INDEPENDENT COMPONENT ANALYSIS

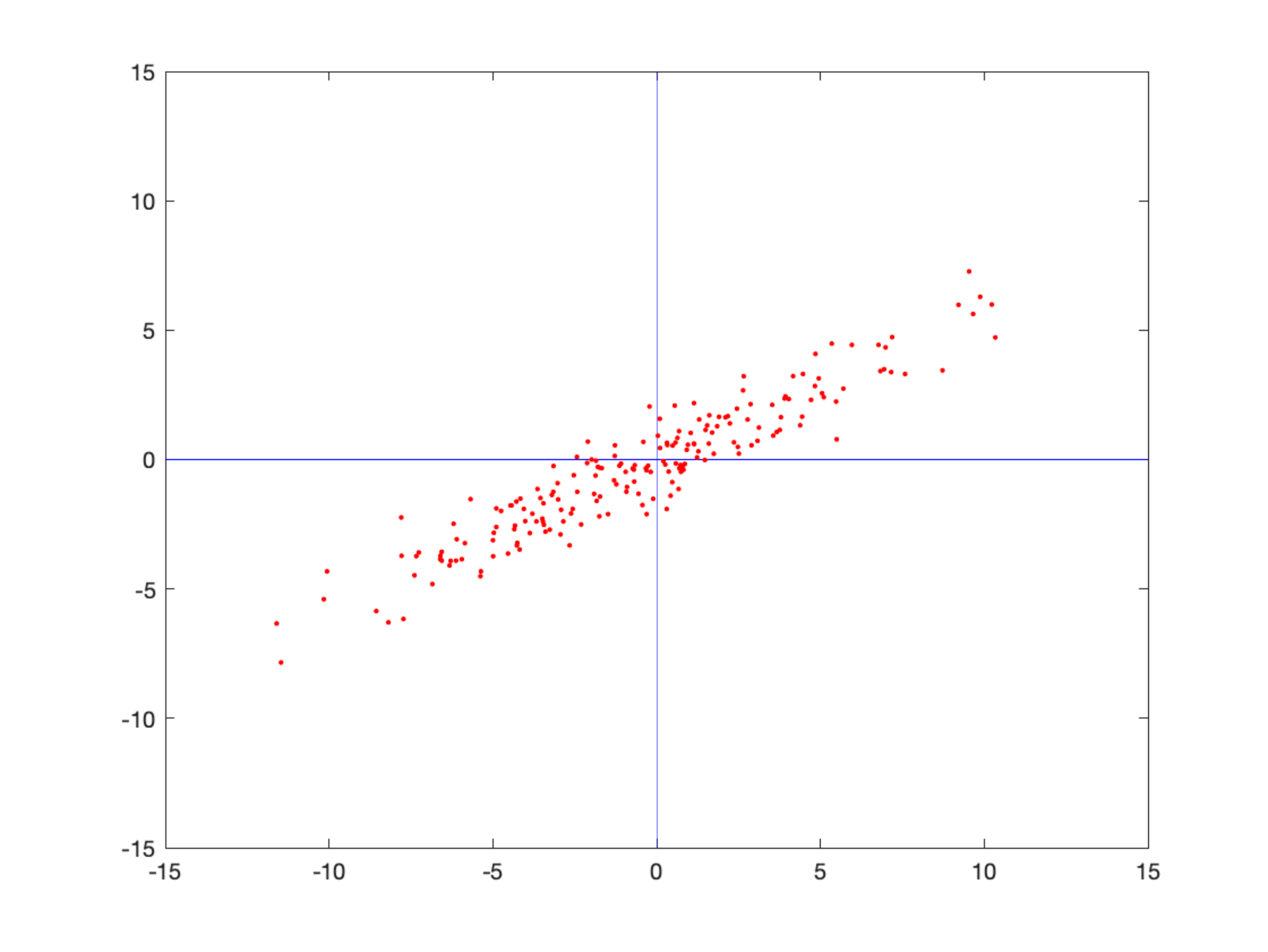
Dr. Keith Evan Schubert

PCA REVIEW

- Independent Gaussian variables
- Linear combinations
- Want:
  - lowest dimensional basis
  - Maximum information/variance
- Solution:
  - Zero mean
  - Take the SVD
  - If variables are in the columns then
    - Right singular vectors will form the correct basis
    - Right singular vectors project to the basis

PCA EXAMPLE

- 2 sensors each take n data points (say x and y)
- Find principal components and rotate the data



