Too Smooth to Be True?
Evaluating Mean-Field Models in Topographic vs. Non-Topographic V1 Circuits

ZNZ Zentrum für Neurowissenschaften Zürich

ETH zürich

Universität Zürich UZH

Maastricht University

Carolina E. Reinhardt[1, 2], Dr. Christopher Lewis[2], Dr. Mark Roberts[1]

1 Faculty of Psychology & Neuroscience, Maastricht University

2 Brain Research Institute, Neuroscience Center Zurich, ETH Zurich & University of Zurich

Background

Mean-field models (MFMs) are a cornerstone of theoretical neuroscience (Wilson & Cowan, 1972) , they reduce massive, spiking neural populations to a handful of averaged 'population-level' variables, making complex circuits tractable and often revealing fundamental dynamical regimes (Dayan & Abbott, 2001; Deco et al., 2008). Here, we ask how well MFMs preserve stimulus information in two biologically inspired instantiations of layer 2/3 in primary visual cortex (V1): a smooth orientation-preference map as observed in carnivores (e.g. cat) (Ohki et al., 2005) and a Salt-and-Pepper layout as seen in rodents (Niell & Stryker, 2008). By measuring how accurately a linear decoder can recover grating orientation (Graf et al., 2011; Jazayeri & Movshon, 2006) both from the full spiking populations and from their low-dimensional, MFM-derived summaries under realistic noise (Touboul & Ermentrout, 2011), we assess the generalizability and limits of mean-field approximations across organizational motifs in V1.

Methods

1. Network Architectures

- Two LIF networks (N = 7 200) of V1 layer 2/3, 80 % excitatory, 20 % inhibitory
- OPM network: neurons with spatially smooth preferred orientations (0–π)
- Salt-and-Pepper: random orientation, each neuron's preferred orientation is drawn uniform [0,π]

2. Simulation Protocol

- Trials: 12 orientations (0–165° in 15° steps), 30 repeats each
- Timing: 200 ms spontaneous → 500 ms stimulus → 300 ms post-stim
- Dynamics: dt = 0.5 ms, LIF equations (C = 0.5 nF, R = 40 MΩ, V_th = –40 mV)
- Noise: OU background (τ=5 ms, σ=8 mV·√ms) + input-drive jitter (σ_in=1.5·√dt)

4. Full-Network Decoding

- Firing rates (200–600 ms window), baseline-subtracted
- Classifier: one-vs-one ECOC SVM with linear learners
- Cross-validation: 5-fold CV, per-fold z-score of features
- Metrics: overall accuracy, confusion matrices

5. Mean-Field Approximation

- Coarse-grain: partition 7 200 neurons into 72 patches of 100 neurons
- Pooling: sum spikes in 200–600 ms window → patch mean counts
- Poisson sampling: treat each patch's mean as λ, draw one sample/trial
- Decode: same SVM pipeline → assess information preserved in low-dimensional summary

