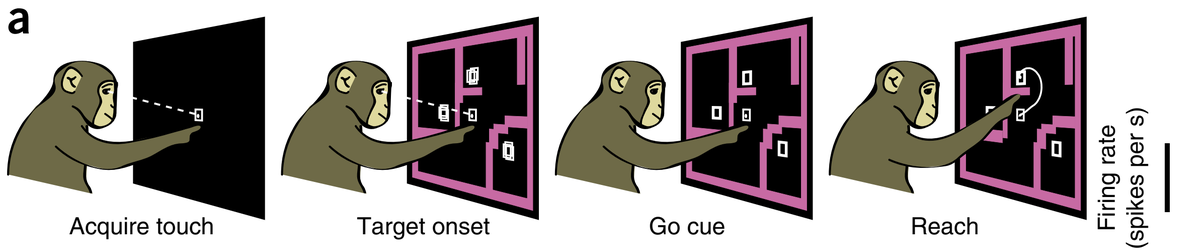
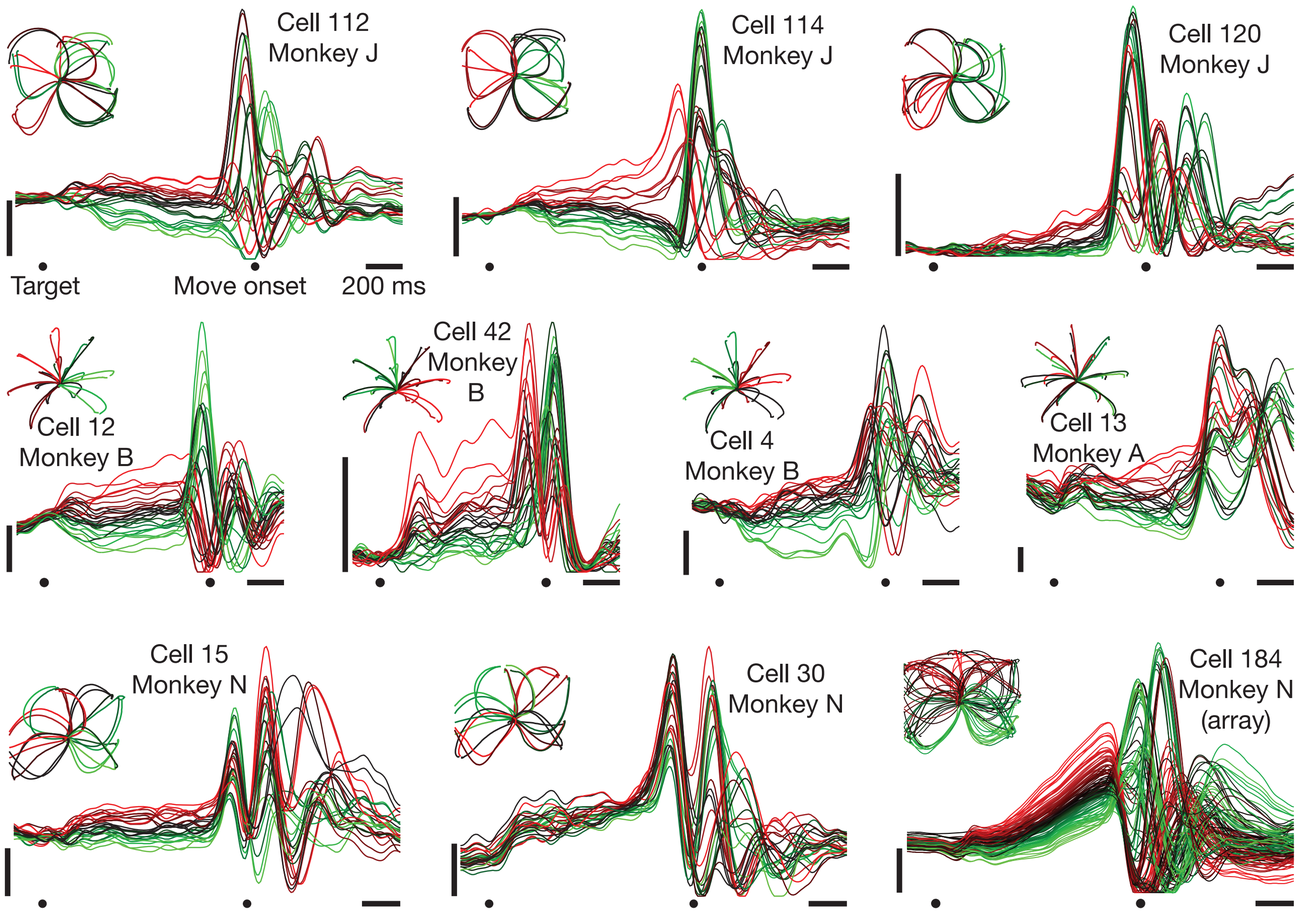
Geometry of Population Neural Activity and Analysis Methods

Quan Wen
Nov 3

Monkey reaching task (revisited)

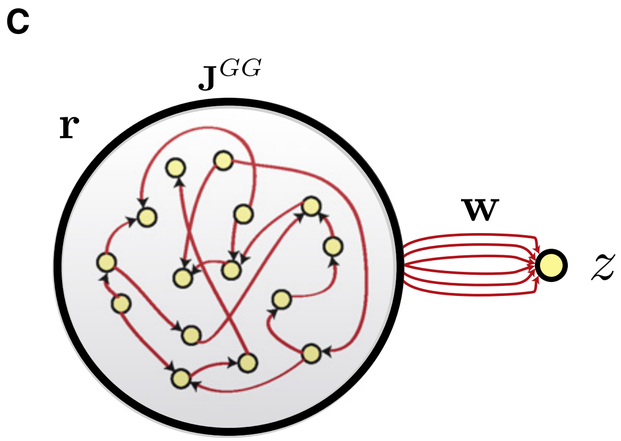


Complicated response patterns



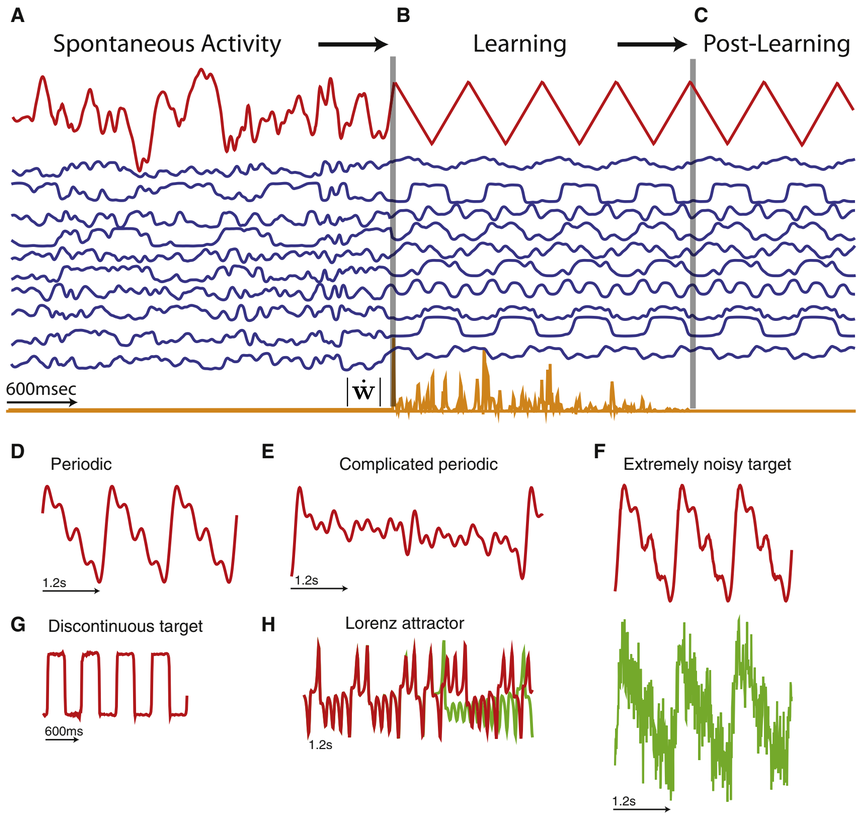
Churchland et al 2012 Nature

Generate coherent motor patterns from chaotic neural network (reservoir computing)

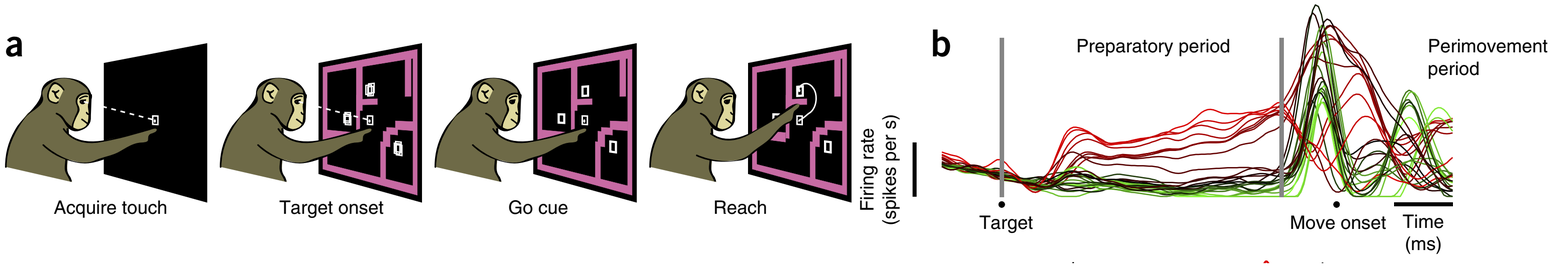


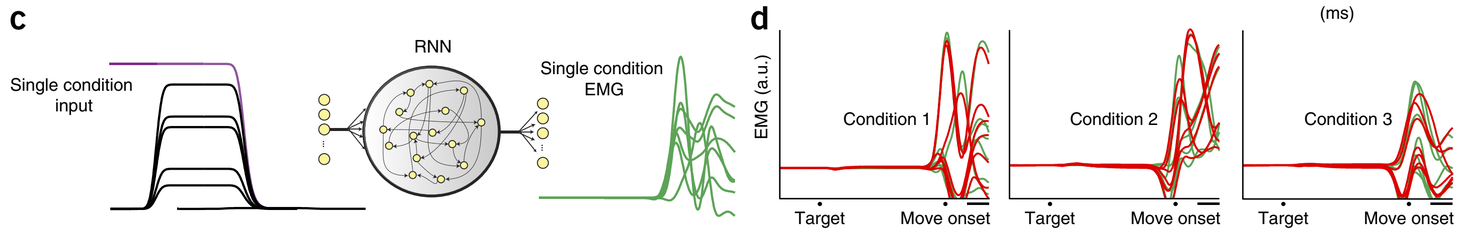
$$\tau \frac{dx_i}{dt} = -x_i + \sum_i J_{ij} r_j$$

