Towards Large-Scale Neural Representation Learning for Calcium Population Dynamics

Xinhong Xu

2026.7

Xu, X.*, Zhang, Y.*, Qian, Q., & Zhang, Y. (2026). CalM: A Self-Supervised Foundation Model for Population Dynamics in Calcium Imaging Data. *ICML26*.

Contents

* Some concept and pipeline figures are generated by AI to provide better illustration.

Contents

Background and Significance

- Large-scale, heterogeneous calcium imaging makes it possible to ask whether reusable neural representations can support multitask prediction, transfer, and biological interpretation.

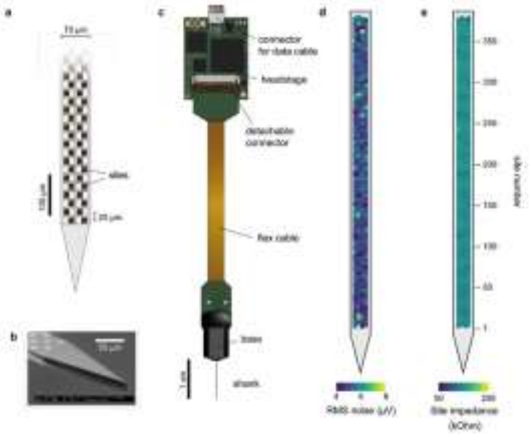
Related Work

Methods and Results

Conclusion and Future Work

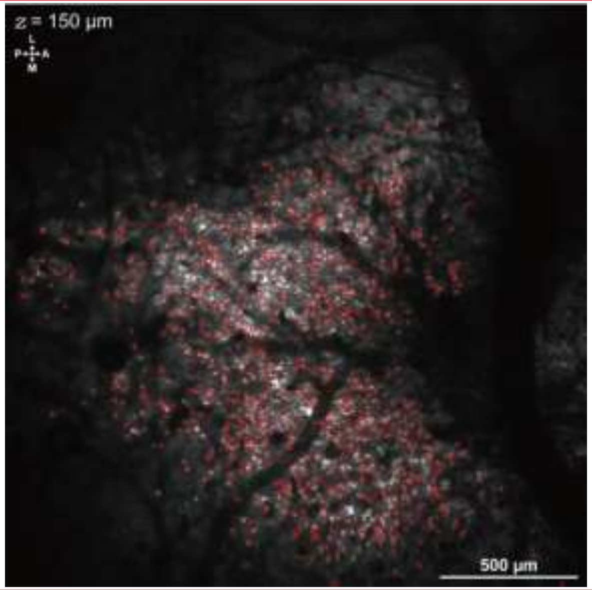
Background

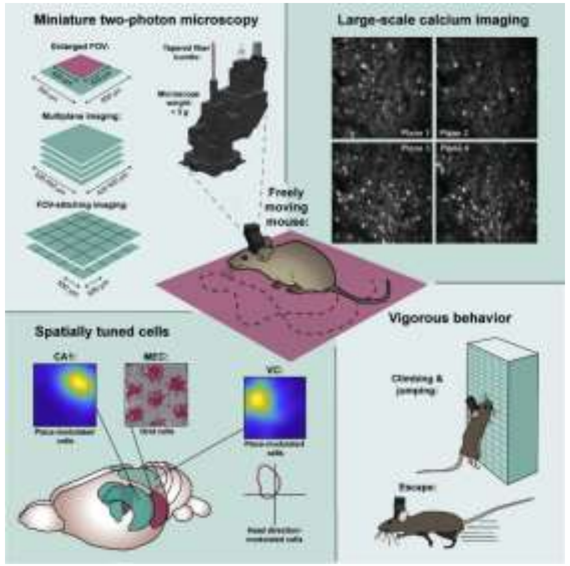
Large-scale neural observation creates a new data regime:

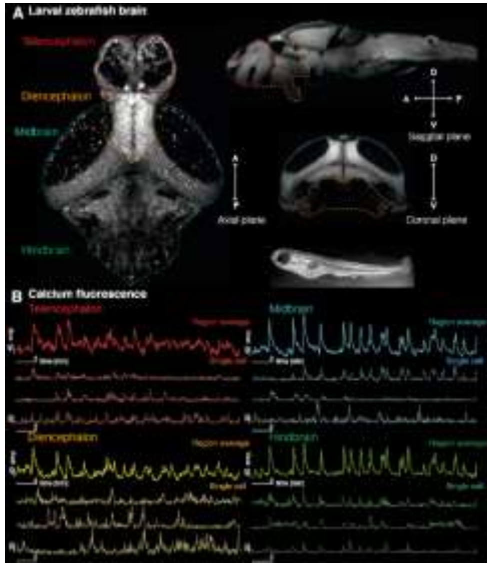


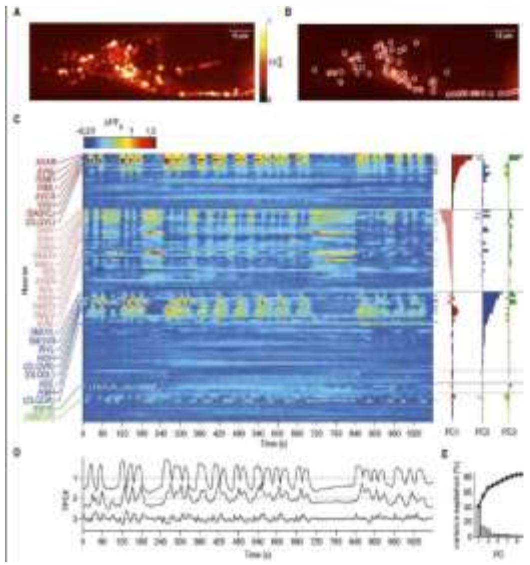
Record tens of thousands of neurons once at a time

Data heterogeneity for animals, experimental paradigms and species



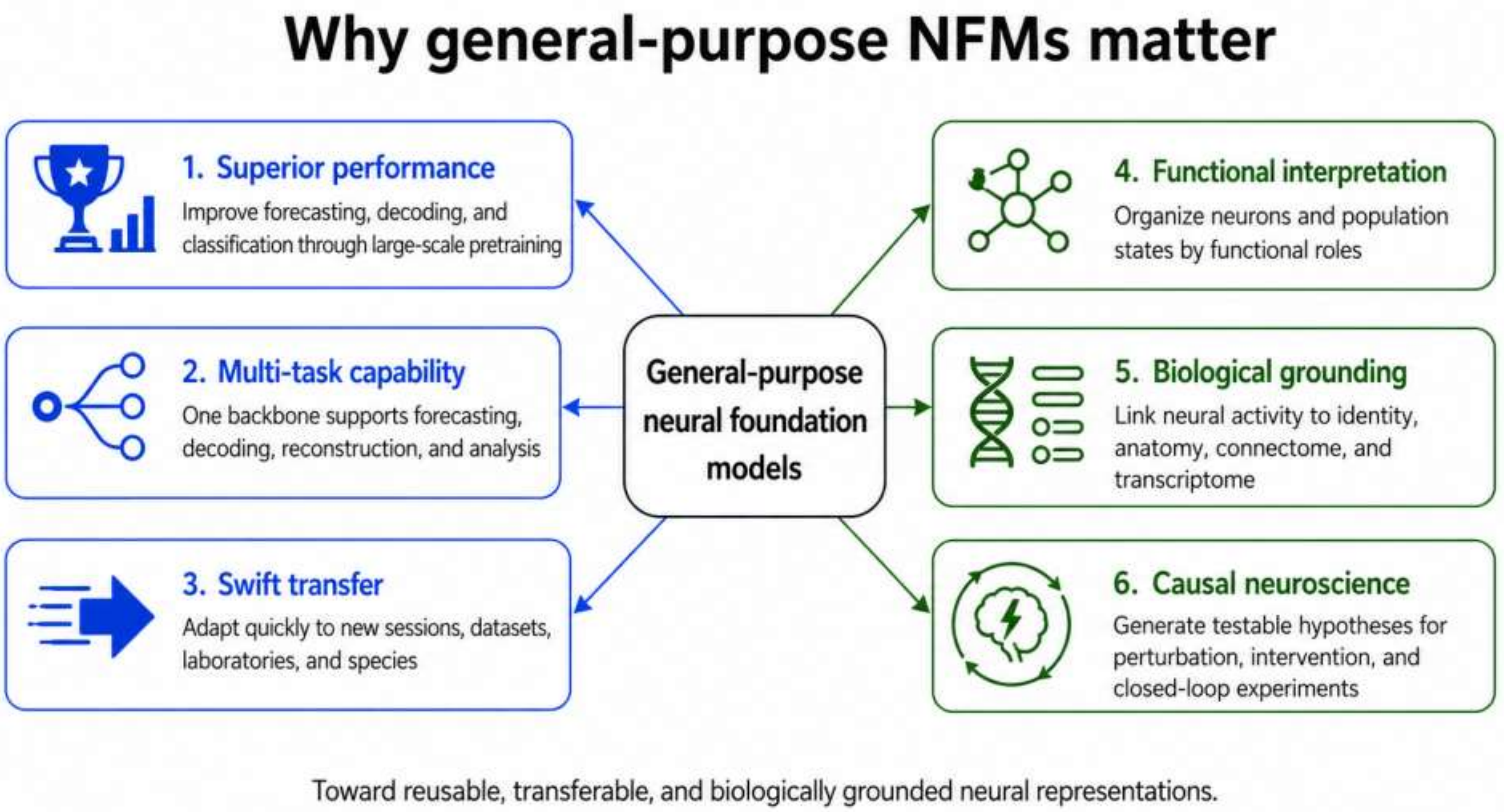






Significance

From prediction tools to functional coordinates:



Contents

Background and Significance

Related Work

- Spike/electrophysiology models provide early templates for foundation-style neural modeling, while calcium-trace models directly motivate multitask modeling, transfer, and biological interpretation.

