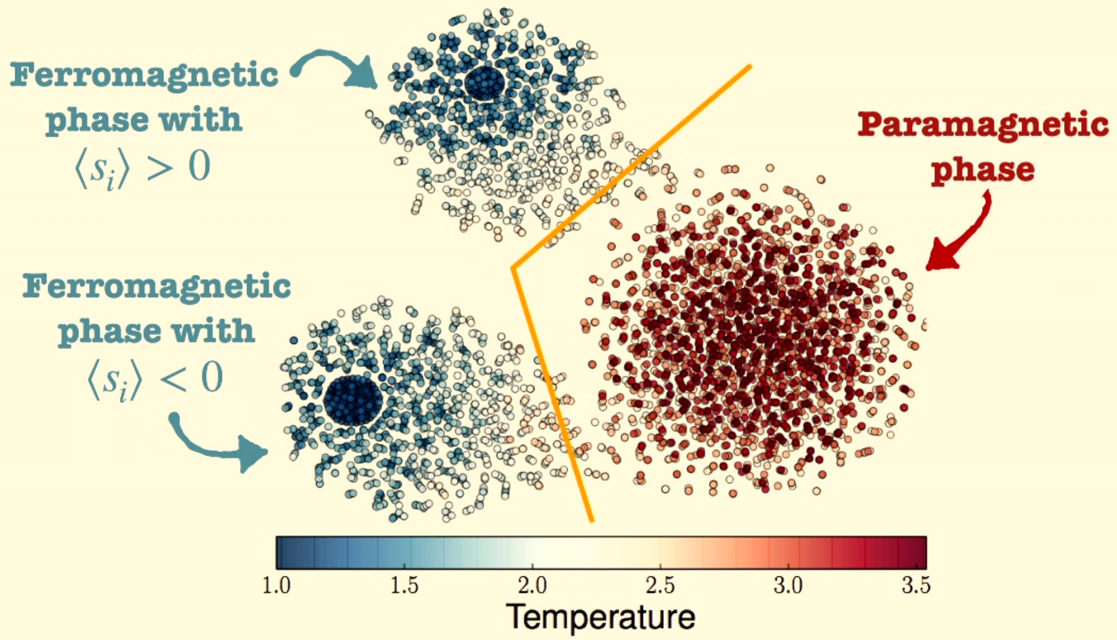
L. Wang, Phys. Rev. B **94**, 195105 (2016)

Credit: Lauren Hayward

PERIMETER SCHOLARS INTERNATIONAL

Identifying phases of matter with t-SNE

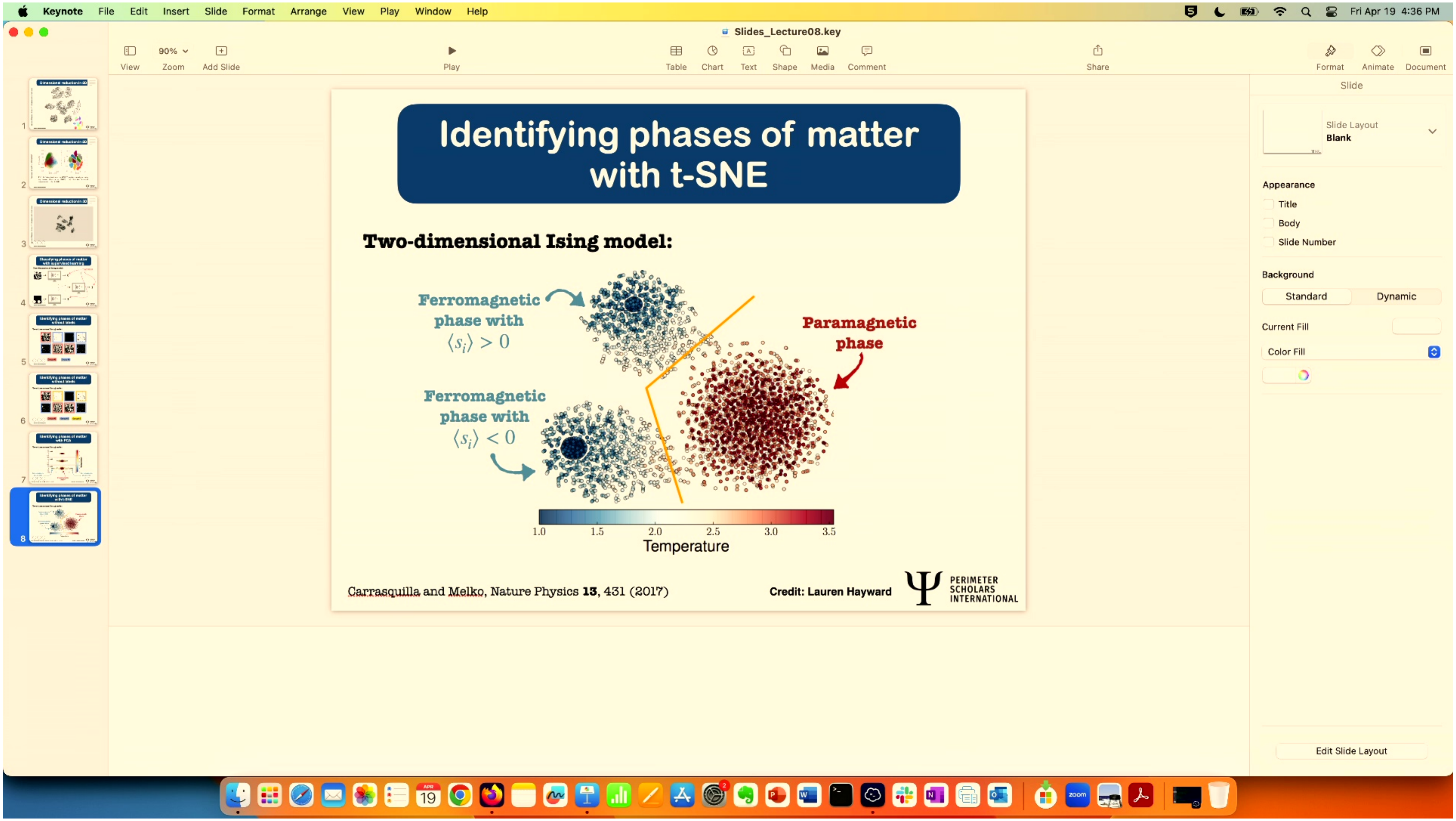
Two-dimensional Ising model:



Carrasquilla and Melko, Nature Physics **13**, 431 (2017)

Credit: Lauren Hayward

PERIMETER SCHOLARS INTERNATIONAL


Keynote File Edit Insert Slide Format Arrange View Play Window Help
Fri Apr 19 4:36 PM
Slides_Lecture08.key
Identifying phases of matter with t-SNE
Two-dimensional Ising model:
Ferromagnetic phase with $\langle s_i \rangle > 0$
Ferromagnetic phase with $\langle s_i \rangle < 0$
Paramagnetic phase
1.0 1.5 2.0 2.5 3.0 3.5
Temperature
Carrasquilla and Melko, Nature Physics 13, 431 (2017)
Credit: Lauren Hayward
PERIMETER SCHOLARS INTERNATIONAL
Slide
Slide Layout
Blank
Appearance
Title
Body
Slide Number
Background
Standard
Dynamic
Current Fill
Color Fill
Edit Slide Layout