$$r_i = f(x_i)$$

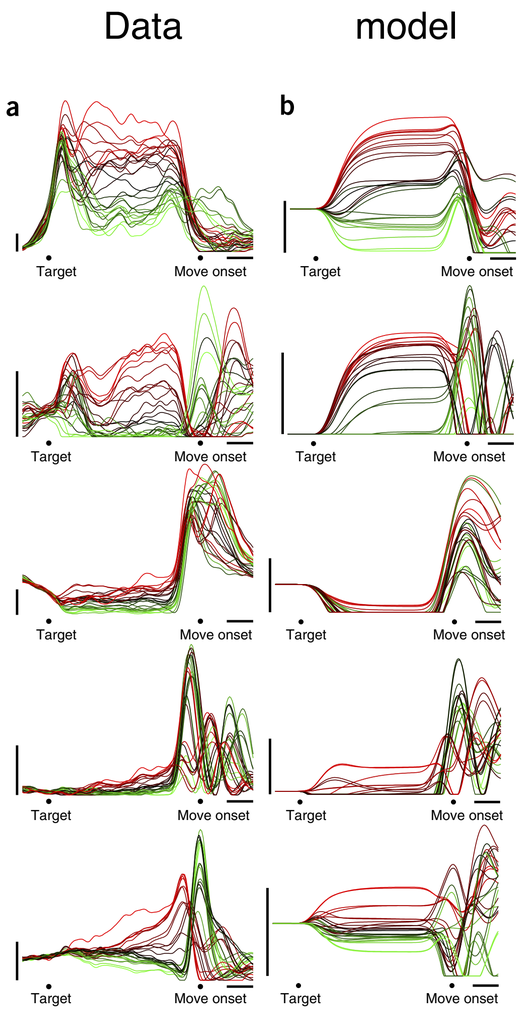


Generate coherent motor patterns from chaotic neural network

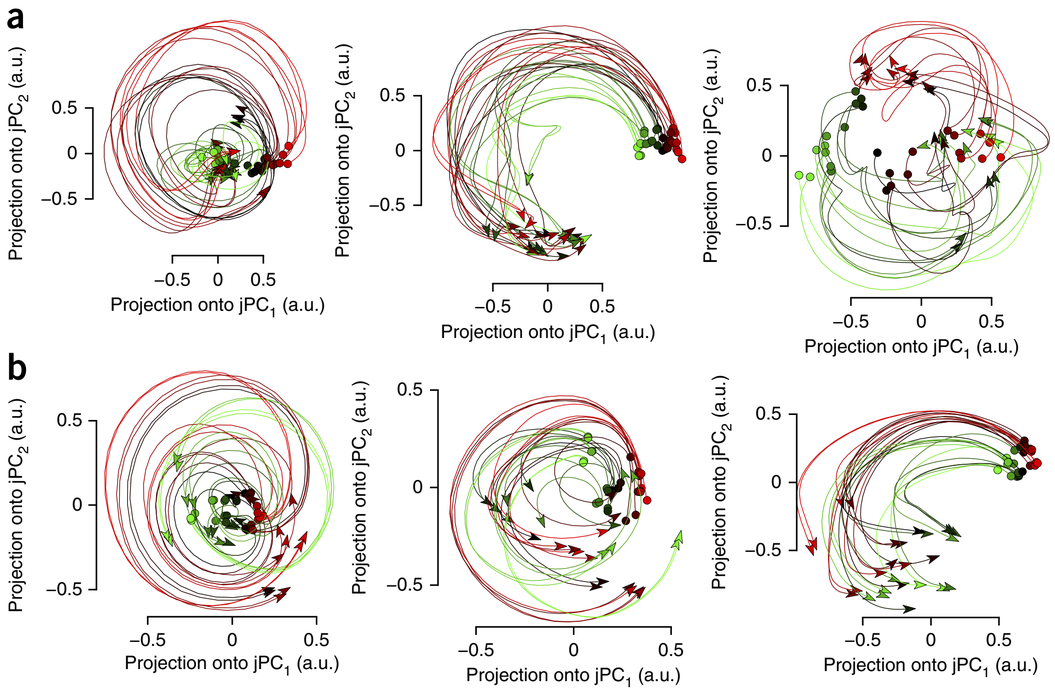




Generate coherent motor patterns from chaotic neural network



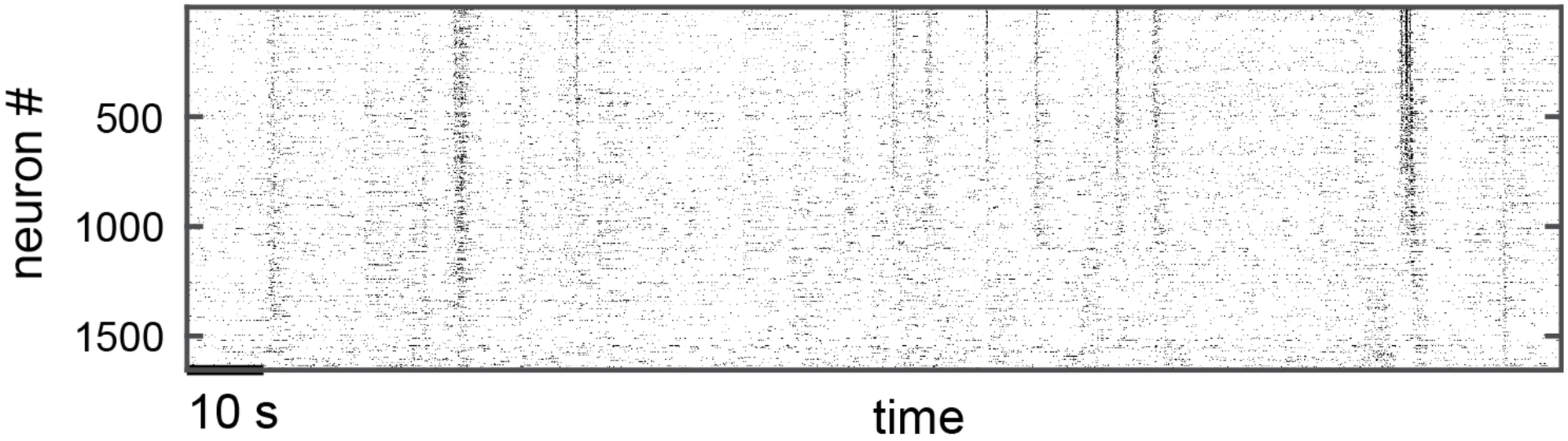
RNN model can reproduce the rotation dynamics on the population level



Two spaces

$$X \in \mathbb{R}^{N \times T}$$

X is always row-centered.



Two spaces

$$X = U\Sigma V^\top \qquad \text{where } U \in \mathbb{R}^N, V \in \mathbb{R}^T$$

Neuron space (dimension N)

column space of X

spanned by U

low-rank truncation

Time space (dimension T)

row space of X

spanned by V

least-squares projection

Orthogonal projectors

$$P^2 = P,\ P^\top = P$$

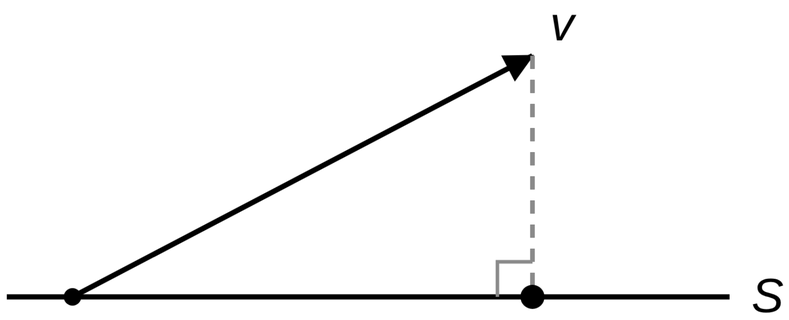
- Idempotent: projecting twice changes nothing (you are already there).
- Symmetric: this is what makes the projection orthogonal rather than merely a projection.
- For a given subspace S the orthogonal projector is unique — the basis you choose is irrelevant.

Orthogonal versus oblique

Drop the symmetry requirement and you still land in S — just not at the closest point.

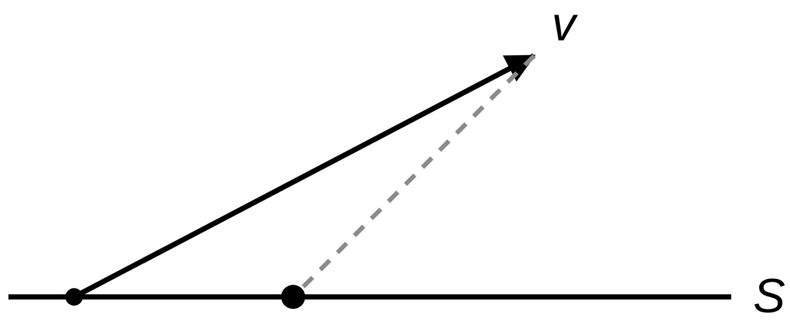
$$P^2 = P, \text{ but } P^\top \neq P$$

Orthogonal projection



residual perpendicular: closest point in S

Oblique projection



lands in S, but not the closest point

Building a projector from a basis

Columns of *A* span subspace *S* with full column rank; *Q* is an orthonormal basis for the same *S*.

$$A \in \mathbb{R}^{n \times k} \; n > k$$

$$P_s = A(A^\top A)^{-1} A^\top, \; P_s = QQ^\top$$

- The two formulas give the same matrix — orthonormality removes the inverse.
- $QQ^\top$ is an $n \times n$ projector; $Q^\top Q$ is the $k \times k$ identity.

What if rows of *A* span *S* with full row rank? (n < k)

Left and right, in the N × T convention

$$\underbrace{U_r U_r^\top X}_{\text{neuron space, left}} \qquad \underbrace{YP}_{\text{time space, right}}$$

- Neuron-space projectors act on the left; time-space projectors act on the right.

Least squares is a projection

One target neuron *y*, predicted from the population *X* by weights *b*.

$$\min_b \|y - bX\|^2 \qquad y \in \mathbb{R}^{1\times T}, b \in \mathbb{R}^{1\times N}$$

$$(y - bX)X^\mathsf{T} = 0 \quad \Longrightarrow \quad b\,X X^\mathsf{T} = y\,X^\mathsf{T}$$

$$\hat{y} = y\,X^\mathsf{T}(XX^\mathsf{T})^{-1}X = y\,P$$

The residual $\epsilon = y - \hat{y}$ must be orthogonal to every source neuron's time course.

Stacking the targets changes nothing

$$\min_B \|Y - BX\|^2 \qquad Y \in \mathbb{R}^{N_y \times T}, B \in \mathbb{R}^{N_y \times N}$$

$$B_{OLS} = YX^\top (XX^\top)^{-1} \qquad \hat{Y} = YP$$

$$X^+ = X^\top (XX^\top)^{-1} \;\Rightarrow\; B_{OLS} = YX^+,\ P = X^+X$$

Notation only: in this regime the pseudoinverse is another name for the expression above.

- Least squares projects each target neuron's activity onto the span of the source neurons' activity.
- Anything outside that span is unreachable, whatever weights you choose.

.

The SVD, read in neuroscience

$$X = U\Sigma V^\top = \sum_i \sigma_i u_i v_i^\top$$

- Each term is a spatial mode u_i — a weighting over neurons — times a temporal profile $v_i^\top$.
- Population activity decomposes into modes, each with its own time course.
- Squared singular values sum to the total variance, so the variance-explained curve is already in hand.

Eckart–Young

Truncating the SVD gives the best rank-r approximation in Frobenius norm.

$$X_r = U_r U_r^\top X, \quad \|X - X_r\|_F^2 = \sum_{i>r} \sigma_i^2$$

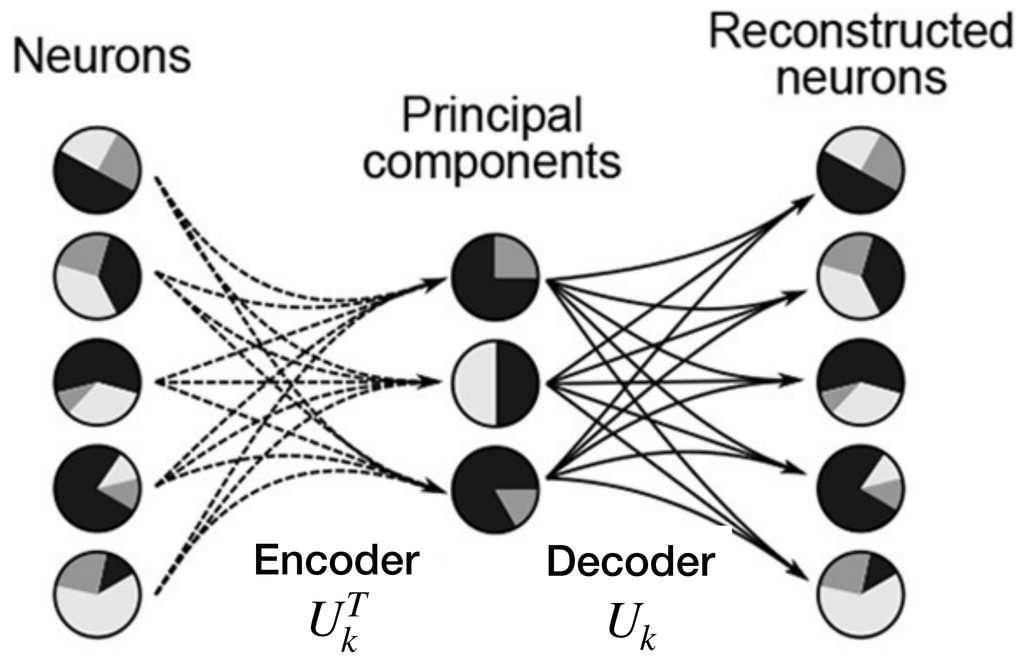
- The truncation is itself a projection, applied on the left — in neuron space.
- The discarded singular values are exactly the error.

PCA, Classical MDS, and Similarity Matching

PCA as reconstruction

A different question: which rank-k subspace loses the least when we project onto it?

$$\min_{U_k} \|X - U_k U_k^\top X\|_F^2$$



Encode the latent state

$$Y = U_K^\top X = \Sigma_k V_k^\top \in \mathbb{R}^{k \times T}$$

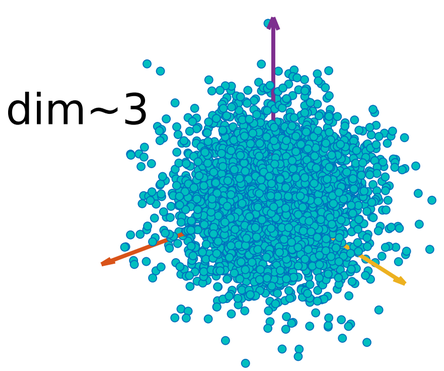
- The latents come out as k × T.
- Equivalently: the top-k row of the SVD, scaled by the singular values σ_k
- Eckart–Young says this subspace is optimal.

How many dimensions does neural activity fill up?

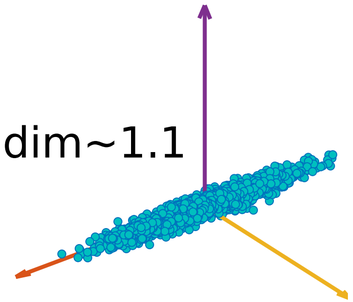
$$C = \frac{1}{T} X X^\top = \frac{1}{T} U \Sigma^2 U^\top = \sum_i \sigma_i^2 u_i u_i^\top$$

$$\lambda_i = \sigma_i^2$$

$$D = \frac{(\sum_i \lambda_i)^2}{\sum_i \lambda_i^2}$$



$\lambda_1 = 1/3$
$\lambda_2 = 1/3$
$\lambda_3 = 1/3$
$\lambda_4 = 0$
$\cdots$
$\lambda_N = 0$



$\lambda_1 = 0.95$
$\lambda_2 = 0.025$
$\lambda_3 = 0.025$
$\lambda_4 = 0$
$\cdots$
$\lambda_N = 0$