Methods and results

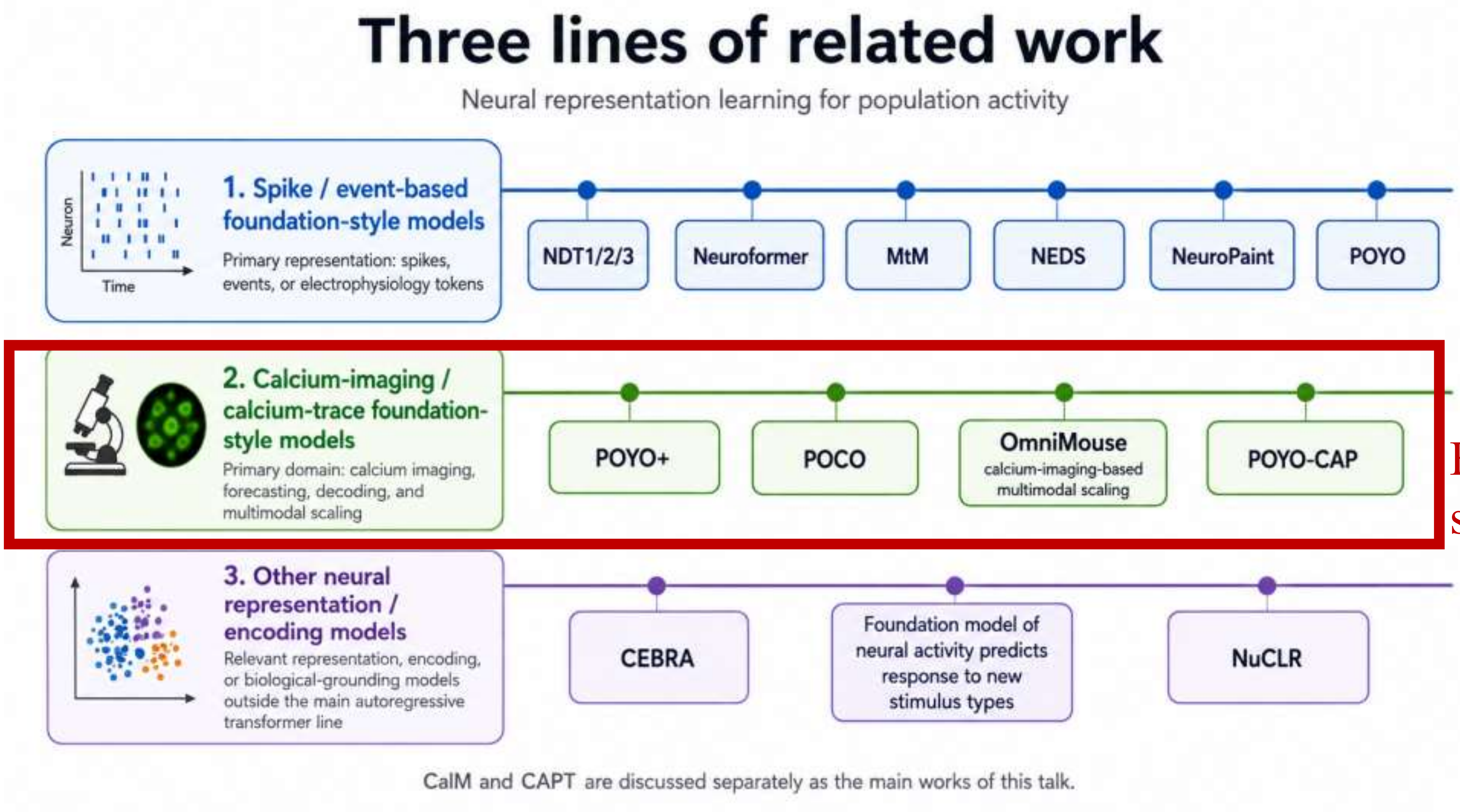
Conclusion and Future work

Liam Paninski

Eva Dyer

Related Work

Overall lines of Transformer for neural data:



Discuss methods on spike/neuropixels and calcium data here.

For computational and systems neuroscience.

Related Work and Roadmap

General technical lines:

- POYO series are widely acknowledged.
- Masked modeling is predominated.
- Adaptive modules (session-specific projection, session / neuron embeddings) are used for cross-session, cross-animal data.

What has been finished:

- Modeling on spike data ==> Masked modeling (MtM, NEDS), decoding (POYO),
- Modeling on calcium trace ==> behavior decoding (POYO+), forecasting (POCO), scaling law (OmniMouse)

What we want to construct for general-purpose NFM:

- A model with standard pre-training, fine-tuning paradigm for diverse functions on calcium trace datasets.
- A model capable of transferring to different sessions, animals, datasets, and even species.
- A model enabling functional interpretation or multimodal biological analysis.

Contents

Background and Significance

Related Work

Methods and results:

- **CalM**: A Self-Supervised Foundation Model for Population Dynamics in Calcium Imaging Data

