DPZ
Deutsches Primatenzentrum
Leibniz-Institut für Primatenforschung

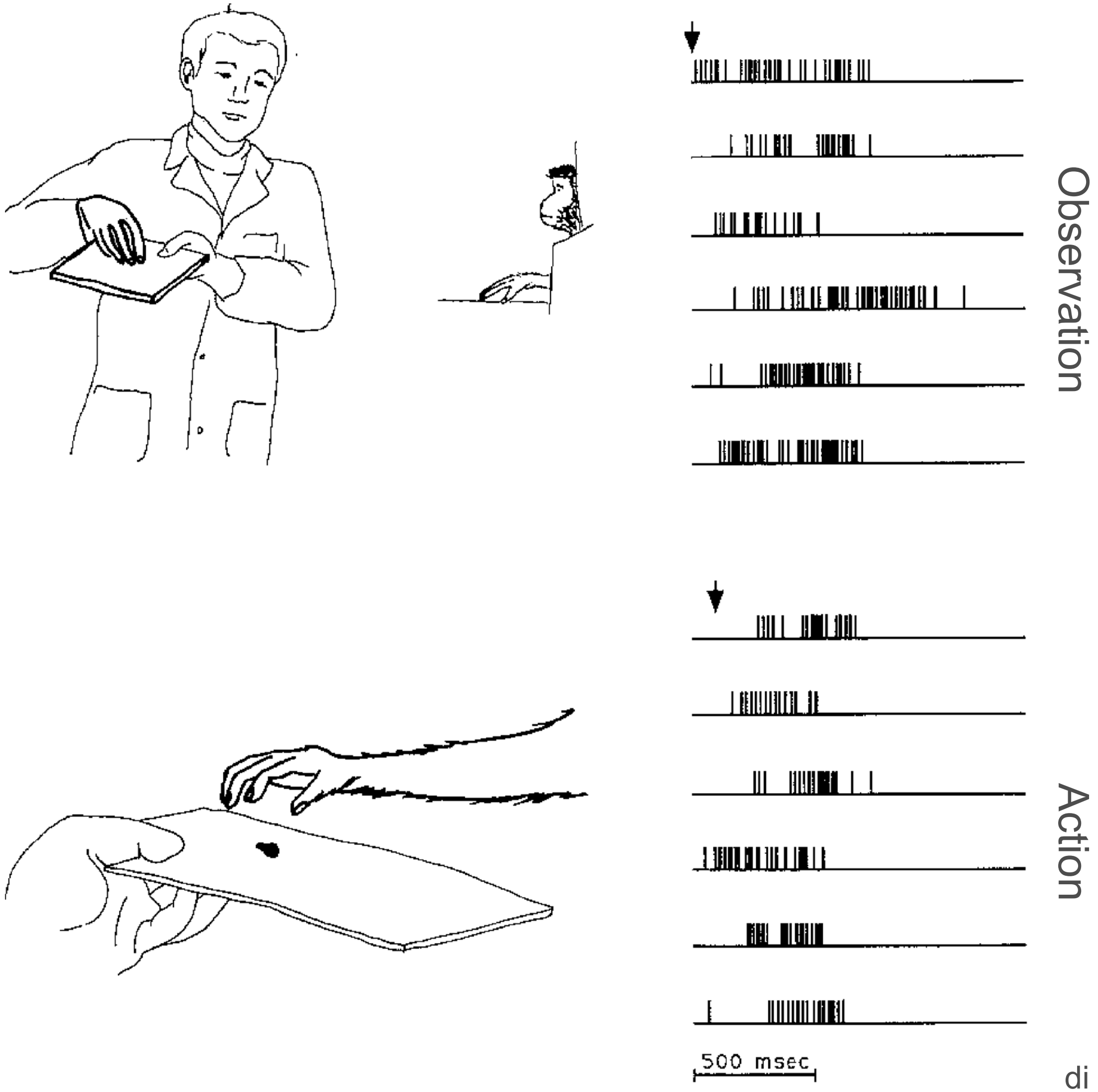
Structure in Neural Activity during Observed and Executed Movements Is Shared at the Neural Population Level, Not in Single Neurons

Jiang X., Saggar H., Ryu S.I., Shenoy K.V., and Kao J.C.

Presented by: James Goodman
NBL Journal Club
08.12.2020

Mitglied der Leibniz-Gemeinschaft

Mirror Neurons respond during action and observation



di Pellegrino et al. 1992 *Exp. Br. Res.*

Mitglied der Leibniz-Gemeinschaft

Mirror neurons: a contentious, possibly overhyped concept

Neuroscience Needs Behavior: Correcting a Reductionist Bias

or behavioral hypothesis is being tested per se. Thus, an interpretation is being mistaken for a result; namely, that the mirror neurons understand the other individual. Additional behavioral evidence that the participant understands the other individual is lacking. This tendency to ascribe psychological properties to

John W. Krakauer,[1,*] Asif A. Ghazanfar,[2] Alex Gomez-Marin,[3] Malcolm A. MacIver,[4] and David Poeppel[5,6]

2017 *Neuron*

Eight Problems for the Mirror Neuron Theory of Action Understanding in Monkeys and Humans

Gregory Hickok

humans. The early hypothesis that these cells underlie action understanding is likewise an interesting and prima facie reasonable idea. However, despite its widespread acceptance, the proposal has never been adequately tested in monkeys, and in humans there is strong empirical evidence, in the form of physiological and neuropsychological (double) dissociations, against the claim.

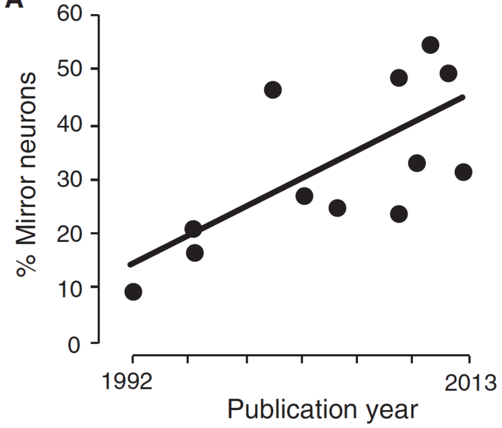
2009 *J Cog. Neuro.*

What We Know Currently about Mirror Neurons

J.M. Kilner and R.N. Lemon

2013 *Curr. Biol.*

A



"Mirror mechanism" may be better understood in neuronal state space

Mirror Neuron Populations Represent Sequences of Behavioral Epochs During Both Execution and Observation

Kevin A. Mazurek,[1,2] Adam G. Rouse,[2,3] and Marc H. Schieber[1,2,3,4]

2018 *J Neurosci.*

Neurons in the Macaque Dorsal Premotor Cortex Respond to Execution and Observation of Actions

Vassilis Papadourakis[1,2] and Vassilis Raos[1,2]

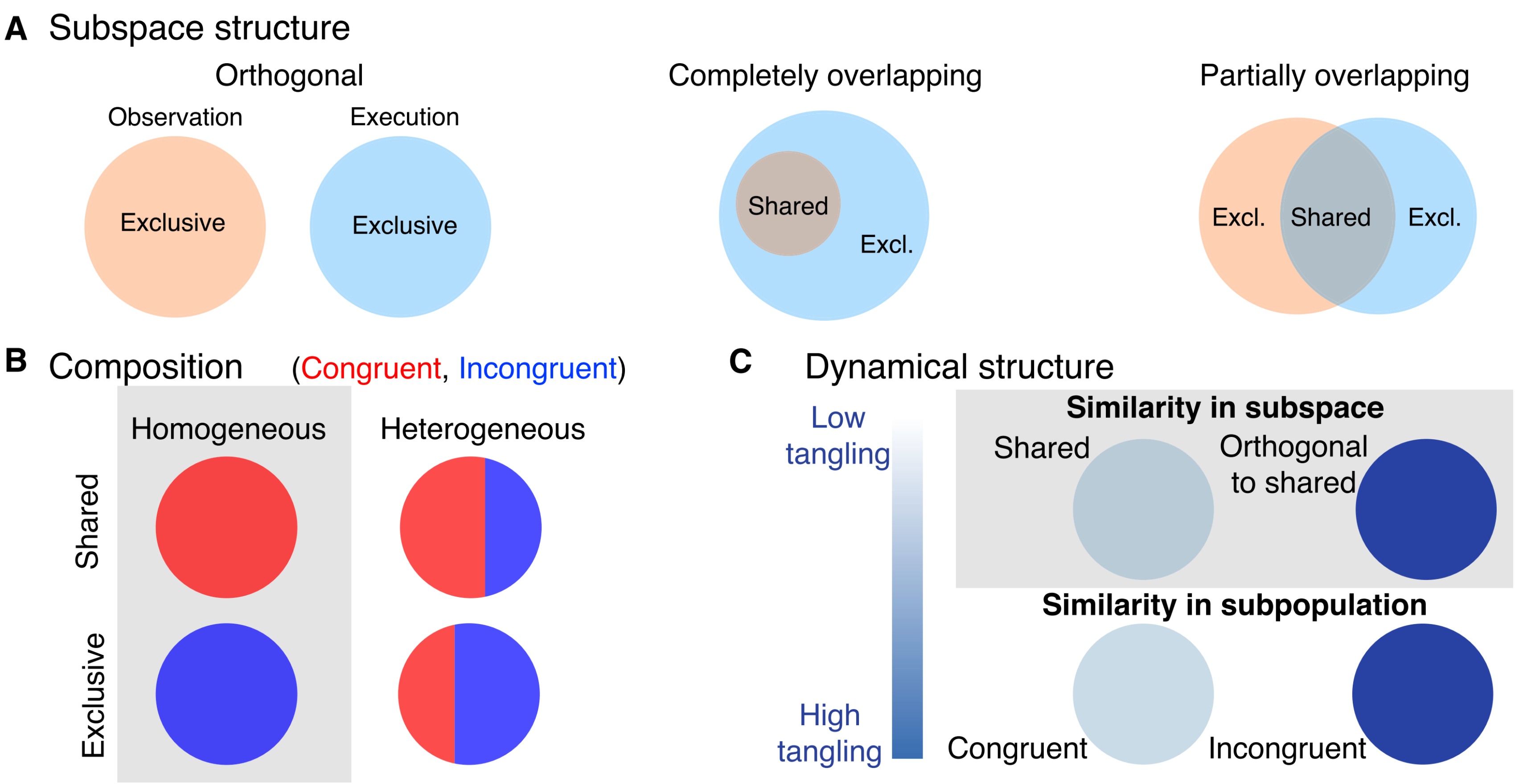
2019 *Cereb. Cortex*

Movement initiation and grasp representation in premotor and primary motor cortex mirror neurons

