

# THE LOW RANK HYPOTHESIS OF COMPLEX SYSTEMS

**Authors: Vincent Thibeault, Antoine Allard, Patrick Desroisiers**  
Article: Nat. Phys. 20, 294–302 (2024).

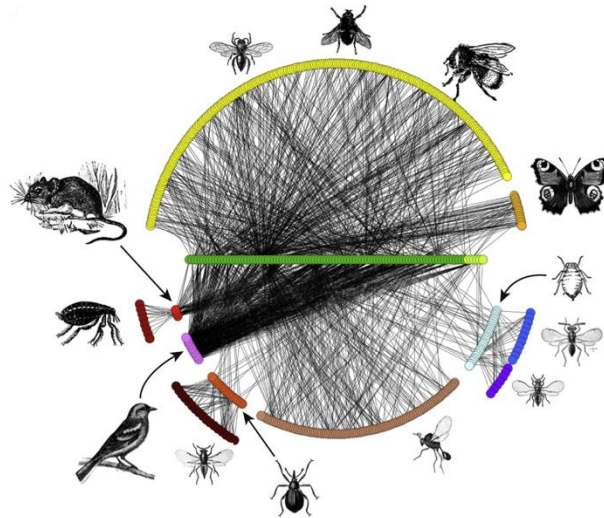
Presenter: Guru Kalyan Jayasingh  
Assisted by: Nigel Goldenfeld  
Department of Physics  
UCSD



**Why is it possible to have simple descriptions of complex phenomena?**

# Low-rank hypothesis

- Complex systems, such as ecosystems, neural networks, and social structures, are often represented by high-dimensional non-linear differential equations.

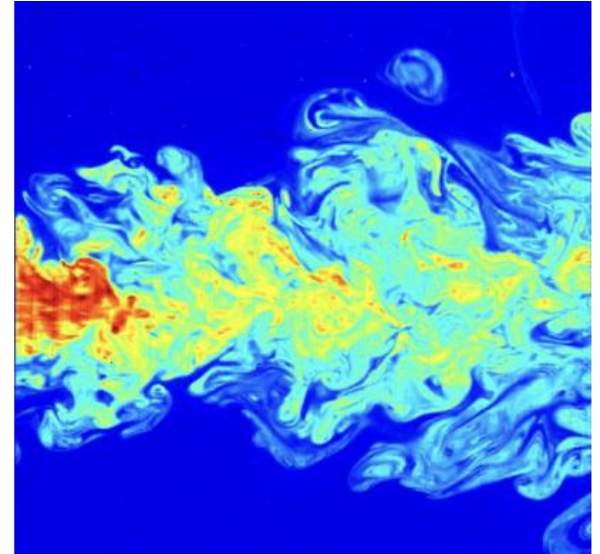
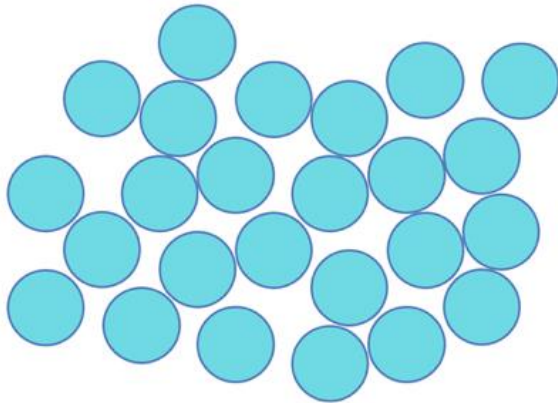


Credit: Quintessence Network DOI: 10.1016/j.j.tree.2015.12.003

- To make interpretable predictions about their large-scale behavior reduce dynamics to a few equations with small number of observables.

# Low-rank hypothesis and dimension reduction

Collection of particles



Navier-Stokes equation

$$\rho \left( \frac{\partial u}{\partial t} + (u \cdot \nabla) u \right) = \rho g - \nabla p + \mu \Delta u$$

- Dimension reduction
  - Reduced system of macrostates or observables
  - Small enough dimension to get an insightful description but large enough to preserve the phenomena of interest.

# In Physics

- Kuramoto-Sivashinsky reaction

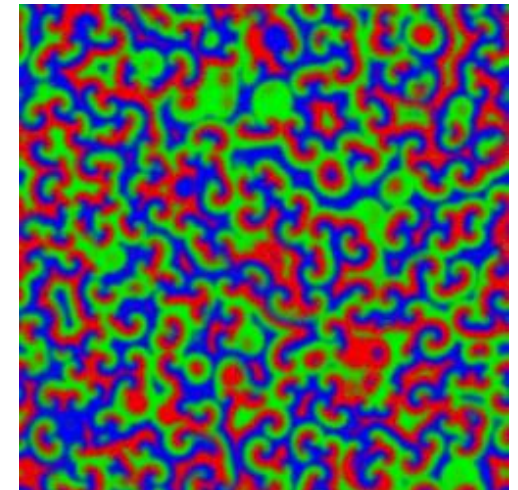
$$u_t + u_{xx} + u_{xxxx} + \frac{1}{2}u_x^2 = 0$$

- For a system of size  $L$ , the long time behavior of solutions approach exponentially fast to a finite dimensional manifold  $M$ .
- *Finite dimensional long time dynamics*



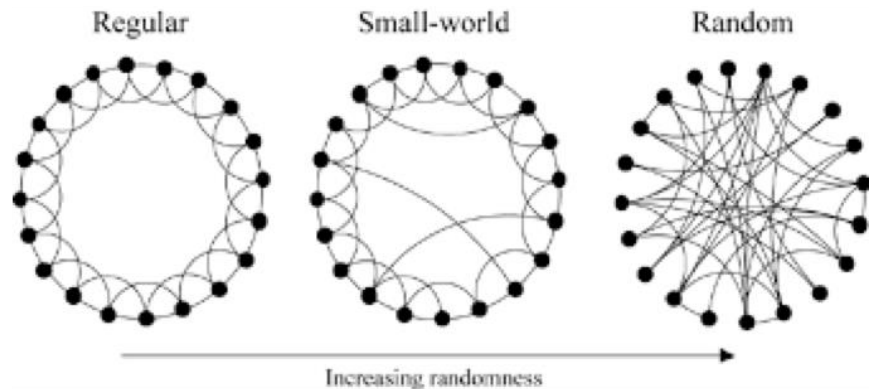
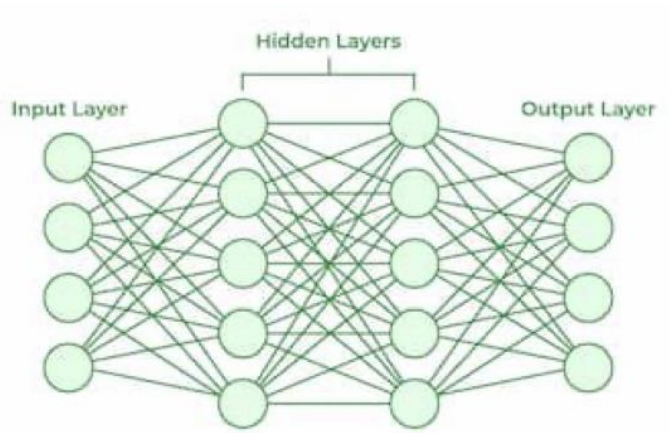
- Belousov–Zhabotinsky reaction

- System of chemical reactions involving over 20 individual reactions amongst 18 chemical species.
- Can be reduced to a system of 3 first order ODEs that reproduce the qualitative features.



# Low rank hypothesis for networks

- In this paper, the authors
  - study the validity of this assumption for dynamics on networks and
  - study the effect of low-rank nature of these networks on dimension reduction of its network dynamics, for e.g. in RNNs.

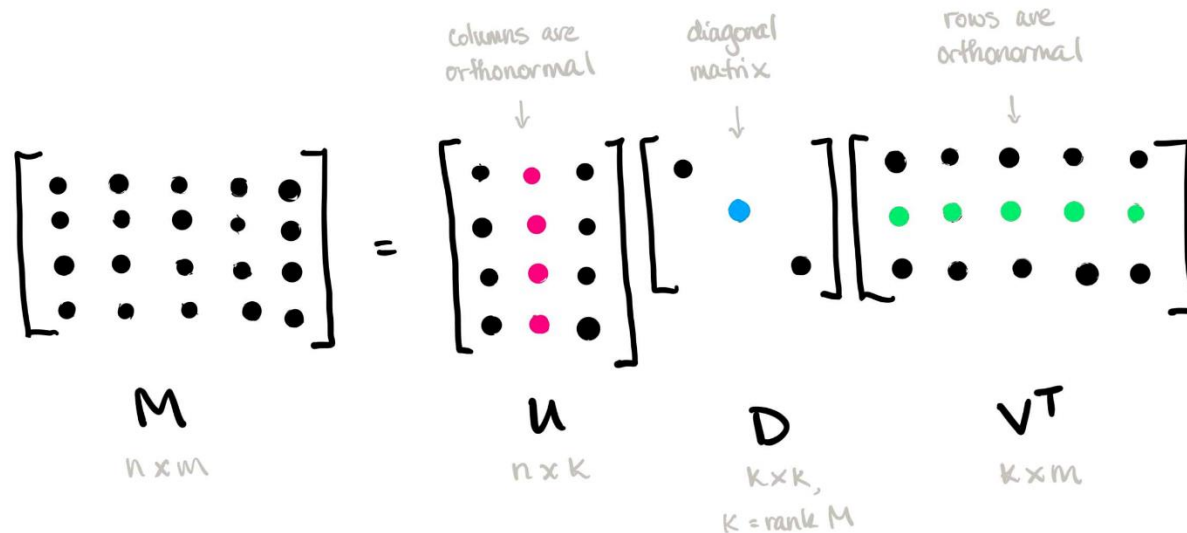


**Rank in low-rank**

# Singular Value Decomposition (SVD)

- What can it do?

$$M = U D V^T$$



- $U$ ,  $D$ ,  $V$  are special

$$UU^T = I \quad VV^T = V^T V = I$$



- Diagonal entries are known as singular values.  $k$  is called the “rank” of the matrix.

$$\begin{matrix}
 \begin{bmatrix} \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \end{bmatrix} & = & \begin{bmatrix} \bullet & \color{pink}\bullet & \bullet \\ \bullet & \color{pink}\bullet & \bullet \\ \bullet & \color{pink}\bullet & \bullet \\ \bullet & \color{pink}\bullet & \bullet \end{bmatrix} & \begin{bmatrix} \bullet & & \\ & \color{blue}\bullet & \\ & & \bullet \end{bmatrix} & \begin{bmatrix} \bullet & \bullet & \bullet & \bullet & \bullet \\ \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet & \color{green}\bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \end{bmatrix} \\
 M & & U & D & V^T \\
 n \times m & & n \times k & k \times k, \quad k = \text{rank } M & k \times m
 \end{matrix}$$

columns are orthonormal

diagonal matrix

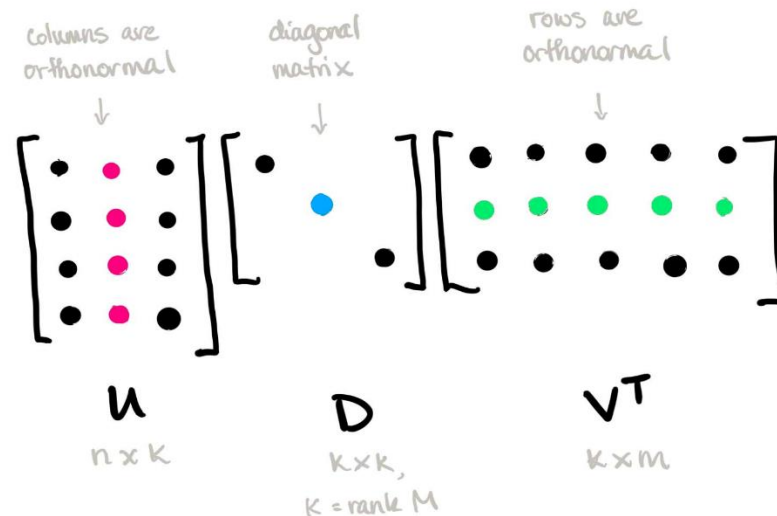
rows are orthonormal

$$D_{11} \geq D_{22} \geq D_{33} \cdots \geq D_{kk}$$

- Can be done for any matrix (even non-square matrices).