Conclusion and Future Work

Zhang

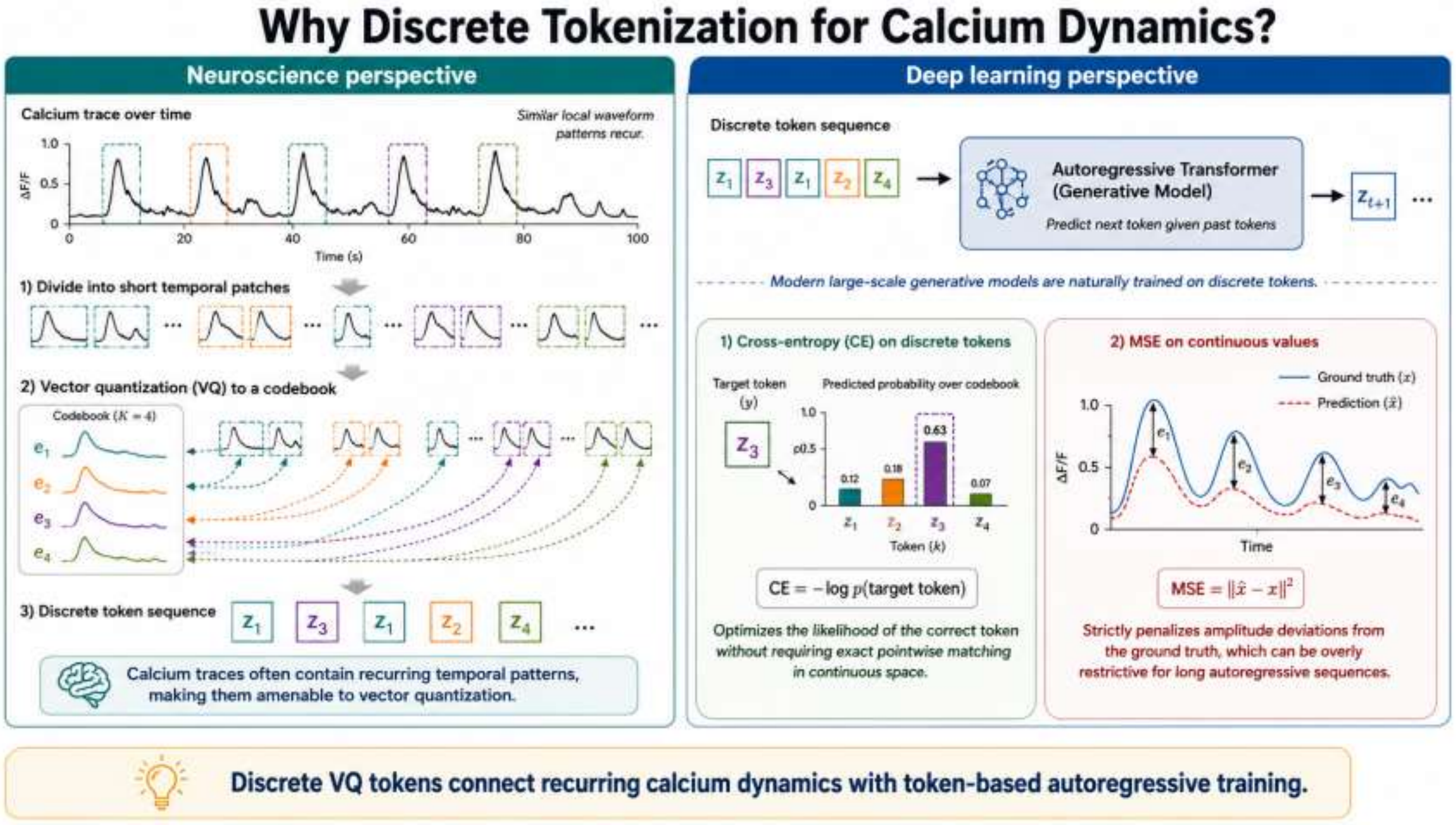
Qian

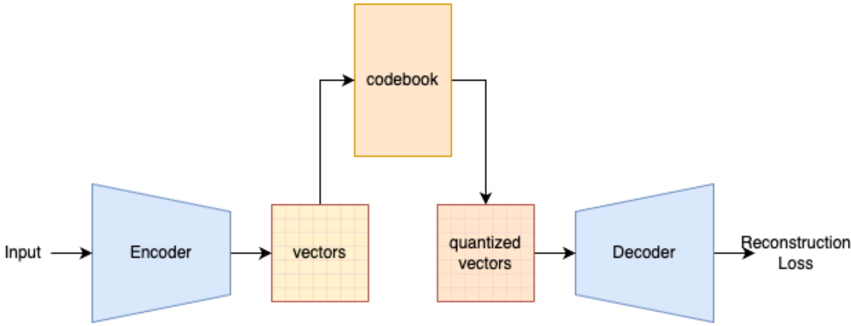
Xu, X.*, Zhang, Y.*, Qian, Q., & Zhang, Y. (2026). CalM: A Self-Supervised Foundation Model for Population Dynamics in Calcium Imaging Data. *ICML26*.

CalM: Method

Intuition from 2 perspectives:

- Neuroscience: calcium traces show periodic pattern, which can possibly be tokenized by VQ net.
- DL: modern large-scale generative models are predominantly trained to predict discrete tokens.





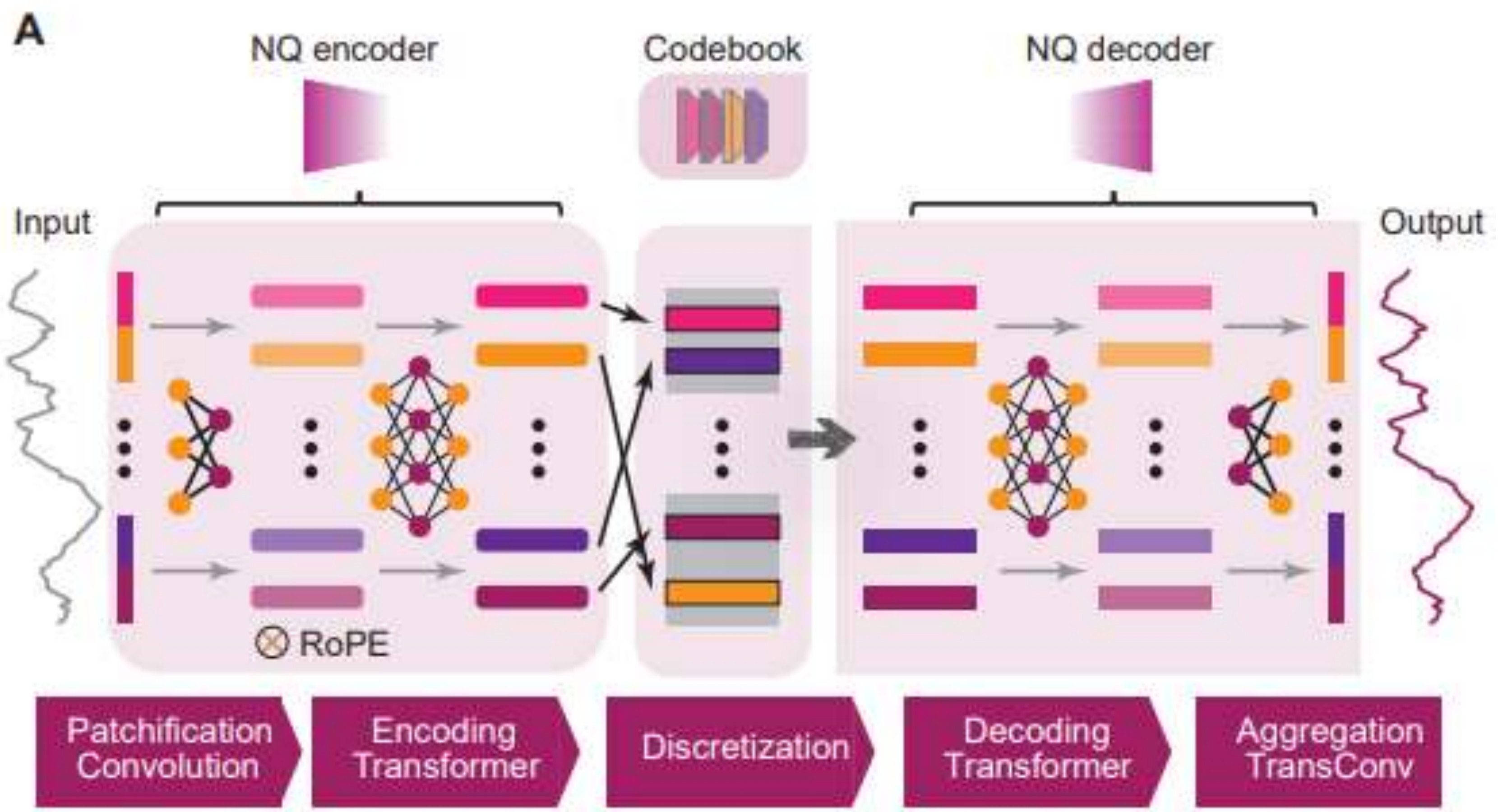
Vector-quantization Variational autoencoder (VQ-VAE)

Explore autoregressive models for calcium data here.

CalM: Method

Neural quantizer (NQ, stage 1):

- NQ is a VQ net that discretizes single neural traces into discrete tokens.



Lossy compression

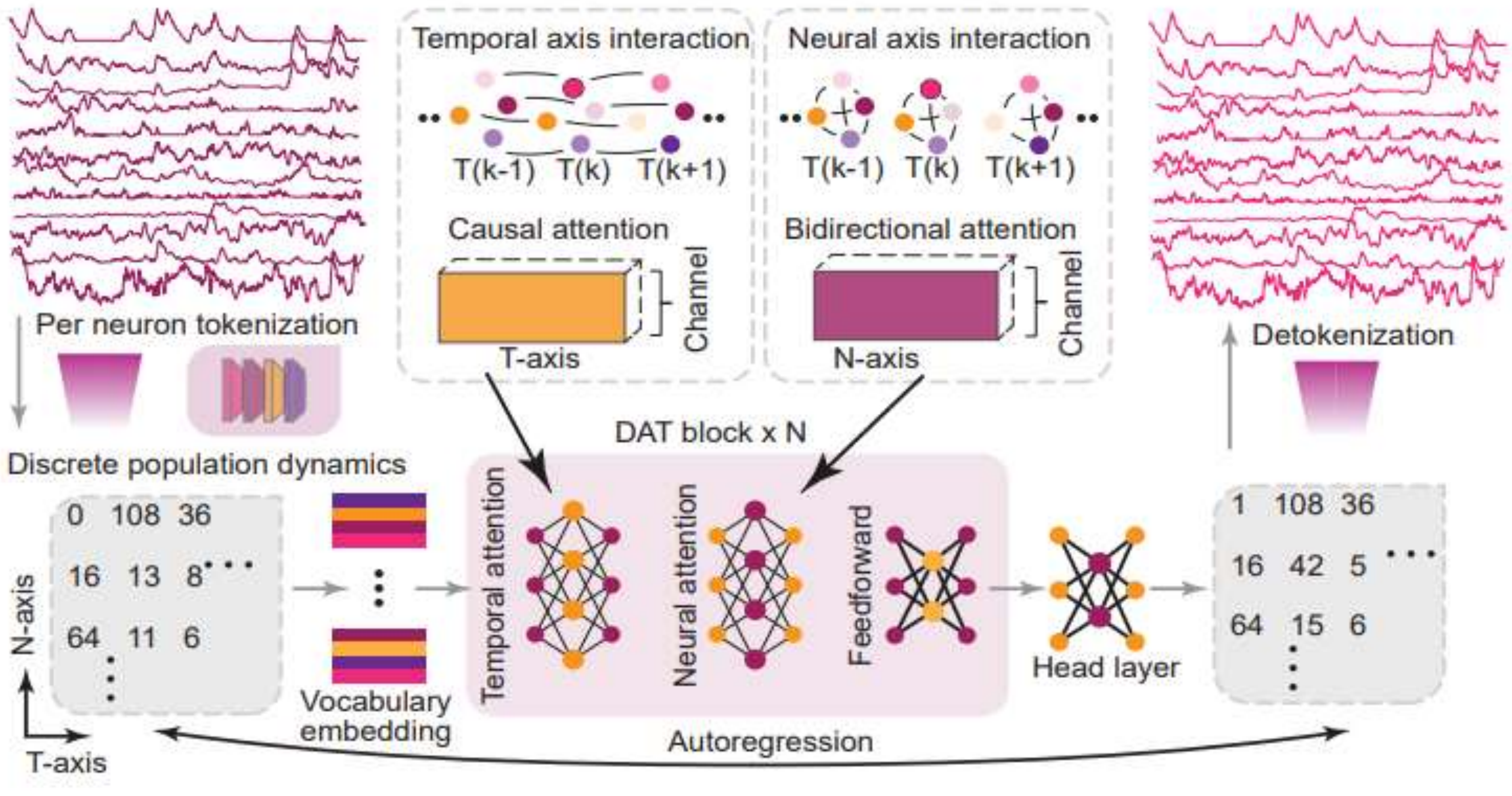
Auxiliary AR head

Gumbel-softmax reparameterization

CalM: Method

Dual-axis Transformer (DAT, stage 2):

- DAT is pretrained to autoregressively predict the next tokens of all neurons given the history.
- Session embeddings and neural embeddings are cooperated into the model for training and fine-tuning.