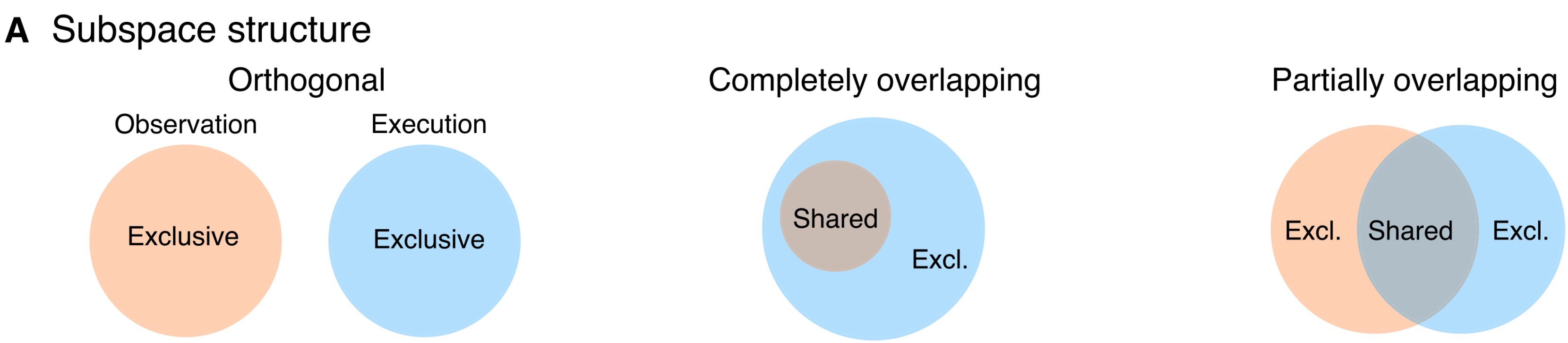
Steven Jack Jerjian[1†], Maneesh Sahani[2], Alexander Kraskov[1*]

2020 *eLife*

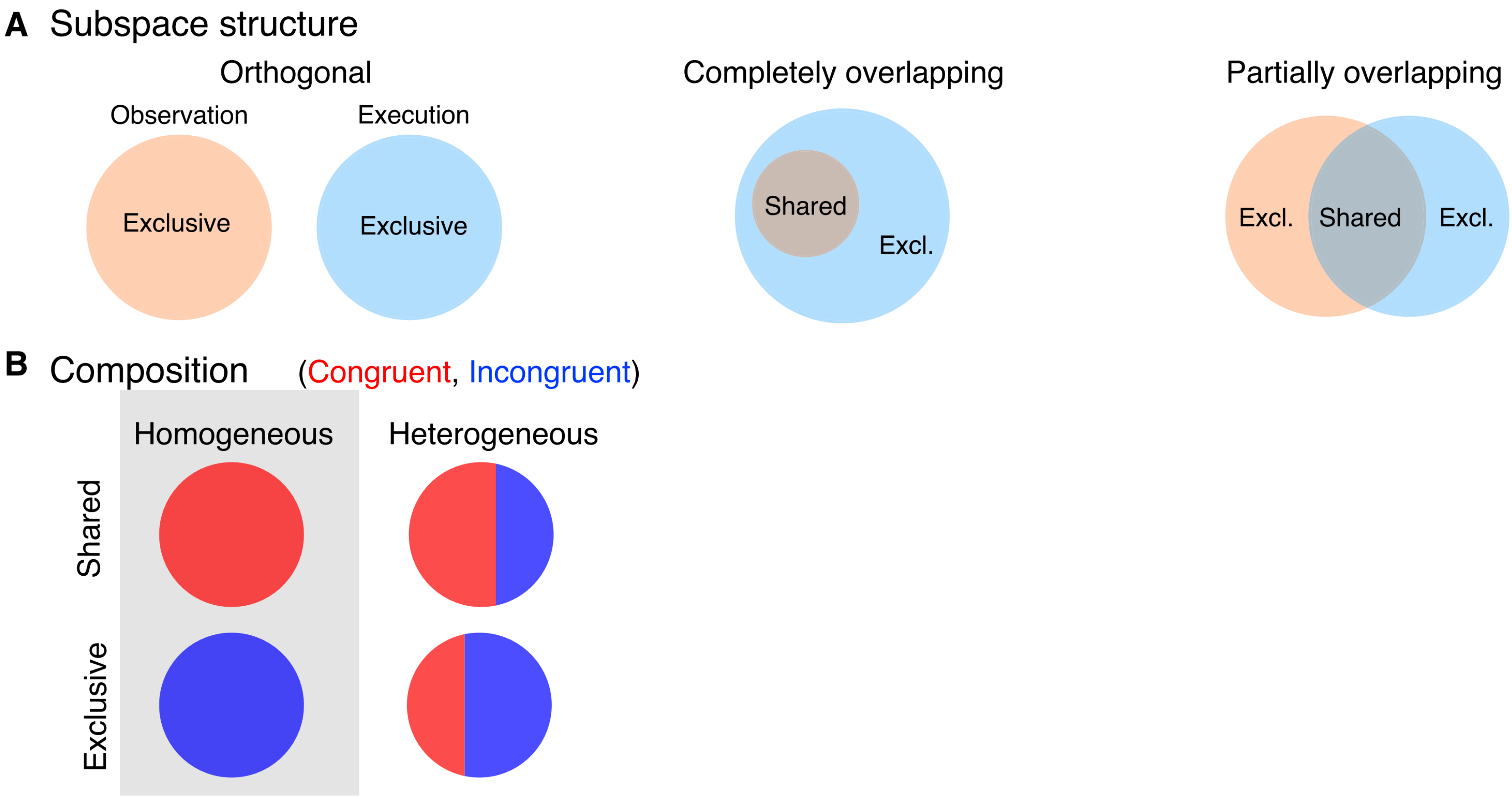
Hypotheses of population structure



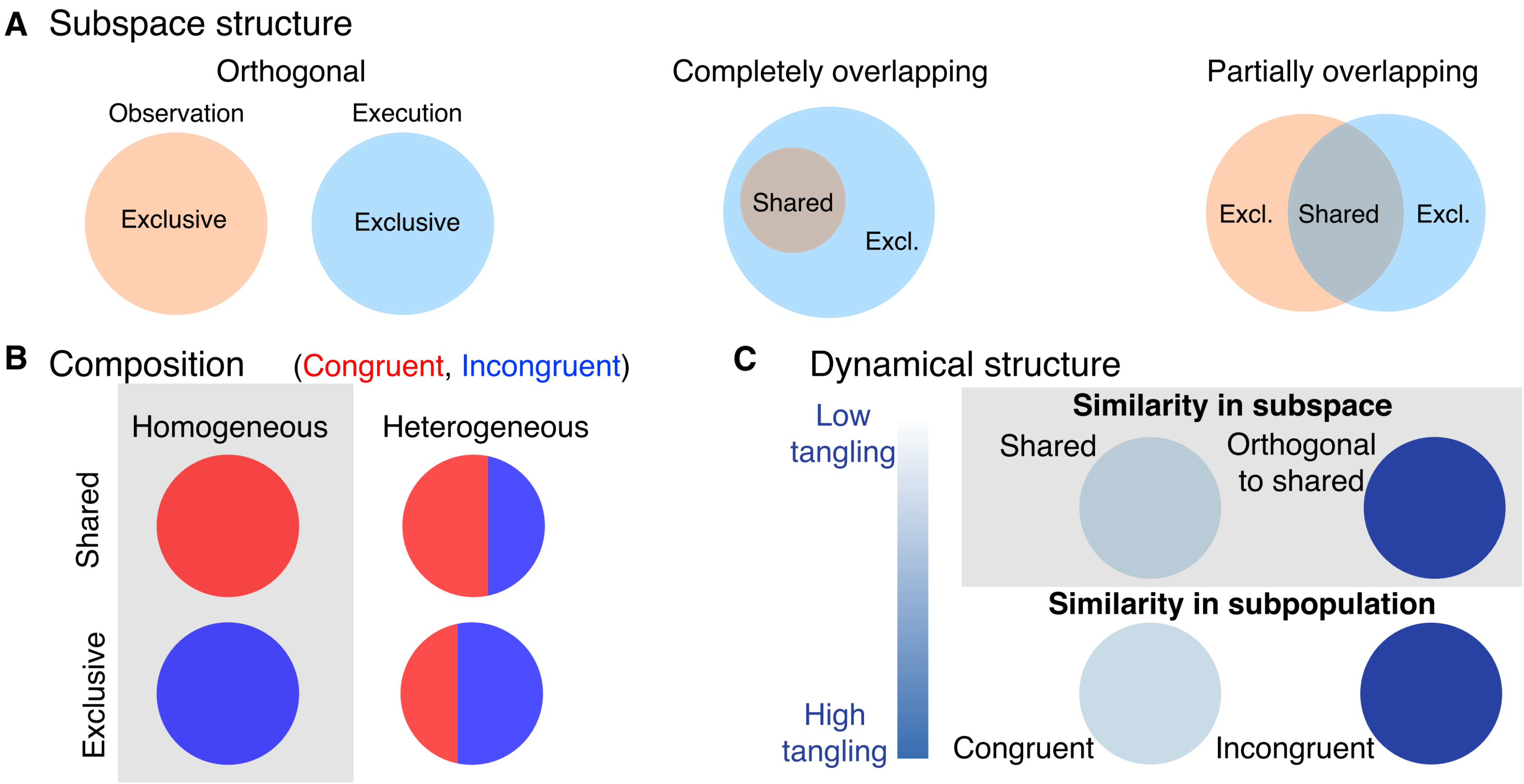
A: Is observation activity orthogonal to, a subspace of, or partially overlapping with execution activity?



B: Given an overlapping subspace, does it comprise a special set of "congruent" neurons? Or a heterogeneous mixture?