PCA EXAMPLE CON'T

- $A = \left[xy\right]$, n=200

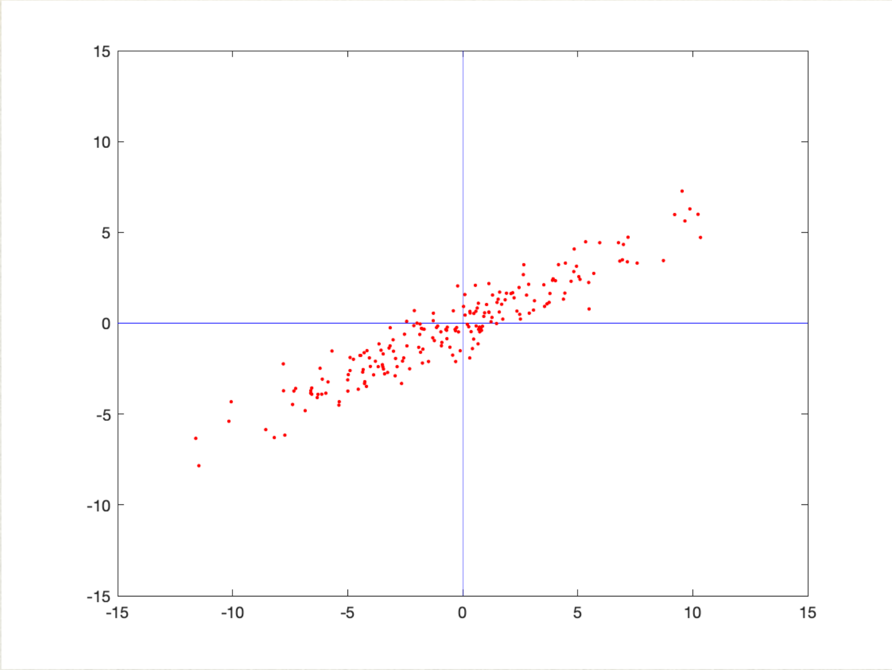
- $$A = \begin{bmatrix} -1.8499 & -0.0439 \\ 1.0413 & 1.0274 \\ -6.5603 & -3.5638 \\ \vdots & \vdots \end{bmatrix}$$

- $U\Sigma V^T = svd(A)$

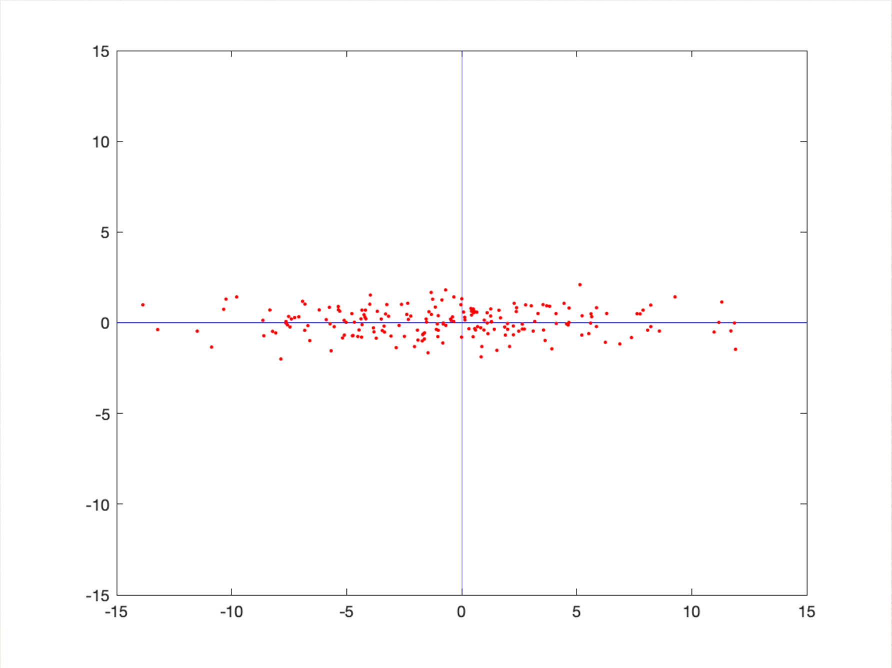
- $V = \begin{bmatrix} 0.8634 & -0.5044 \\ 0.5044 & 0.8634 \end{bmatrix}, \theta \approx 30.29^o$