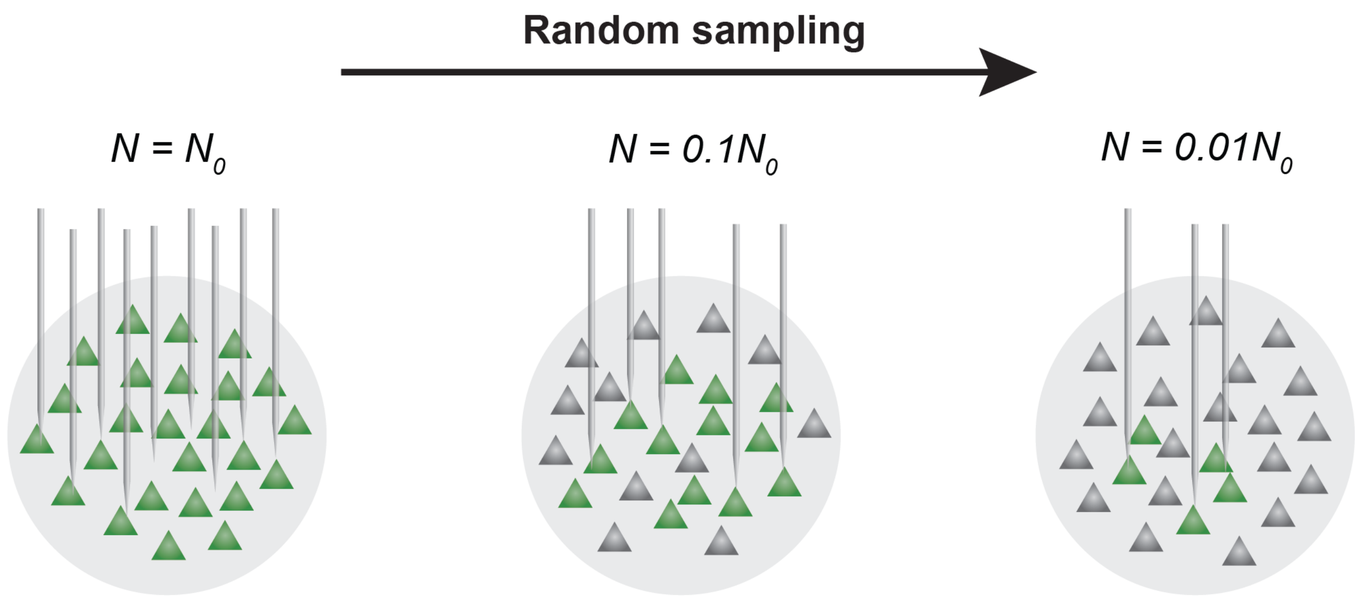
Modified from Eric Shea-Brown

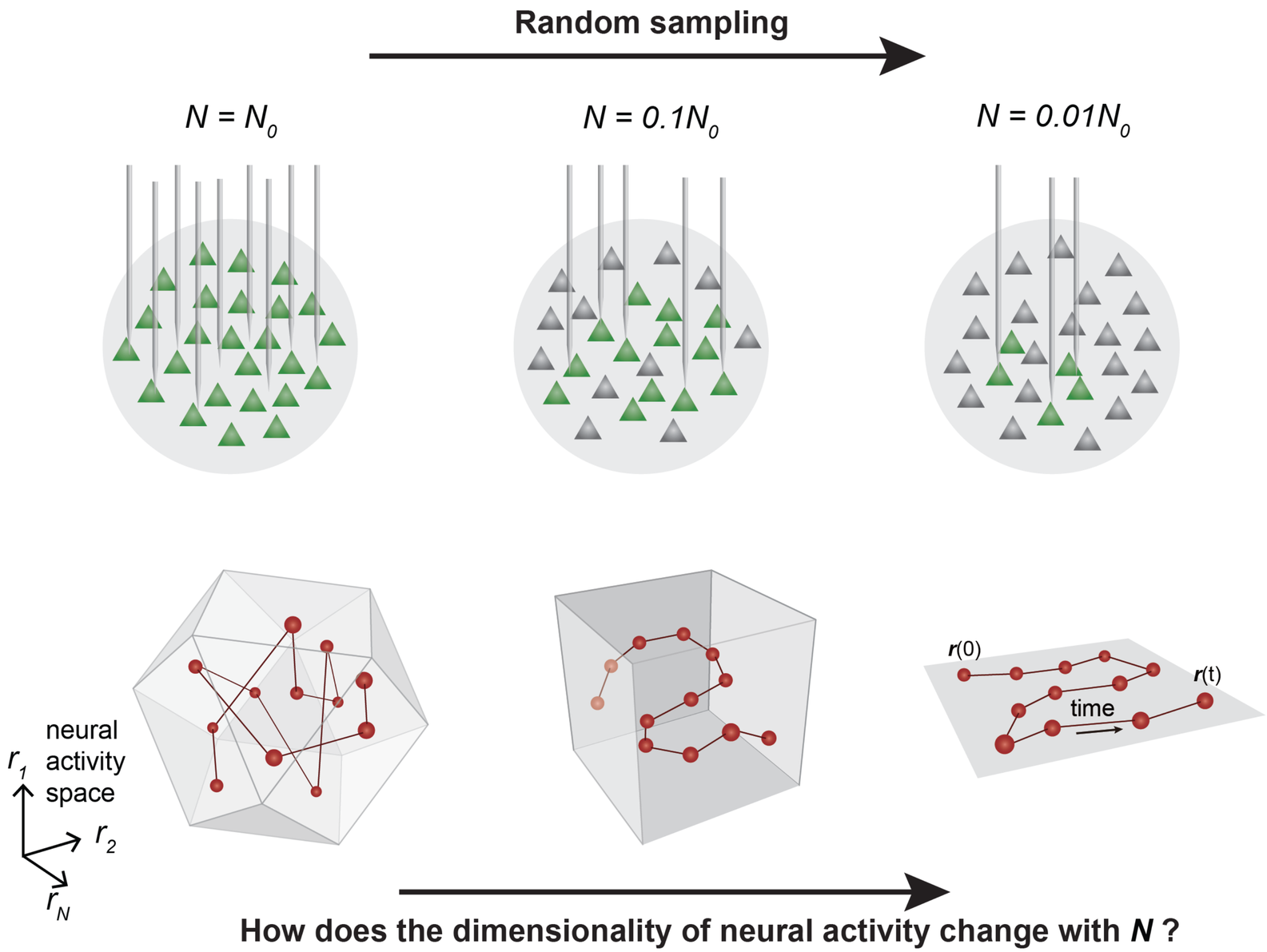
Neuron
Viewpoint

Can One Concurrently Record Electrical Spikes from Every Neuron in a Mammalian Brain?

David Kleinfeld,[1,2,13,*] Lan Luan,[3] Partha P. Mitra,[4] Jacob T. Robinson,[5] Rahul Sarpeshkar,[6,7,8,9] Kenneth Shepard,[10] Chong Xie,[3] and Timothy D. Harris[11,12,*]

Kleinfeld et al ‘19


Random sampling
$N = N_0$
$N = 0.1N_0$
$N = 0.01N_0$


Random sampling
$N = N_0$
$N = 0.1N_0$
$N = 0.01N_0$
neural activity space
r_1
r_2
r_N
r(0)
r(t)
time
How does the dimensionality of neural activity change with N ?

How does the dimensionality *D* change with *N?*

$$D = \frac{N^2 \mathbb{E}(\sigma_i^2)^2}{N\mathbb{E}(\sigma_i^4) + N(N-1)\mathbb{E}_{i \neq j}(C_{ij}^2)}$$

$$\lim_{N \to \infty} D = \frac{\mathbb{E}(\sigma_i^2)^2}{\mathbb{E}_{i \neq j}(C_{ij}^2)}$$

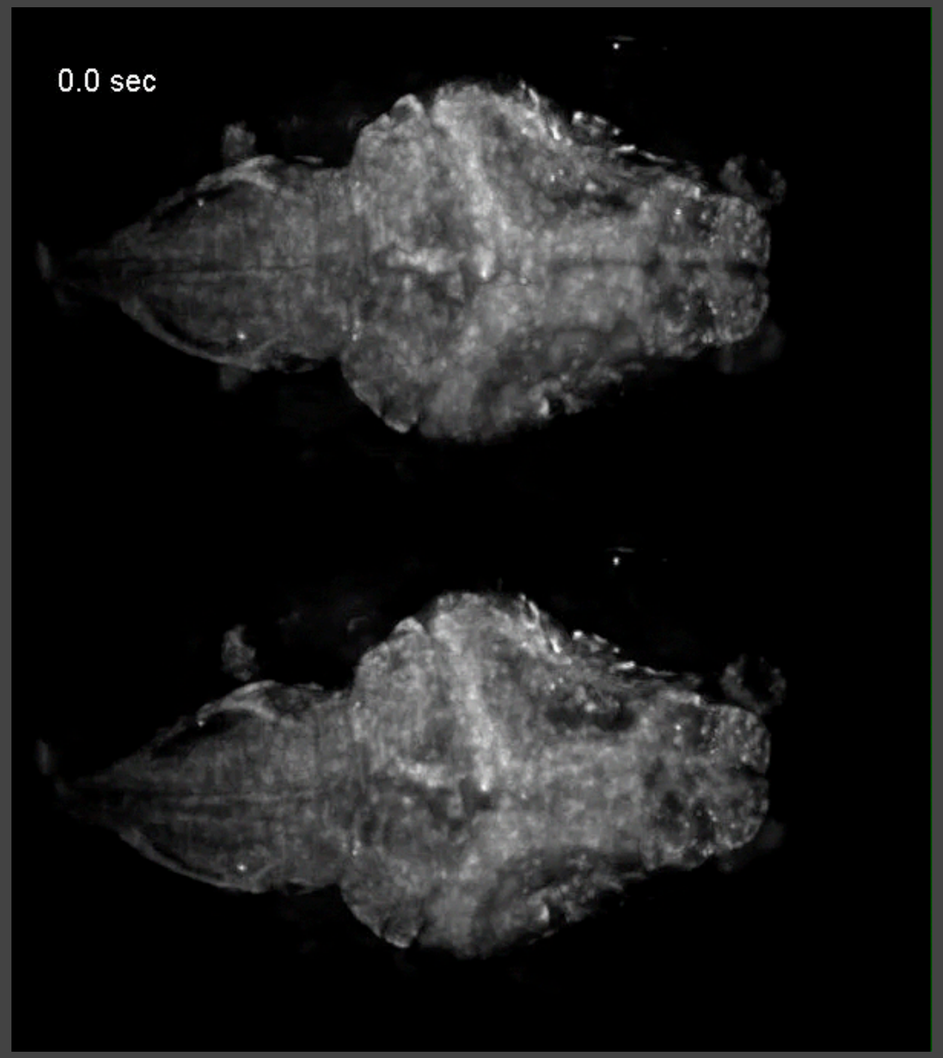
σ_i^2 : **variance of neural activity of a single neuron**

C_{ij} : **covariance between two neurons**

N : **the number of neurons in the population**

The dimensionality *D* is invariant under random subsampling for large *N*

Also See Dahmen et al ’19 & ’20 for a relevant discussion, but Renart et al ’10 for a completely different scenario!



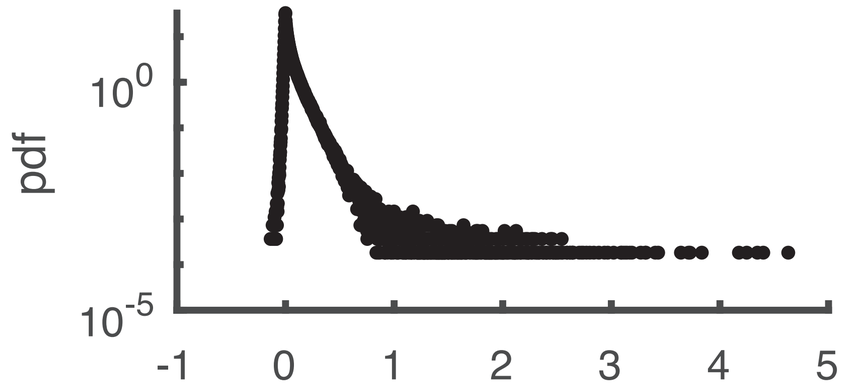
800 um x 800 um x 200 um FOV

3.4 um x 3.4 um x 5 um spatial resolution for sparsely labelled sample

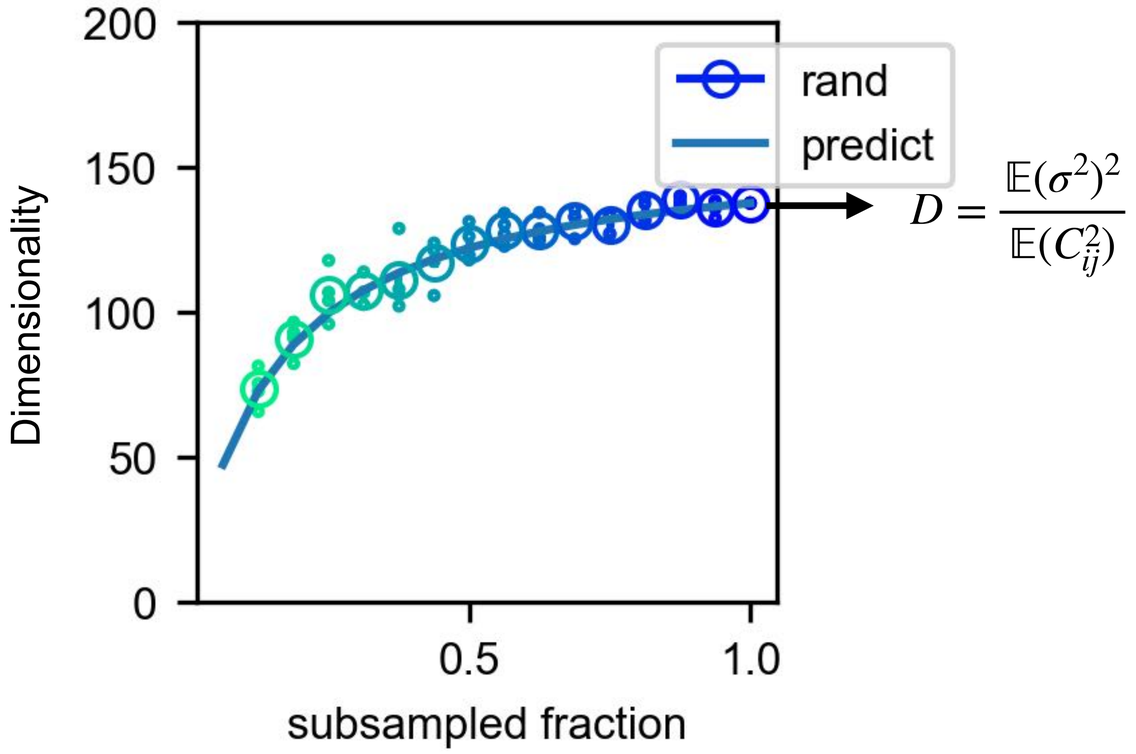
100 Hz volume rate

Cong et. al., 2017

The bounded dimensionality of brain-wide activity in larval zebrafish



$$C_{ij} = \mathbb{E}\left[(r_i - \mathbb{E}(r_i))(r_j - \mathbb{E}(r_j))\right]$$



Wang and Mai et al., 2025

Classical MDS starts somewhere else

No coordinates — only pairwise distances between time points.

$$D_{st}^2 = \|x_s - x_t\|^2,\ D^2 \in \mathbb{R}^{T \times T}$$

- The object is now T × T, living entirely in time space.
- Nothing here refers to neurons, axes, or a covariance matrix.