- Useful? Say if we want to approximate the matrix  $M$  by a pair vectors. The best known approximation\* that minimizes least squares error from  $M$  would be

$$M \sim U_1 D_{11} V_1^T$$

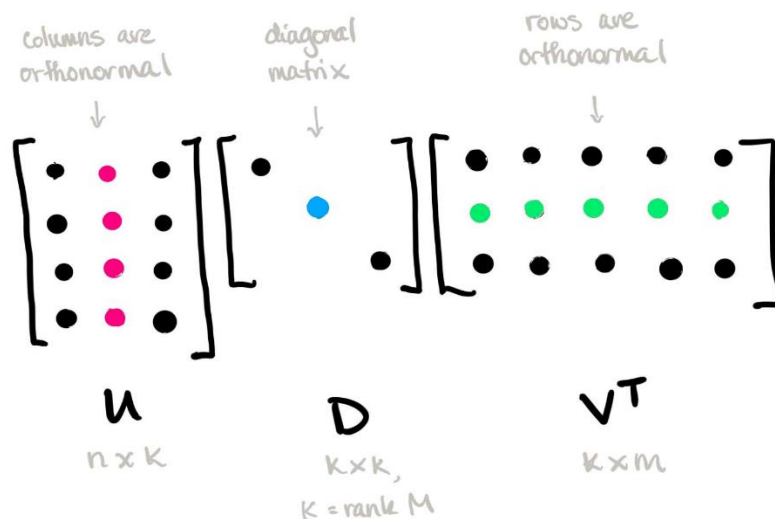


- How about improving with using 2 pairs of vectors? Or more? Just keep adding more terms

$$M \sim U_1 D_{11} V_1^T + U_2 D_{22} V_2^T + U_3 D_{33} V_3^T + \dots$$

- Useful? Say if we want to approximate the matrix  $M$  by a pair vectors. The best known approximation\* that minimizes least squares error from  $M$  would be

$$M \sim U_1 D_{11} V_1^T$$



- How about improving with using 2 pairs of vectors? Or more? Just keep adding more terms

$$M \sim \underbrace{U_1 D_{11} V_1^T}_{\text{Rank 1}} + \underbrace{U_2 D_{22} V_2^T}_{\text{Rank 2}} + \underbrace{U_3 D_{33} V_3^T}_{\text{Rank 3}} + \dots$$

- How to get  $U, V, D$  ?

Use

$$M \cdot M^T = U D^2 U^T$$

- Columns of  $U$  are left eigenvectors of  $MM^T$

Eigenvalues are singular values squared

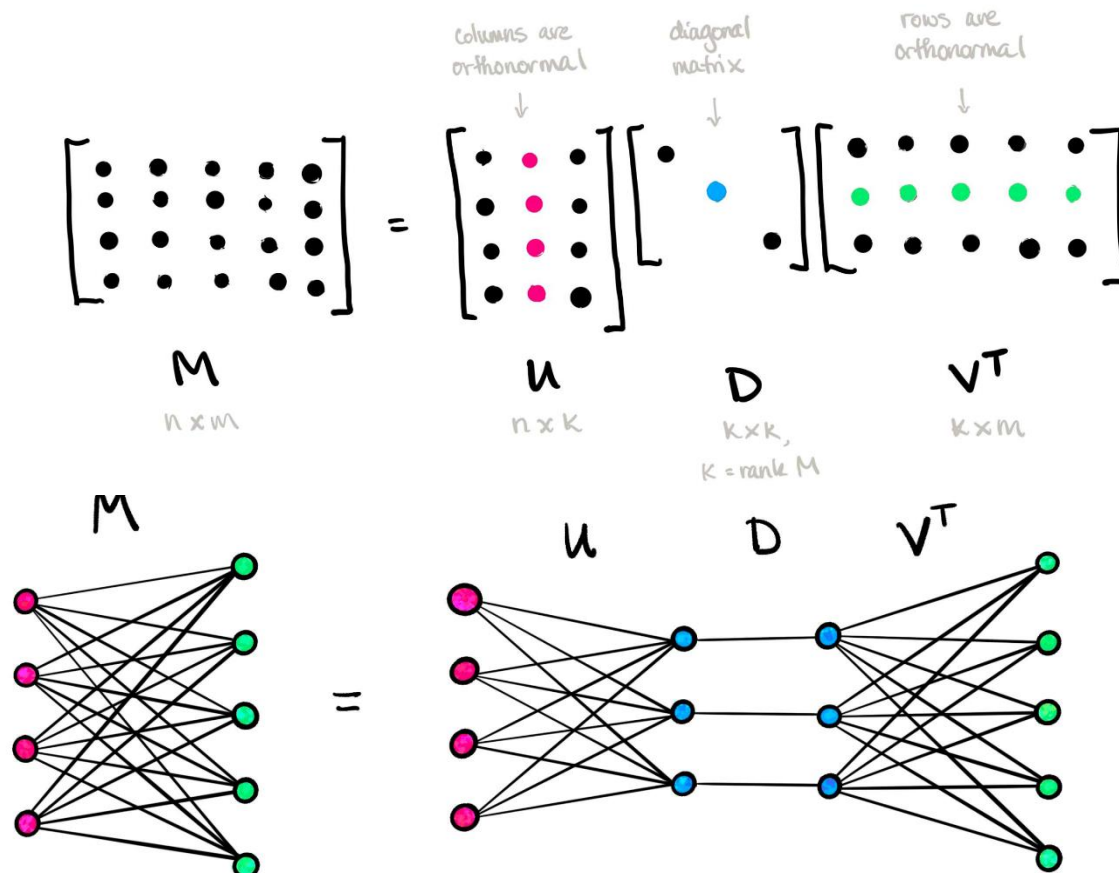
$$\det(UD^2U^T - \lambda I) = 0$$

$$\implies \lambda_i = (\sigma_i)^2$$

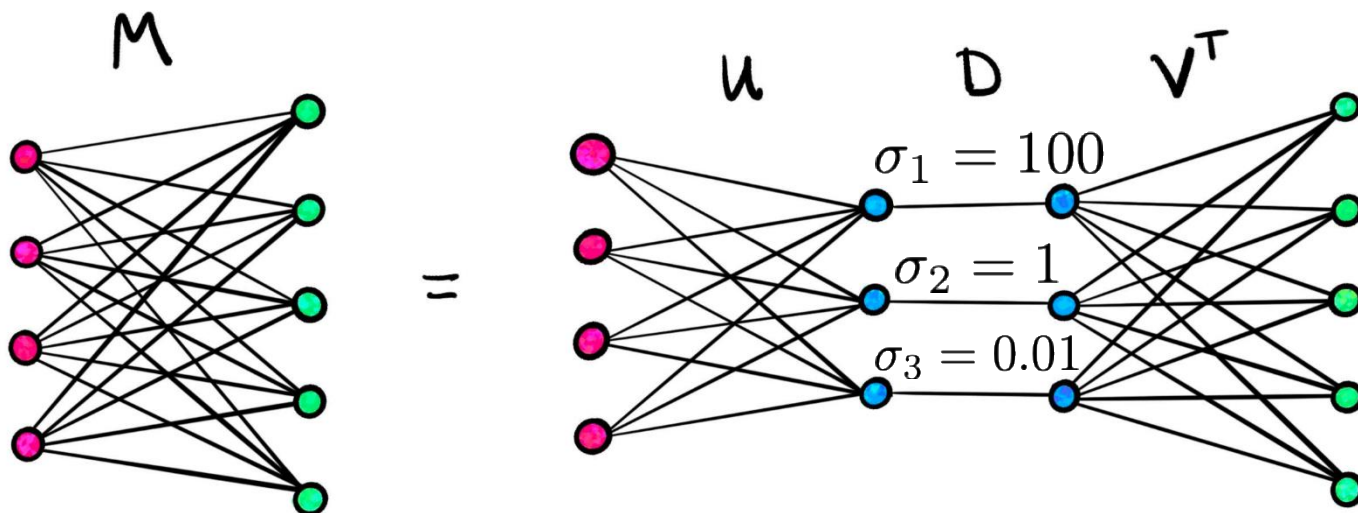
- Repeat with  $M^T M$  and compute the right eigenvector to get  $V$ .

# SVD: Visualization

- The matrix  $D$  can be thought of as a bridge “joining” the rows to the columns. If rank of  $M$  is small, then it means there are a small number of such connections.

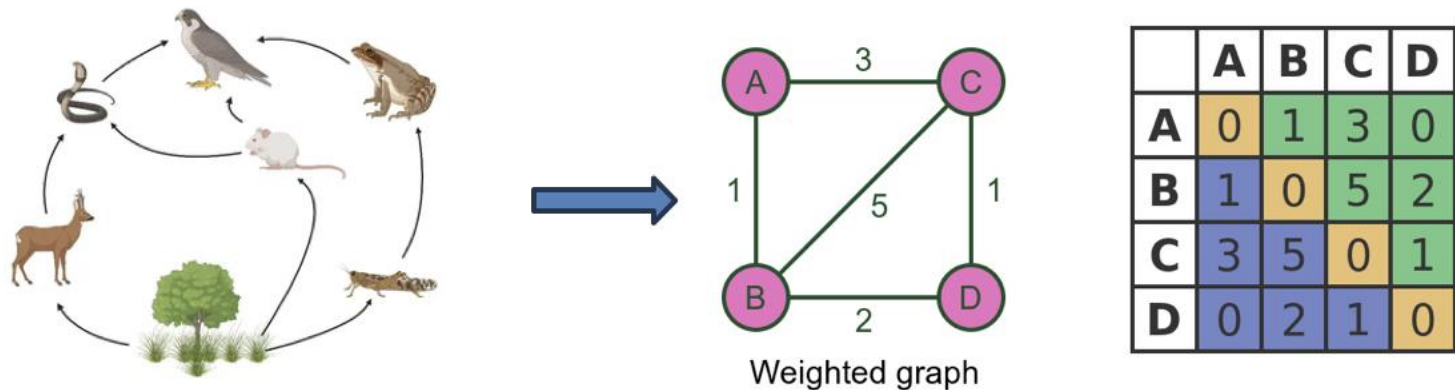


- Intuitively, singular values characterize “interaction” between  $U$  (rows) and  $V$  (column) spaces and how that contributes to total information contained in  $M$ .
- Singular values denote how much information is present in a particular bridge.
- A low rank implies that majority of information is concentrated in first few bridges.