PCA EXAMPLE CON'T

- Plot A

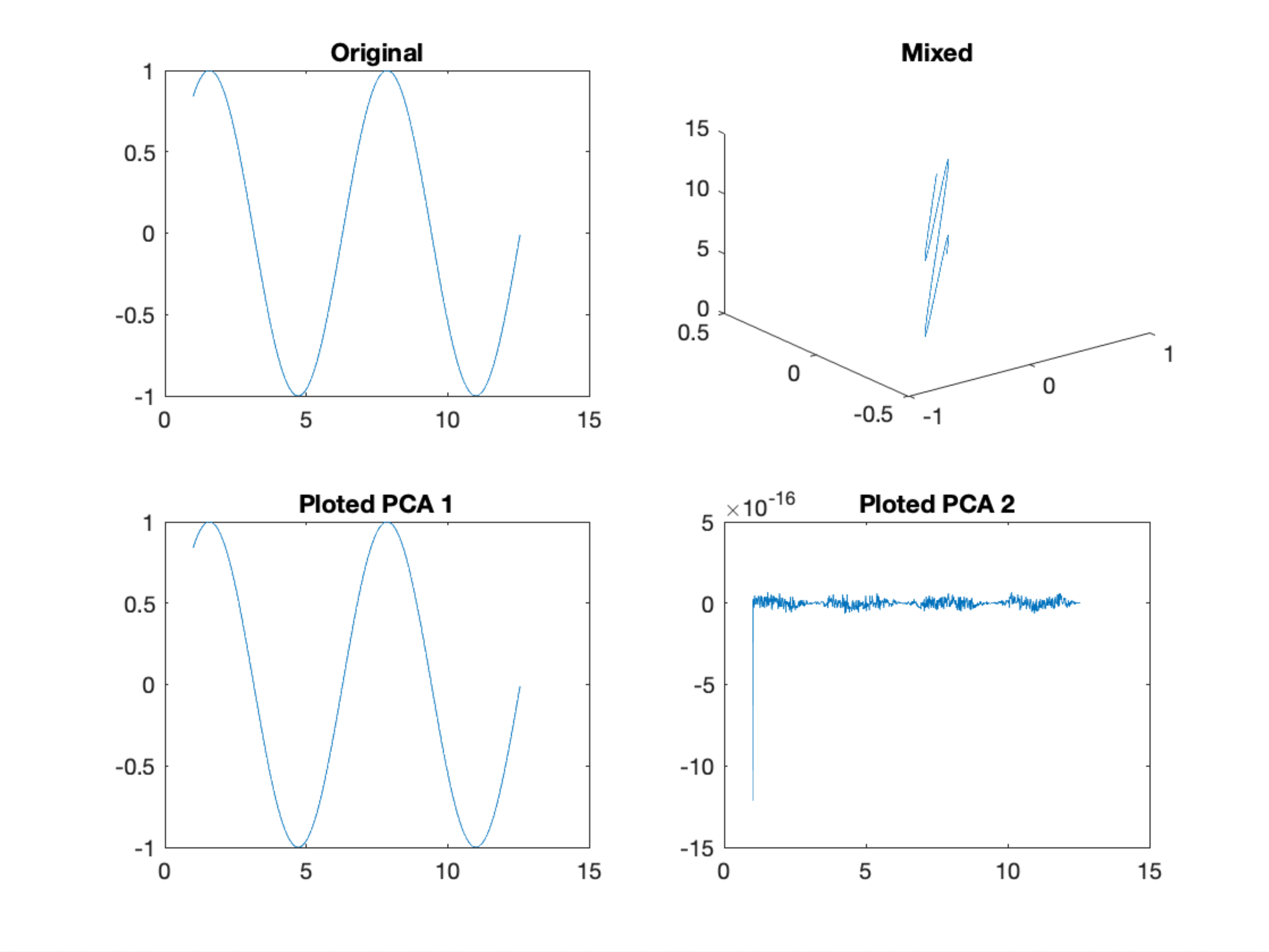


- Plot $Av = U\Sigma$



SECOND PCA EXAMPLE

- Rotated sinusoid



INDEPENDENT COMPONENTS

- Data sources are independent random variables
- Non-gaussian distributions
- Measurements are (nonlinear) sums of sources
- Want sources and mixing

HOW?

- Zero mean
- Whiten the data
- Reduce the dimension of data

say with PCA

- Pre-conditioners
- Iterative algorithm to separate based on a cost function
- Two families:
  - Minimize mutual information
  - Maximize non-gaussian