Double centering recovers the Gram matrix

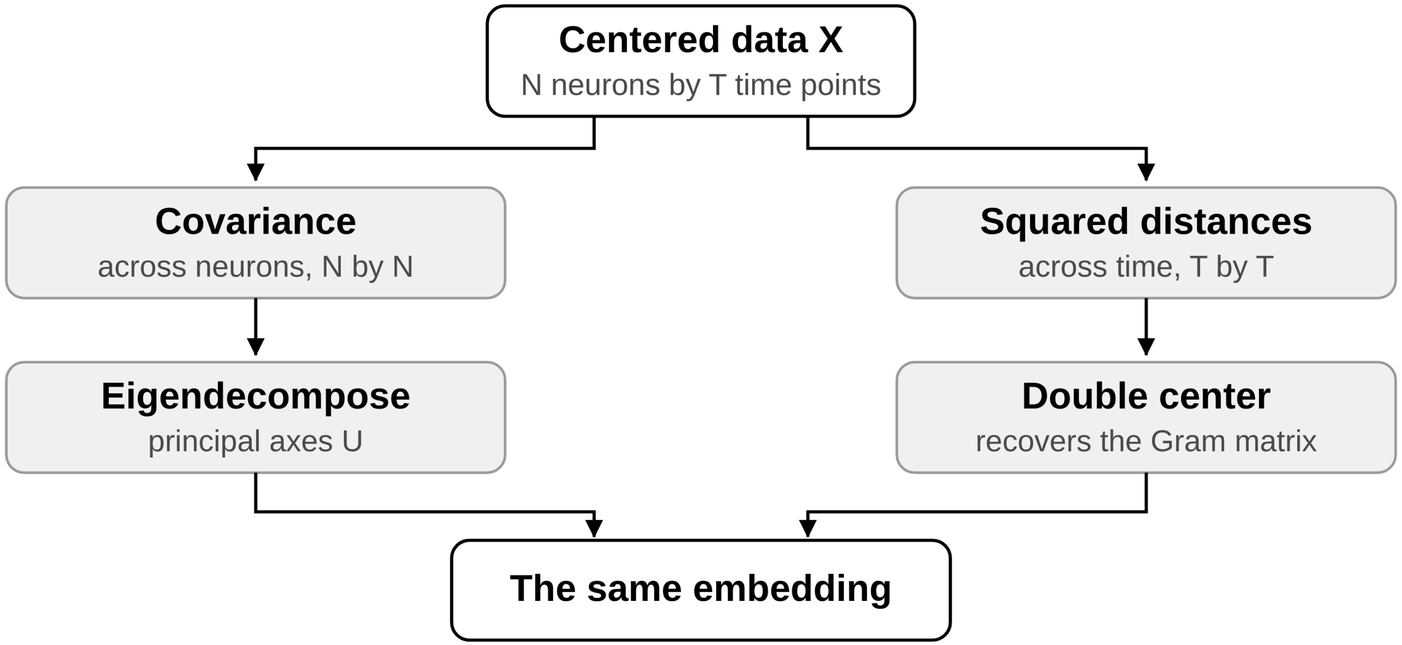
$$D_{st}^2 = \|x_s\|^2 + \|x_t\|^2 - 2x_s^T x_t$$

$$D^2 = c\mathbf{1}^T + \mathbf{1}c^T - 2G \qquad G = X^T X$$

- Double centering eliminates the two rank-1 terms in the distance matrix.

The two routes give the same embedding

$$X^\top X = V\Sigma^2 V^\top \qquad\qquad U_K^\top X = \Sigma_k V_k^\top$$



PCA is the covariance-side reading of the SVD; classical MDS is the distance-side reading.

Which Matrix to decompose?

- 130 neurons by 3000 frames: the covariance wins.
- 1000 neurons and 8 stimuli (or conditions): the Gram wins instead.

The useful consequences

- Classical MDS needs only a distance matrix — so changing the distance metric gives new methods for free (e.g., Isomap, diffusion maps, UMAP, TSNE).
- An arbitrary kernel in place of the Gram matrix: kernel PCA.

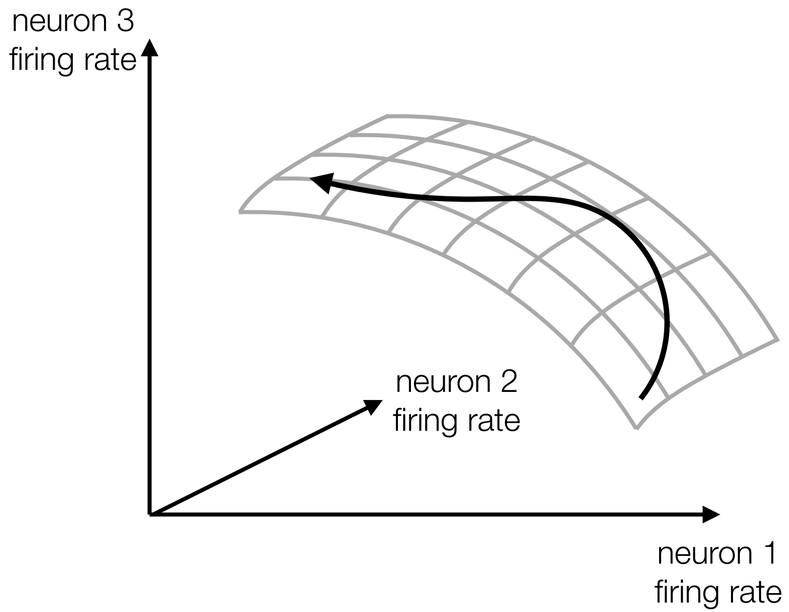
Demixed neural activity

Variance is not importance.

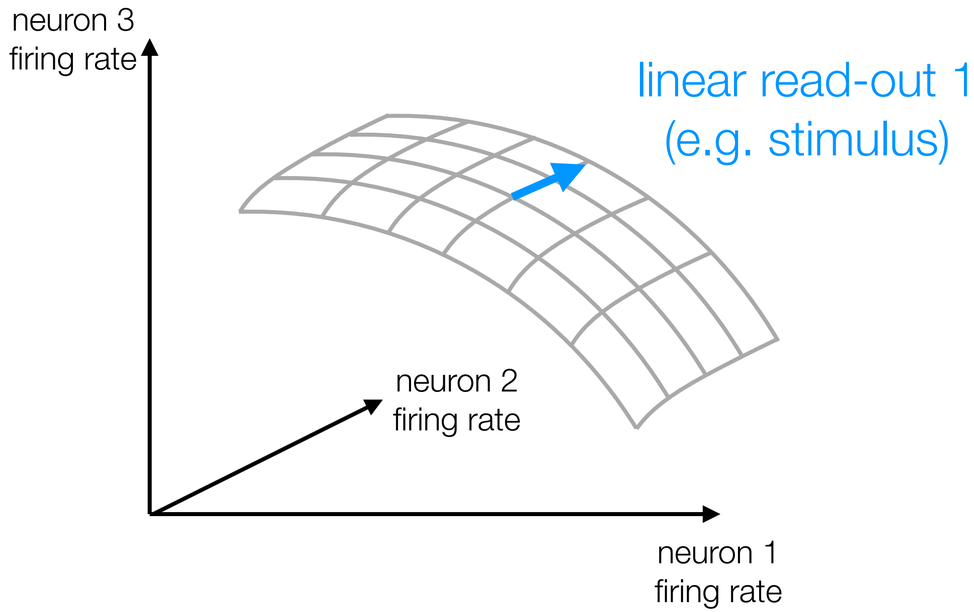
A dimension can carry little variance and great behavioral relevance.

PCA optimizes a criterion nobody claimed the brain uses.

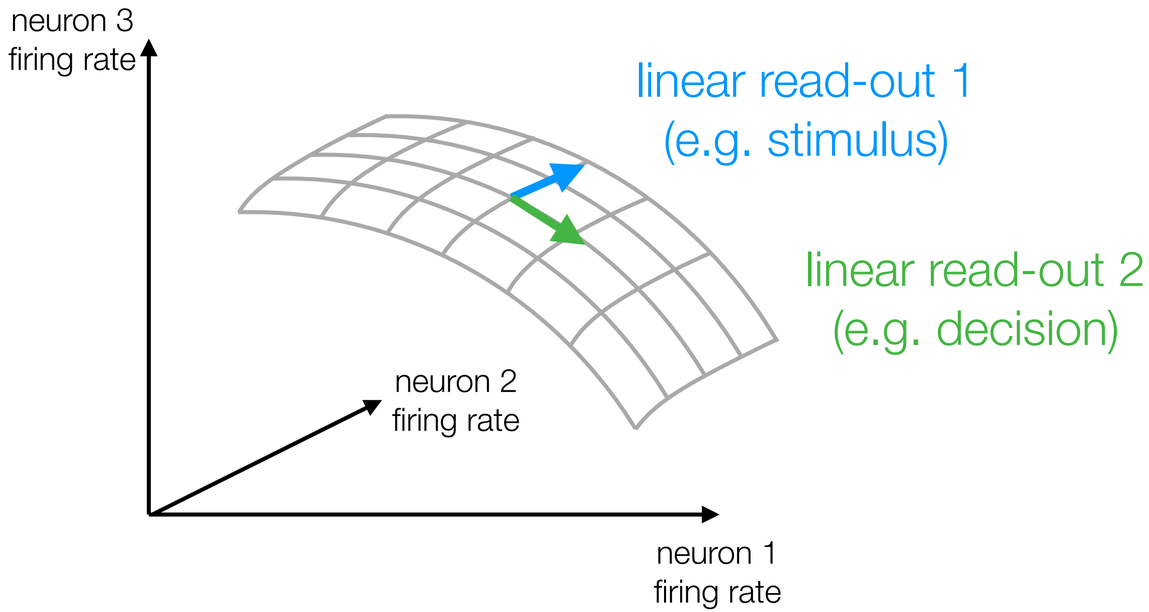
Can we assign meaning to the neural manifold?

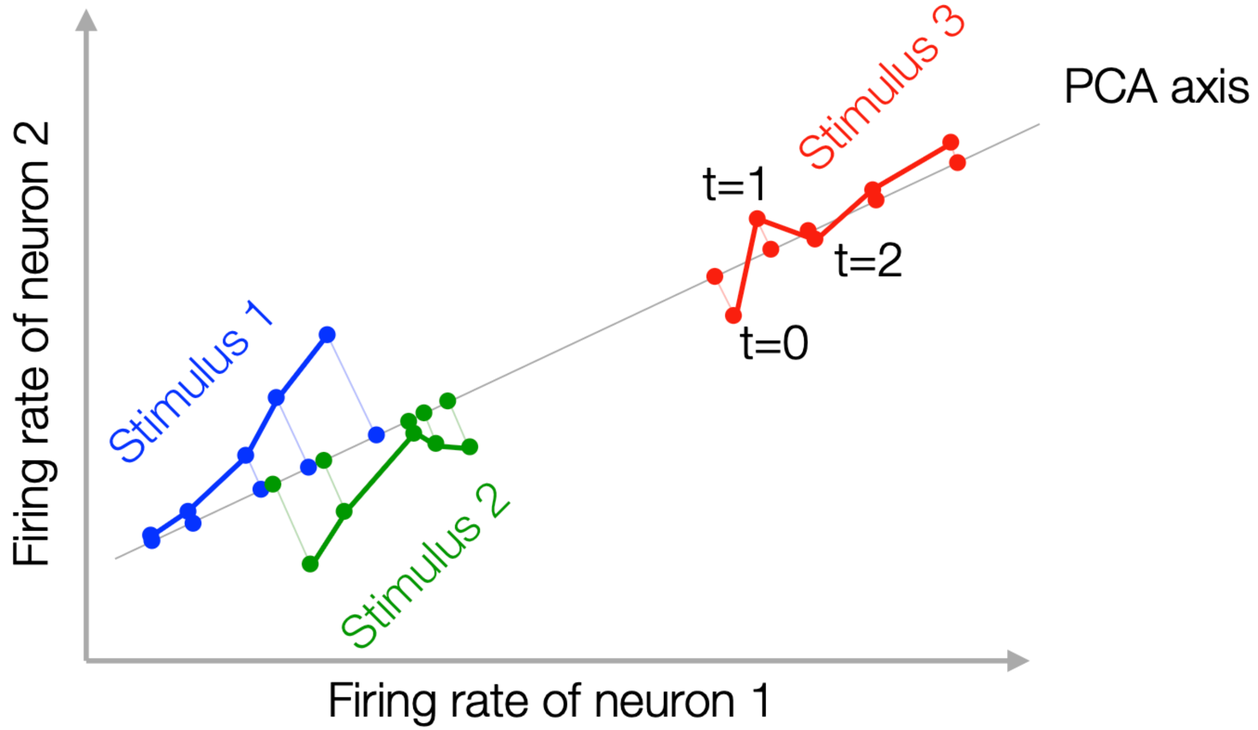


Can we assign meaning to the neural manifold?

