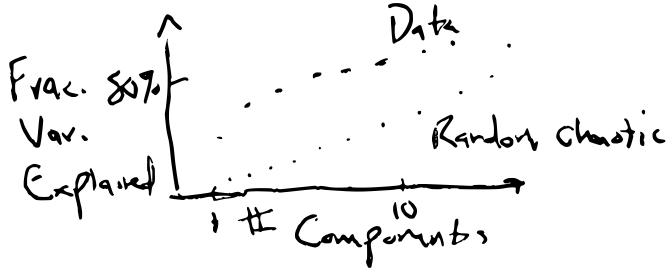
Last time: Chaotic RNNs

This time: Low-rank RNNs (Mastrogiuseppe & Ostojic 2018)

Observations:

1) Low-d dynamics in motor cortex during tasks:



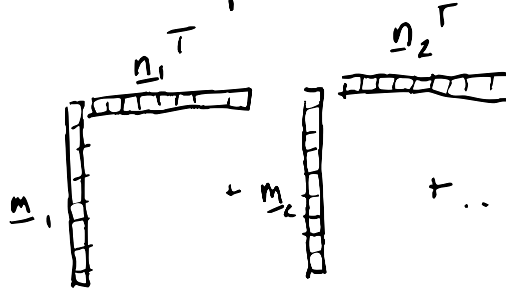
2) Trained RNNs learn (approx.) low-rank updates to initial random connectivity (Schuessler et al. 2020)

$J = J_0 + P$ ← learned

Rank of a matrix:

$$P_{ij} = \sum_{r=1}^{R} m_{r,i}\, n_{r,j} \quad \text{(2 vectors per rank)}$$

$$P = \sum_{r=1}^{R} \underline{m}_r \underline{n}_r^T$$



$$P = M N^T$$

P: #row × #col

M: #row × R

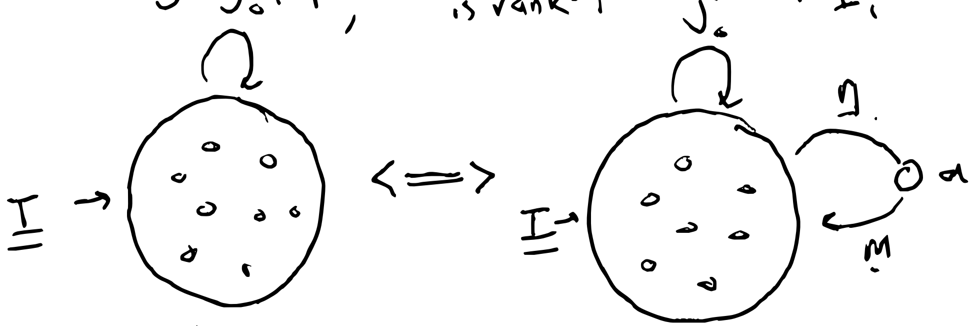
N^T: #col × R

See SVD.

RNN with low-rank component:

$$\frac{dx_i}{dt} = -x_i + \sum_{j=1}^{N} J_{ij}\, \phi_j + I_i$$

$J = J_0 + P$, P is rank-1



$$a = \sum_{j=1}^{N} n_j\, \phi_j$$

$$\frac{dx_i}{dt} = -x_i + \sum_{j=1}^{N} J_{ij,0}\, \phi_j + m_i a$$

See e.g. Sussillo & Abbott 2009 (FORCE learning)

a "listens to" $\underline{n}$ direction and "talks to" $\underline{m}$ direction.

$\underline{m}_r$, $\underline{n}_r$, $\underline{I}$ are "special" directions in activity space.

Example: $\underline{I} = 0$, $R = 1$.

$$J_{ij,0} \sim \mathcal{N}(0, g^2/N) \qquad O(1/\sqrt{N})$$

$$P_{ij} = \frac{1}{N} m_i n_j \qquad O(1/N)$$

If $\underline{\phi}$ is independent (& independent of J),

$$\mathrm{Var}\left(\sum_{j=1}^{N} J_{ij} \phi_j\right) = N \, \mathrm{Var}(J_{ij}) \, \mathrm{Var}(\phi_j)$$

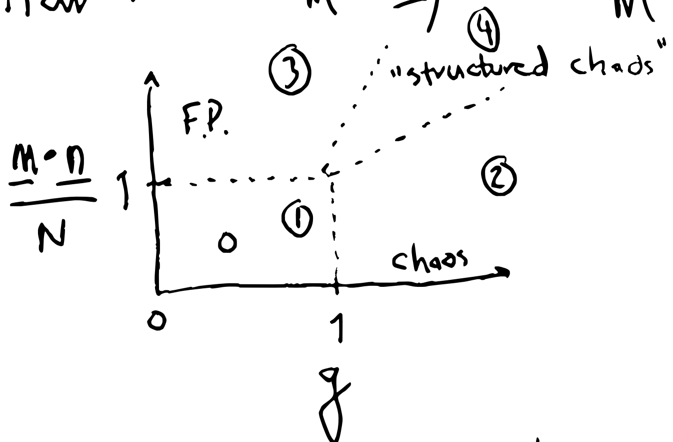
$$= \frac{g^2}{N} \cdot N \cdot \mathrm{Var}(\phi_j)$$