MINIMIZE MUTUAL INFORMATION

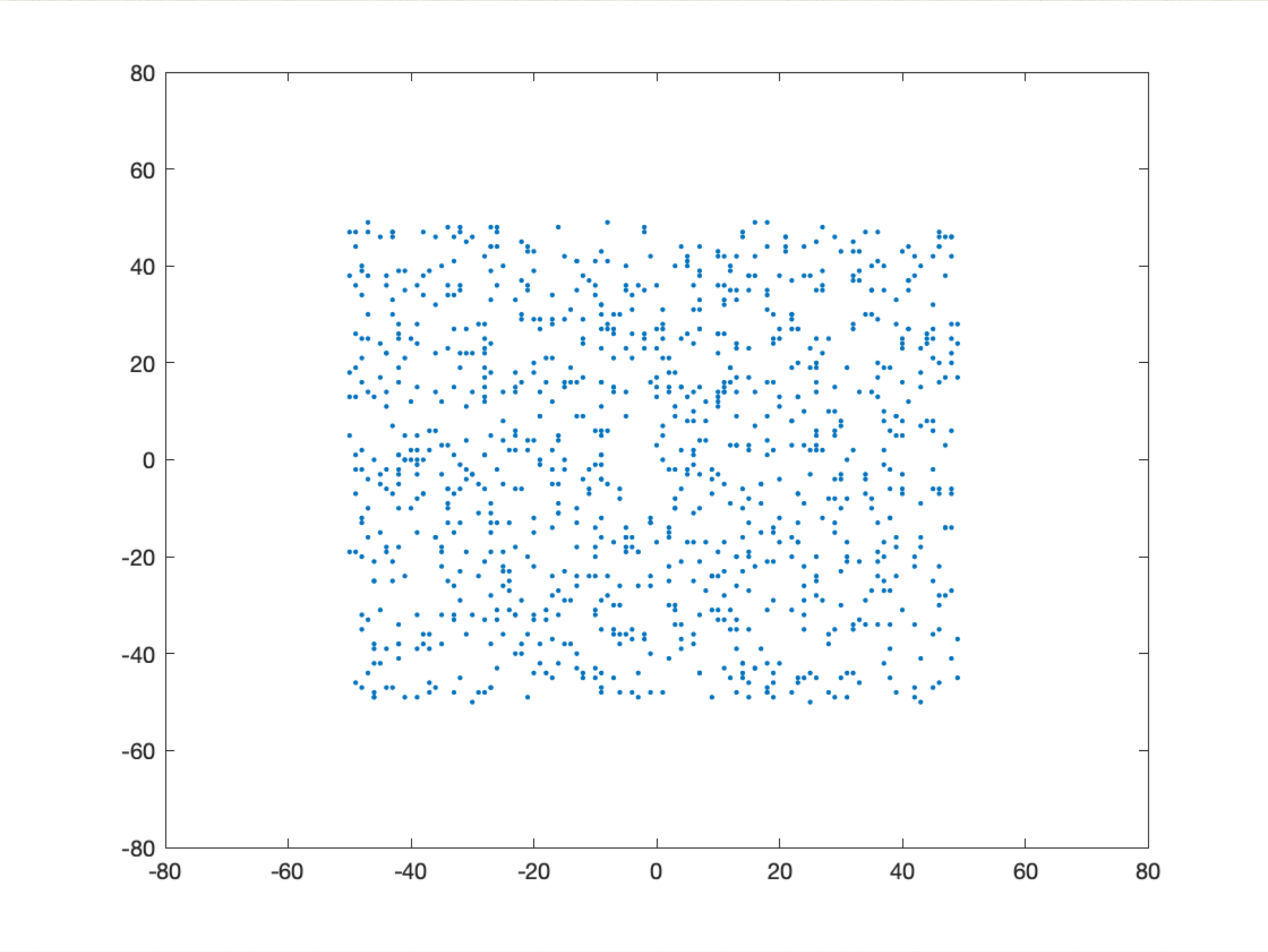
- Available data
  - Distributions that might fit
  - Find the one that maximizes either
- Kullback-Leiber Divergence (Relative Entropy)
  - $D_{KL}(P \| Q) = -\sum_{x \in X} P(x) \log \left(\frac{Q(x)}{P(x)} \right)$
- Maximum Entropy
  - $S = -\sum_{i} P_i \log P_i$

MAXIMIZE NON-GAUSSIAN

- Kurtosis
  - scaled fourth moment
  - Fatness (positive) or thinness (negative) of tail
  - $Kurt[X] = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right]$
- Negentropy
  - Distance to normal (Gaussian)
  - Entropy of Gaussian with same mean and variance minus entropy of distribution