Computational complexity ?

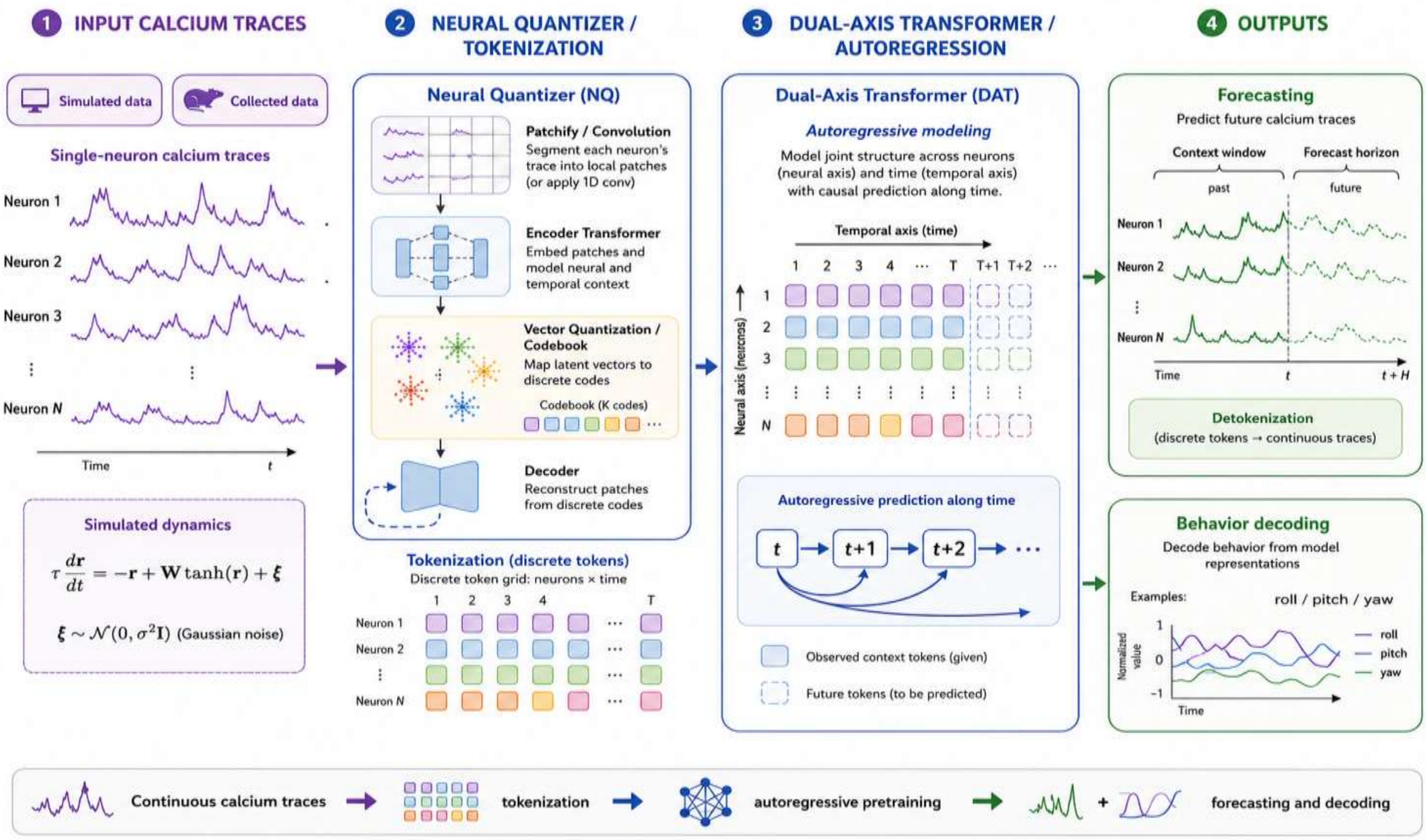
Forecasting task ?

Scheduled Sampling (SS)

Neighborhood Replacement (NR)

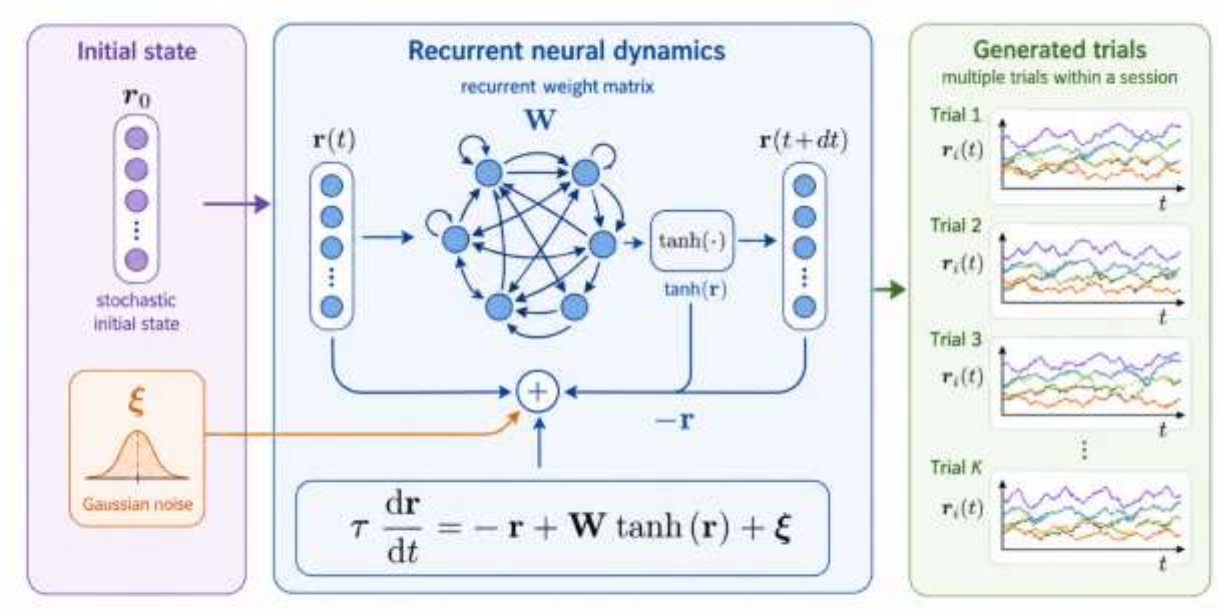
CalM: Method

Overall Pipeline:

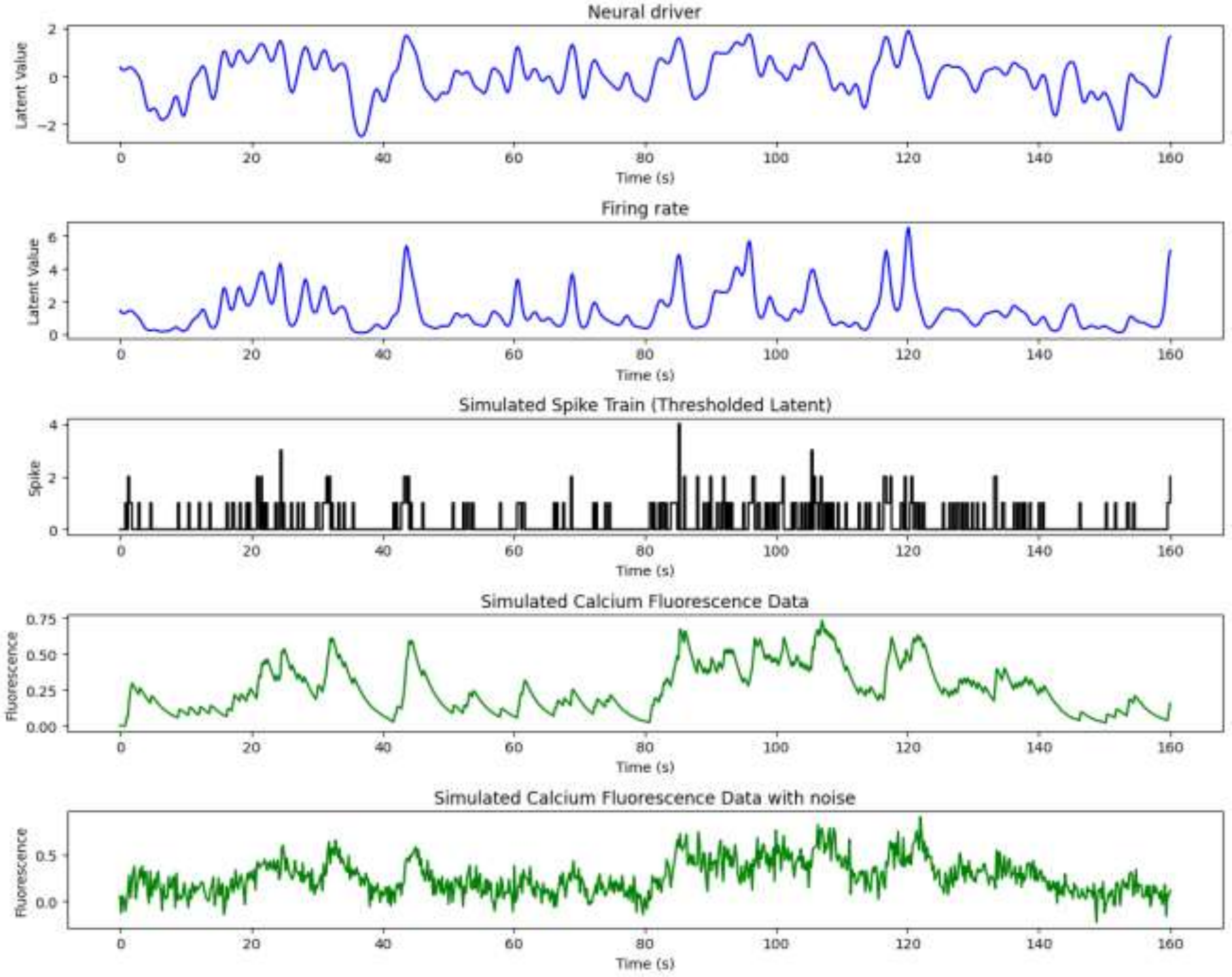


CalM: Datasets

Simulation with linear dynamical system:



Simulated data: recurrent neural dynamics generate multiple trials within a session.



Fixed connectivity W within a session

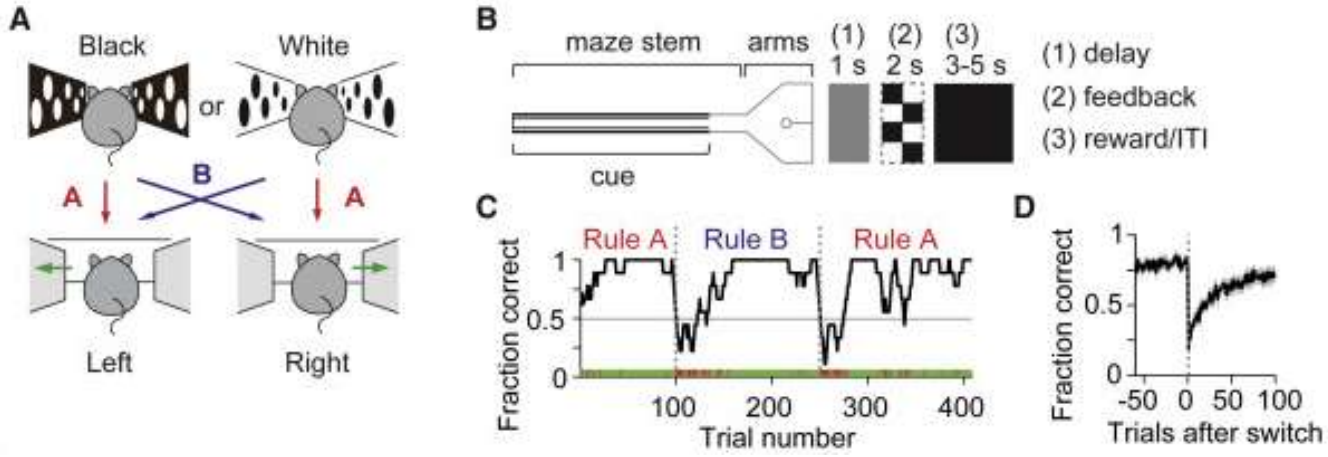
200 Neurons with 400 trials

100 time points at 5 Hz

CalM: Datasets

Collected large-scale open-source datasets:

- Tseng. Dataset: Contains calcium activity from 6 different cortical areas in mouse posterior cortex.



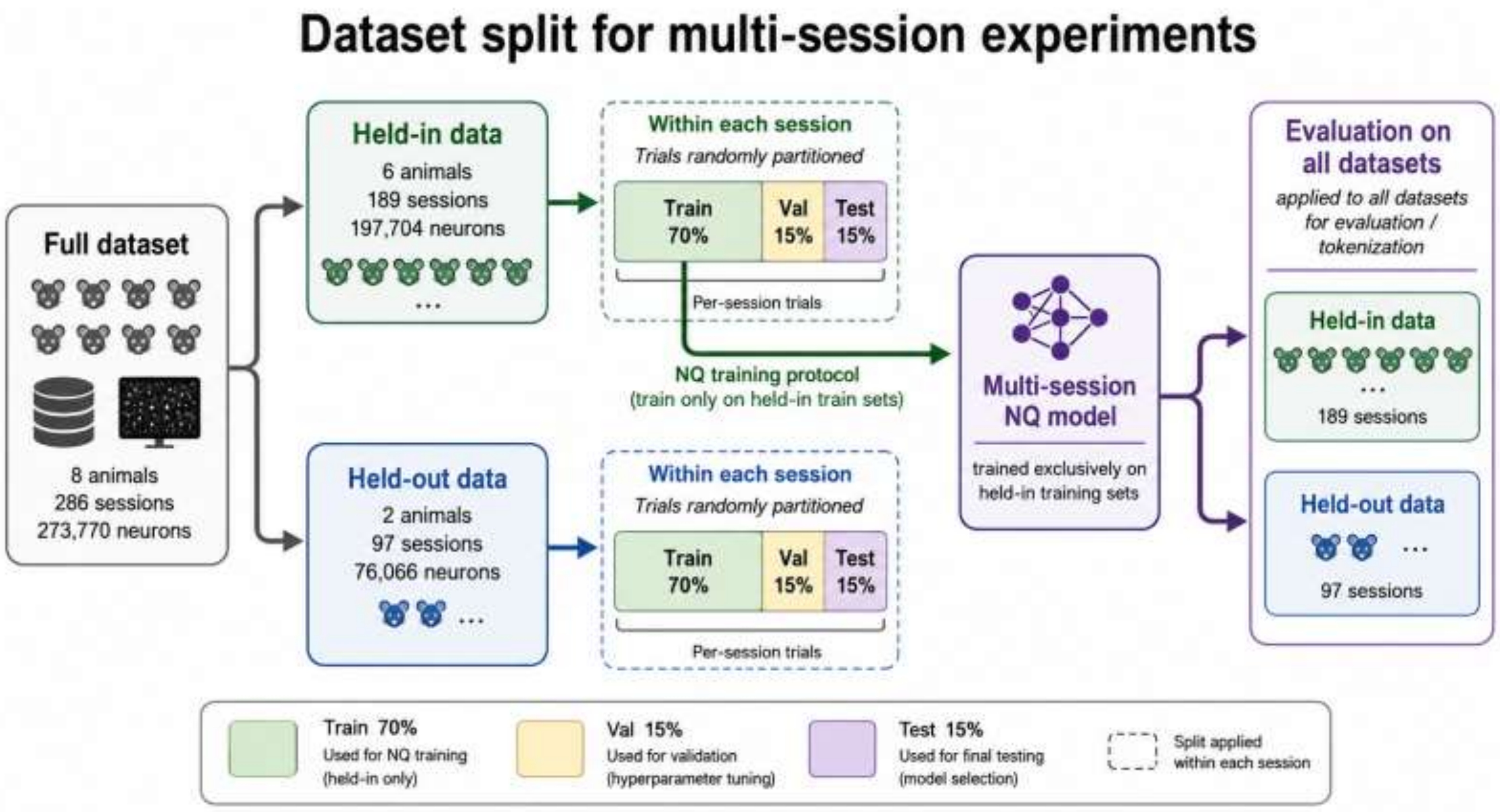
Model	Subjects	Sessions	Neurons / Units
NEDS	73	74	27,380
POCO (Harvey mice)	4	12	~1.6K
POYO+	256	1,335	>110,000
NeuroPaint (IBL)	20	20	21,568
CalM	8	286	273,770

