

Question:

Brain responses are dynamic



Most vision models are purely feedforward.

A handful of modern DNN models now have dynamics

Vary in how:

- Where recurrence originates
- How signals are integrated
- What training objective

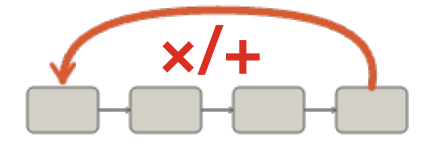
How do category representations change over time across recurrent DNN models?

Methods

Model Catalogue

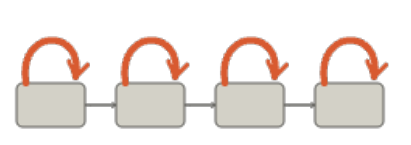
| Model | Recurrence Type | Mechanism | Loss Timing |
|-----------|-------------------|----------------|-----------------|
| LRM3 | Long-range | Mult. | Both passes |
| LRA3 | Long-range | Add. | Both passes |
| BL | Within-layer | Add.(separate) | All steps |
| CORnet-RT | Within-layer | Add.(shared) | Final step only |
| ConvRNN | Within + top-down | Mult.(gated) | Final step only |

LRM/LRA (Konkle & Alvarez, 2023)



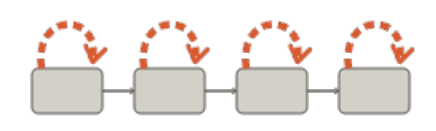
Feedforward Counterpart: AlexNet

BL (Spoerer et al., 2020)

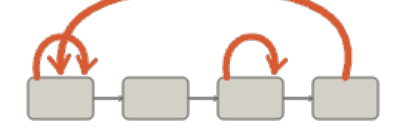


Feedforward Counterpart: B

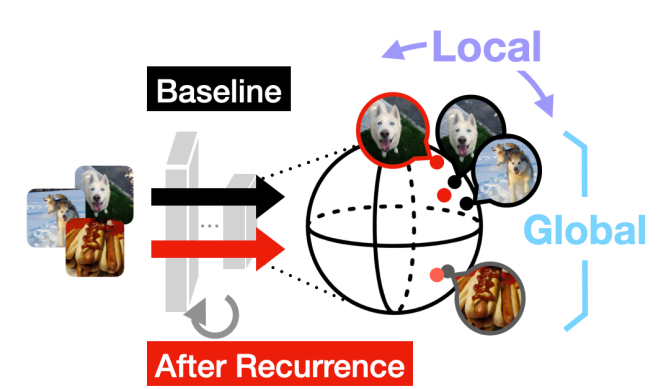
CORnet-RT (Kubilius et al., 2018)



ConvRNN (Nayebi et al., 2018)