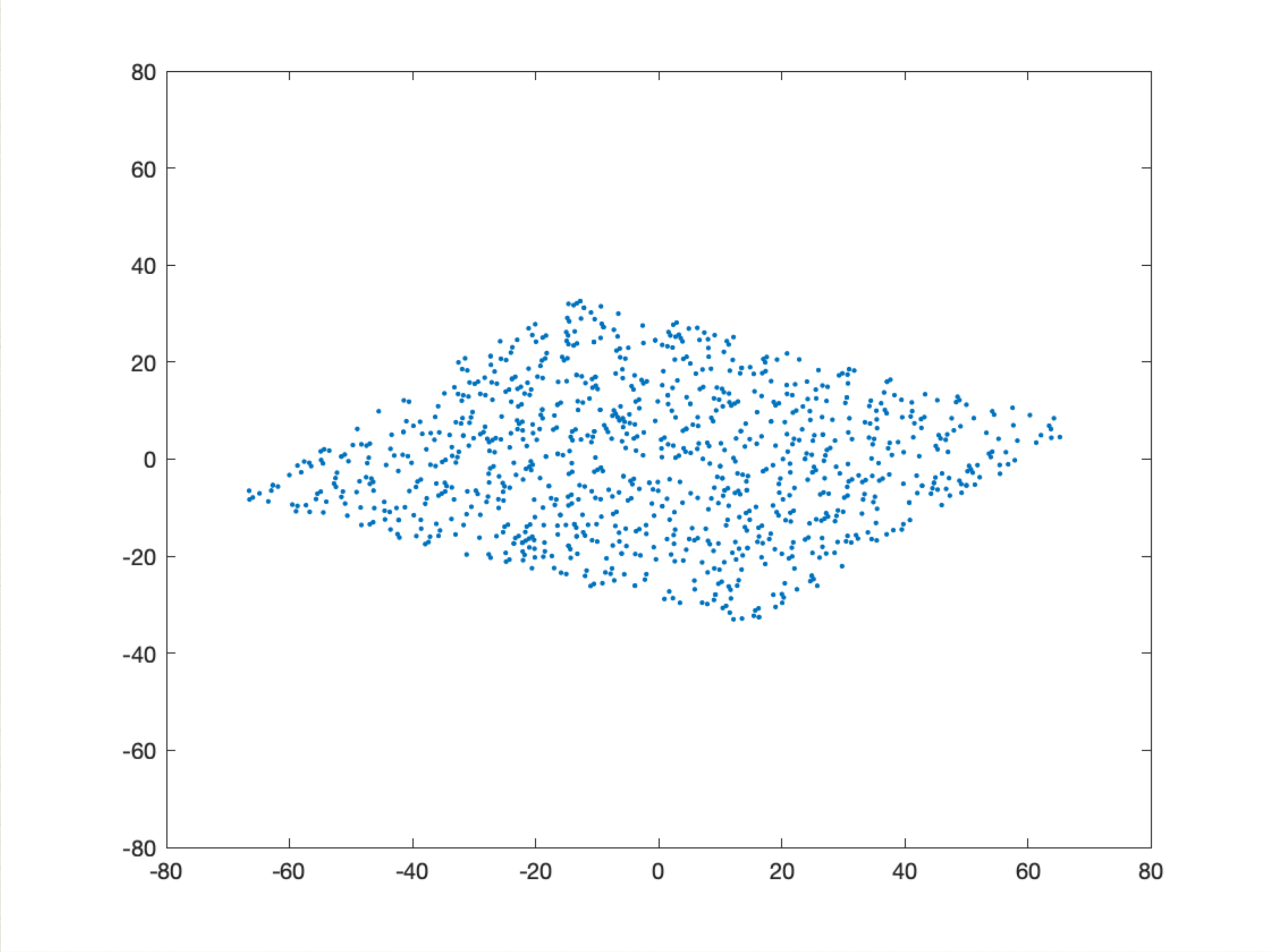
SLOW WALKTHROUGH

- Take x and y from independent uniform variables