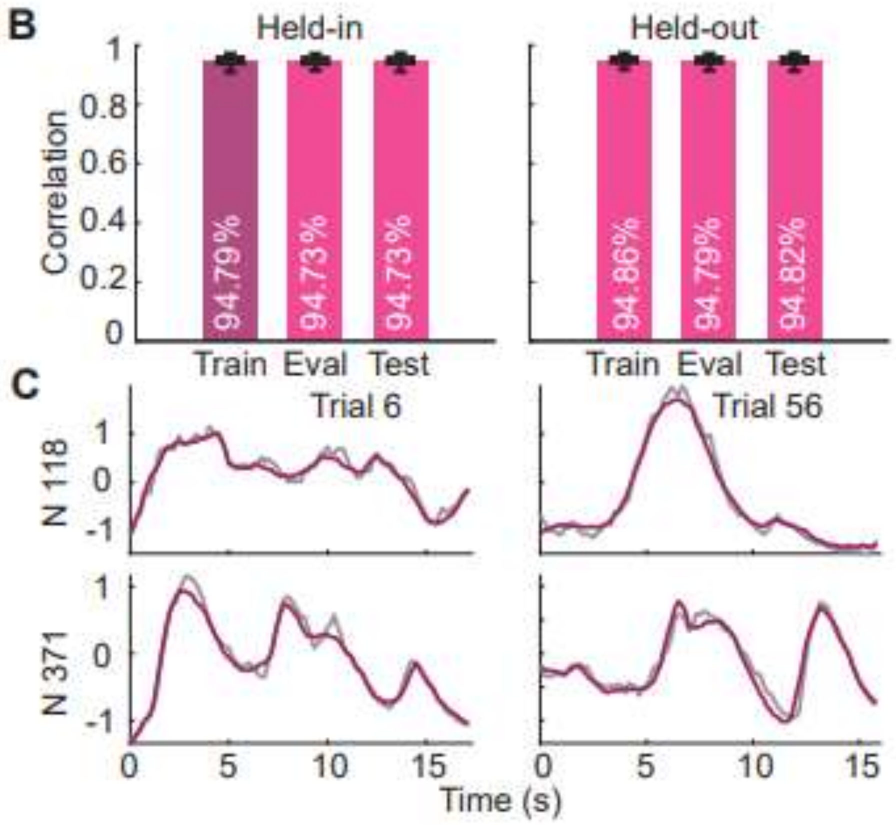
Comparison with other methods

CalM: Predictive Performance

Data splitting and NQ performance:

- Reconstruction quality and generalization are ensured.

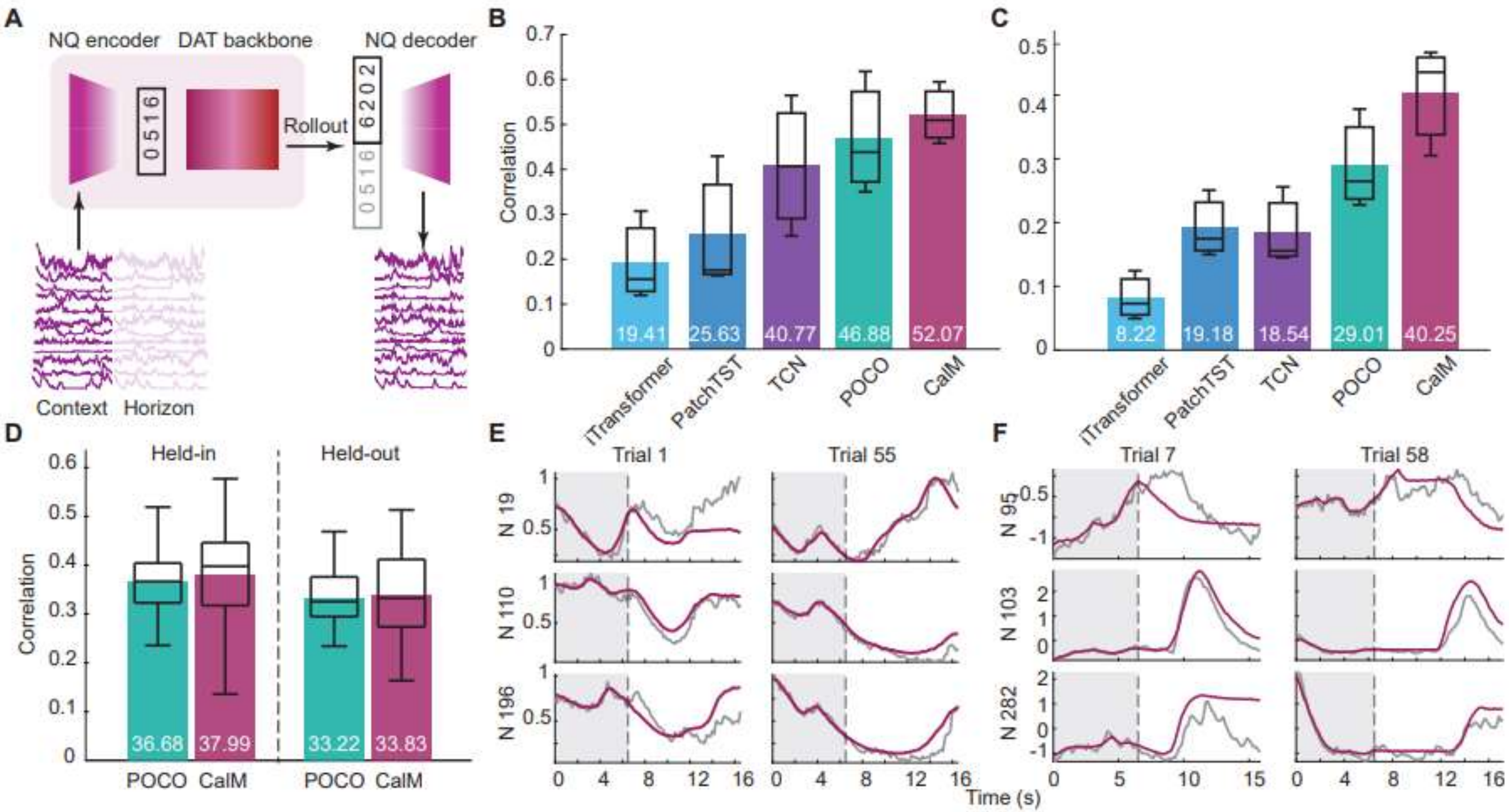




CalM: Predictive Performance

Neural dynamics forecasting:

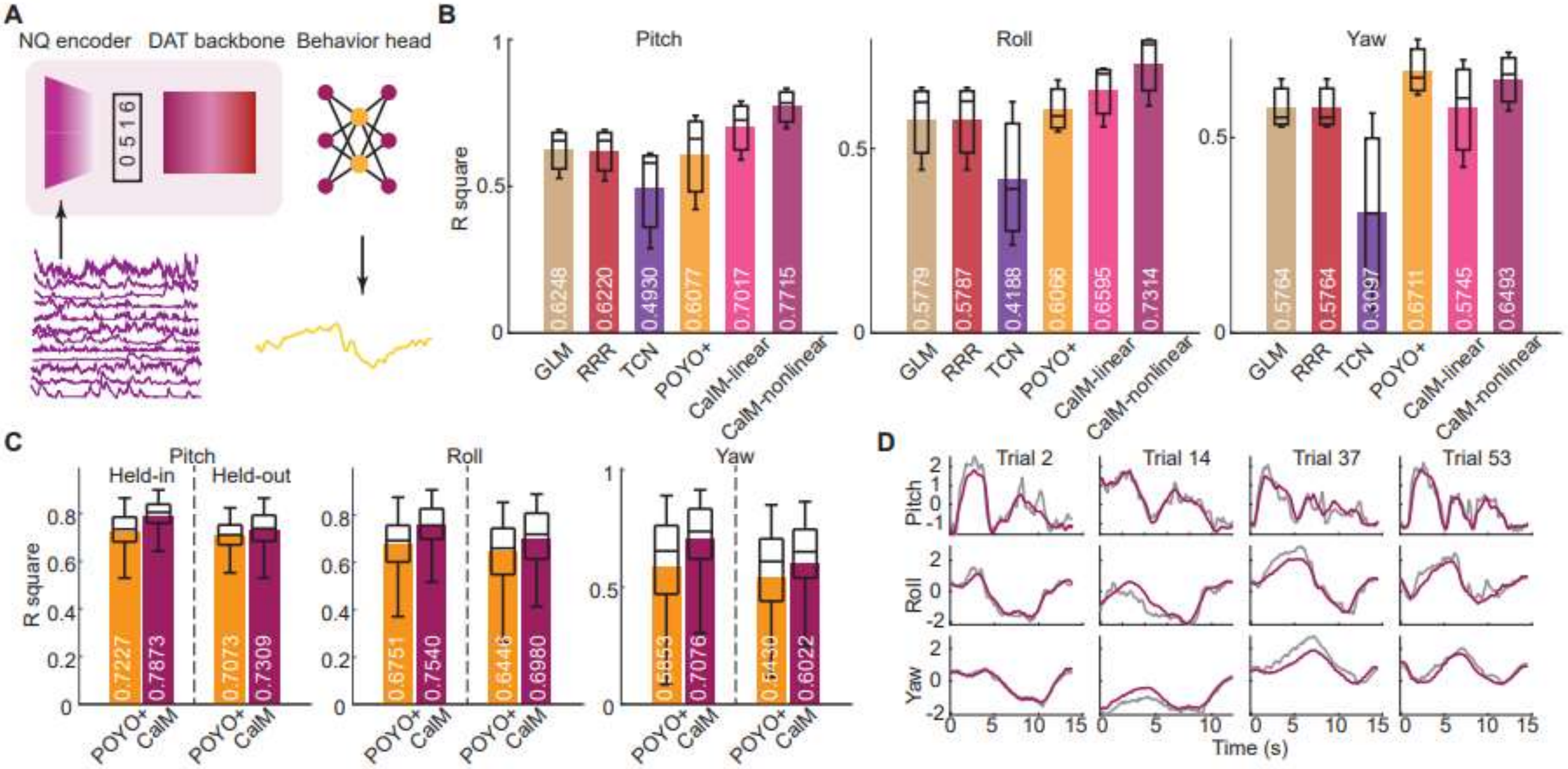
- Fine-tuning just requires adding extra session and neuron embeddings with backbone frozen.