# Complex systems and SVD

- Topology of interactions amongst different constituents of a complex system can be represented as a graph.



- Low-rank of a complex system  $\rightarrow$  low rank of matrix
- Do we keep B & C? or C and A?
  - No! U,V have columns/rows that are superposition of basis vectors
  - Whole point of SVD : creates collective coordinates based on information content.

# Testing low-rank hypothesis on random networks

- The authors test the low-rank hypothesis for several random networks.

$$W = \langle W \rangle + R$$

Where  $W$  is the weight matrix,  $\langle W \rangle$  is the mean weight matrix and  $R$  is the noise matrix.



# Rank-perturbed Gaussian

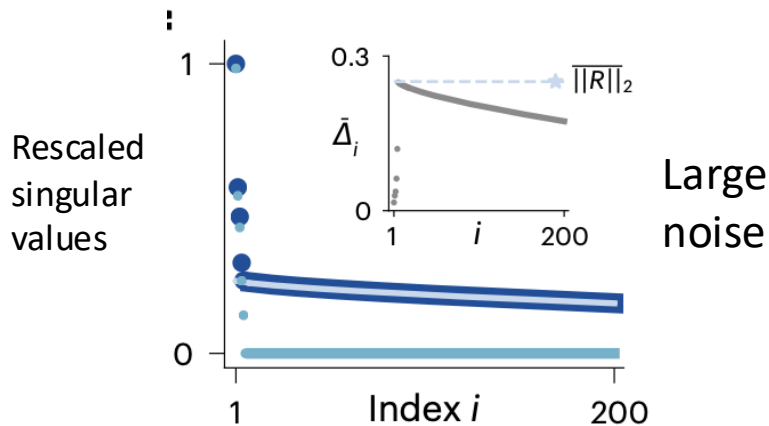
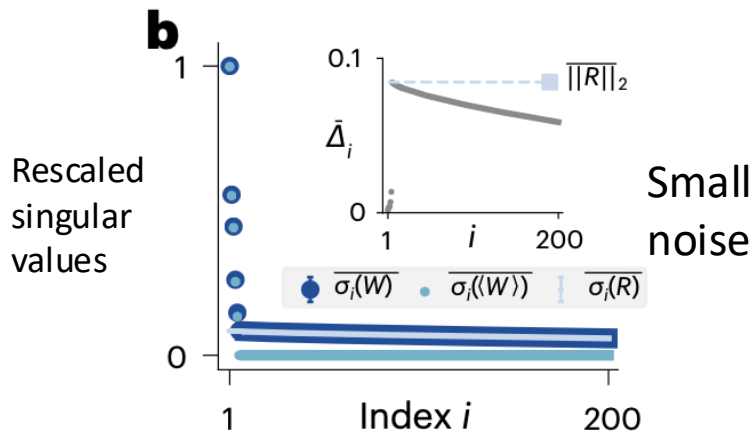
$$\langle W \rangle = \phi(L) = L$$

$$\text{rank}(L) \ll N$$

$R$ : Gaussian

## Indicators of low-rank hypothesis

- Avg weight matrix  $\langle W \rangle$  is a function  $\phi(L)$  of some low rank matrix  $L$ .
- Few dominant singular values, quick decay.
- Low rank: Effective rank should scale sub-linearly with system size.



# Singular values of random networks

**a**

$$\text{Random weight matrix } W = \underbrace{\phi(\text{Low-rank matrix } L)}_{\text{Expected weight matrix } \langle W \rangle} + \text{Random noise matrix } R$$

First indicator

Rank-perturbed Gaussian

$$\langle W \rangle = \phi(L) = L$$

rank  $(L) \ll N$   
R: Gaussian

Degree-corrected stochastic block

$$\langle W \rangle = \phi(L) = L$$

rank  $(L) \leq \text{\#blocks}$   
R: Poisson

Directed  $S^1$  random geometric

$$\langle W \rangle = \phi_{\text{FD}}(L) = \frac{1}{1 + L^{\beta/2}}$$

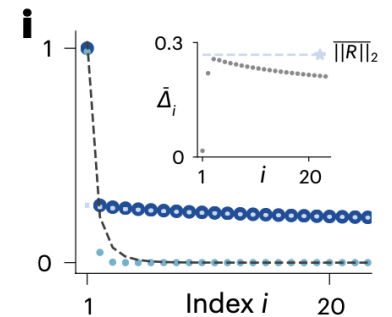
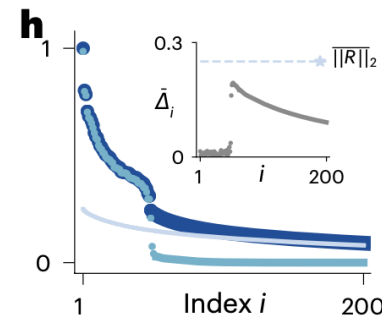
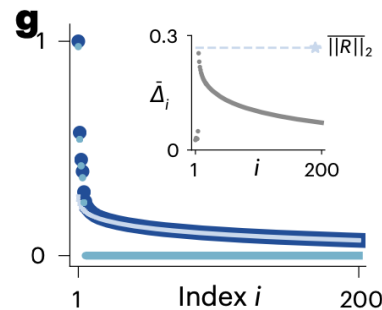
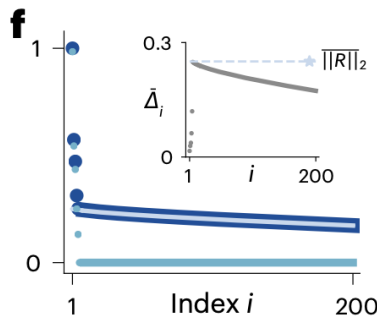
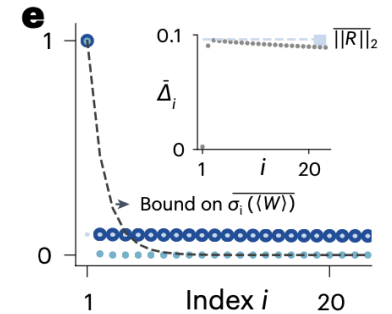
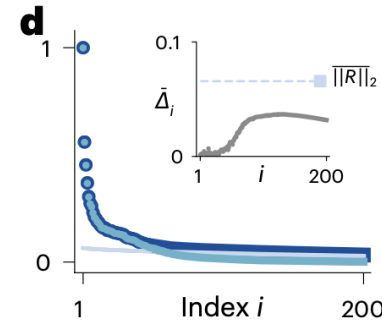
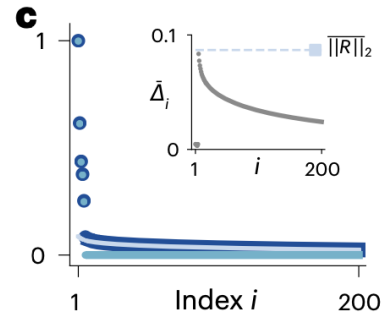
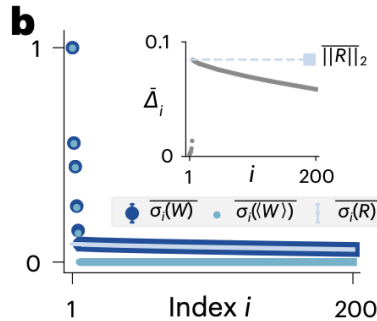
rank  $(L) \leq 3$   
R: Bernoulli

Weighted directed soft configuration

$$\langle W \rangle = \phi_{\text{BE}}(L) = \frac{L}{1 - L}$$

rank  $(L) = 1$   
R: Geometric

Second indicator



$\bullet \overline{\sigma_i(W)}$   $\bullet \overline{\sigma_i(\langle W \rangle)}$   $\text{I} \overline{\sigma_i(R)}$

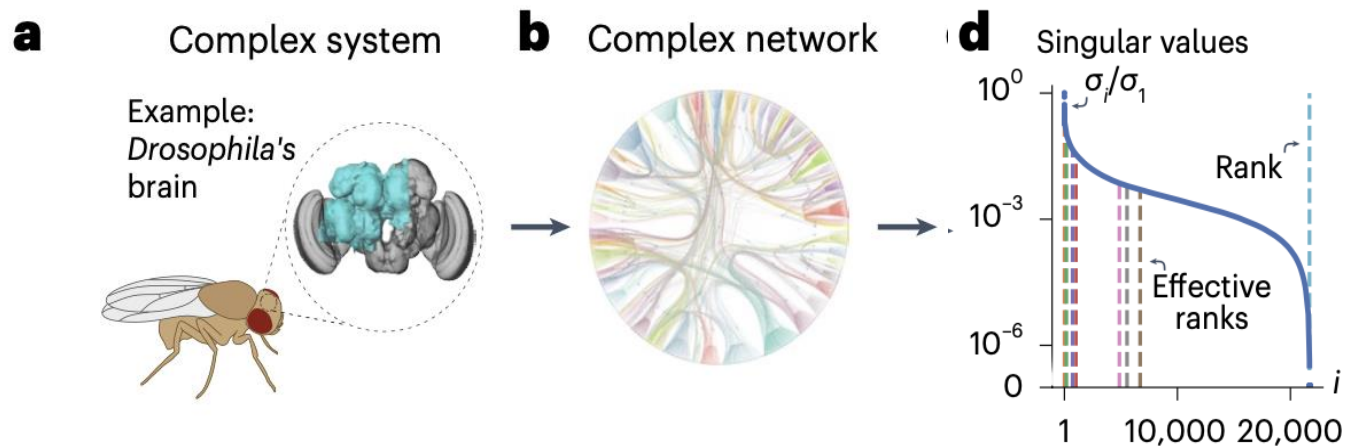
## Different measures of effective ranks