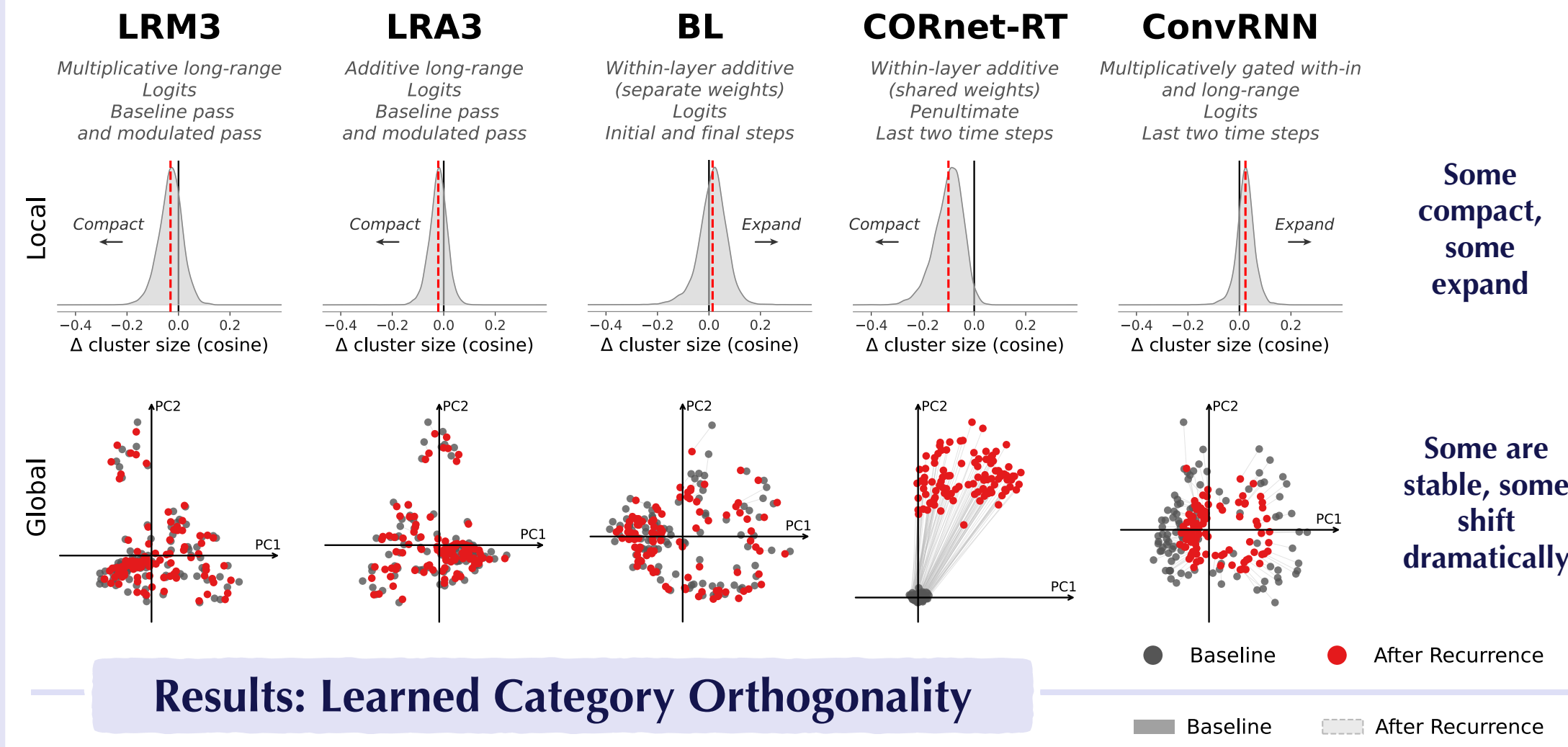
Probe Category Representational Geometry



- ImageNet Val:
 - 100 categories
 - 50 images per category
- Prototype: Category mean
- Local: exemplar-to-prototype cosine distance
- Global: pairwise cosine distance between category prototypes