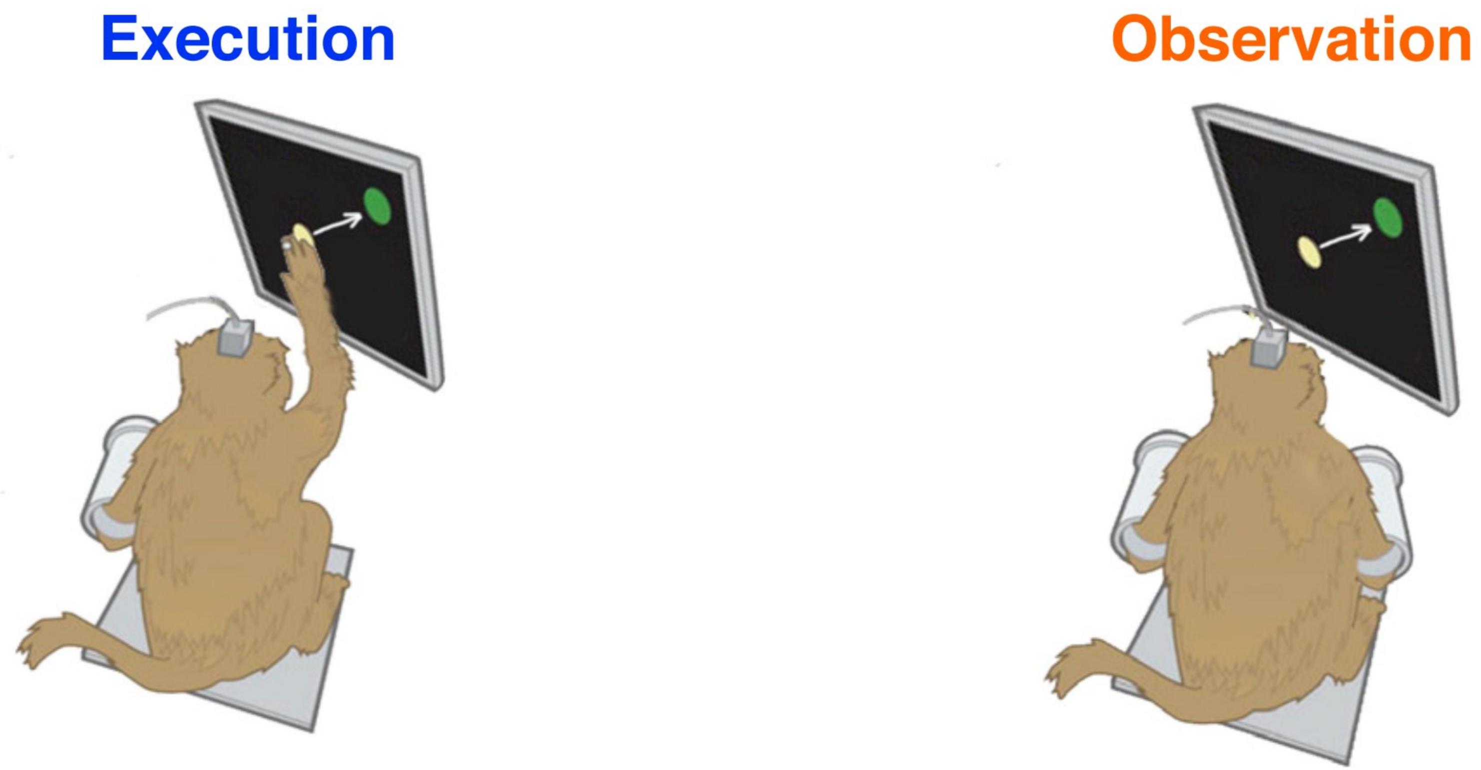


C: Is dynamical structure preferentially preserved among "congruent" neurons? Or in a general shared subspace?

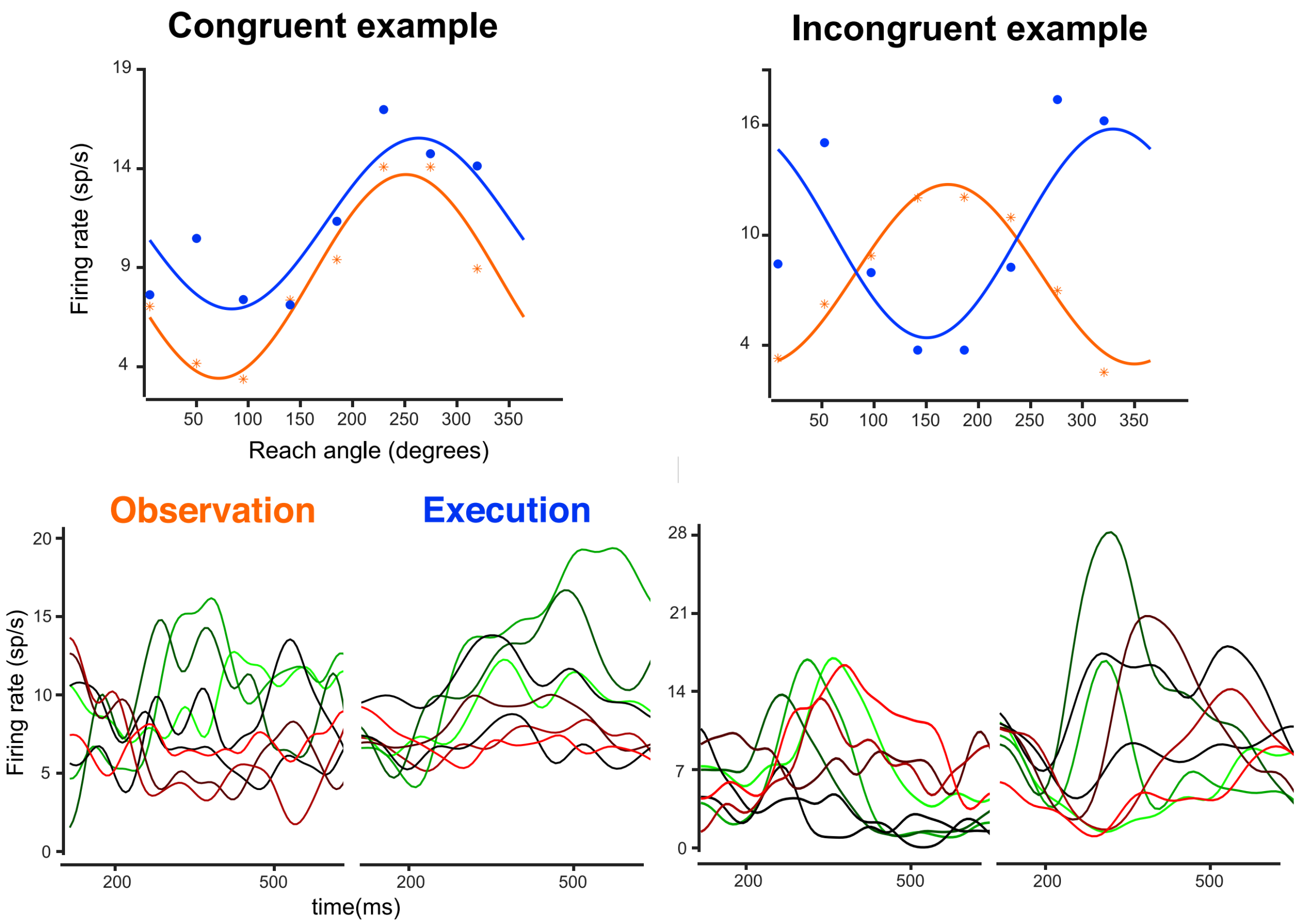


Mitglied der Leibniz-Gemeinschaft

Movement execution and observation experiments



Some neurons *could* be clustered into congruent and incongruent varieties...

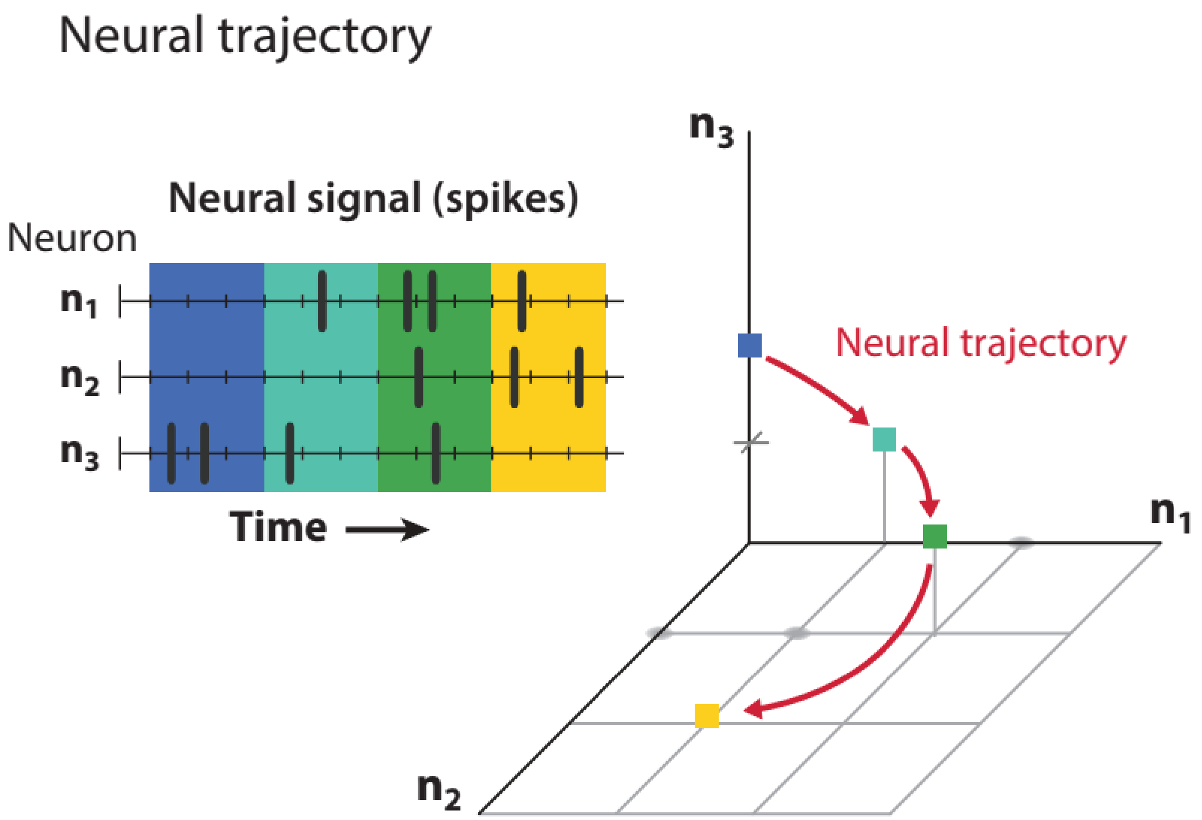


Mitglied der Leibniz-Gemeinschaft

…but there's little evidence for distinct functional classes

sition. To test this, we calculated each neuron's preference index for neuron i as $B(i) = B_{obs}(i)/\overline{B}_{obs} - B_{ex}(i)/\overline{B}_{ex}$, where B is tuning strength and $\overline{B}$ is average tuning strength across all neurons (Elsayed et al., 2016). We then performed Hartigan's dip test on these preference indices to test if there was a bimodal distribution of observation versus execution-preferring neurons. We did not find that these neuron preference indices were bimodal ($p = 0.974$ [0.978] in Monkey J [L]). Together, these results support the concept that a heterogeneous population is responsible for shared structure between observed and executed movements, rather than a subpopulation of neurons.

Mixed, messy population justifies state space analysis!