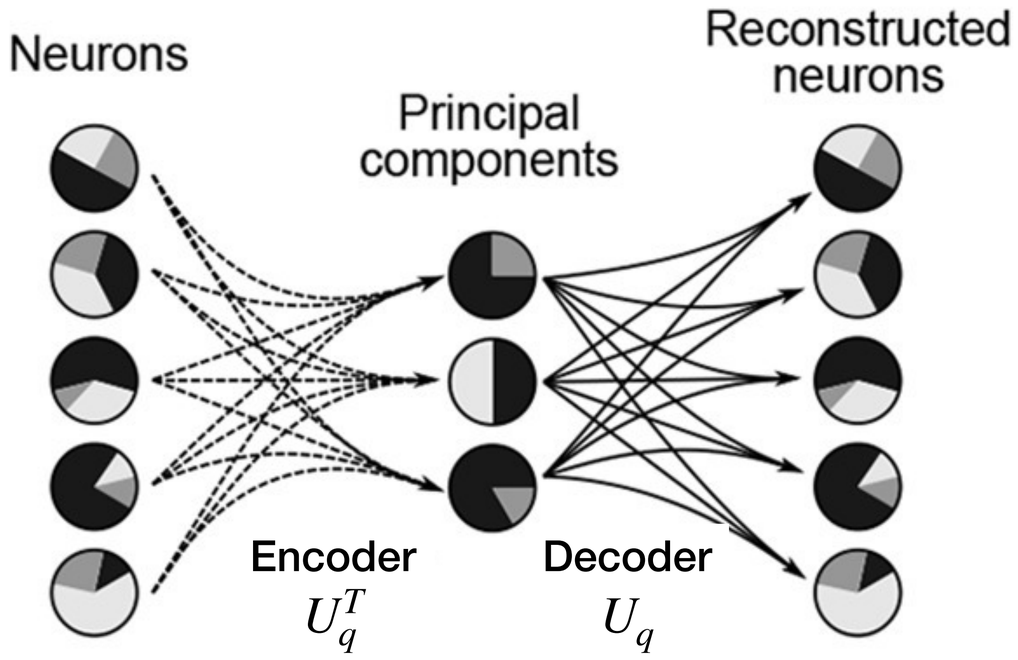


Can we assign meaning to the neural manifold?

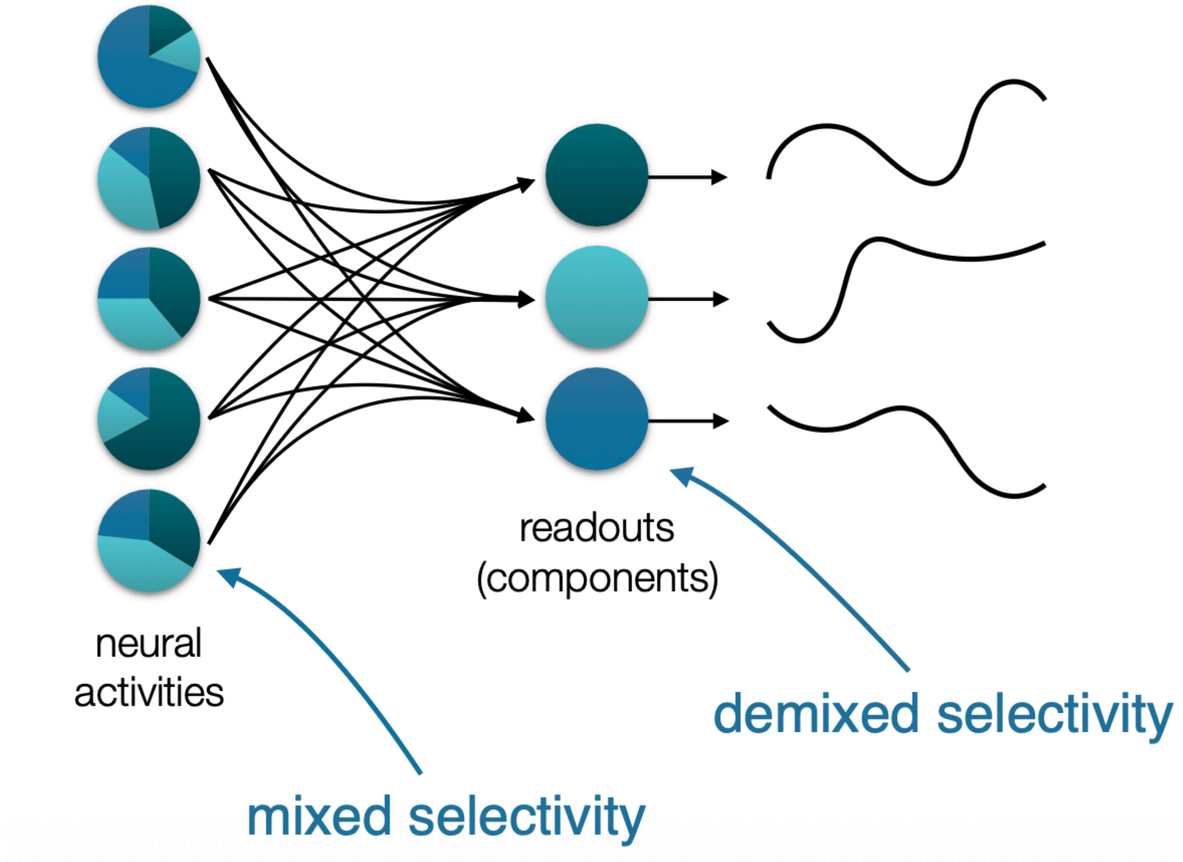



Firing rate of neuron 2
Firing rate of neuron 1
PCA axis
Stimulus 1
Stimulus 2
Stimulus 3
t=0
t=1
t=2

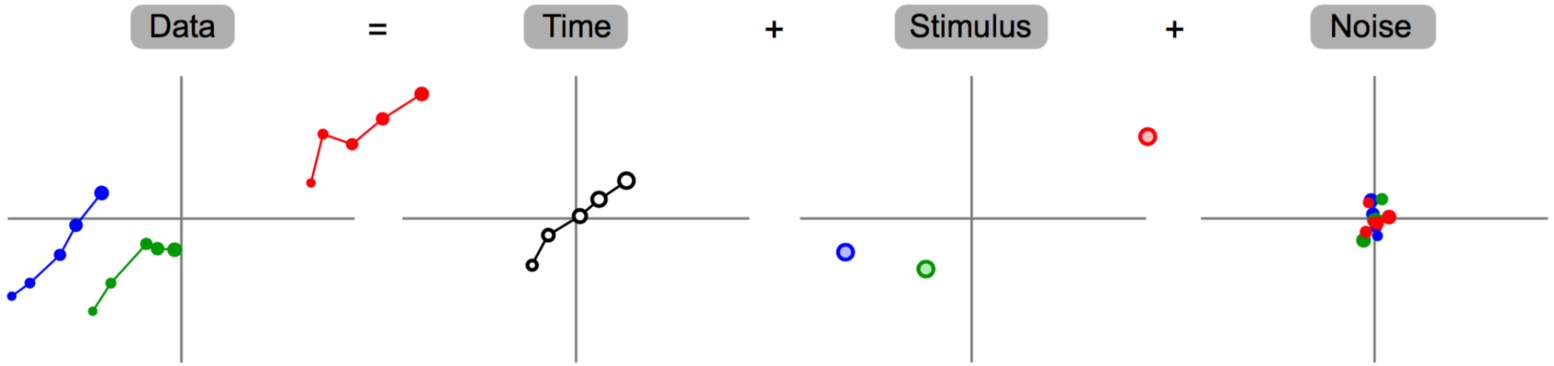
PCA


Neurons
Principal components
Reconstructed neurons
Encoder
U_q^T
Decoder
U_q

Demixed PCA


readouts
(components)
neural
activities
demixed selectivity
mixed selectivity

Decompose data into marginalized averages

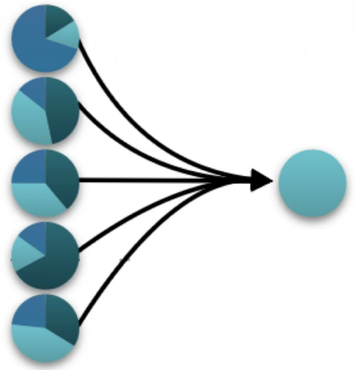


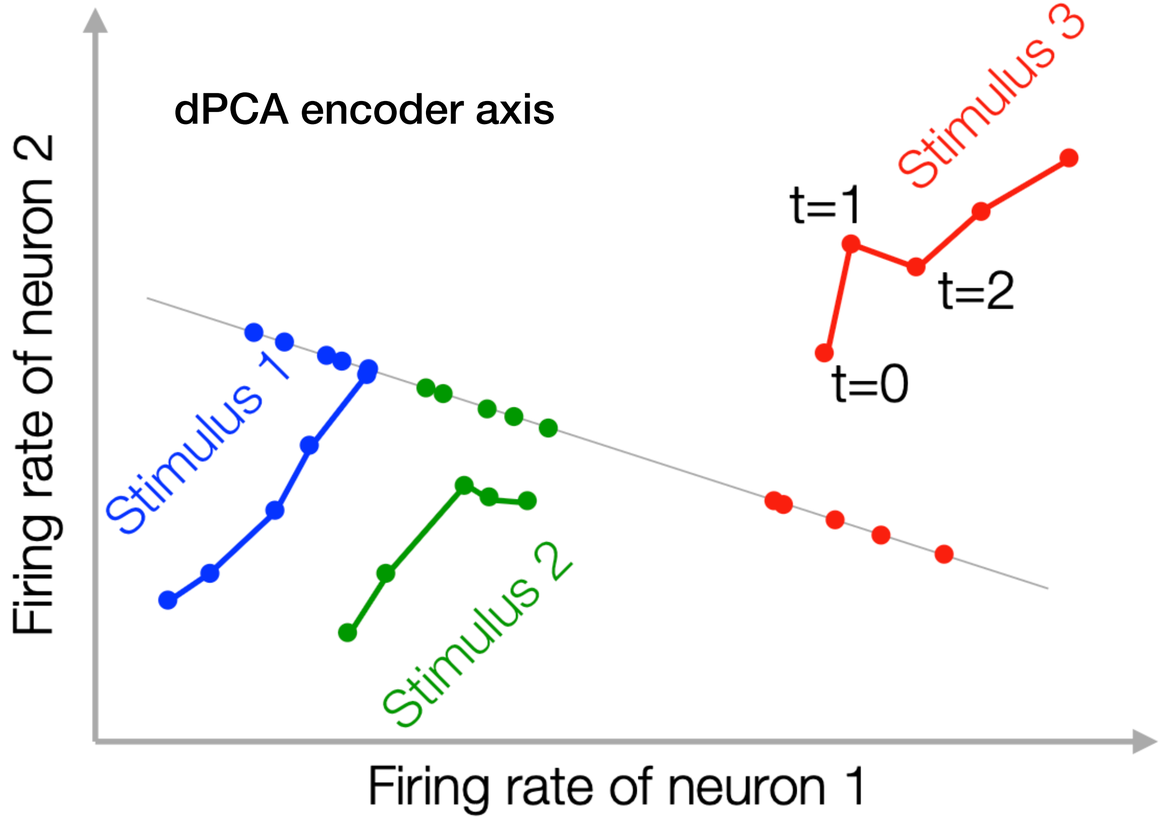
$$\mathbf{X} = \mathbf{X}_t + \mathbf{X}_s + \mathbf{X}_{st}$$