CalM: Predictive Performance

Behavior decoding:

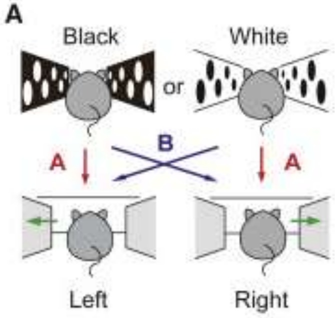
- We freeze the backbone and replace final layer with a behavior decoding head.



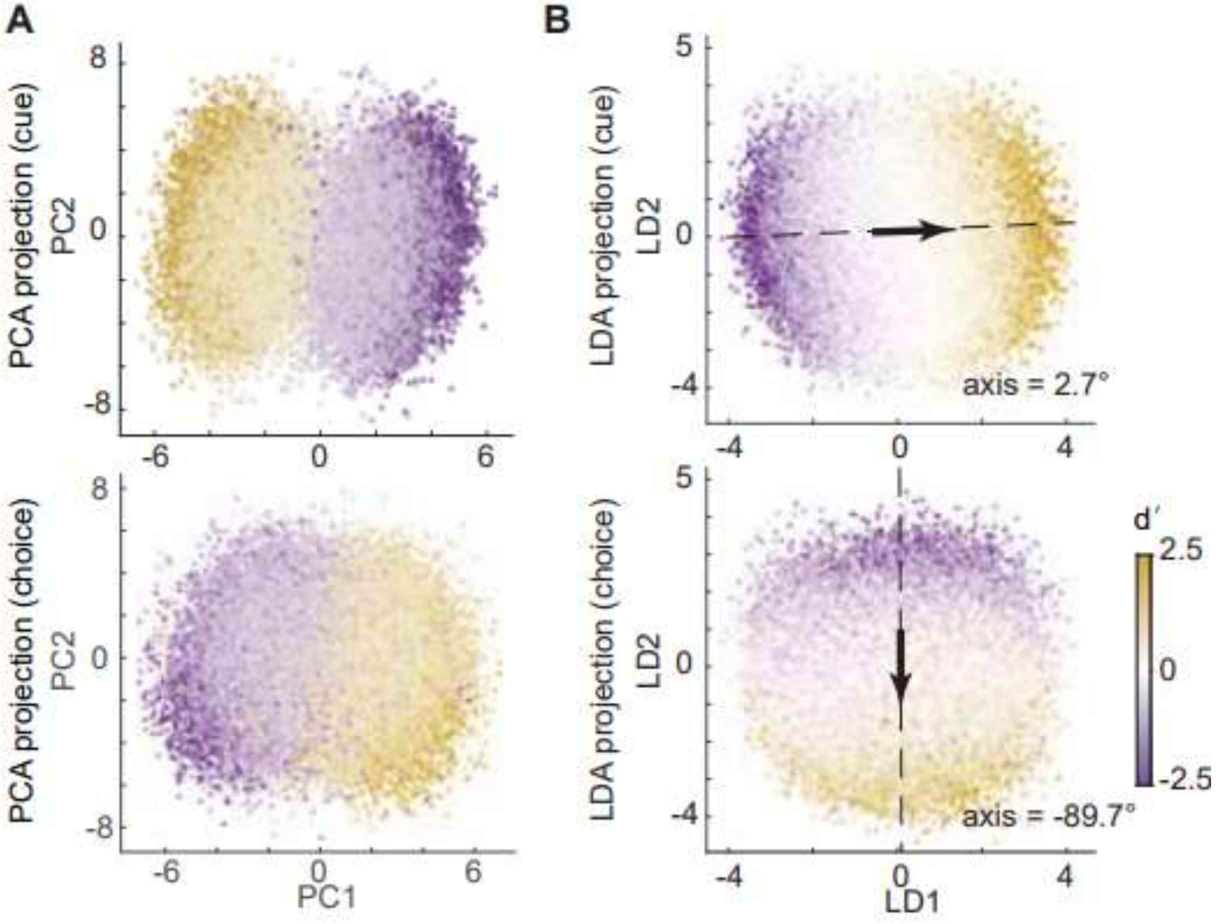
CalM: Linear Functional Study

Interpretability:

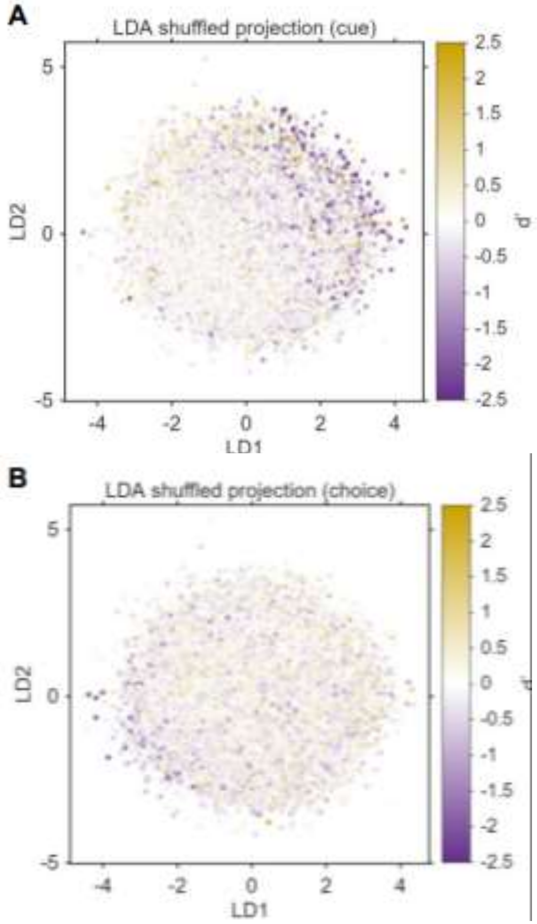
- Unsupervised and supervised dimensional reduction demonstrates clear function separation for neurons.