Vyas et al. 2020 *Annu. Rev. Neurosci.*

Normalized variance captured in a subspace

$$A = \frac{\mathrm{Tr}(\mathbf{Q}^T \mathbf{C} \mathbf{Q})}{\sum_{i=1}^{d} \lambda_i},$$

Mitglied der Leibniz-Gemeinschaft

Normalized variance captured in a subspace

Q: 10-D orthonormal basis

Can be:

- Top 10 Execution PCs
- Top 10 Observation PCs
- Randomly-generated from a uniform spherical distribution

$$A = \frac{\mathrm{Tr}(\mathbf{Q}^T \mathbf{C} \mathbf{Q})}{\sum_{i=1}^{d} \lambda_i},$$

Normalized variance captured in a subspace

$$A = \frac{\mathrm{Tr}(\mathbf{Q}^T\mathbf{C}\mathbf{Q})}{\sum_{i=1}^{d}\lambda_i},$$

C: Firing rate covariance matrix

Can be:

- Observation covariance
- Execution covariance

Normalized variance captured in a subspace

$$A = \frac{\mathrm{Tr}(\mathbf{Q}^T \mathbf{C} \mathbf{Q})}{\sum_{i=1}^{d} \lambda_i},$$

λ: Eigenvalue of **C**

(d = 10 here)

Normalized variance captured in a subspace

Q: 10-D orthonormal basis

Can be:

- Top 10 Execution PCs
- Top 10 Observation PCs
- Randomly-generated from a uniform spherical distribution

C: Firing rate covariance matrix

Can be:

- Observation covariance
- Execution covariance

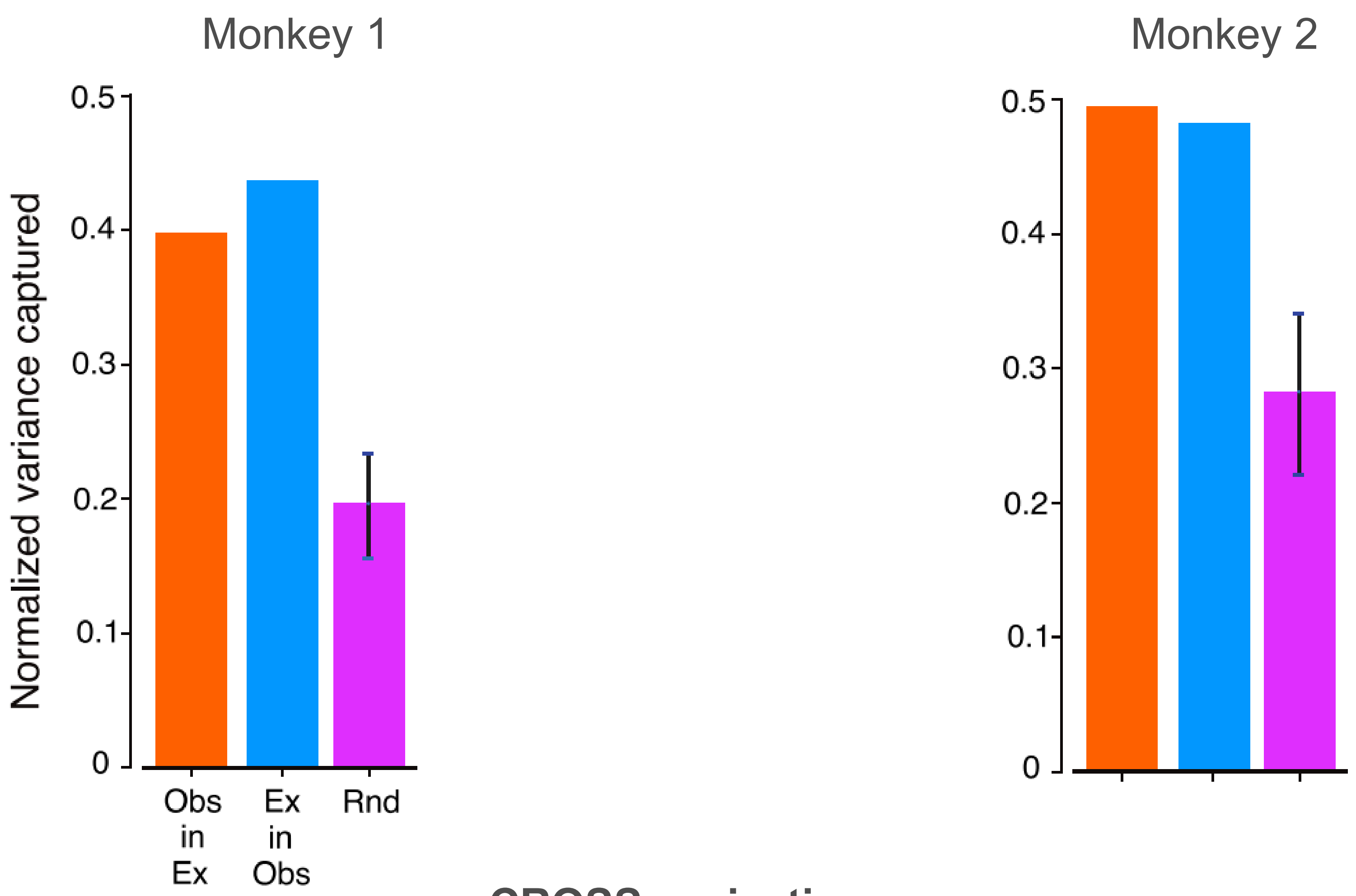
$$A = \frac{\mathrm{Tr}(\mathbf{Q}^T \mathbf{C} \mathbf{Q})}{\sum_{i=1}^{d} \lambda_i},$$

λ: Eigenvalue of **C**

(d = 10 here)

Mitglied der Leibniz-Gemeinschaft

Observation and execution subspaces are not orthogonal



CROSS-projections

Are we *sure* about that? Explicitly seeking orthogonal subspaces

$$\underset{\mathbf{Q}\in M_{d_{\text{Obs}}+d_{\text{Ex}}}\left(\mathbb{R}^N\right)}{\text{maximize}} \quad \frac{\text{Tr}\left(\mathbf{Q}_{\text{Orth-Obs}}^T \mathbf{C}_{\text{Obs}} \mathbf{Q}_{\text{Orth-Obs}}\right)}{\sum_{i=1}^{d_{\text{Obs}}} \lambda_i^{\text{Obs}}} + \frac{\text{Tr}\left(\mathbf{Q}_{\text{Orth-Ex}}^T \mathbf{C}_{\text{Ex}} \mathbf{Q}_{\text{Orth-Ex}}\right)}{\sum_{i=1}^{d_{\text{Ex}}} \lambda_i^{\text{Ex}}}$$

Subject to: $\quad \mathbf{Q}_{\text{Orth-Obs}}^T \mathbf{Q}_{\text{Orth-Ex}} = 0.$

Are we *sure* about that? Explicitly seeking orthogonal subspaces

$$\underset{\mathbf{Q}\in M_{d_{\text{Obs}}+d_{\text{Ex}}}(\mathbb{R}^N)}{\text{maximize}} \quad \frac{\text{Tr}\left(\mathbf{Q}_{\text{Orth-Obs}}^T \mathbf{C}_{\text{Obs}} \mathbf{Q}_{\text{Orth-Obs}}\right)}{\sum_{i=1}^{d_{\text{Obs}}} \lambda_i^{\text{Obs}}} + \frac{\text{Tr}\left(\mathbf{Q}_{\text{Orth-Ex}}^T \mathbf{C}_{\text{Ex}} \mathbf{Q}_{\text{Orth-Ex}}\right)}{\sum_{i=1}^{d_{\text{Ex}}} \lambda_i^{\text{Ex}}}$$

Subject to: $\mathbf{Q}_{\text{Orth-Obs}}^T \mathbf{Q}_{\text{Orth-Ex}} = 0.$

$$A = \frac{\text{Tr}(\mathbf{Q}^T \mathbf{C} \mathbf{Q})}{\sum_{i=1}^{d} \lambda_i},$$

Are we *sure* about that? Explicitly seeking orthogonal subspaces

$$\underset{\mathbf{Q}\in M_{d_{\mathrm{Obs}}+d_{\mathrm{Ex}}}(\mathbb{R}^N)}{\text{maximize}} \quad \frac{\mathrm{Tr}\left(\mathbf{Q}_{\mathrm{Orth-Obs}}^T \mathbf{C}_{\mathrm{Obs}} \mathbf{Q}_{\mathrm{Orth-Obs}}\right)}{\sum_{i=1}^{d_{\mathrm{Obs}}} \lambda_i^{\mathrm{Obs}}} + \frac{\mathrm{Tr}\left(\mathbf{Q}_{\mathrm{Orth-Ex}}^T \mathbf{C}_{\mathrm{Ex}} \mathbf{Q}_{\mathrm{Orth-Ex}}\right)}{\sum_{i=1}^{d_{\mathrm{Ex}}} \lambda_i^{\mathrm{Ex}}}$$

Subject to: $\mathbf{Q}_{\mathrm{Orth-Obs}}^T \mathbf{Q}_{\mathrm{Orth-Ex}} = 0.$

$$A = \frac{\mathrm{Tr}(\mathbf{Q}^T \mathbf{C} \mathbf{Q})}{\sum_{i=1}^{d} \lambda_i},$$

(a) Observation on observation : $\frac{\mathrm{Tr}\left(\mathbf{Q}_{\mathrm{Orth-Obs}}^T \mathbf{C}_{\mathrm{Obs}} \mathbf{Q}_{\mathrm{Orth-Obs}}\right)}{\sum_{i=1}^{d_{\mathrm{Obs}}} \lambda_i^{\mathrm{Obs}}}$

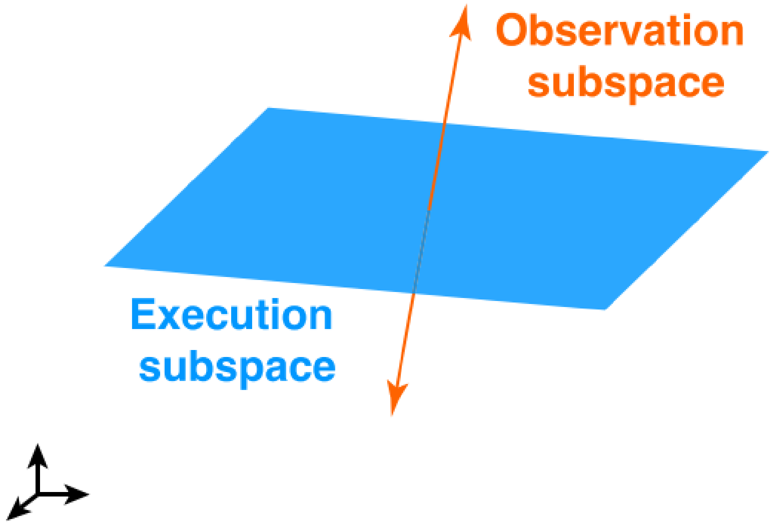
(b) Observation on execution : $\frac{\mathrm{Tr}\left(\mathbf{Q}_{\mathrm{Orth-Ex}}^T \mathbf{C}_{\mathrm{Obs}} \mathbf{Q}_{\mathrm{Orth-Ex}}\right)}{\sum_{i=1}^{d_{\mathrm{Ex}}} \lambda_i^{\mathrm{Obs}}}$