**Table 2 | Different effective ranks of a matrix of dimension  $N \times N$  and of rank  $r$  expressed in terms of its singular values  $\sigma_1 \geq \dots \geq \sigma_N$**

| Abbreviation | Expression  |
|--------------|---|
| srank        | $\sum_{i=1}^r \sigma_i^2 / \sigma_1^2$  |
| nrank        | $\sum_{i=1}^r \sigma_i / \sigma_1$  |
| energy       | $\min \left[ \arg \max_{\ell \in \{1, \dots, N\}} \left( \sum_{i=1}^{\ell} \sigma_i^2 / \sum_{j=1}^r \sigma_j^2 > \tau \right) \right]$   |
| elbow        | $\frac{1}{\sqrt{2}} \arg \max_{i \in \{1, \dots, N\}} \left  \frac{i-1}{N-1} + \frac{\sigma_i - \sigma_N}{\sigma_1 - \sigma_N} - 1 \right  - 1$                                 |
| erank        | $\exp \left[ - \sum_{i=1}^r \frac{\sigma_i}{\sum_{j=1}^r \sigma_j} \log \frac{\sigma_i}{\sum_{j=1}^r \sigma_j} \right]$ <span style="color: red;">e<sup>^(entropy)</sup></span> |
| thrank       | $\# \left\{ \sigma_i \mid i \in \{1, \dots, N\} \text{ and } \sigma_i > \frac{4\sigma_{\text{med}}}{\sqrt{3}\mu_{\text{med}}} \right\}$   |
| shrank       | $\# \{ s^*(\sigma_i) \mid i \in \{1, \dots, N\} \text{ and } s^*(\sigma_i) > 0 \}$  |

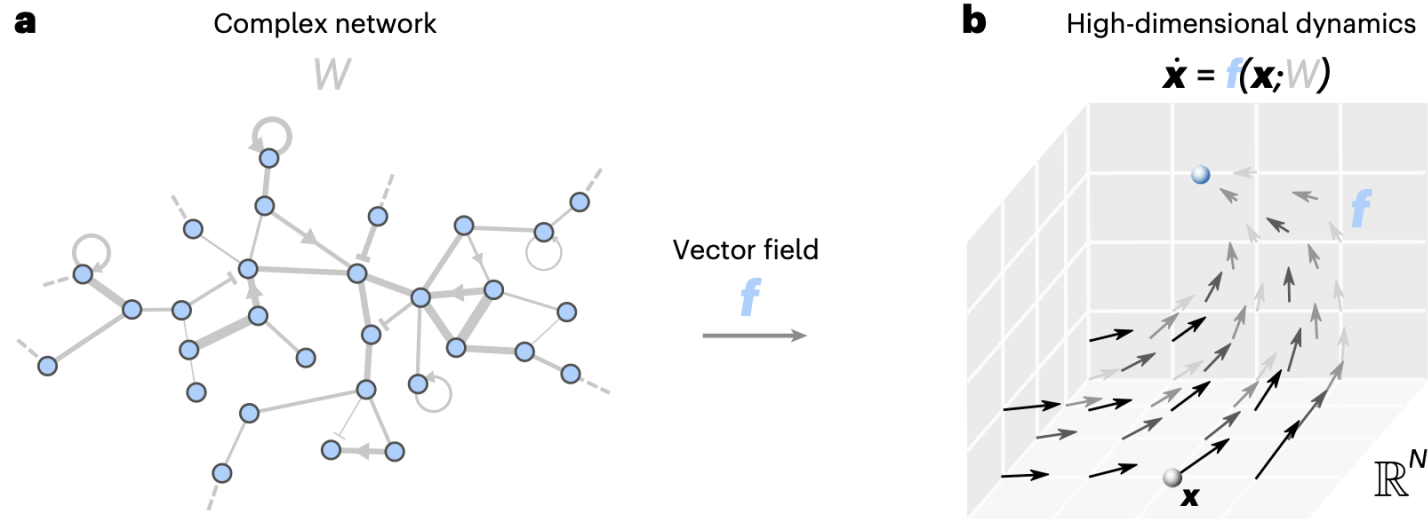
# Testing the low rank hypothesis on real networks

- Looked at singular-value profile of the connectome of *Drosophila melanogaster* for 679 real networks from ten different origins.

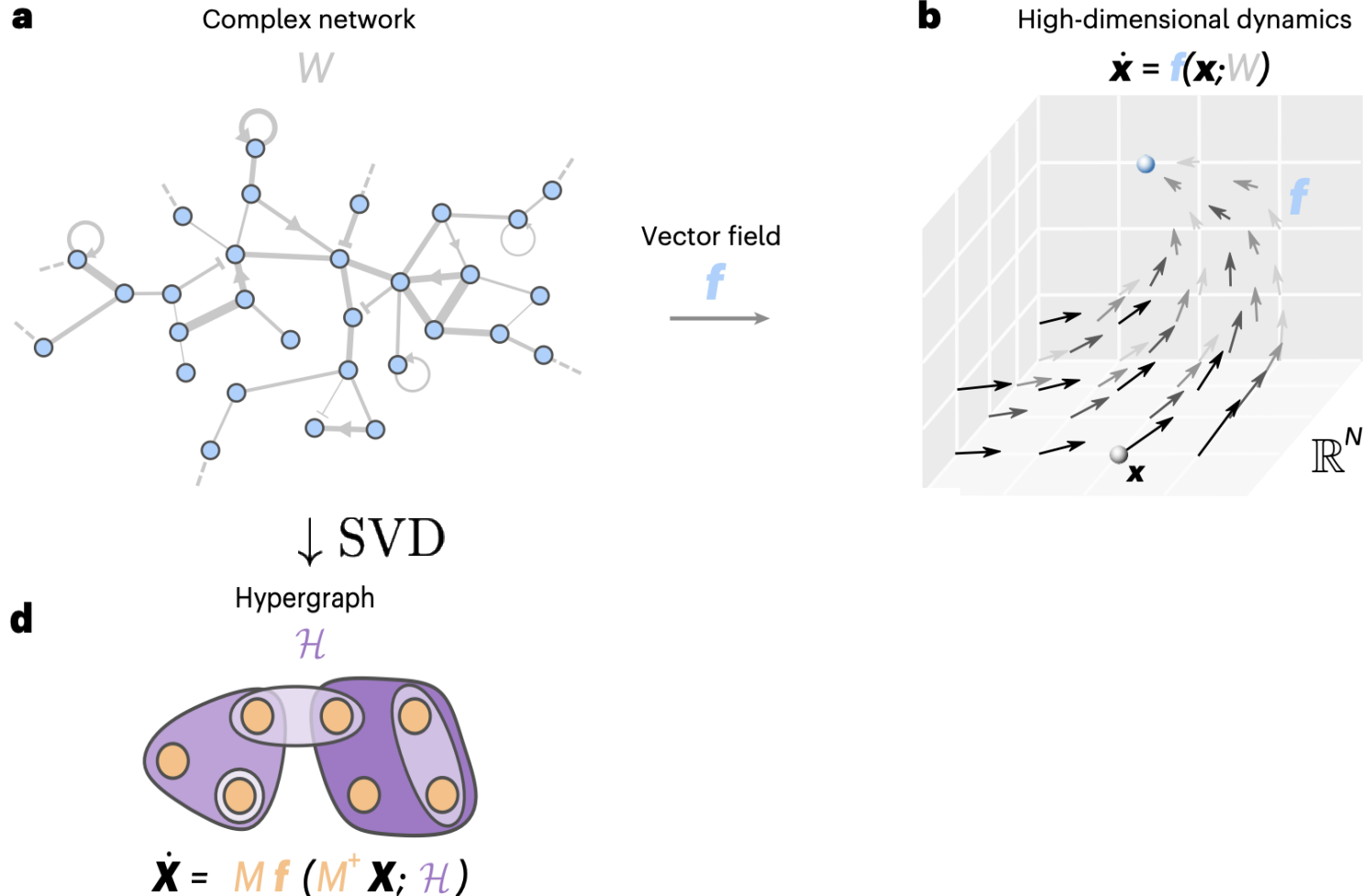


- Such observations seem to be widespread for big data matrices, but they remain puzzling.

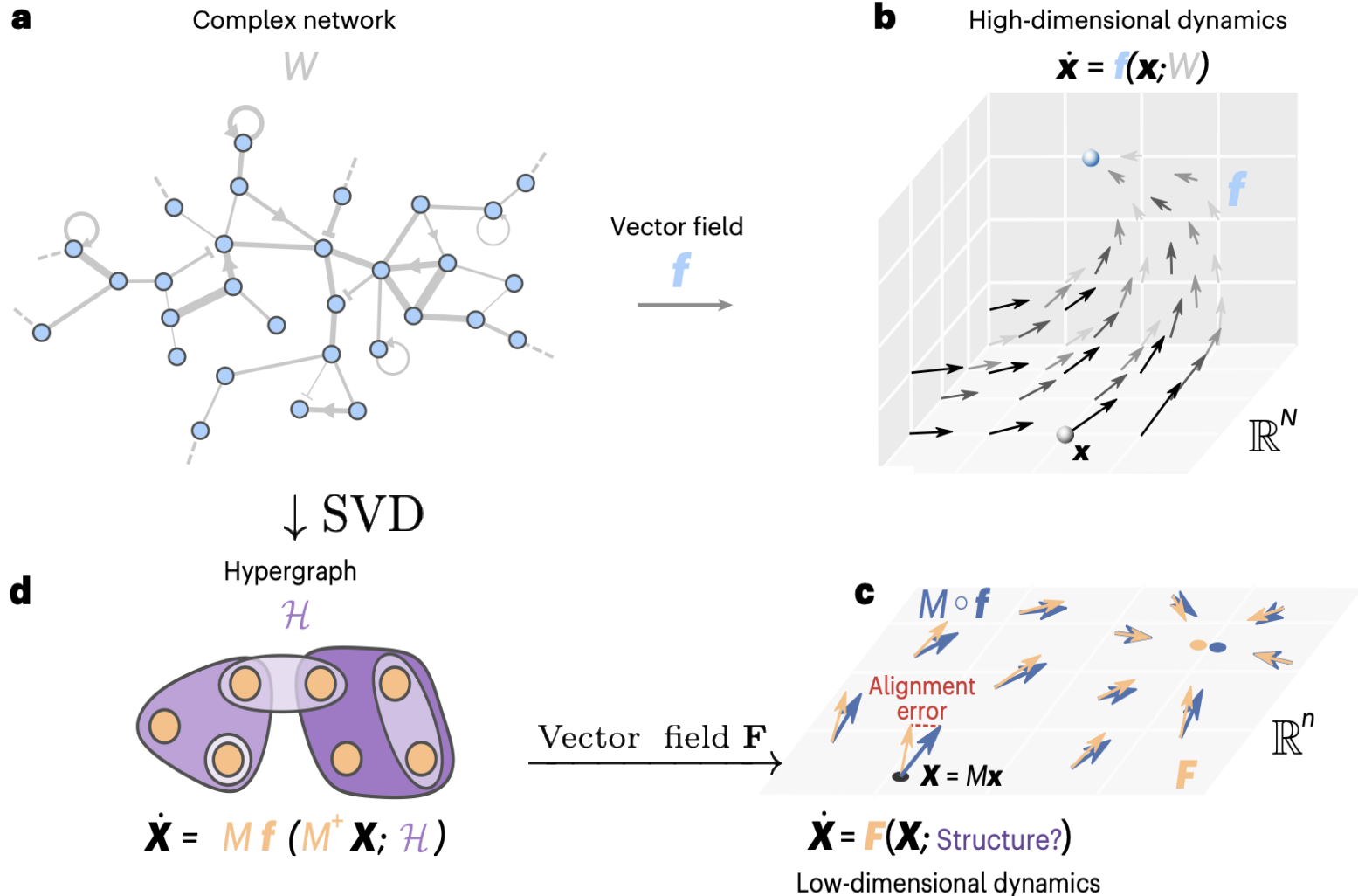
# How does low rank affect dynamics?