Use d’ as functional labels



Shuffled control



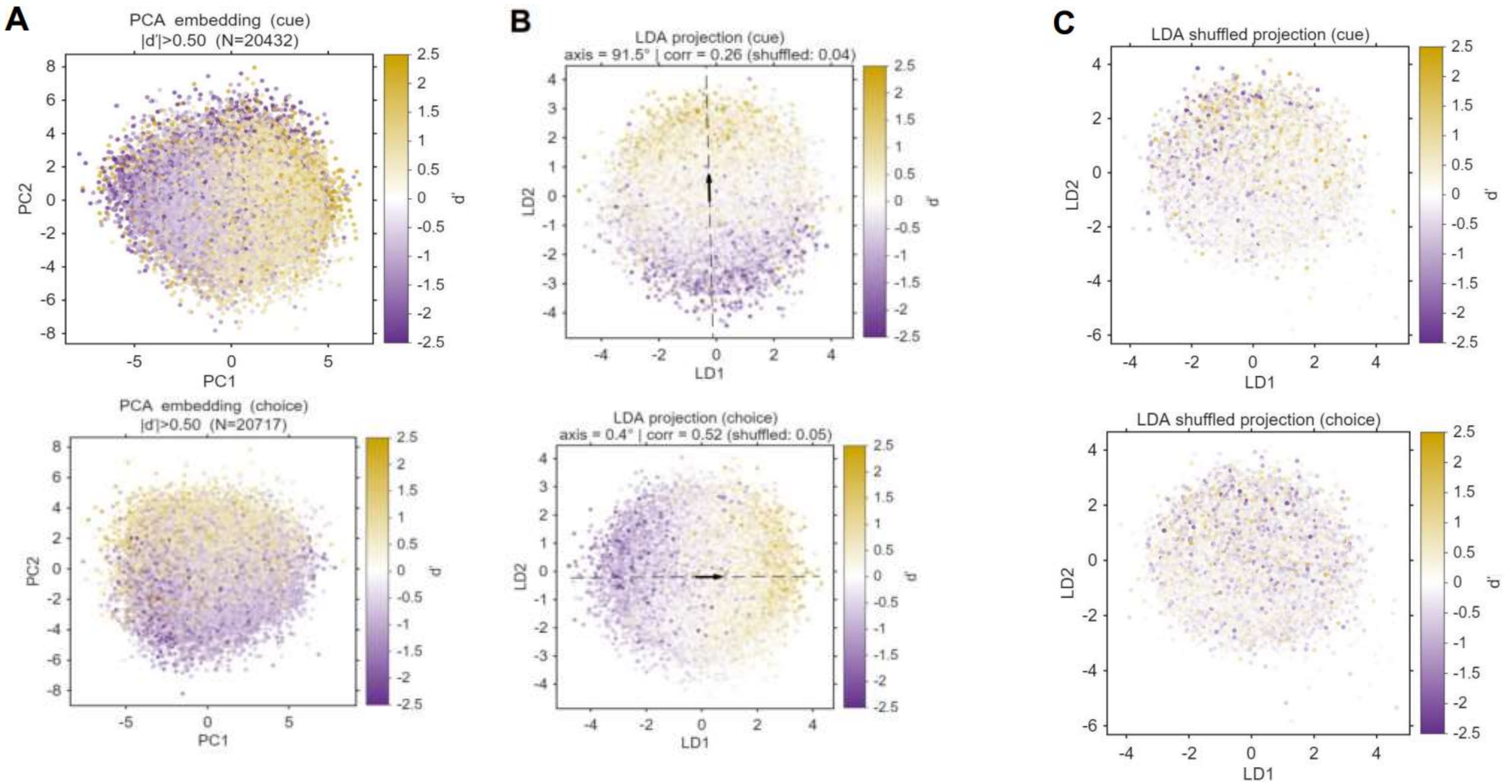
Shared functional coordinate

CalM: Linear Functional Study

Interpretability:

- The same results are revealed in held-out part.

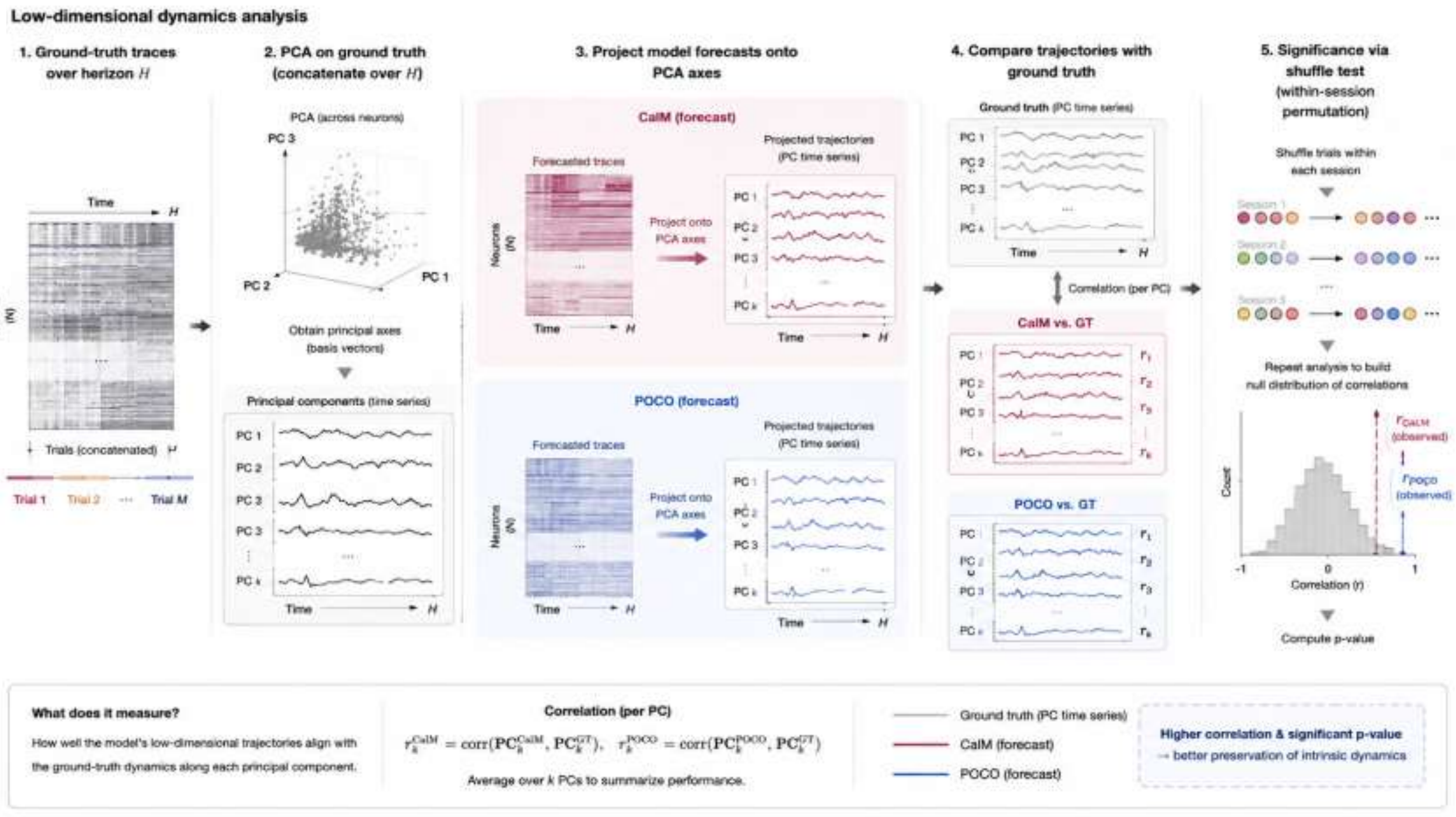
Shuffled control

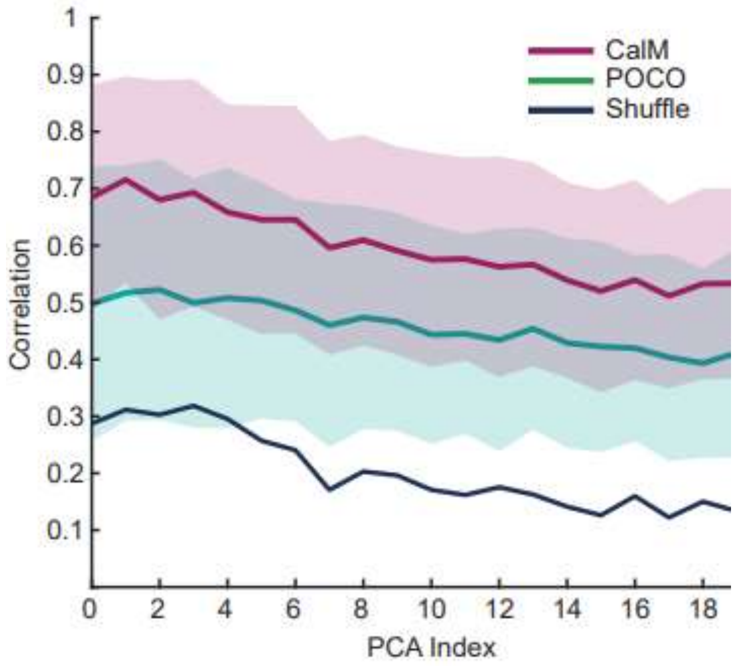


CalM: Linear Functional Study

Interpretability:

- Linear dynamics preservation despite pointwise forecasting error.





Contents

Background and Significance

Related Work

Methods and results

Conclusion and Future Work

- CalM suggests that calcium NFMs should be multi-task capable, transferable and biologically interpretable, but general-purpose calcium NFMs will require broader pretraining, multimodal grounding, and causally validated biological interpretation.

Conclusion

Take-home messages:

CalM:

- Leverages DAT together with NQ tokenization strategy with autoregressive training.
- Effectively bridges generative and discriminative tasks.
- Linear analyses for interpretability demonstrate that the learned representations exhibit clear functional segregation and preserve dynamics.

Future Work

Also serves as the limitations for the work !

I: Broader calcium corpus pretraining:

- General-purpose models require broader pretraining across: brain regions, calcium indicators, sampling rates, laboratories, experimental paradigms, species.
- Current models are still limited by source-domain bias and relatively narrow pretraining distributions.

Key questions:

- Does performance improve with more sessions, neurons, species, and tasks ?
- Is transfer limited by model size, data diversity, or source-domain bias ?
- Can scaling laws be observed for calcium neural data ?
- Can few-shot / zero-shot learning be implemented ?

The next stage is not only larger models, but larger and more diverse calcium pretraining corpora.

Future Work

II: Functional coordinates and multimodal grounding:

- A functional coordinate is a learned representation space where neurons, sessions, or population states can be compared by functional and biological structure.

Key questions:

- Do embeddings align with known cell identities or anatomical classes ?
- Can embeddings predict neurotransmitter type, connectomic, or transcriptomic similarity ?
- Can functional coordinates reveal new biologically meaningful neuron groups?

The goal is not only to improve prediction, but to align learned neural representations with biological structure.

清華大學 Tsinghua University

Thanks for the attention !

Xinhong Xu

2026.7