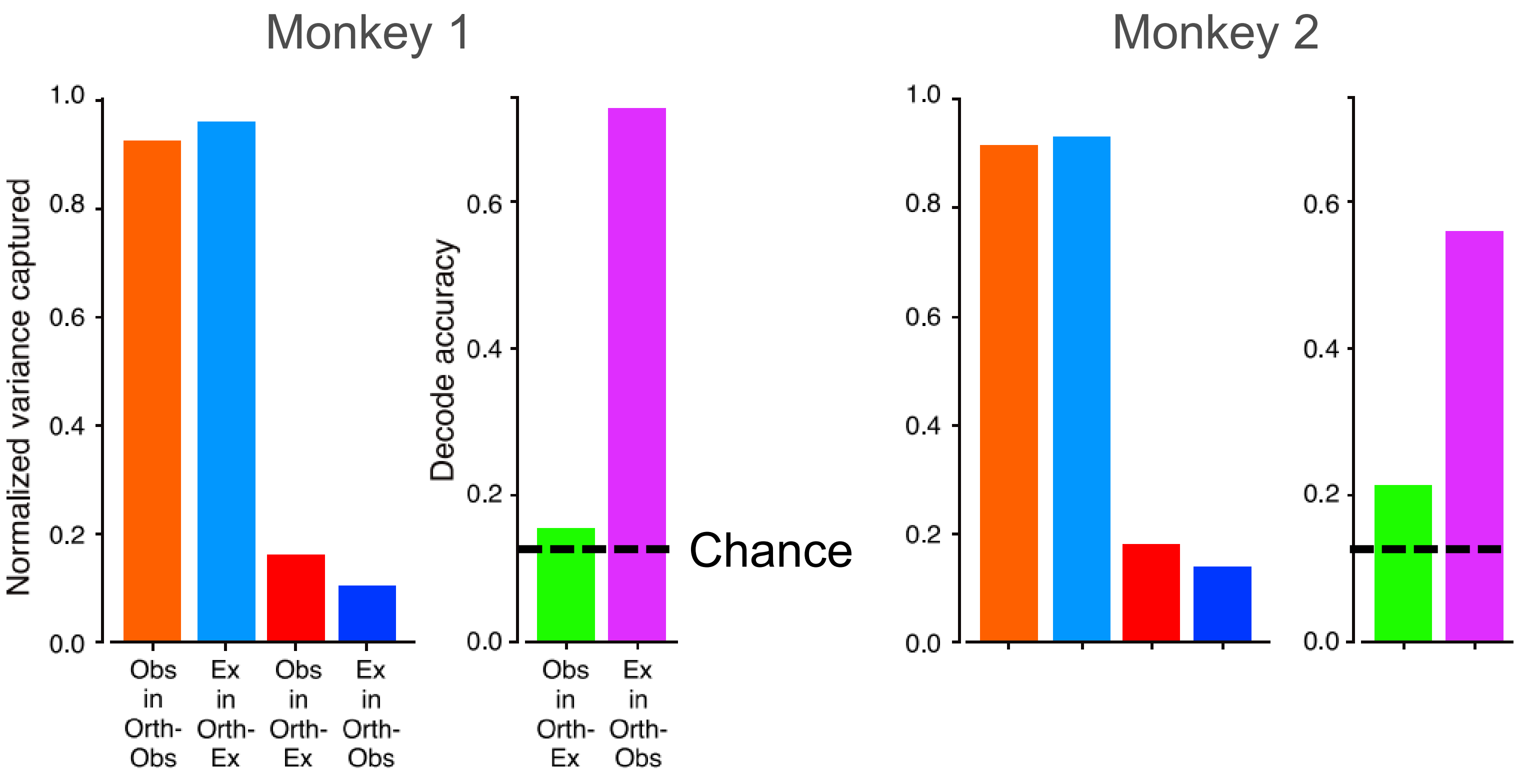
(c) Execution on observation : $\frac{\mathrm{Tr}\left(\mathbf{Q}_{\mathrm{Orth-Obs}}^T \mathbf{C}_{\mathrm{Ex}} \mathbf{Q}_{\mathrm{Orth-Obs}}\right)}{\sum_{i=1}^{d_{\mathrm{Obs}}} \lambda_i^{\mathrm{Ex}}}$

(d) Execution on execution : $\frac{\mathrm{Tr}\left(\mathbf{Q}_{\mathrm{Orth-Ex}}^T \mathbf{C}_{\mathrm{Ex}} \mathbf{Q}_{\mathrm{Orth-Ex}}\right)}{\sum_{i=1}^{d_{\mathrm{Ex}}} \lambda_i^{\mathrm{Ex}}}$

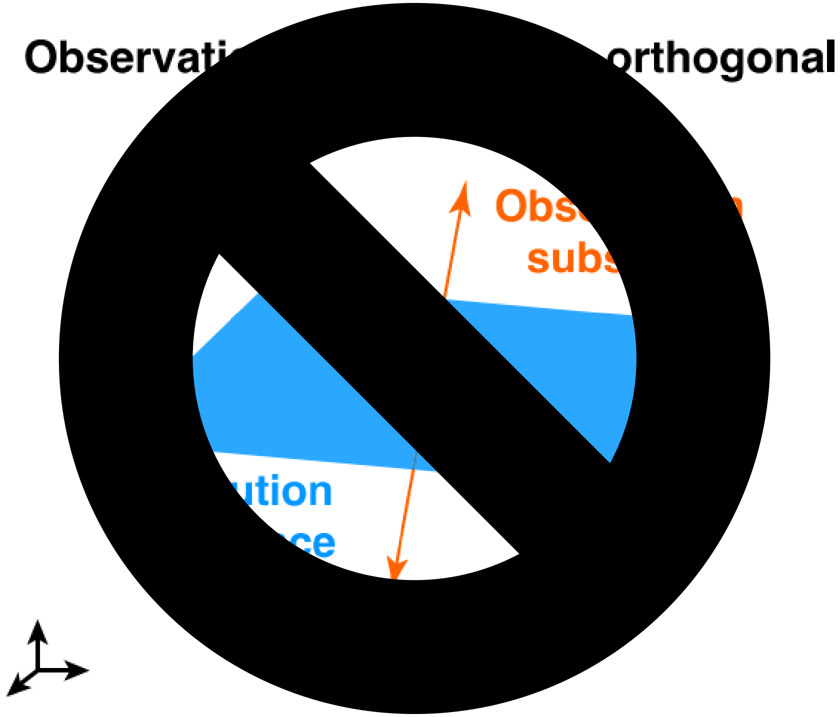
"Orthogonal" subspaces contain nonzero, information-rich neural variance

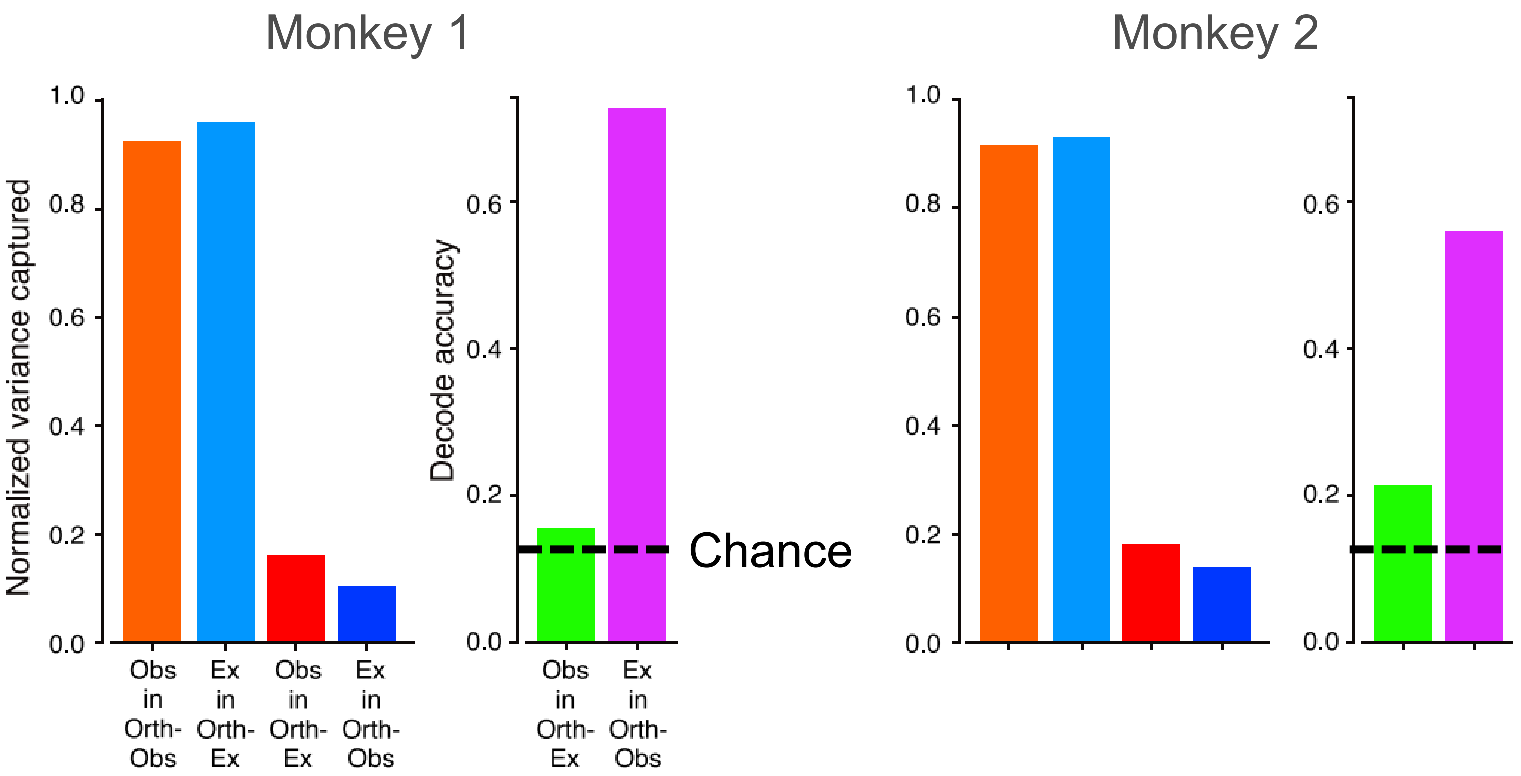
Observation and execution orthogonal





"Orthogonal" subspaces contain nonzero, information-rich neural variance





Mitglied der Leibniz-Gemeinschaft

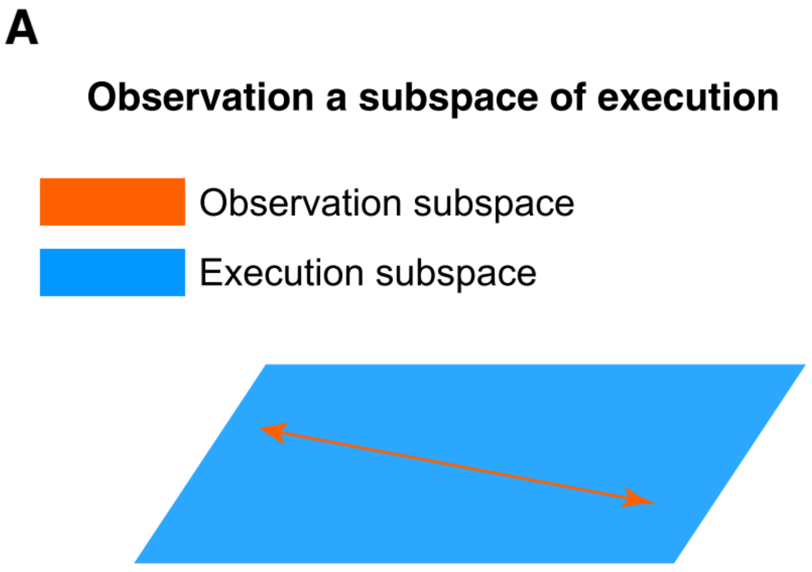
Is observation activity a subset of execution activity? Seeking "exclusive" subspaces