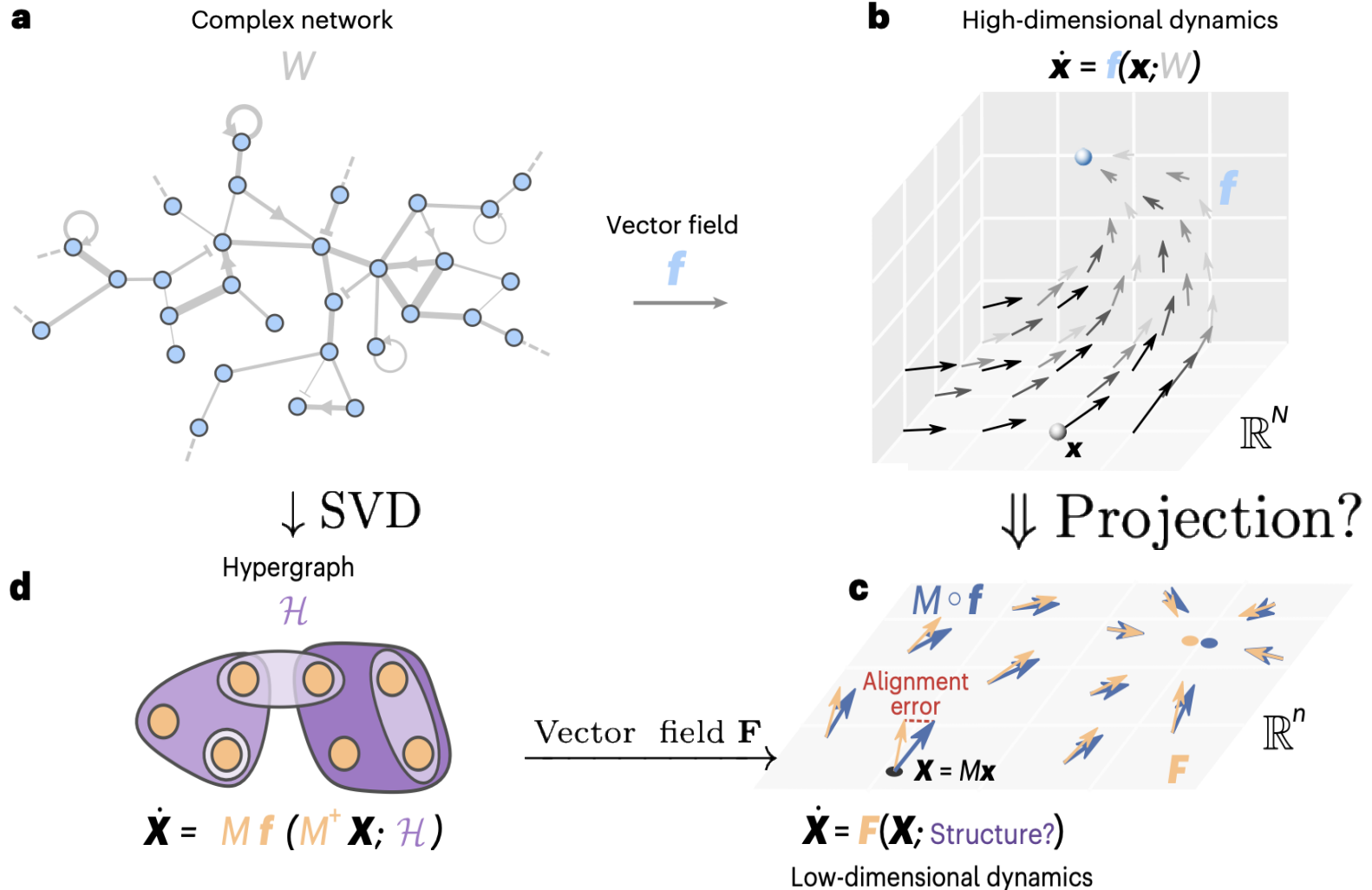
# How does low rank affect dynamics?



# How does low rank affect dynamics?



# How does low rank affect dynamics?





# How does low rank affect dynamics?

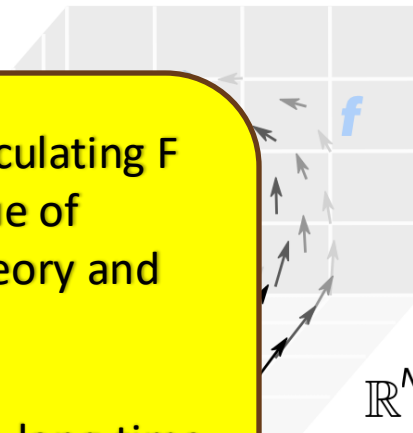
**a** Complex network

$W$



**b** High-dimensional dynamics

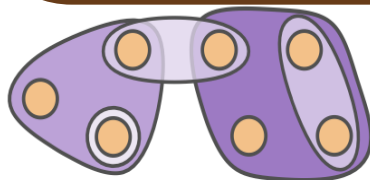
$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}; W)$$



While the authors don't give a recipe for calculating  $\mathbf{F}$  explicitly, it can be done using technique of renormalization group, center manifold theory and similar techniques.

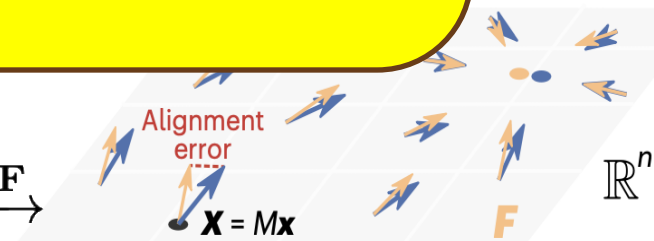
These techniques allow us to extract effective long-time dynamics/reduced system description from the original dynamics.

**d**



$$\dot{\mathbf{X}} = M \mathbf{f}(M^+ \mathbf{X}; \mathcal{H})$$

Vector field  $\mathbf{F}$



$$\dot{\mathbf{X}} = \mathbf{F}(\mathbf{X}; \text{Structure?})$$

Low-dimensional dynamics

- Alignment error:  
Error when comparing observables from reduced dynamics and its higher dimensional representation.
- Good dimensional reduction?
  - Preserve qualitative/quantitative features depending on particular question
- Does low rank  $\rightarrow$  low dimensions? No!
  - Depends on the observables you track.
- The authors consider dynamics of the kind

$$\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}, W) \quad \mathbf{g} : \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}^N$$

where  $W$  describes the  $N \times N$  network. Calculate bounds on optimal alignment error.

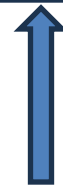
# Alignment error is bounded by singular values

The authors choose projection using SVD to study the effect of low-rank nature on reduced dynamics. Upper bound

$$\sqrt{n} \mathcal{E}(\mathbf{x}) \leq \|V_n^T J'_x (I - P) \mathbf{x}\| + \boxed{\sigma_{n+1}} \|V_n^T J'_y\|_2 \|\mathbf{x}\|$$



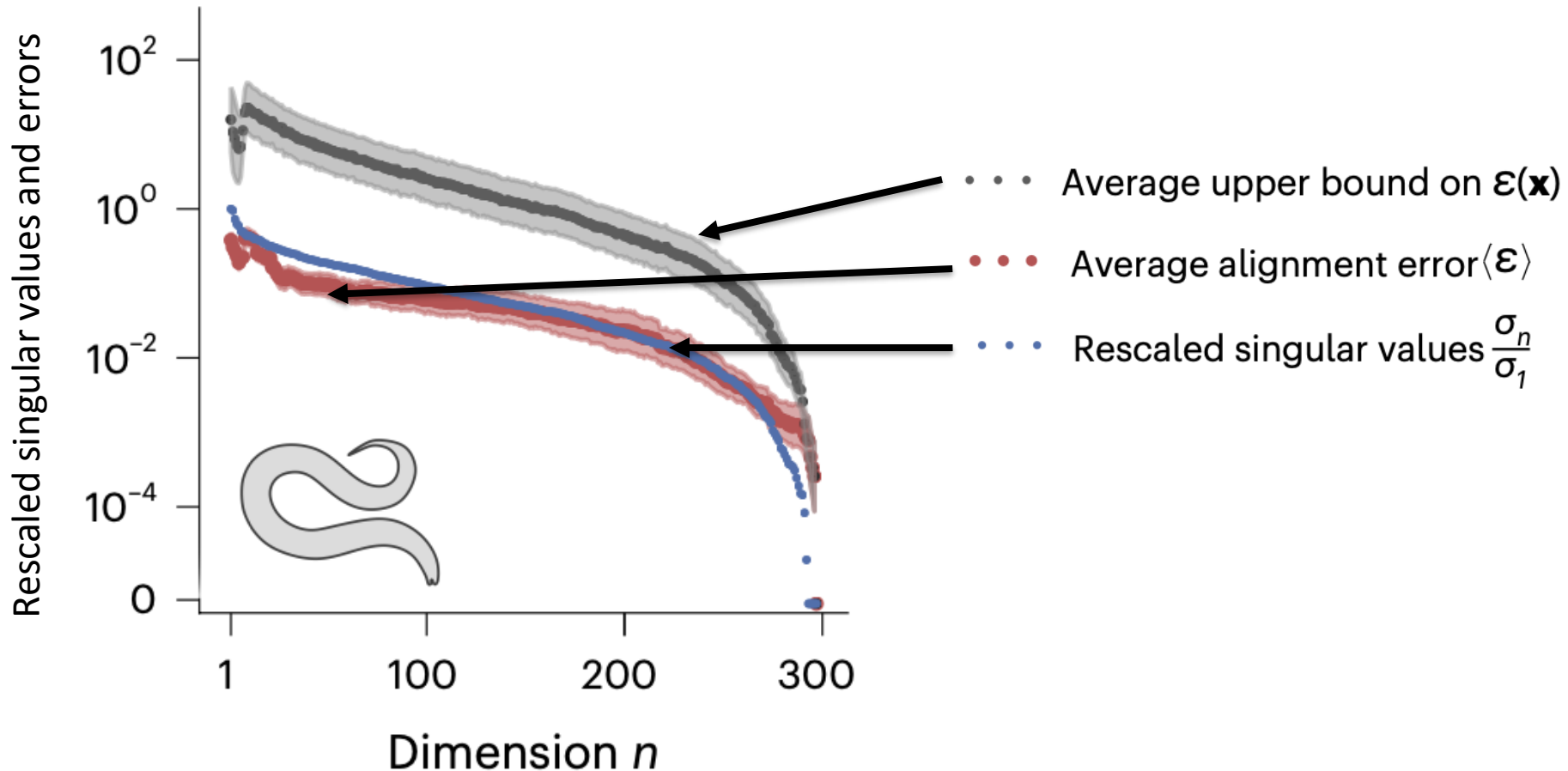
Alignment error



Next order singular value for a rank-n SVD

Vanishes for a class of dynamics (e.g. RNNs and neuronal dynamics).  
Dimension reduction is exact for these models.