Results: Distinct Local and Global Dynamics

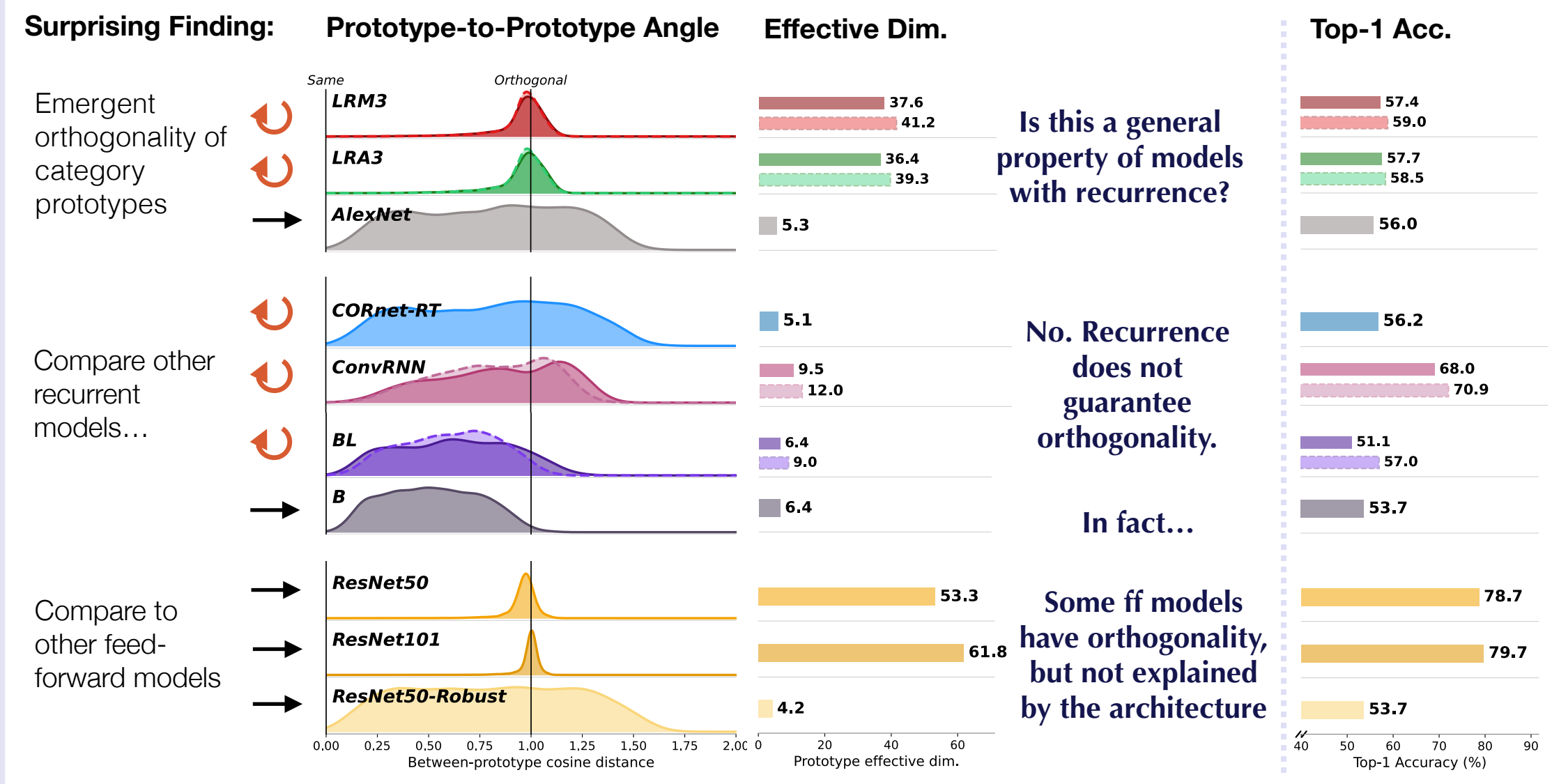
How do local and global representational geometry change across recurrent steps?