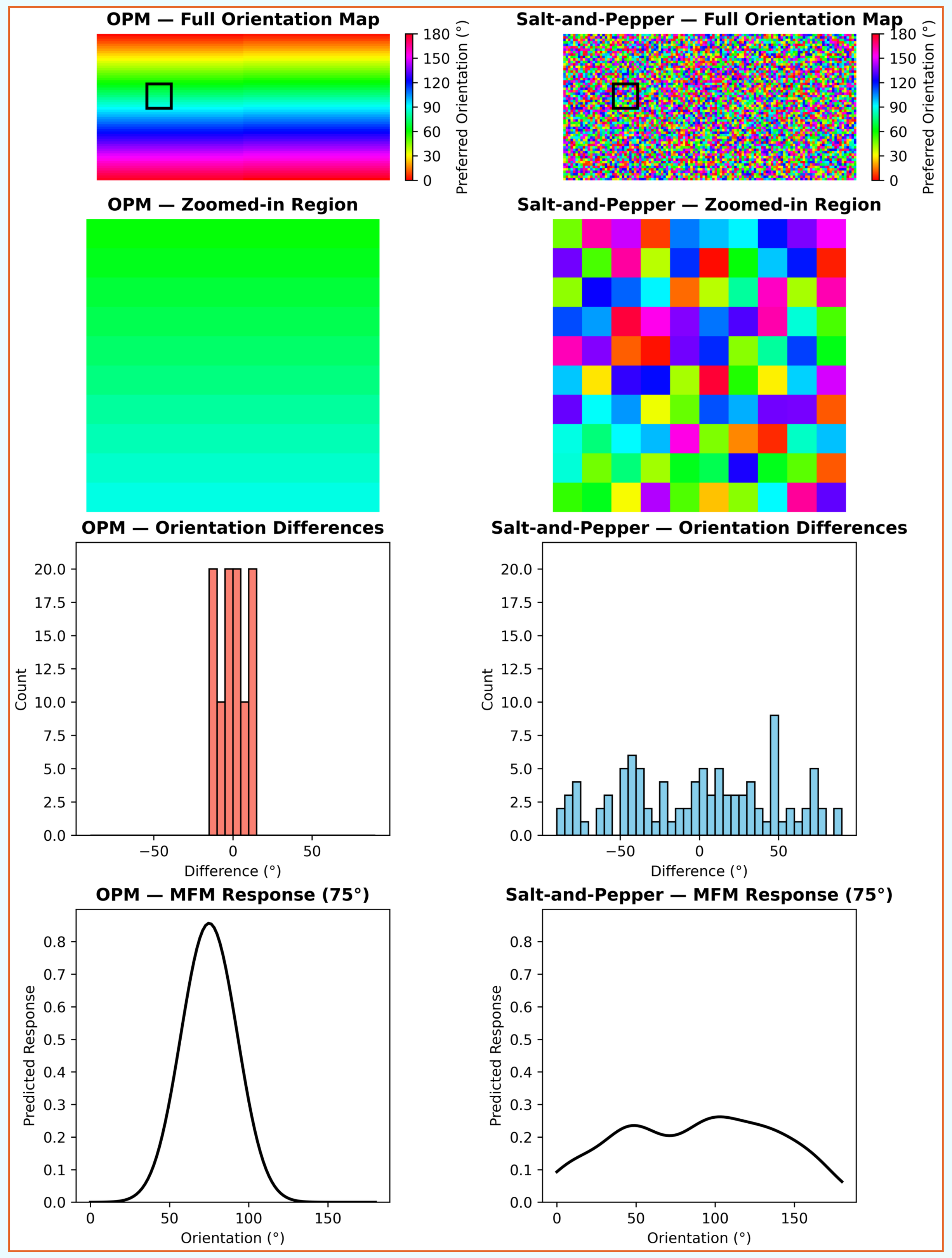

Figure 1. *Comparison of organized (OPM) vs. random (Salt-and-Pepper) orientation networks. Columns correspond to network type; rows (from top to bottom) show:*

1. Full 2D preference maps (color = preferred angle) with a 10x10 patch outlined.
2. Zoomed-in view of that patch.
3. Histogram of pairwise orientation differences within the patch.
4. Mean-field–model prediction of the population response to a 75° stimulus.