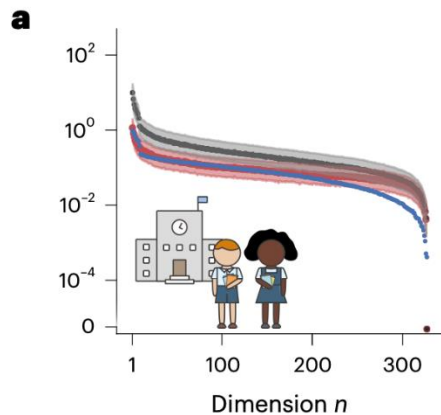
# Alignment error is bounded by singular values



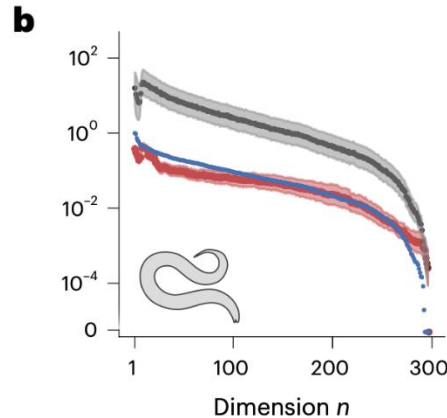
Wilson-Cowan model of neuronal dynamics on the *C. elegans* connectome

# Alignment error is bounded by singular values

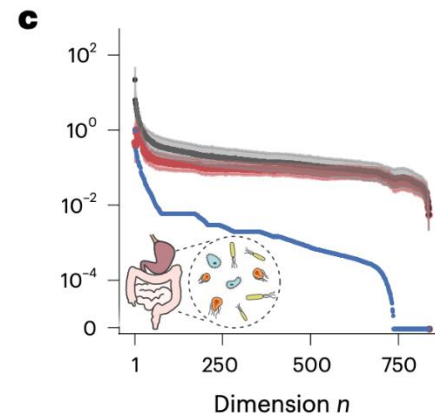
• • • Average alignment error  $\langle \mathcal{E} \rangle$ 
• • • Average upper bound on  $\mathcal{E}(\mathbf{x})$ 
• • • Rescaled singular values  $\frac{\sigma_n}{\sigma_1}$



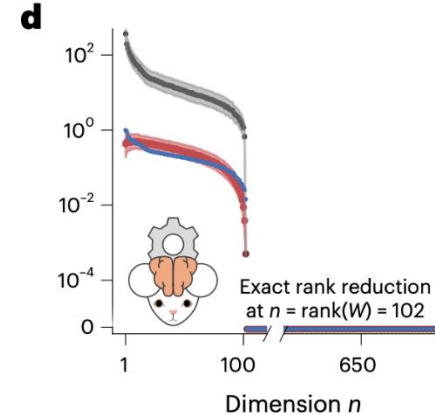
Susceptible-Infected-Susceptible  
 Model of  
 disease spreading  
 On a high school  
 network



Wilson-Cowan model  
 Of neuronal dynamics  
 On the *C.elegans*  
 connectome

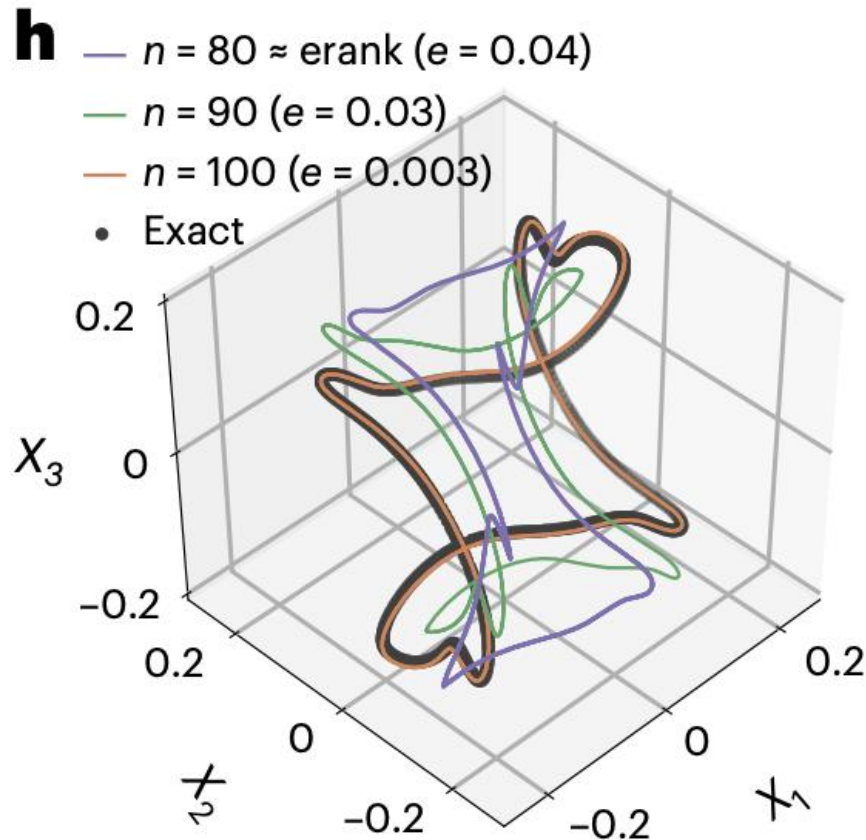


Microbial population  
 dynamics on a human  
 gut microbiome  
 network

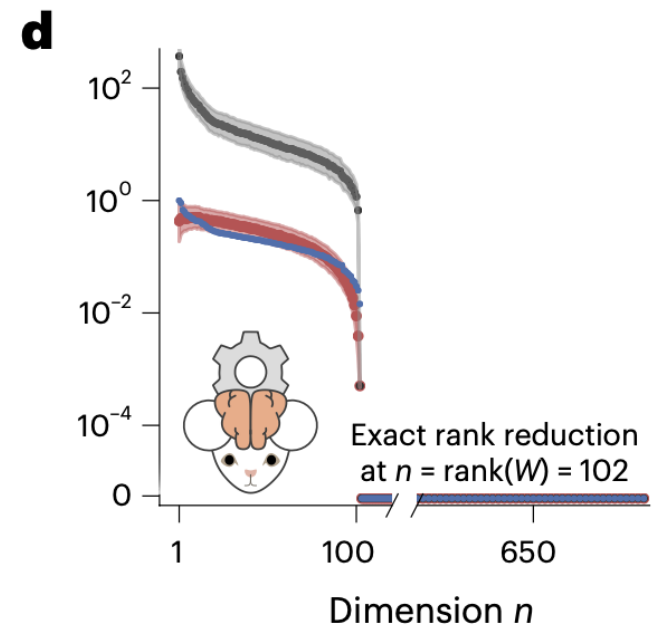


RNN dynamics on  
 a learned network  
 (N = 669, signed,  
 weighted,  
 directed)

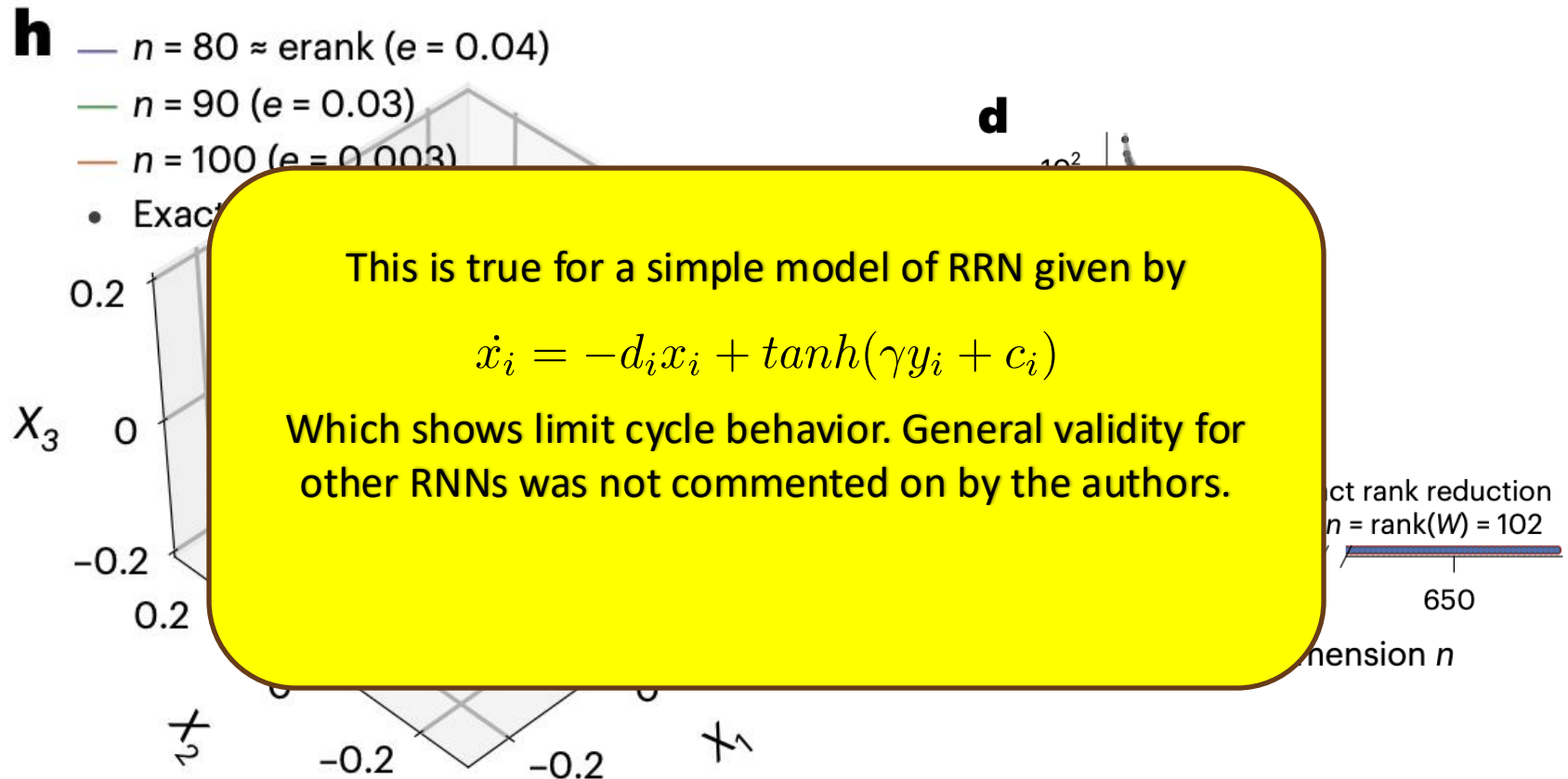
# Dynamics is well captured by low-rank



Limit cycle of RNN dynamics on a learned network ( $N = 669$ , signed, weighted, directed)