$$X_t = \langle X \rangle_s$$

$$X_s = \langle X \rangle_t$$

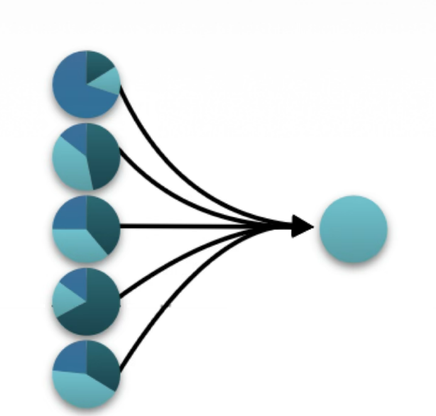
Linear readout with dPCA

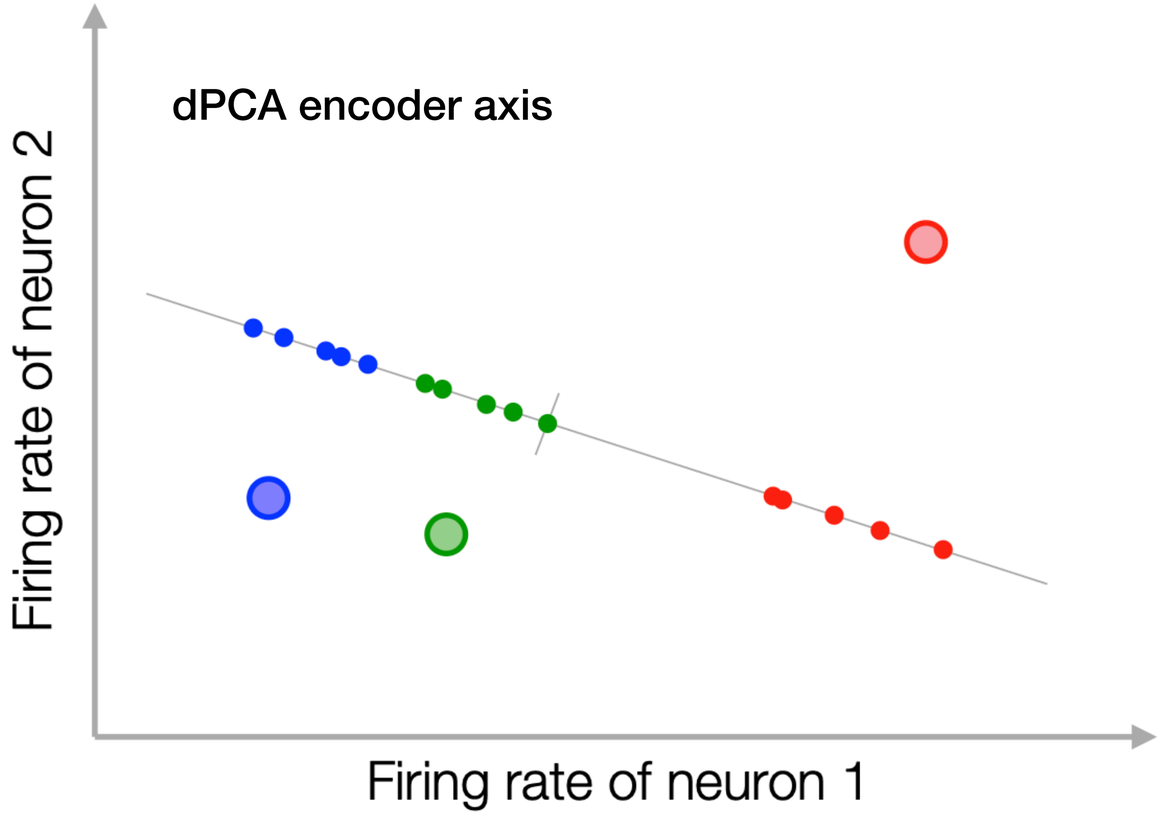




Kobak et al. *eLife* 2016

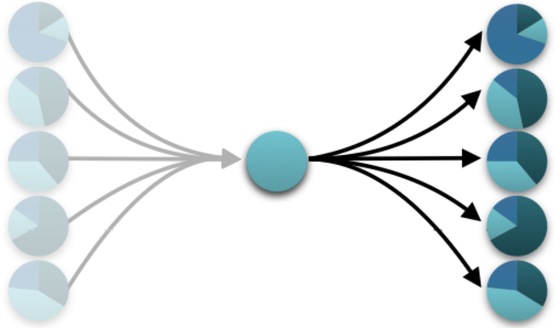
Linear readout with dPCA

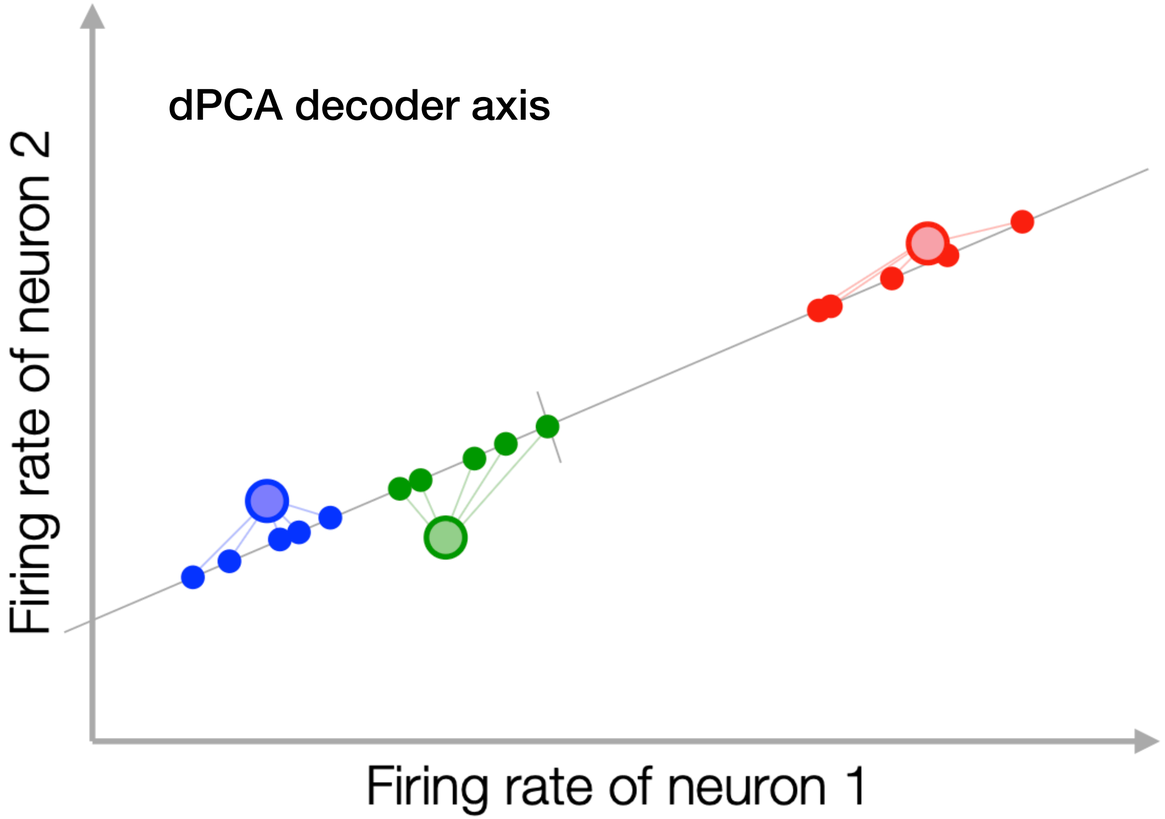




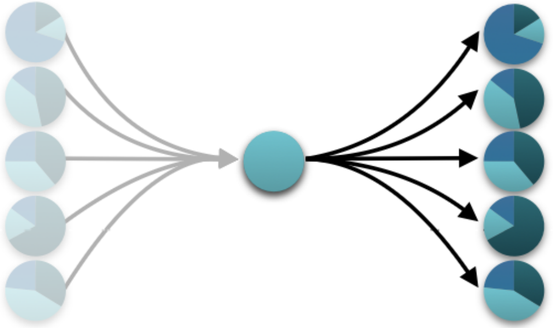
Kobak et al. *eLife* 2016

Linear readout with dPCA





Kobak et al. *eLife* 2016



Loss function:

$$L = \|\mathbf{X}_s - \mathbf{FDX}\|^2$$

‘marginalized’ averages

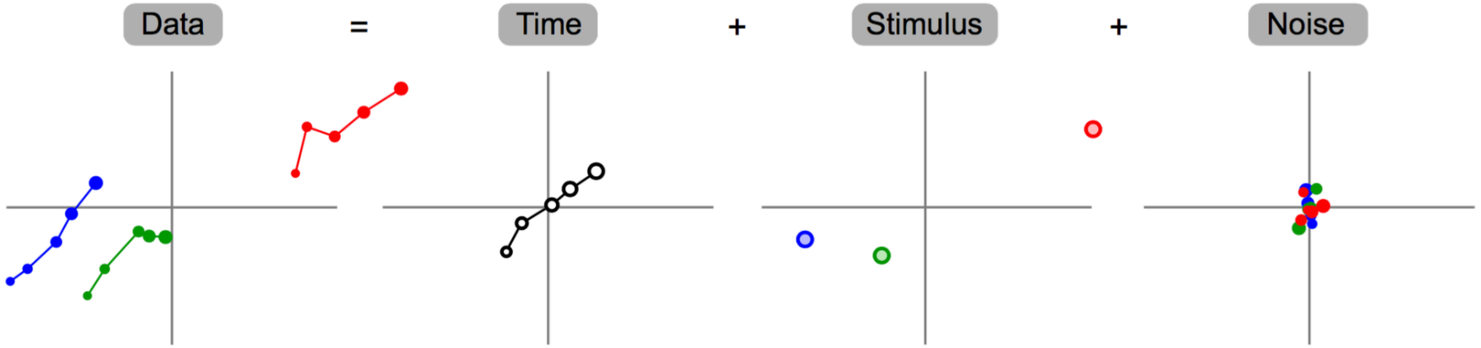
neural activities

decoder

encoder

reduced rank regression!

Separate encoders/decoders for marginalized averages



$$\mathbf{X} = \mathbf{X}_t + \mathbf{X}_s + \mathbf{X}_{st}$$

$$L = \|X_t - F_t D_t X\|^2 + \|X_s - F_s D_s X\|^2 + \|X_{ts} - F_{ts} D_{ts} X\|^2 = \sum_{\phi} \|X_\phi - F_\phi D_\phi X\|^2$$

One rank-constrained regression per marginalization: source X, target X_ϕ, rank q.

dPCA is reduced-rank regression, term by term

$$\|X_\phi - F_\phi D_\phi X\|^2 = \|X_\phi - \hat{X}_\phi\|^2 + \|\hat{X}_\phi - B_q X\|^2, \quad B_q = F_\phi D_\phi$$

$$\hat{X}_\phi = X_\phi P = X_\phi X^\top (XX^\top)^{-1} X = B_{OLS} X$$

$$\min_{B_q} \|B_{OLS} X - B_q X\|^2$$

Fit with least square regression, then find the optimal q-rank truncation.

dPCA is reduced-rank regression, term by term

$$\|X_\phi - F_\phi D_\phi X\|^2 = \|X_\phi - \hat{X}_\phi\|^2 + \|\hat{X}_\phi - B_q X\|^2, \quad B_q = F_\phi D_\phi$$

$$\hat{X}_\phi = X_\phi P = X_\phi X^\top (XX^\top)^{-1} X = B_{OLS} X$$

$$\min_{B_q} \|\hat{X}_\phi - B_q X\|^2$$

$$\hat{X}_\phi = U \Sigma V^\top$$

$$B_q X = U_q U_q^\top B_{OLS} X$$ Apply Eckart-Young theorem!

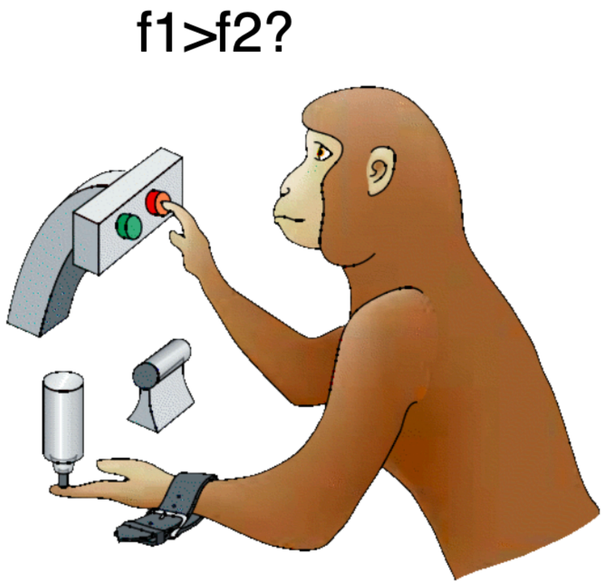
$$F_\phi = U_q \, , \, D_\phi = U_q^\top B_{OLS}$$

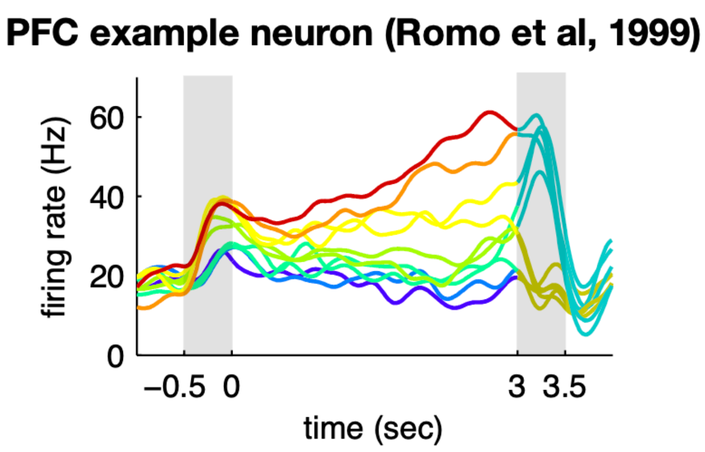
Two projections again, one on each side

$$U_q U_q^\top \; X_\phi \; P$$

- The right projector does the regression: keep only what the source can explain.
- The left projector does the dimensionality reduction: keep the top q directions of what is left.