Results

Decoding Accuracy Summary

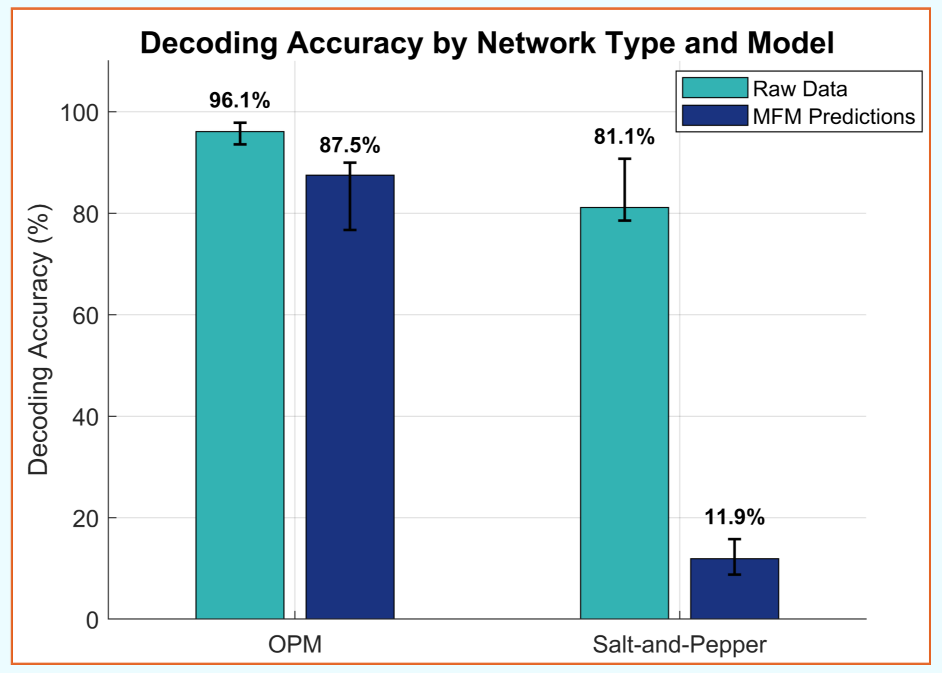

Figure 2. *Mean orientation-decoding accuracy (±95 % CI) for OPM vs. Salt-and-Pepper, comparing full ("Raw Data") and mean-field ("MFM Predictions") models, N = 360 trials.*

- OPM (full): 96.1 % [93.6, 97.9] %
- Salt-and-Pepper (full): 81.1 % [76.7, 85.0] %
- Mean-field: OPM retains 85.3 % vs. Salt-and-Pepper collapses to 11.9 %