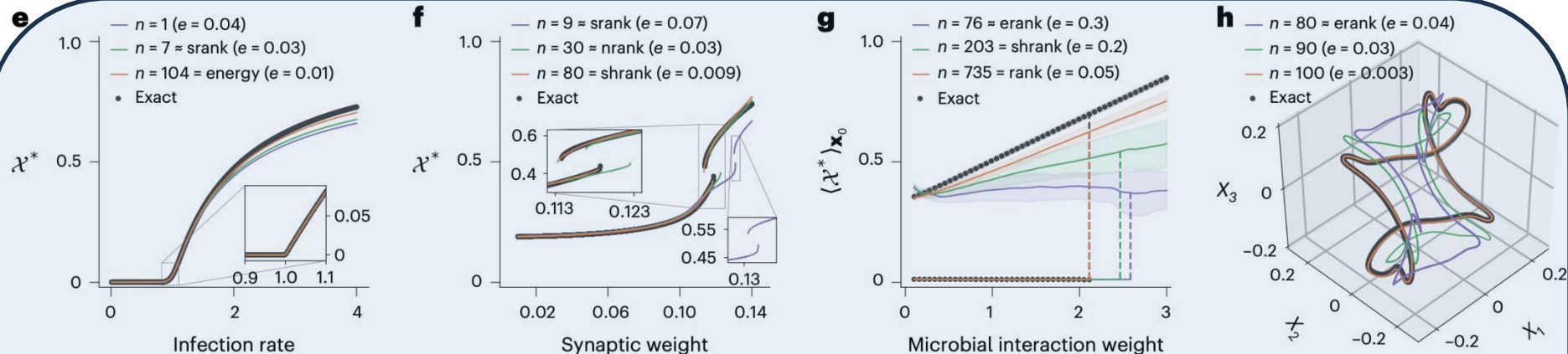
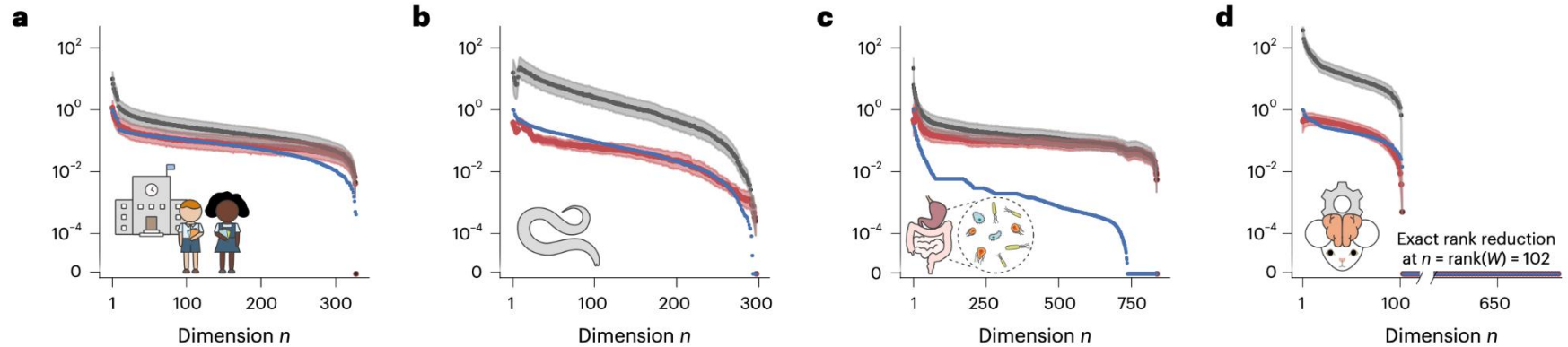
# Dynamics is well captured by low-rank



Limit cycle of RNN dynamics on a learned network (N = 669, signed, weighted, directed)

# Dynamics is well captured by low-rank

• • • Average alignment error  $\langle \mathcal{E} \rangle$     
 • • • Average upper bound on  $\mathcal{E}(\mathbf{x})$     
 • • • Rescaled singular values  $\frac{\sigma_n}{\sigma_1}$



Susceptible-Infected-Susceptible  
 Model of  
 disease spreading  
 On a high school  
 network

Wilson-Cowan model  
 Of neuronal dynamics  
 On the *C.elegans*  
 connectome

Microbial population  
 dynamics on a human  
 gut microbiome  
 network

RNN dynamics on  
 a learned network  
 (N = 669, signed,  
 weighted,  
 directed)



# Emergence of higher order interactions

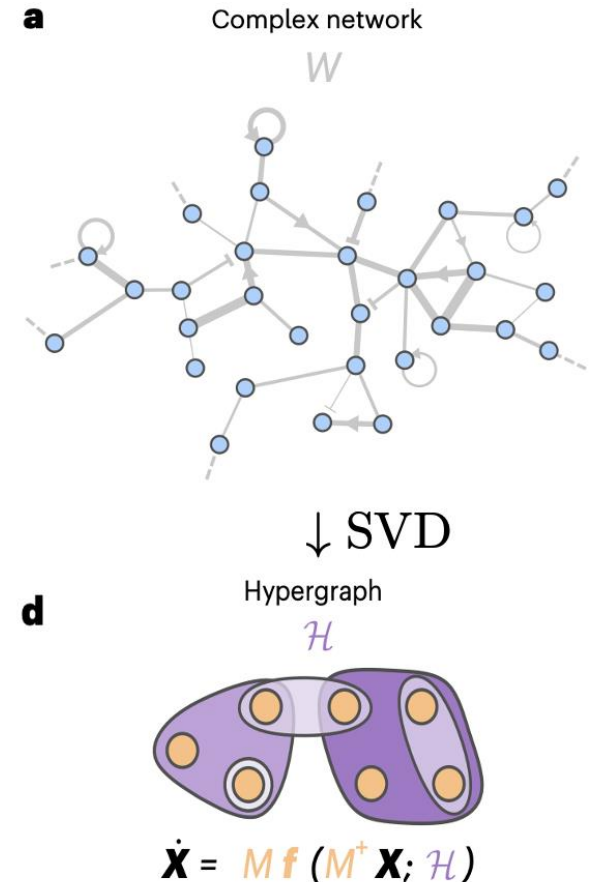
- Reducing the dimensionality produces a higher order interaction in the reduced system.

$$\dot{x}_i = -d_i x_i + \gamma(1 - x_i)y_i$$



$$\dot{X}_\mu = \sum_{v=1}^n (\mathcal{D}_{\mu v} + \mathcal{W}_{\mu v}) X_v + \sum_{v,k=1}^n \mathcal{T}_{\mu v k} X_v X_k$$

- Reason? Clustering: Reducing dimensions creates clusters of vertices from the original network. Observables depend on all vertices in a cluster.
- More generally, non-linear dynamics.



**Are there systems which are not low-rank?**

**Yes! The brain**

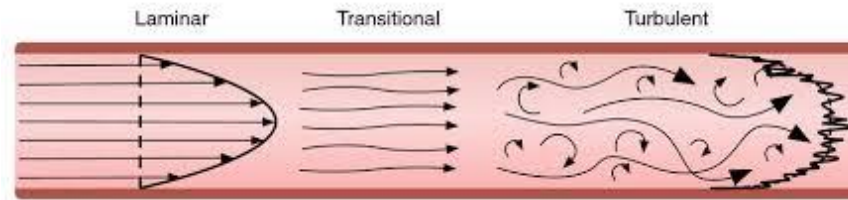
# The brain: Criticality conjecture

- Case for brain at criticality
  - Claim: Brain operates near a critical point of a phase transition
  - Neural networks self-organize near a phase transition between subcritical (low activity) and supercritical (overactive) regimes.
  - Not low rank, need all scales.
- Evidence for
  - Neuronal Avalanches follow power-law distribution<sup>1</sup>
  - High sensitivity to perturbations, optimal information processing
  - Critical slowing down
- Evidence against
  - Multiple neural states, transition b/w different regimes. Not at critical point
  - Non-power law fits, features could arise from different mechanisms
  - Local criticality, not global

1. Beggs, J. M.; Plenz, D. (2003). ["Neuronal avalanches in neocortical circuits"](#). *The Journal of Neuroscience*. **23** (35): 11167–11177.<sup>39</sup>

# Are there systems which are not low-rank?

- Phase transitions
- Fluid transition to turbulence:
  - Low rank approach: study most unstable modes and their linear instability as a function of Reynolds number.
  - Linear instability
  - not only doesn't predict "ANY" transition, but in cases where it does, it predicts an incorrect critical  $Re^1$ .



- Economics

1. Shih, HY., Hsieh, TL. & Goldenfeld, N. Ecological collapse and the emergence of travelling waves at the onset of shear turbulence. *Nature Phys* 12, 245–248 (2016).

2. Lemoult, G. et al. Directed percolation phase transition to sustained turbulence in Couette flow. *Nat. Phys.* 12, 254–258 (2016).

Picture credit: Heritability of haemodynamics in the ascending aorta. *Scientific Reports*. 2020. 14356. 10.1038/s41598-020-71354-7.

# My Perspective

- The authors demonstrate low-rank quantifiers and validate dimension reduction across various models and data but lack predictive tools for identifying low-rank systems during modeling.
- Their analysis is post-dictive, justifying low-rank structures retrospectively, but extending it to new problems requires additional signatures.

# Conclusion

- Low-rank networks seem ubiquitous in nature. They allow for dimension reduction, modelling of reduced dynamics and are easier to interpret.
- The authors quantify metrics of effective ranks for networks and give sufficient (but not necessary) conditions for a complex network to be low rank.
- Give a mechanism for emergence of higher order interactions .  
As the system is probed at a coarser scale, these higher order interactions would likely manifest themselves.
- They show exact dimension reduction for models of RNNs, neuronal dynamics, disease spreading and microbial population dynamics.