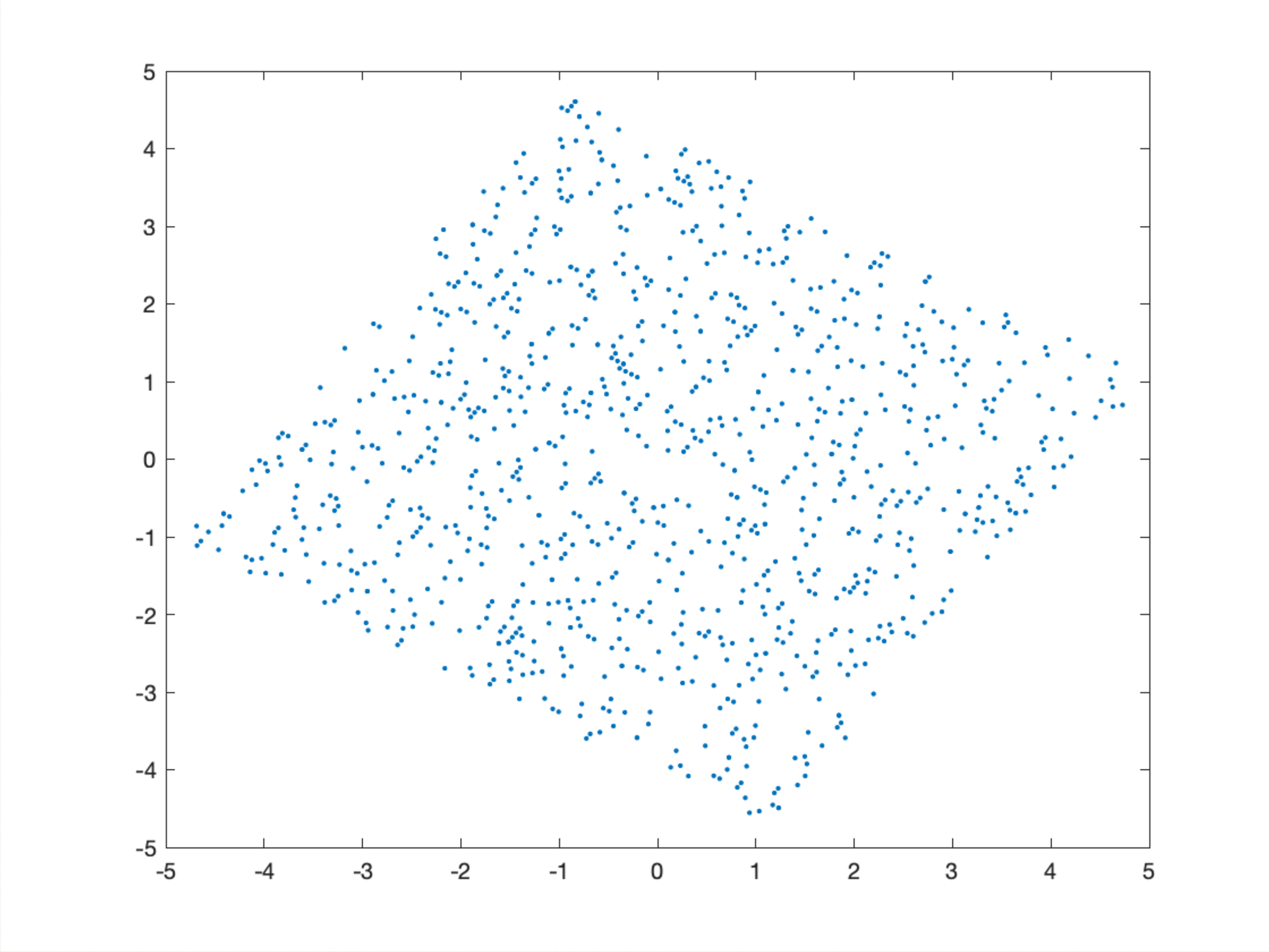
SLOW WALKTHROUGH

- Mix
- How do we recover



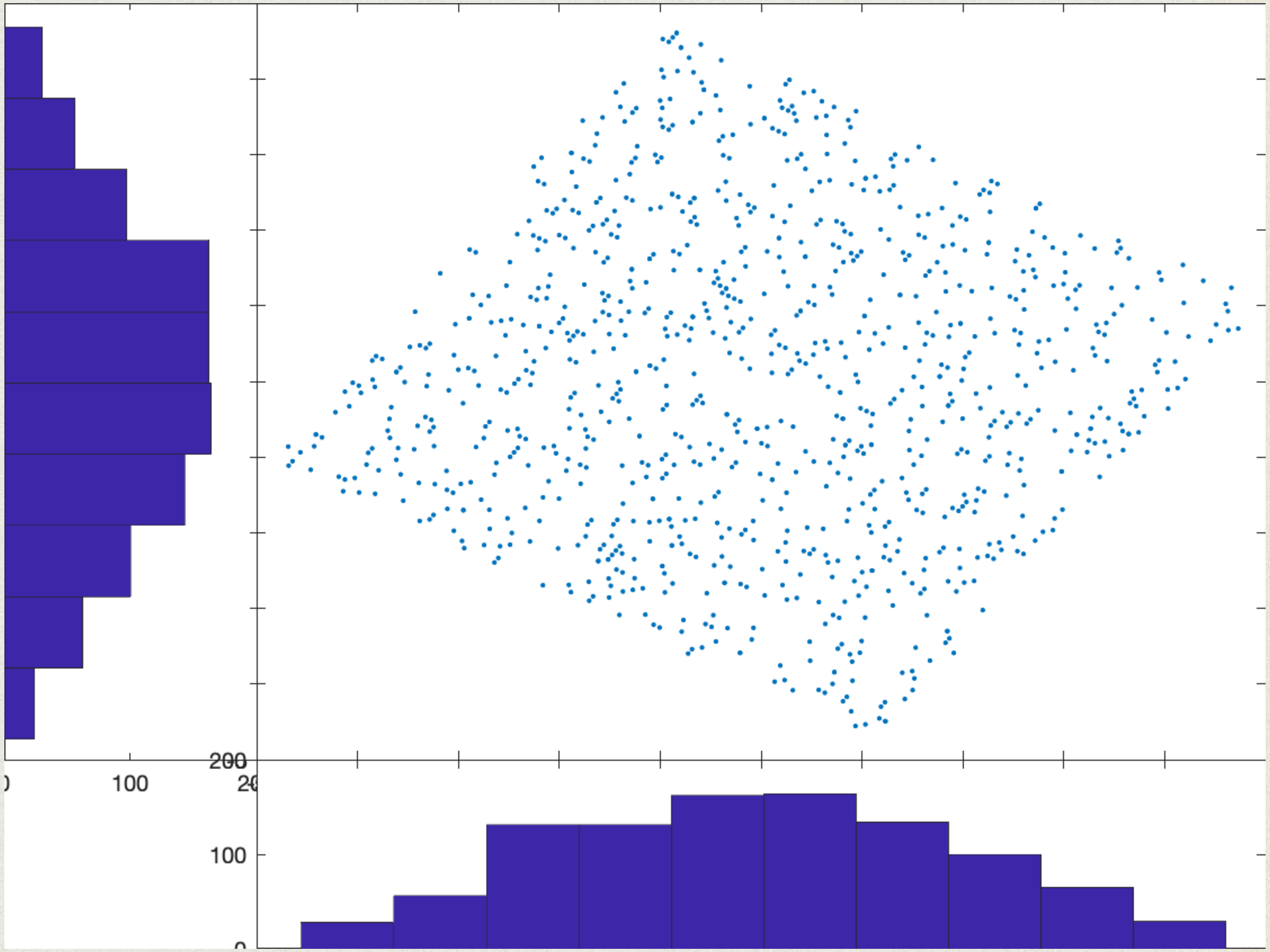
SLOW WALKTHROUGH