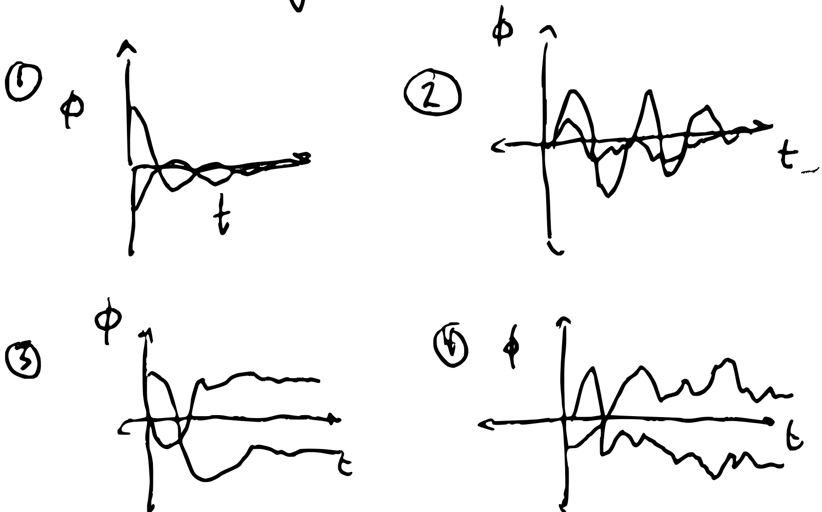
$$\sum_{j=1}^{N} P_{ij} \phi_j = \frac{1}{N} m_i \sum_j n_j \phi_j$$

If $\underline{n}$ and $\underline{\phi}$ are independent, small.

If $\frac{1}{N}(\underline{n}\cdot\underline{\phi}) > 0$, $O(1)$.

How to make $\frac{1}{N}(\underline{n}\cdot\underline{\phi}) > 0$? $\frac{1}{N}(\underline{m}\cdot\underline{n}) > 0$.





DMFT: $K = \frac{1}{N}\left(\underline{n} \cdot \underline{\phi}\right)$

Last class: $K = 0$.

Self-consistent equations for K, $C(\tau)$

$$\frac{dx_i}{dt} = -x_i + \eta_i(t)$$

At steady state:

$$\bar{x}_i = \langle \eta_i \rangle_t = m_i K$$

Let $\Delta_x(\tau) = \langle (x_i(t) - \bar{x}_i)(x_i(t+\tau) - \bar{x}_i) \rangle_i$

$$\mu = \langle \eta \rangle_{t,i} = \langle m_i \rangle K$$

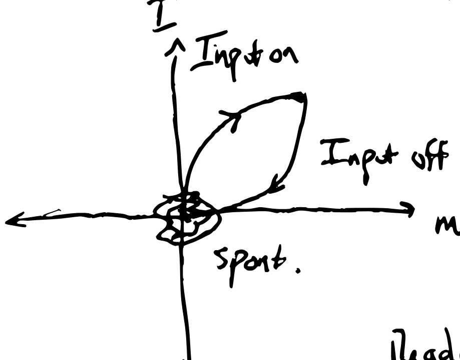
$$\Delta_x(\tau) = \underbrace{g^2 \Delta_\phi(\tau)}_{C(\tau) \text{ from last class}} + \underbrace{\left(\langle m_i^2 \rangle - \langle m_i \rangle^2\right) K^2}_{\text{extra variance from structure}}$$

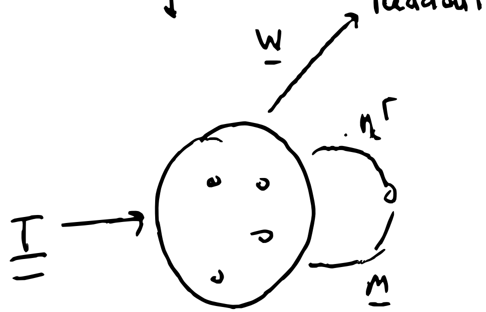
If statistics of m, n specified, can calculate μ, $\Delta_x(\tau)$ self-consistently

For higher-rank networks, interactions described by $\kappa_1 \ldots \kappa_r = \frac{1}{N}(\underline{n}_r \cdot \underline{\phi})$

Activity lies in subspace of $\underline{m}_1 \ldots \underline{m}_r$ and $\underline{I}$.

Example: $\underline{m}$, $\underline{n}$ orthogonal but $\frac{1}{N}(\underline{n} \cdot \underline{I}) > 0$.





If $\frac{1}{N}(\underline{m} \cdot \underline{w}) > 0$, can selectively respond to $\underline{I}$ & not other inputs.

Rank 2+: 1) Oscillations

2) Context-dependent responses