# Reserve Slides

- Good for approximation. What else?
- Used to extract relevant subspace of a dataset, commonly under the technique of principal component analysis (PCA).
- For a low rank matrix i.e. small  $k$ , it can impute missing data values of a matrix. Replace the missing value with the column mean, and compute the SVD up to required rank. Repeat this until the entries stop changing.
- Also a number of other applications.

| Name            | Complete vector field $h_i(x_i, y_i)$  | Reduced vector field $H_\mu(X_1, \dots, X_n)$   |
|-----------------|--|---|
| Epidemiological | $-d_i x_i + \gamma(1 - x_i) y_i$   | $\sum_{\nu=1}^n (\mathcal{D}_{\mu\nu} + \mathcal{W}_{\mu\nu}) X_\nu + \sum_{\nu, \kappa=1}^n \mathcal{T}_{\mu\nu\kappa} X_\nu X_\kappa$   |
| Microbial       | $a - d_i x_i + b x_i^2 - c x_i^3 + \gamma x_i y_i$                             | $\mathcal{C}_\mu + \sum_{\nu=1}^n \mathcal{D}_{\mu\nu} X_\nu + \sum_{\nu, \kappa=1}^n (\mathcal{D}_{\mu(\nu, \kappa)} + \mathcal{T}_{\mu\nu\kappa}) X_\nu X_\kappa + \sum_{\nu, \kappa, \tau=1}^n \mathcal{D}_{\mu(\nu, \tau, \kappa)} X_\nu X_\kappa X_\tau$ |
| Oscillator      | $i\omega_i x_i + \gamma e^{-i\alpha} y_i - \gamma e^{i\alpha} x_i^2 \bar{y}_i$ | $\sum_{\nu=1}^n (\mathcal{D}_{\mu\nu} + \mathcal{W}_{\mu\nu}) X_\nu + \sum_{\nu, \kappa, \tau=1}^n \mathcal{T}_{\mu(\nu, \kappa)\tau} X_\nu X_\kappa \bar{X}_\tau$  |
| RNN             | $-d_i x_i + \tanh(\gamma y_i + c_i)$   | $\sum_{\nu=1}^n \mathcal{D}_{\mu\nu} X_\nu + \sum_{i=1}^N M_{\mu i} \tanh \left( \gamma \sum_{\nu=1}^n \mathcal{W}_{j\nu} X_\nu + c_i \right)$  |
| Neuronal        | $-d_i x_i + (1 - a x_i) \mathcal{S}[b(\gamma y_i - c_i)]$                      | $\sum_{\nu=1}^n \mathcal{D}_{\mu\nu} X_\nu + \sum_{j=1}^N M_{\mu j} \left( 1 - a \sum_{\nu=1}^n M_{j\nu}^+ X_\nu \right) \mathcal{S} \left[ b \left( \gamma \sum_{\kappa=1}^n \mathcal{W}_{j\kappa} X_\kappa - c_i \right) \right]$                           |

# Alignment error is bounded by singular values

- If one focuses on observables  $X_\mu$  in a lower dimension  $R^n$  ( $n < N$ ) then alignment error is defined (in a least squares sense)

$$\|F \circ X - M \circ g\|$$

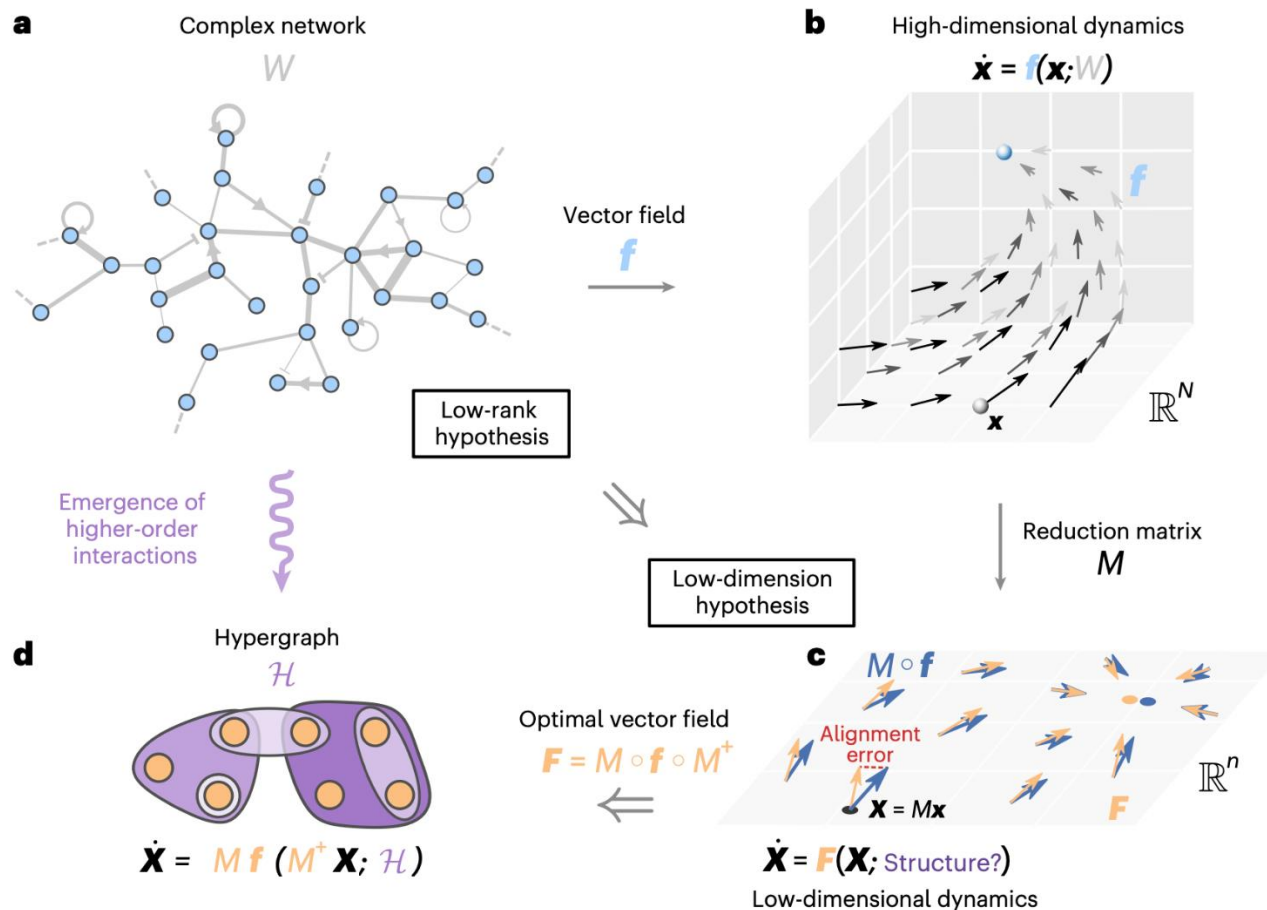
where  $F$  denotes the new dynamics in the reduced dimensions and  $M$  denotes a mapping  $R^N \rightarrow R^n$  between the complete system and the reduced system.

- The choice of  $(F, M)$  depends on the modeler's objective. The authors choose  $M$  as  $V^T$  using the SVD decomposition to study the effect of low-rank nature on reduced dynamics. This allows the

$$\sqrt{n} \mathcal{E}(\mathbf{x}) \leq \|V_n^T J'_x (I - P) \mathbf{x}\| + \sigma_{n+1} \|V_n^T J'_y\|_2 \|\mathbf{x}\|.$$

# How does low rank affect dynamics?

- Intuitively, low rank would help ..... Reduced dimensions.











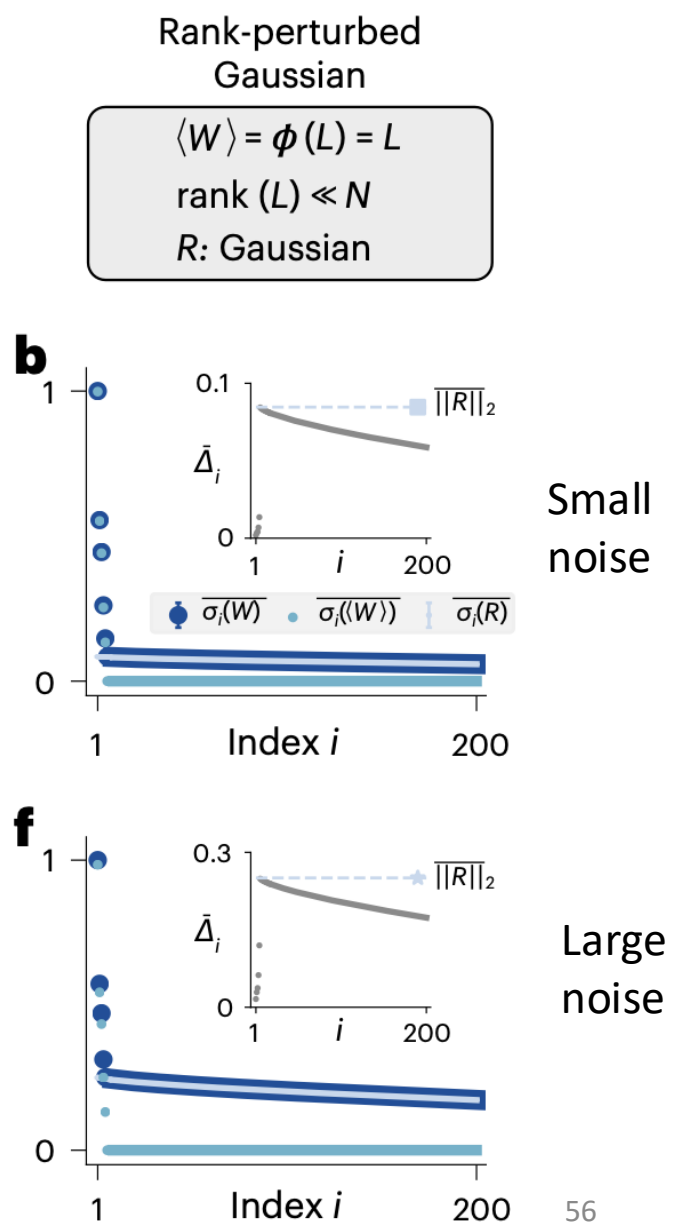


# Prep

- What is the low-rank hypothesis?
  - Emphasize the complexity – simplicity duality
  - What do we know before?
- What is SVD and why?
- Network models are low rank?
  - Quantify low-ness
  - Verification for real networks
- Low rank dynamics:
  - Induced low dimensions from low rank graphs (paper)
  - Not low rank data but low rank learned dynamics (Litwin's paper)
- Emergence of higher order interactions
- Low rank can be incorrect: examples
- Another example of same theme: Sloppy models
- Takeaways and relevance for this class
  - RNNs, dimensional reduction and other things
  - What do we know after?
  - What did it set out to do and where does it lack?



- Indicators of low-rank hypothesis
  - Avg weight matrix  $\langle W \rangle$  is a function  $\phi(L)$  of some low rank matrix  $L$ .
  - Few dominant singular values, quick decay, subdominant ones are affected by noise.
  - Should have a low effective rank i.e. at mo with large  $N$  (# of vertices) as
 
$$O(N^{1-\epsilon}) \quad \epsilon > 0$$
  - Implies that effective rank over system size vanishes in the large  $N$  limit.





# Dimension reduction in physics

- Sloppy models
- Renormalization group

- Procaccia-Grassberger algorithm