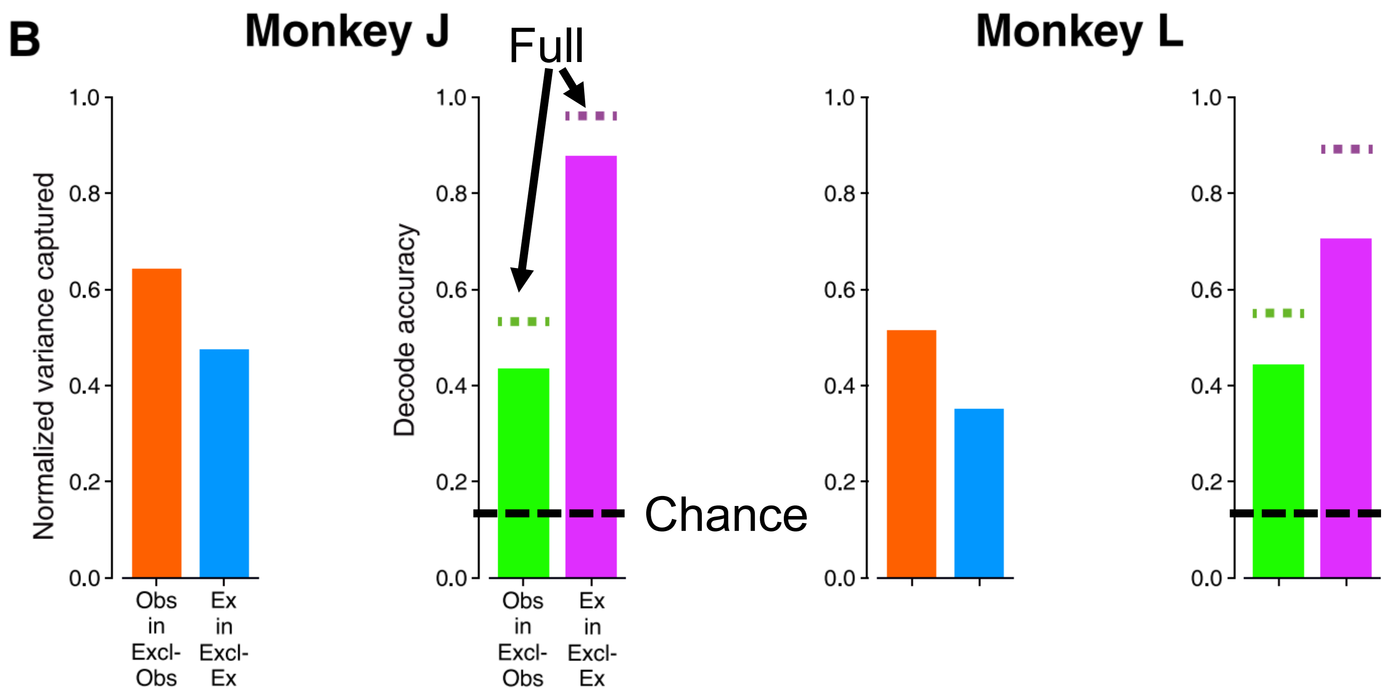
$$\underset{\mathbf{Q}_{\text{Excl-Ex}} \in M_{d_{\text{Ex}}}\left(\mathbb{R}^N\right)}{\text{maximize}} \quad \frac{\operatorname{Tr}\left(\mathbf{Q}_{\text{Excl-Ex}}^T \mathbf{C}_{\text{Ex}} \mathbf{Q}_{\text{Excl-Ex}}\right)}{\sum_{i=1}^{d_{\text{Ex}}} \lambda_i^{\text{Ex}}}$$

$$\text{subject to} \quad \frac{\operatorname{Tr}\left(\mathbf{Q}_{\text{Excl-Ex}}^T \mathbf{C}_{\text{Obs}} \mathbf{Q}_{\text{Excl-Ex}}\right)}{\sum_{i=1}^{d_{\text{Ex}}} \lambda_i^{\text{Obs}}} \leq v$$

(and vice-versa)

Substantial "exclusive" subspaces show that observation activity is not a subset of execution activity



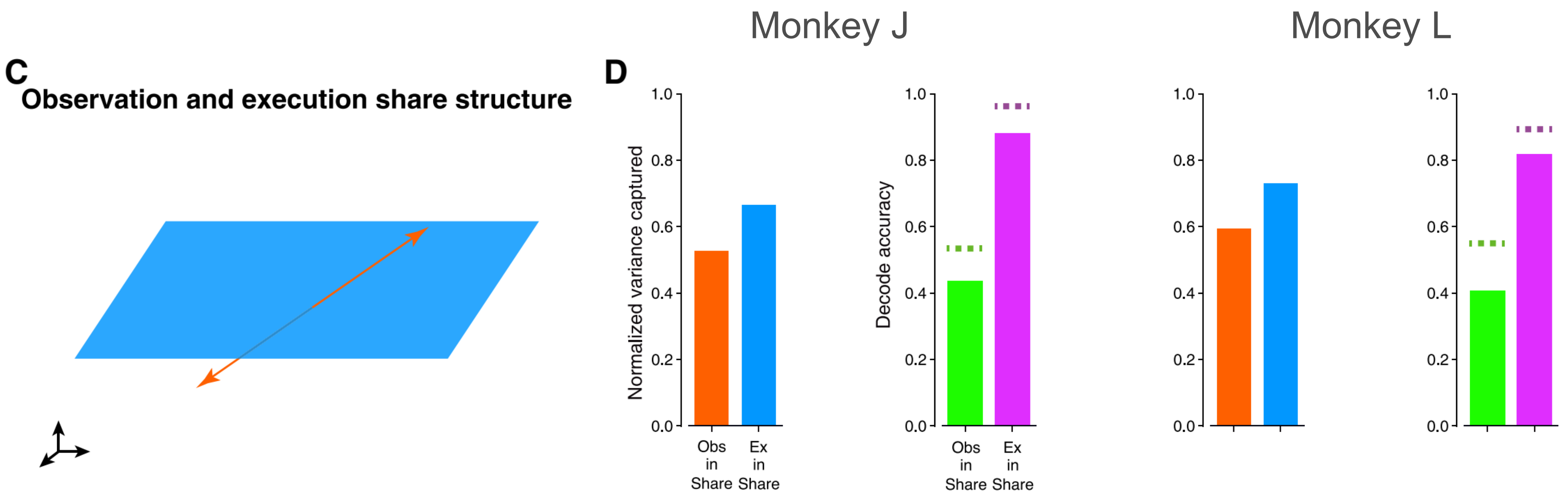


Is the shared subspace meaningful? Seeking the "shared" subspace

$$\underset{\mathbf{Q}_{\text{Shared}} \in M_{d_{\text{Shared}}}\left(\mathbb{R}^N\right)}{\text{maximize}} \quad \frac{\mathrm{Tr}\left(\mathbf{Q}_{\text{Shared}}^T \mathbf{C}_{\text{Ex}} \mathbf{Q}_{\text{Shared}}\right)}{\sum_{i=1}^{d_{\text{Shared}}} \lambda_i^{\text{Ex}}} + \frac{\mathrm{Tr}\left(\mathbf{Q}_{\text{Shared}}^T \mathbf{C}_{\text{Obs}} \mathbf{Q}_{\text{Shared}}\right)}{\sum_{i=1}^{d_{\text{Shared}}} \lambda_i^{\text{Obs}}}$$

$$\text{subject to} \quad \mathbf{Q}_{\text{Shared}} \perp \mathbf{Q}_{\text{Excl-Obs}}, \quad \mathbf{Q}_{\text{Shared}} \perp \mathbf{Q}_{\text{Excl-Ex}}$$

The shared subspace captures substantial, information-rich variance



Mitglied der Leibniz-Gemeinschaft

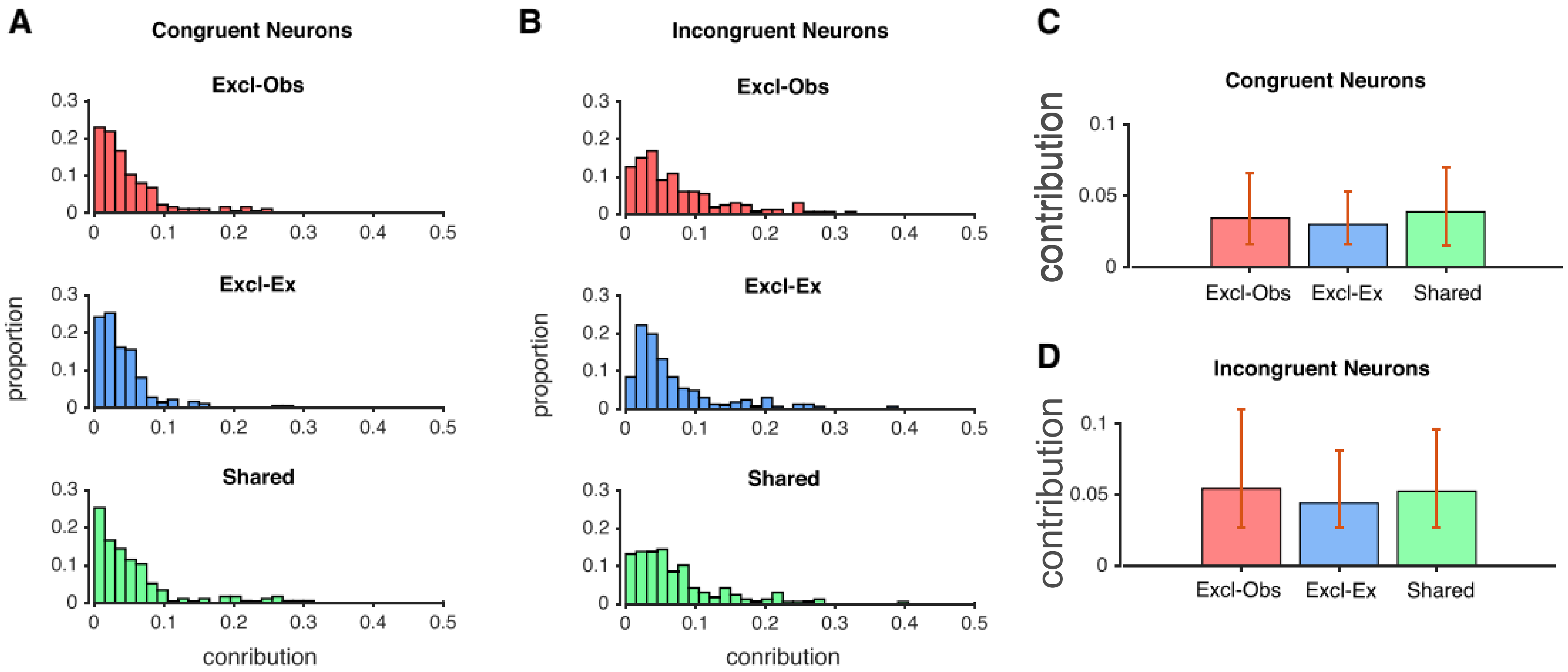
How much does each neuron contribute to each subspace?