Some compact, some expand

Some are stable, some shift dramatically

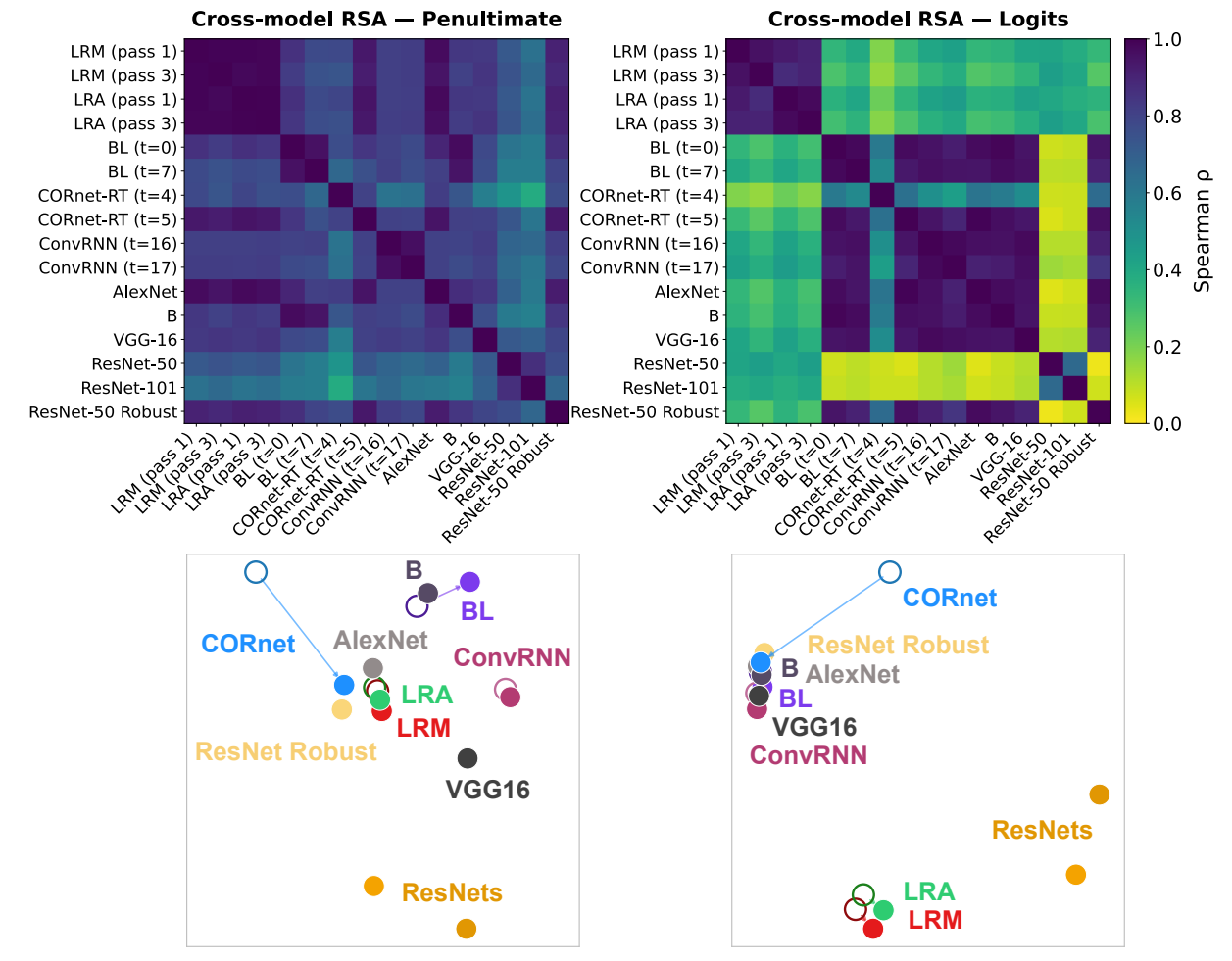
Results: Learned Category Orthogonality



There is no simple single factor that predicts emergent orthogonality.

Cross-Model Representational Similarity

How do different model geometries relate to each other?



- Models tend to converge at the feature stage but diverge at the decision stage.
- Adding recurrence affects different model families differently.

Conclusions

- Different recurrent model families show highly distinct representational dynamics
 - Do these dynamic models have different behavioral capacities for clutter, occlusion, etc?
 - What is happening in the brain?
- When we talk about “recurrent processing” in models and brains, what do we really mean?!?
 - Too simple to say “we have a recurrent model”...
 - Need for more clarity on terminology and characterization of representational consequences...
- Answers offered by NeuroAI inferences may be... complex
 - Not just the main effects of Data, Architecture, Task, and Learning Rules
 - Rather, highly interactive inductive biasing influences