- Zero mean
- Whiten

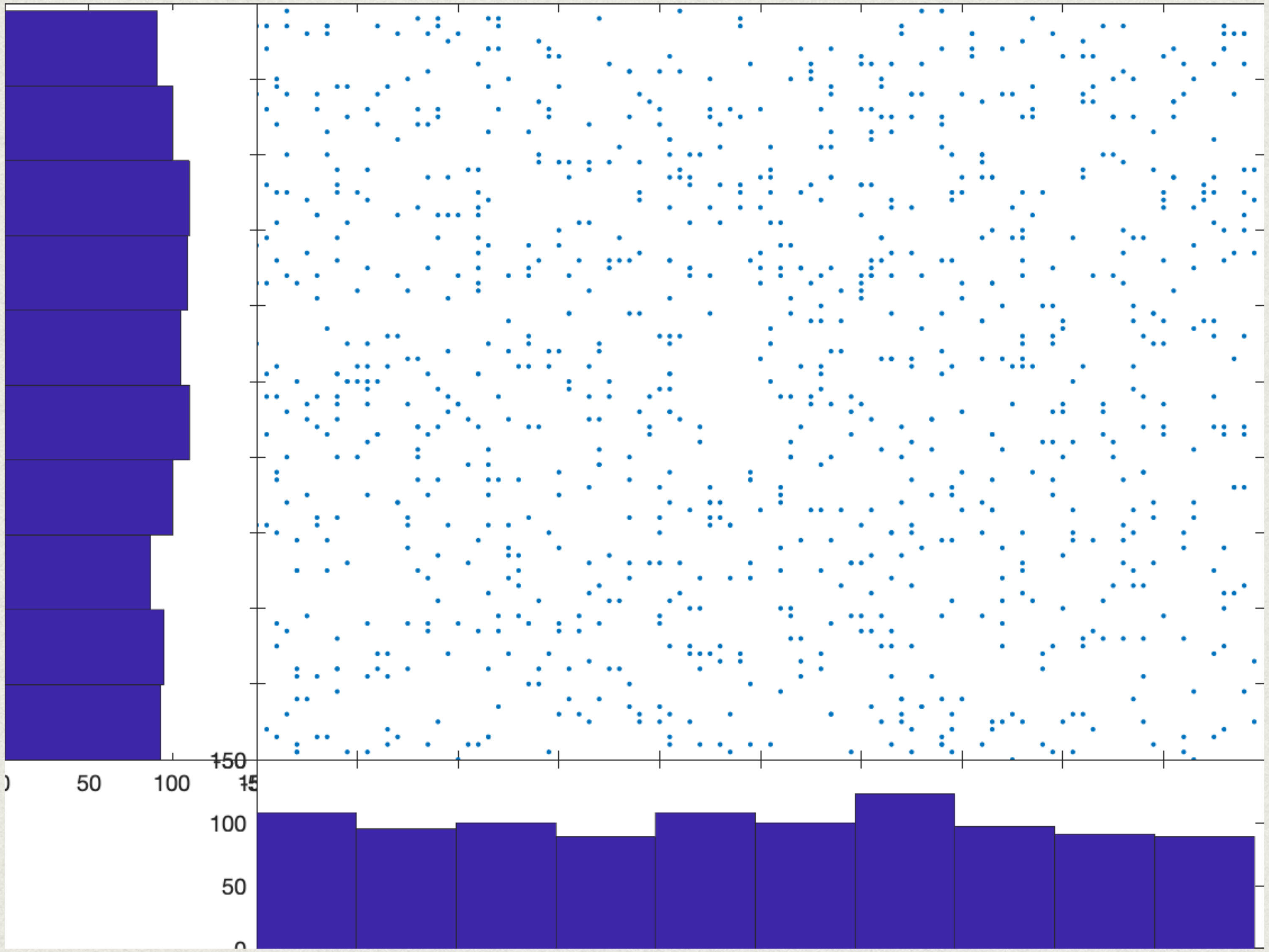


SLOW WALKTHROUGH

- Make non-gaussian
- What does it look like now?