$$c_i = \sqrt{\frac{\mathrm{Tr}\left(\mathbf{Q}^T \mathbf{X}_i \mathbf{X}_i^T \mathbf{Q}\right)}{\mathrm{Tr}\left(\mathbf{X}_i \mathbf{X}_i^T\right)}} \cdot FR = \sqrt{\frac{\mathrm{Tr}\left(\mathbf{X}_i \mathbf{X}_i^T \mathbf{Q}\mathbf{Q}^T\right)}{\mathrm{Tr}\left(\mathbf{X}_i \mathbf{X}_i^T\right)}} \cdot FR = \sqrt{\frac{var(x_i) \|\mathbf{w}_i\|^2}{var(x_i)}} \cdot FR = \|\mathbf{w}_i\| \cdot FR$$

where $var(x_i)$ is neuron i's total variance, and $\|\mathbf{w}_i\|^2$ is the squared sum of $\mathbf{Q}$'s i^{th} row.

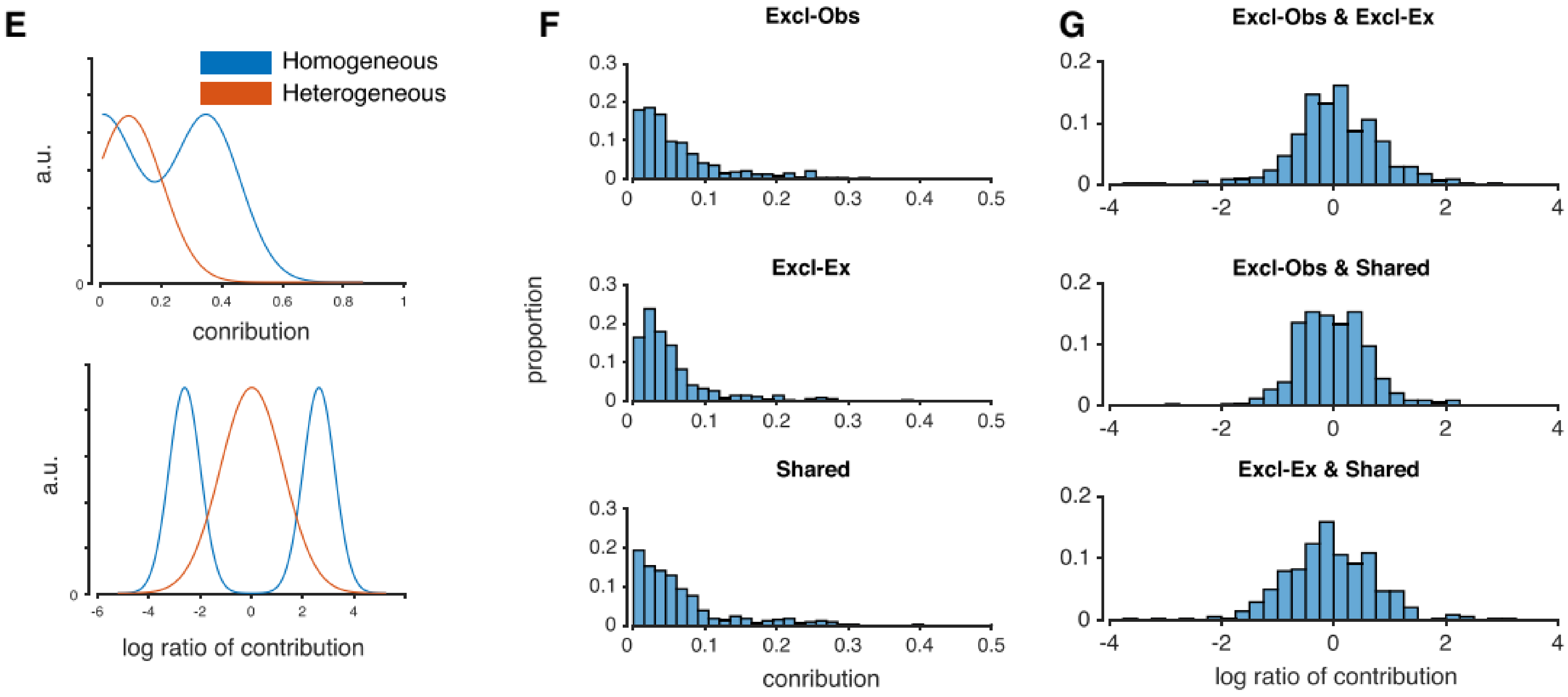
Mitglied der Leibniz-Gemeinschaft

The shared subspace is not merely a collection of congruent neurons



Mitglied der Leibniz-Gemeinschaft

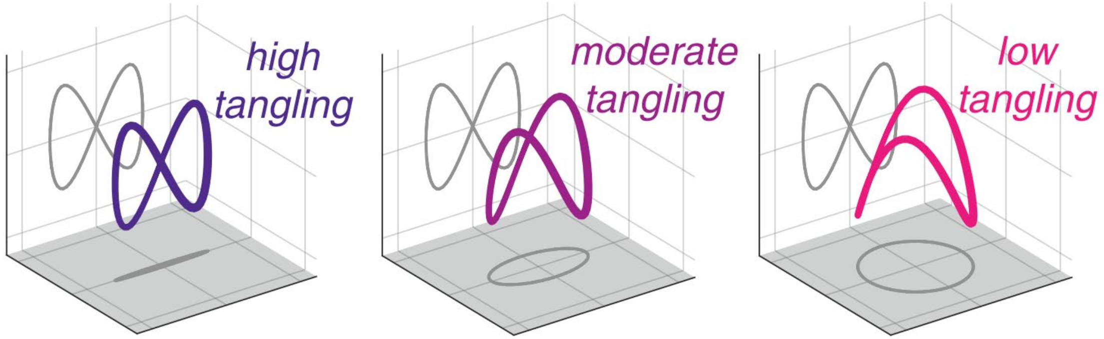
No subspace defines a particular class of neurons – each subspace is a heterogeneous mix

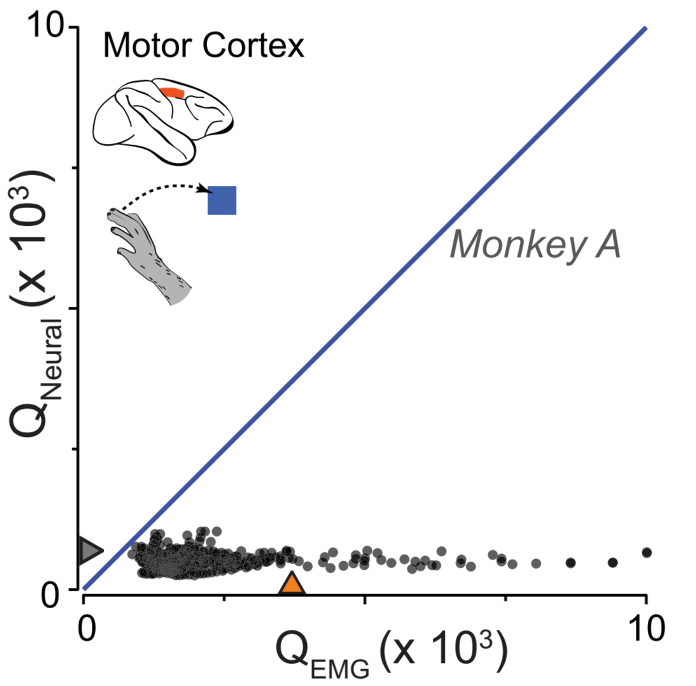


Mitglied der Leibniz-Gemeinschaft

Not just a representation: quantifying dynamical "tangling"

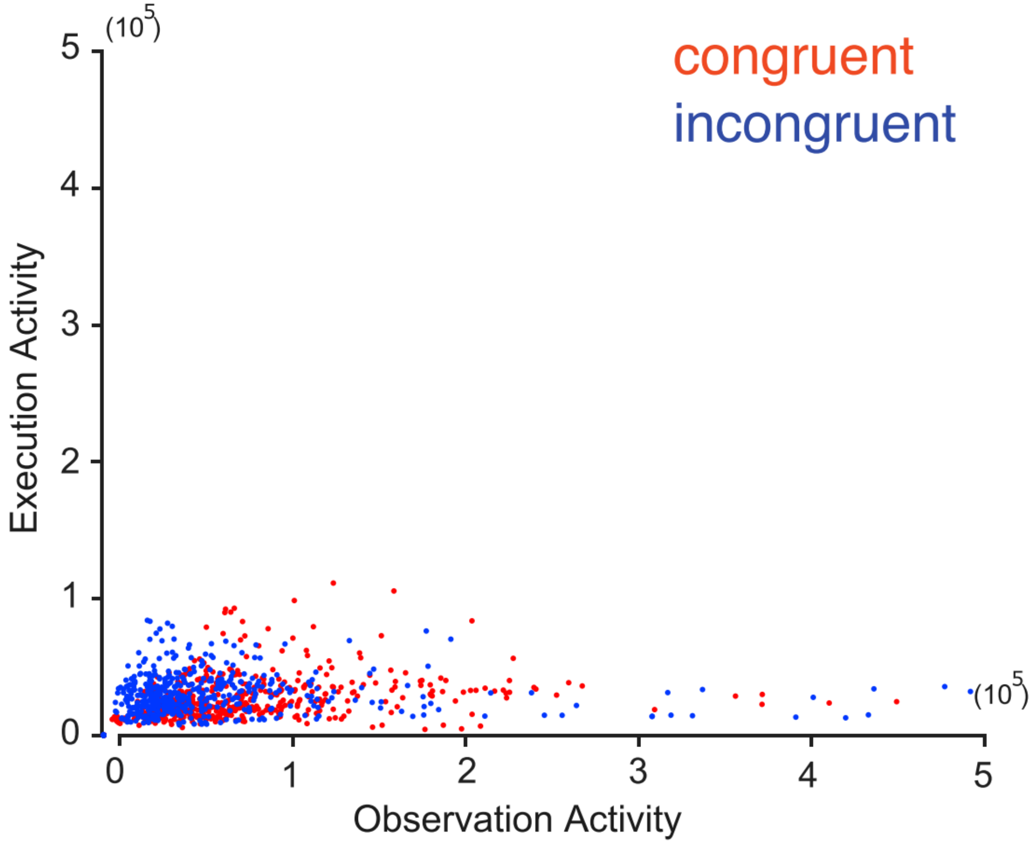
$$Q(t) = \max_{t'} \frac{\|\dot{\mathbf{x}}_t - \dot{\mathbf{x}}_{t'}\|^2}{\|\mathbf{x}_t - \mathbf{x}_{t'}\|^2 + \varepsilon},$$