A PFC working memory task





task parameters

stimulus f1

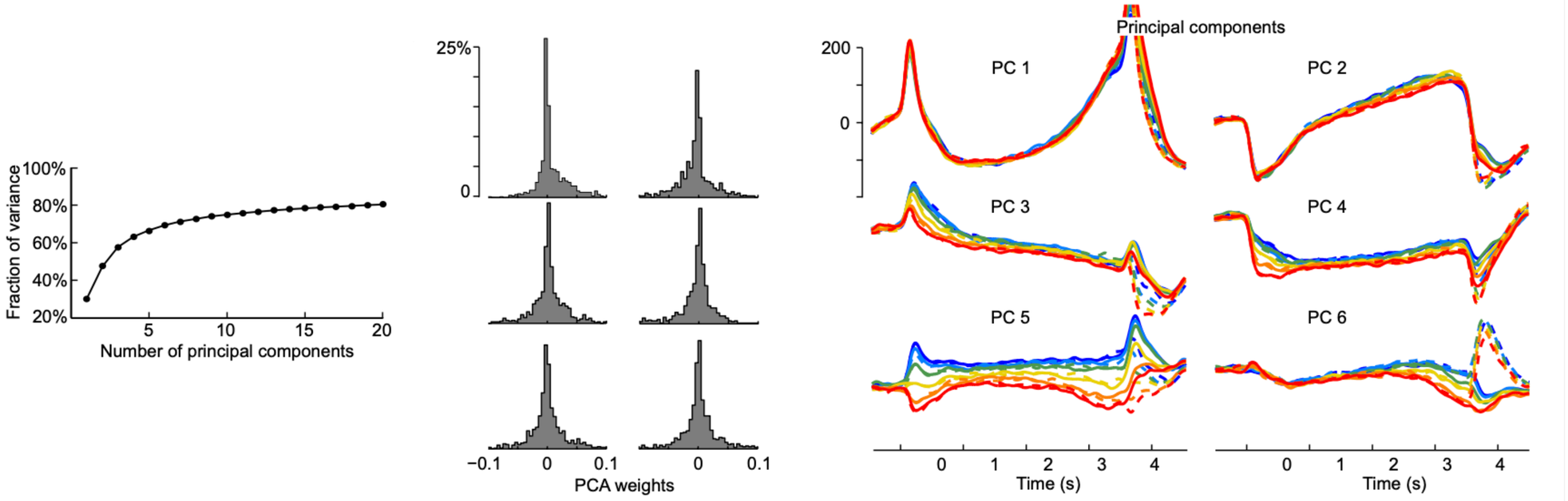
10 Hz 34 Hz

decision

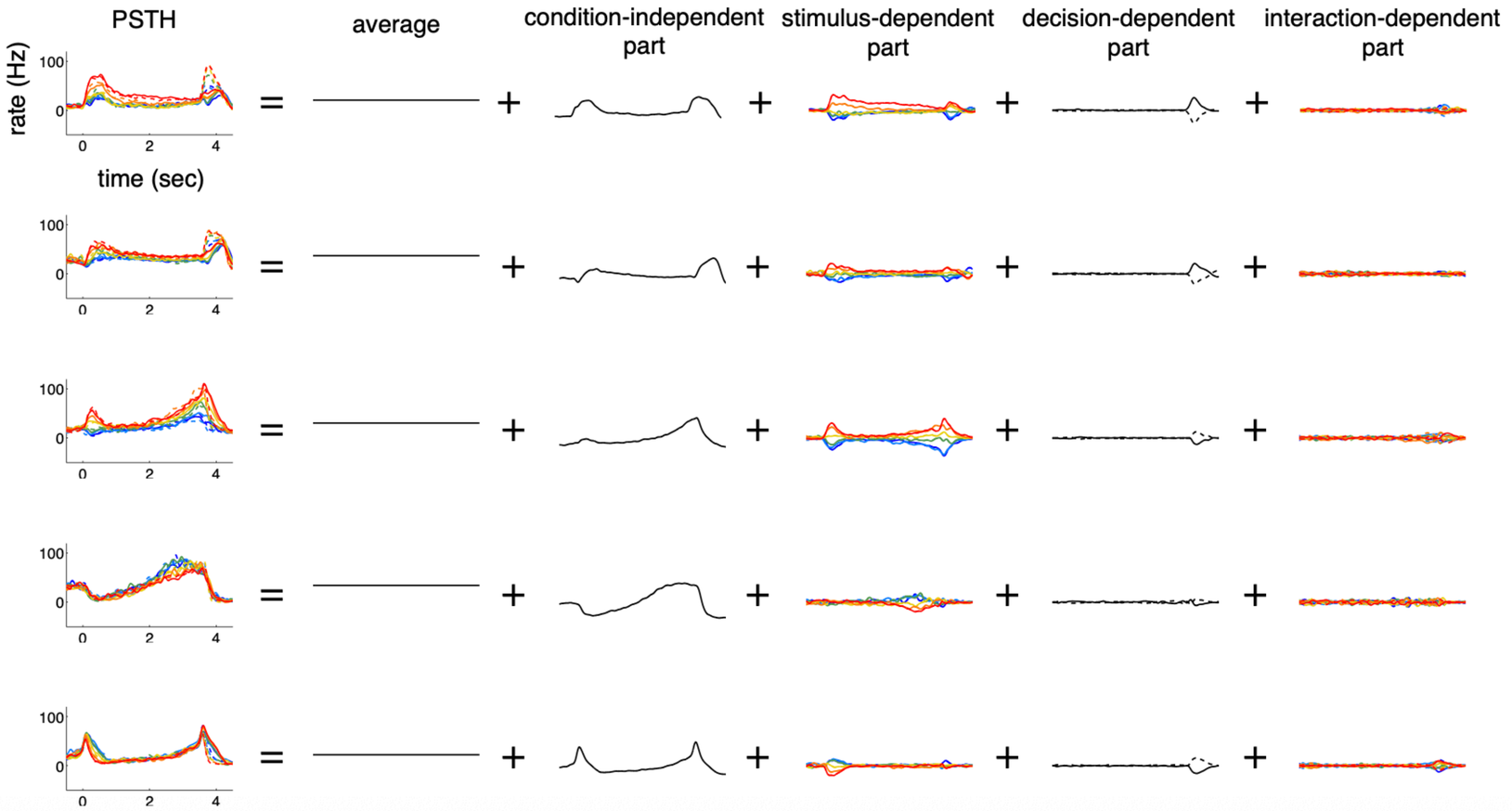
Yes No

Kobak et al. *eLife* 2016

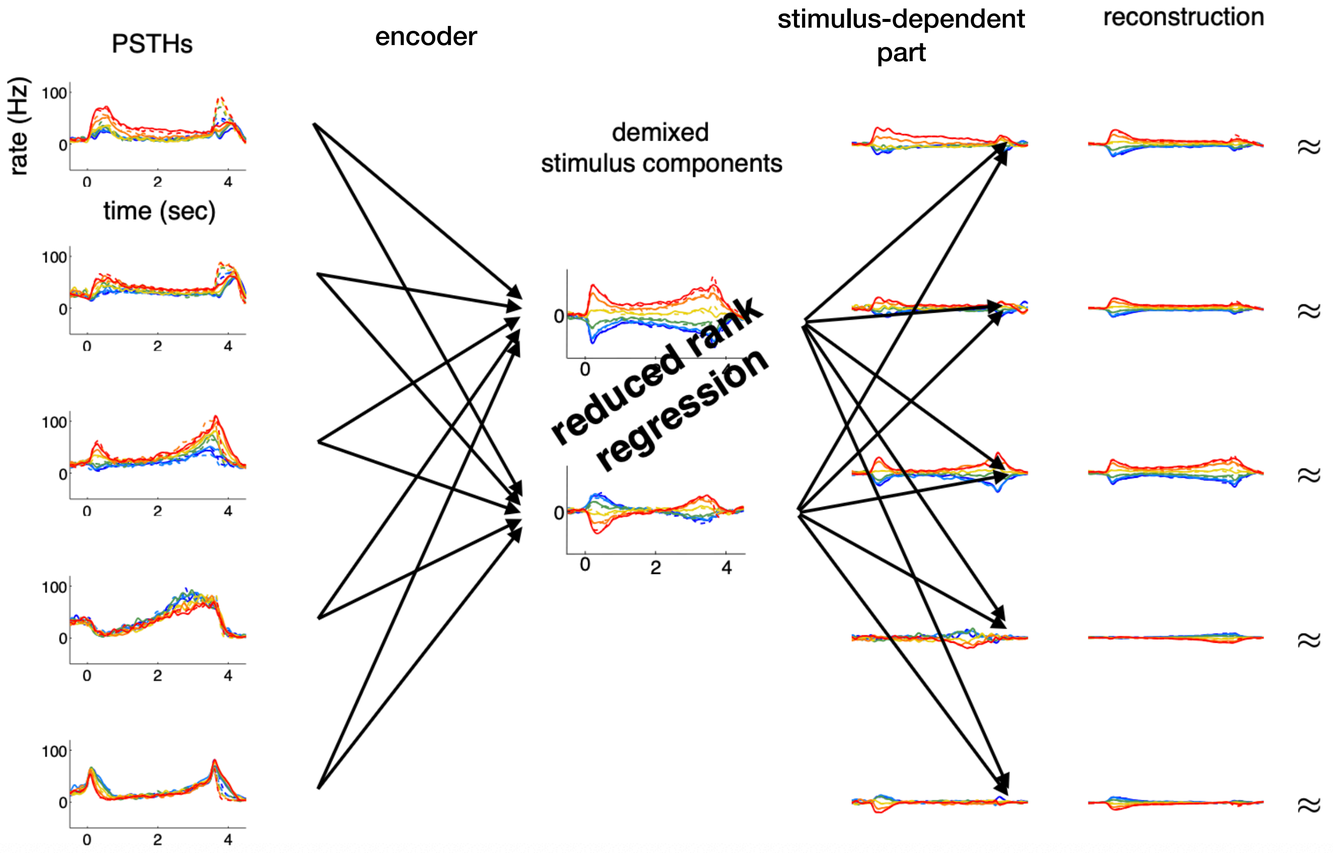
A PFC working memory task

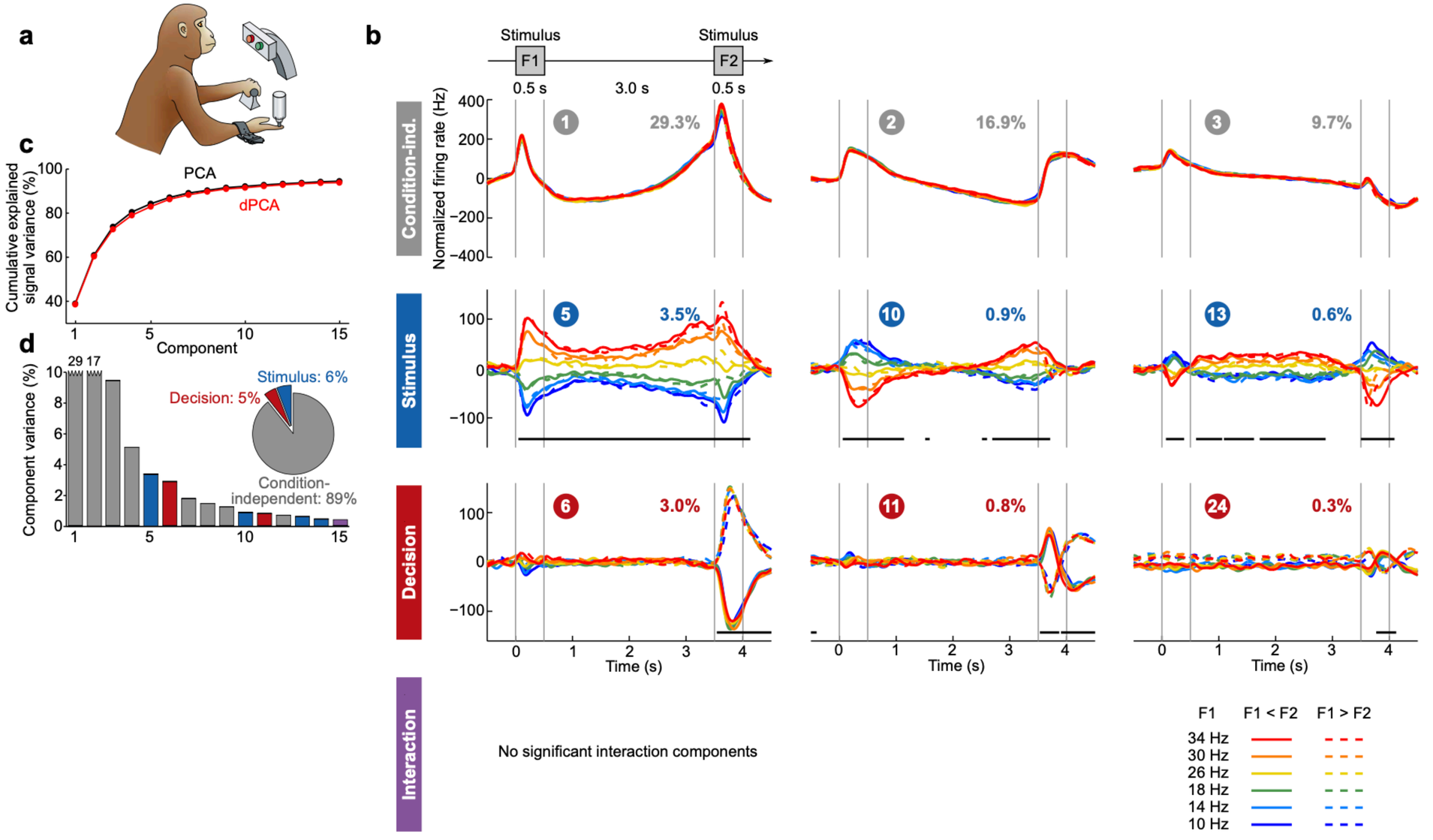


dPCA Algorithm, step 1



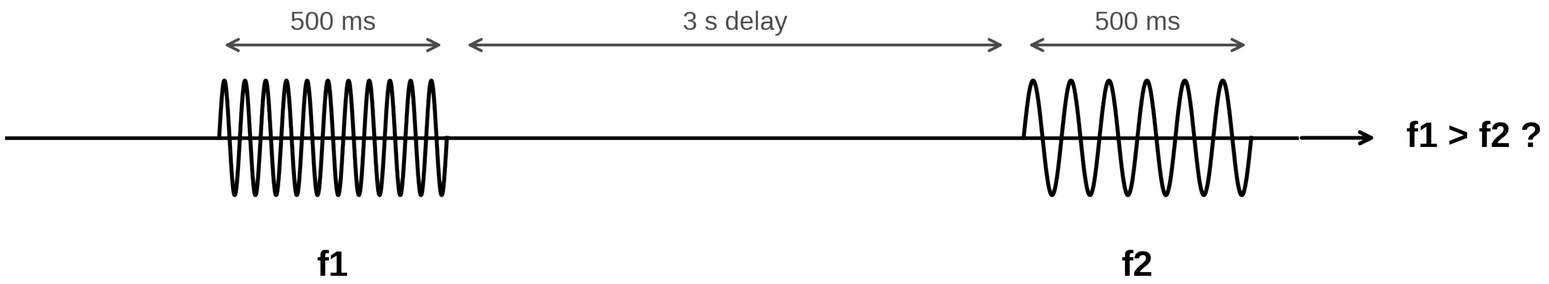
dPCA Algorithm, step 2





Kobak et al. *eLife* 2016

Exercise: dPCA on the Romo dataset



Steps

1. Load romo_allpsth.mat
2. Read the help for dpca.m
3. Set the parameters, separating time from condition
4. Explained variance with dpca_explainedVariance
5. Plot with dpca_plot

Advanced

1. Reorganize the data as neurons × stimuli × decisions × time
2. Re-run dPCA on the full factorial design

Communication subspace

What does area A send to area B?