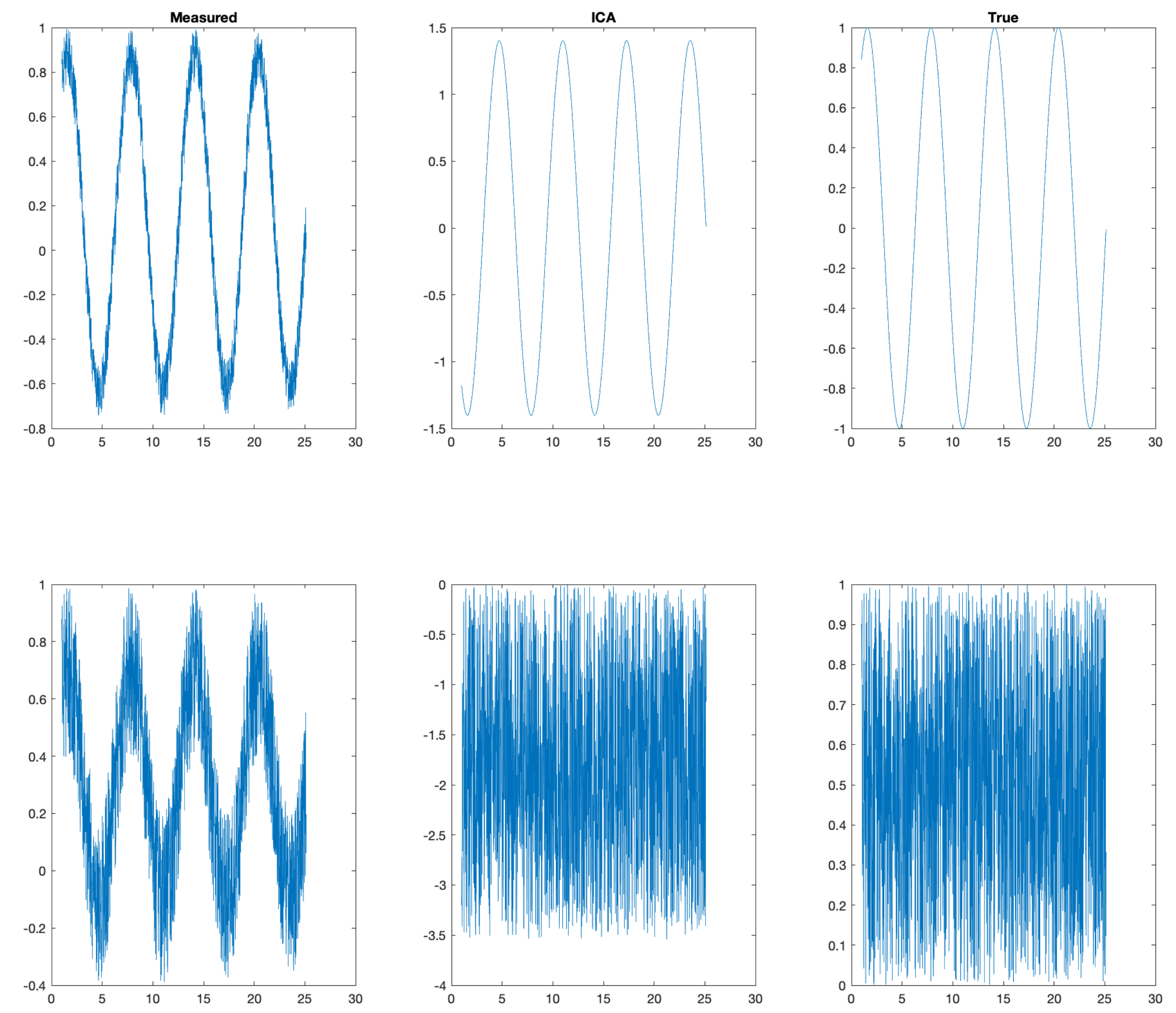


- Rotate till you get the most non-gaussian



SIMPLE ICA EXAMPLE

- Two signals get mixed



ONE LAST EXAMPLE

- Sinusoid and sawtooth