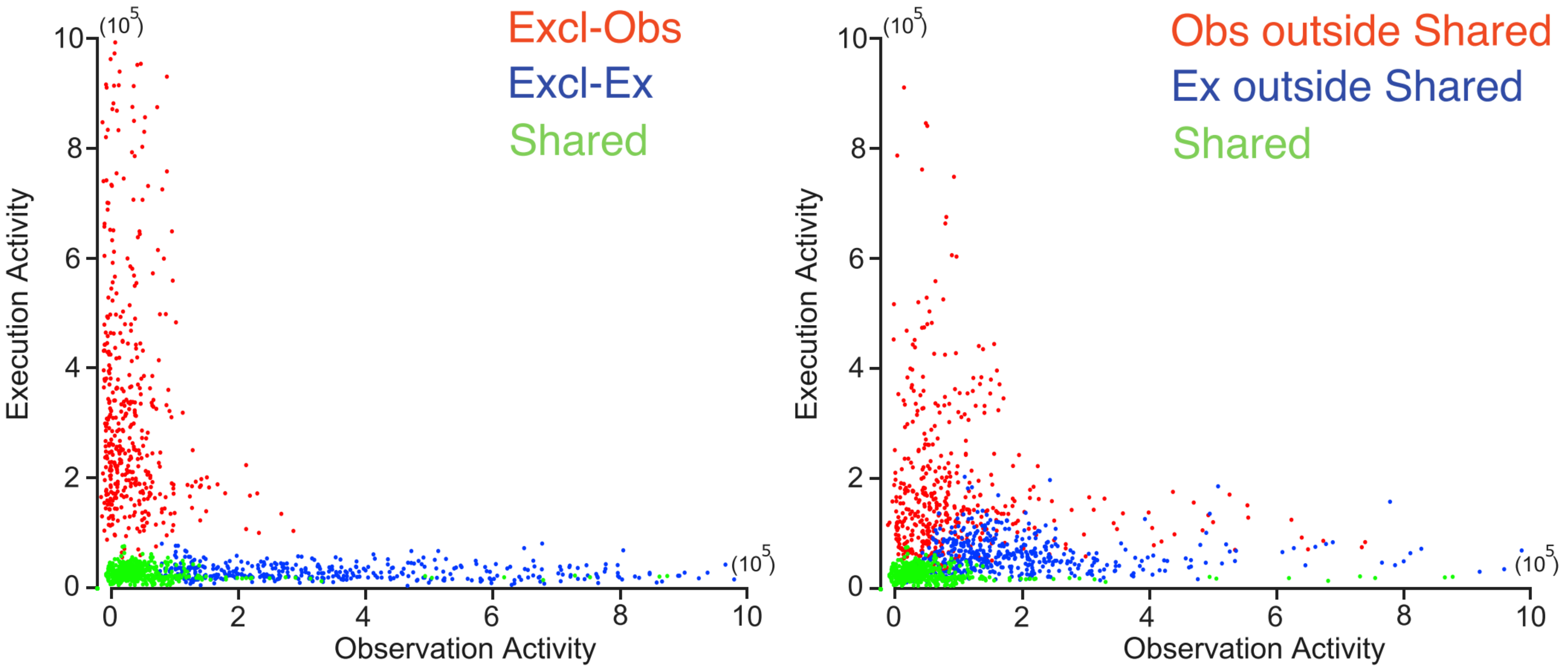


Russo et al. 2018 *Neuron*

Congruent and incongruent neurons do not define dynamically distinct subspaces

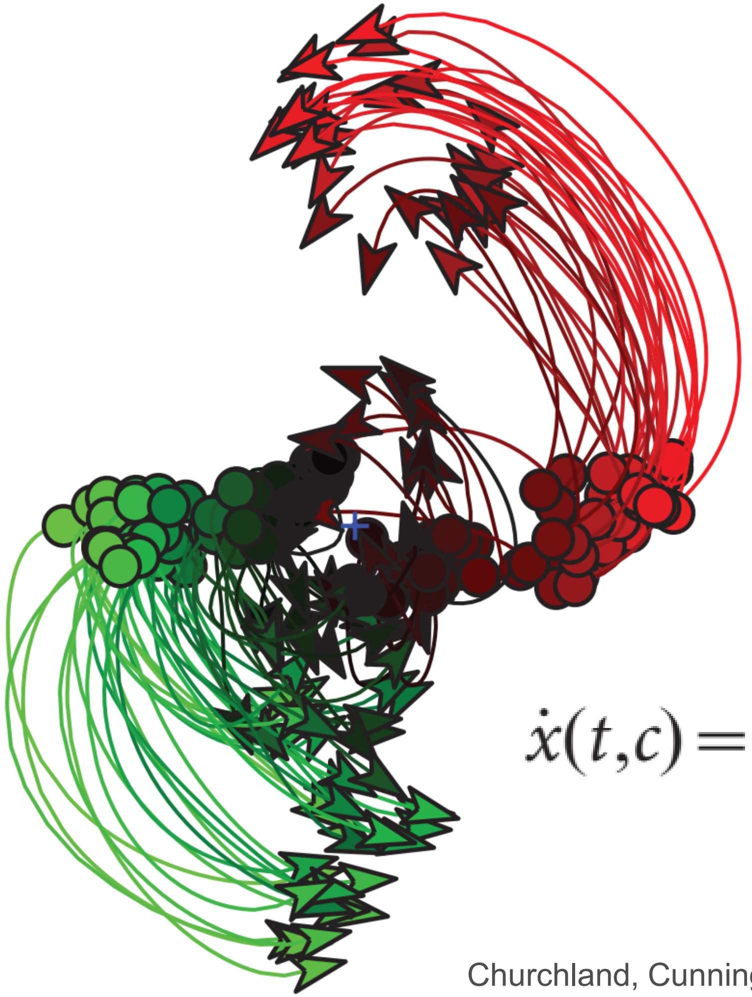


Subspace decomposition reveals where observation activity exhibits low tangling



Mitglied der Leibniz-Gemeinschaft

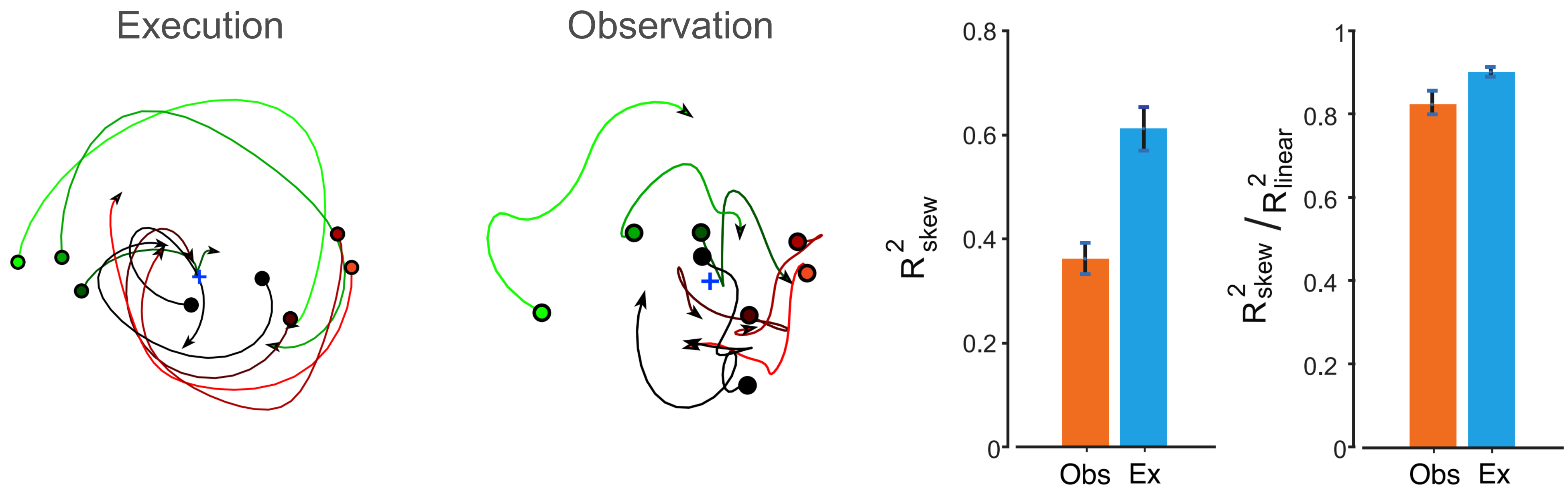
What do these dynamics look like? Returning to jPCA



$$\dot{x}(t,c) = M_{\text{skew}}\, x(t,c)$$

Churchland, Cunningham, et al. 2012 *Nature*

Observation dynamics are weaker, but just as rotational, as execution dynamics



Conclusions

- There is no specific class of congruent mirror neurons
- Neural state spaces during action and observation are neither orthogonal nor totally overlapping
  - A "shared" subspace
  - Two "exclusive" subspaces
- Movement information is contained in both the shared and two exclusive subspaces
- Activity across all subspaces exhibits rotational dynamics, to the extent those subspaces are explored in a given context
- Overall, observation seems noisier than action
  - Less movement information
  - Weaker dynamics

Caveats!

- No reporting of overall variance / firing rate / SNR within each condition
  - Only *differences* in baseline firing rate & modulation depth; both favor execution
  - How close to zero are we during observation?
- These animals have been trained on BCI cursor control
- Once more, PMd and M1 are pooled here
- No evidence that activity during observation is important for any kind of behavior

Mitglied der Leibniz-Gemeinschaft