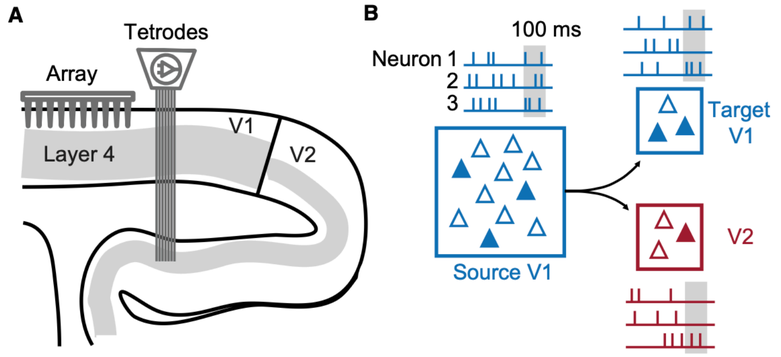
Setup

Two simultaneously recorded populations, both row-centered.

$$X \in \mathbb{R}^{N_X \times T}, Y \in \mathbb{R}^{N_Y \times T}$$

$$B_{OLS} = YX^T(XX^\top)^{-1}$$

$$\hat{Y} = B_{OLS}X = YP$$

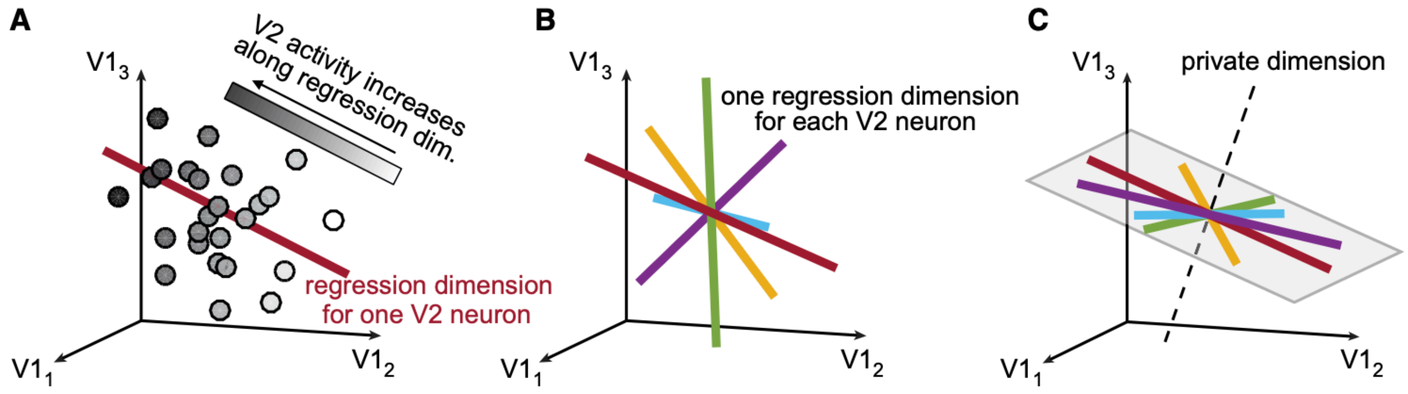


$\hat{Y}$: the best prediction for target from source

Semedo et al., 2019

Is the interaction high or low dimensional?

Again, reduced-rank regression!



$$\min_{\text{rank } B \leq r} \left\| \hat{Y} - BX \right\|_F^2$$

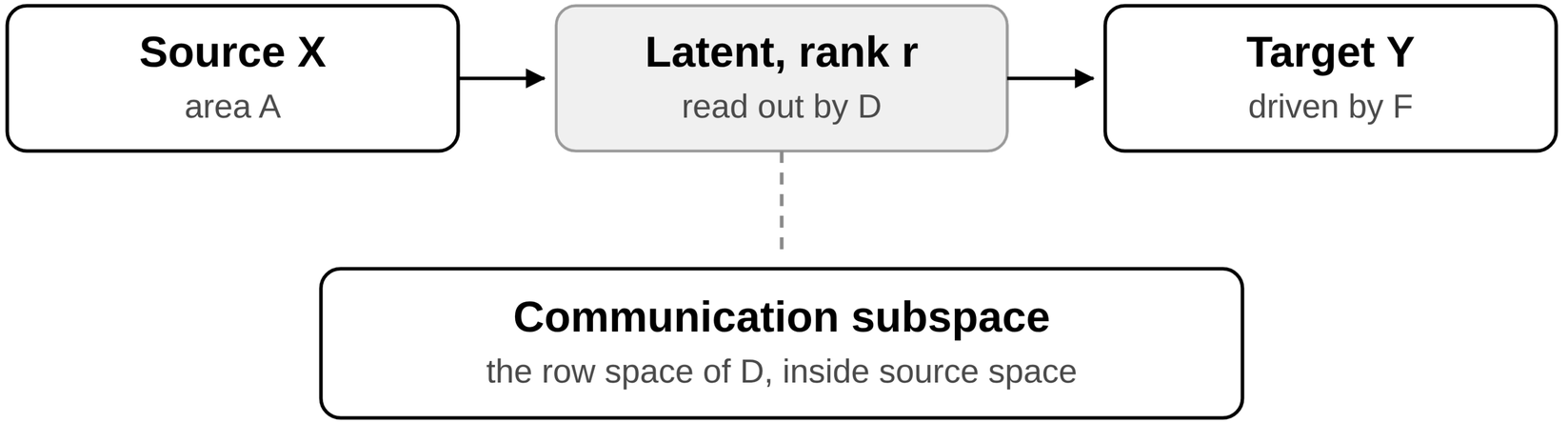
$$\hat{Y} = U \Sigma V^\top$$

$$BX = U_r U_r^\top B_{OLS} X$$

$$B = U_r U_r^\top B_{OLS}$$

Two spaces of $B \in \mathbb{R}^{N_Y \times N_x}$

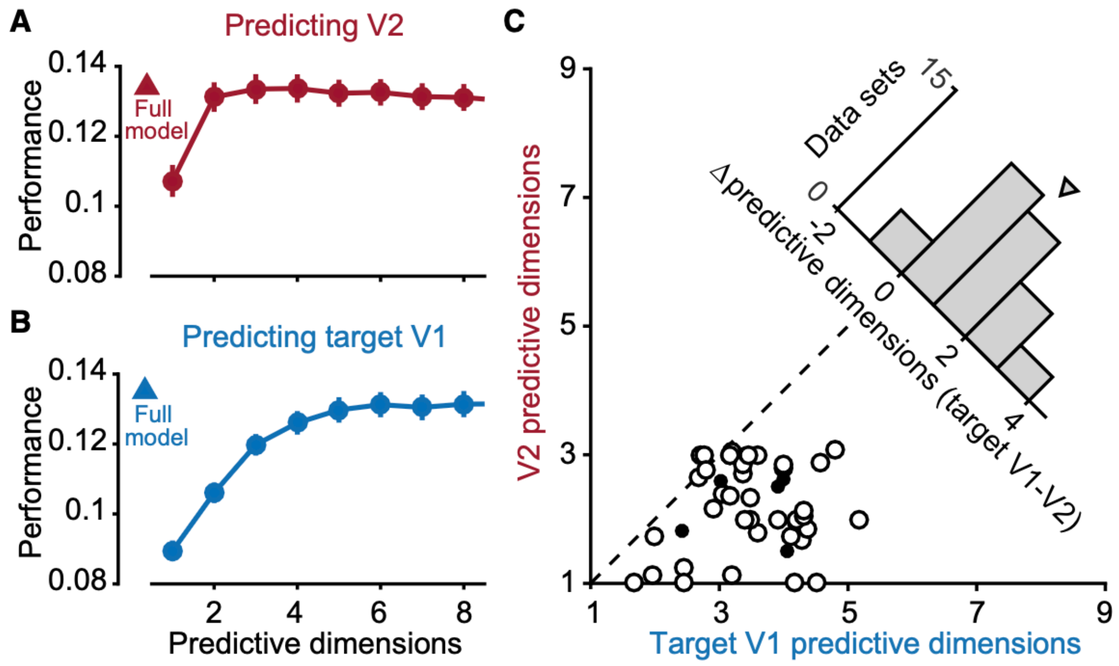
$$B = U_r U_r^\top B_{OLS} \qquad D = U_r^\top B_{OLS} \in \mathbb{R}^{r \times N_X} \qquad F = U_r \in \mathbb{R}^{N_Y \times r}$$



- Row space in source space: which dimensions are read out.
- Column space in target space: which dimensions get driven.
- $\text{null}(D) =$ private subspace (does not involve in the interaction)

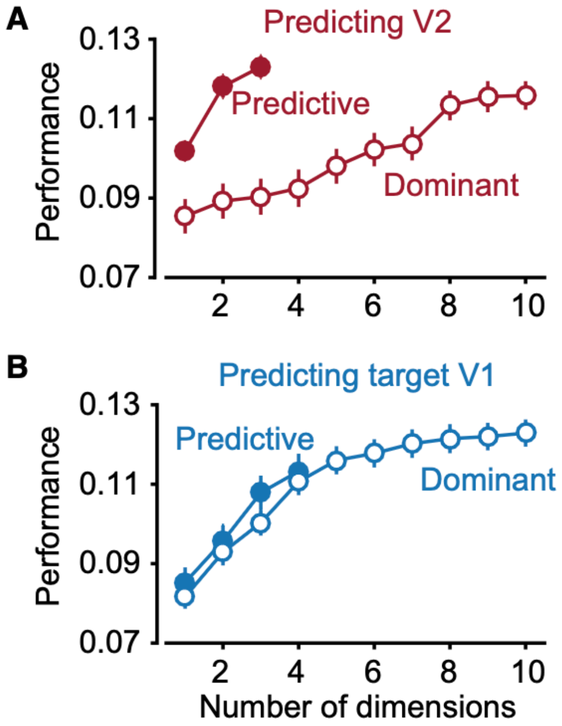
Low dimensional interaction between V1 and V2

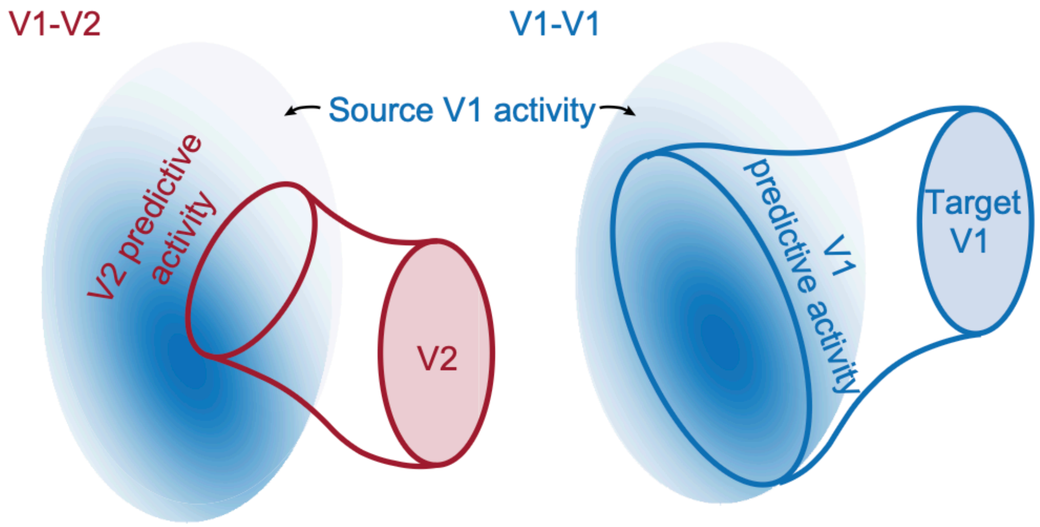
$r \approx 2$



Semedo et al., 2019

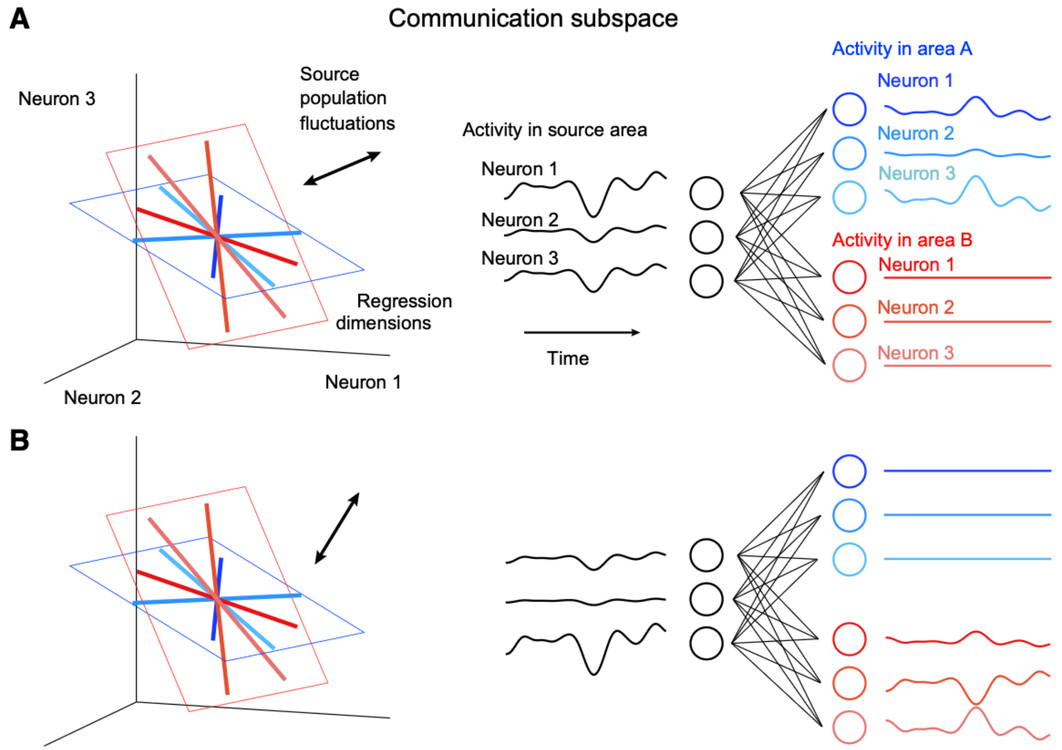
The predictive dimensions for V2 is not the dominant dimensions in V1





Semedo et al., 2019

Selective communication between different areas



Semedo et al., 2019

Limitations

- Not causal. Common input to both areas produces exactly the same signature as directed communication.
- Instantaneous by default. Time-lagged RRR is the natural extension.
- Linear. What about nonlinear mapping?

Regress on the right.

Truncate on the left.

Change what sits on the left-hand side and the same estimator becomes PCA, demixed PCA, or a communication subspace.