Confusion Matrices: Raw vs. Mean-Field

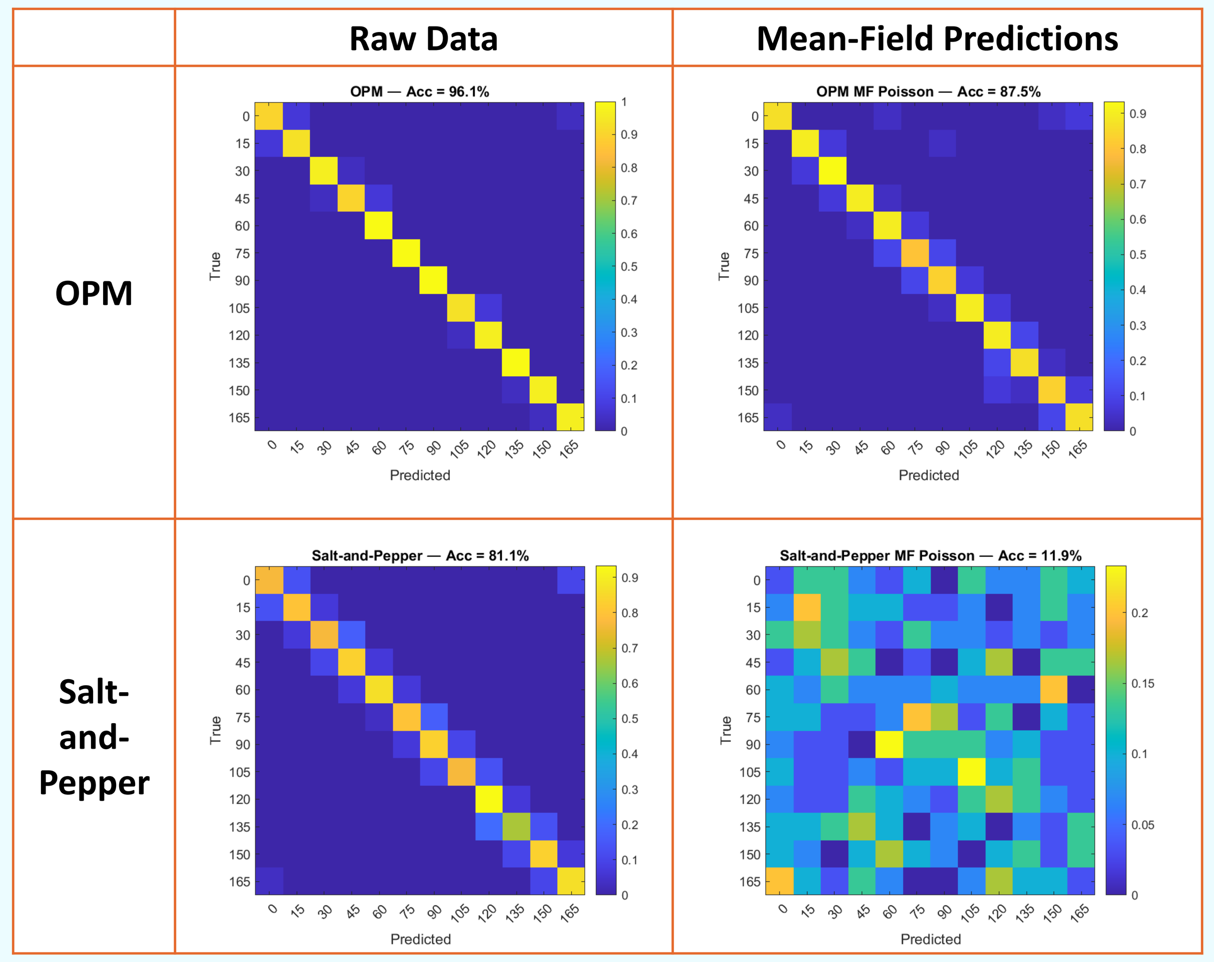

Figure 3. *Normalized confusion matrices (rows = true orientation, columns = predicted) for OPM (top) and Salt–and-Pepper (bottom), using Raw Data (left) or Mean-Field (right) decoding in the 200–600 ms window.*

- OPM: High diagonals under both raw and MFM, showing robust, low-dimensional coding.
- Salt-and-Pepper: Strong raw diagonals collapse under MFM, indicating information loss.

Conclusions & Outlook

Key Finding: Spatially structured orientation maps (OPM) preserve the vast majority of stimulus information when reduced to a low-dimensional mean-field model (96 → 85 % decoding accuracy), whereas unstructured, Salt-and-Pepper networks lose almost all information upon coarse-graining (81 → 12 %).

Broader Implication: Mean-field approximations are highly effective in brain areas with clear topographic or functional organization (e.g. carnivore/primate V1), but will likely fail in regions where neurons are intermingled or where the underlying spatial structure is unknown.

Cautionary Note: Because many cortical and subcortical areas lack well–characterized maps, applying mean-field models there without validating the presence of coherent clustering risks overlooking critical information loss.

Future Directions:

1. Patch size dependence: Systematically vary group sizes (e.g. 20–200 neurons) to chart the trade-off between dimensionality reduction and information retention.
2. Temporal dynamics: Extend to sliding-window analyses of decoding over time and explore how mean-field fidelity evolves during stimulus onset and offset.

References

Dayan, P., & Abbott, L. F. (2001). *Theoretical Neuroscience*. MIT Press.

Deco, G., Jirsa, V. K., Robinson, P. A., Breakspear, M., & Friston, K. (2008). The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS Computational Biology*, 4(8), e1000092.

Graf, A. B. A., Kohn, A., Jazayeri, M., & Movshon, J. A. (2011). Decoding the activity of neuronal populations in macaque primary visual cortex. *Nature Neuroscience*, 14(2), 239–245.

Jazayeri, M., & Movshon, J. A. (2006). Optimal representation of sensory information by neural populations. *Nature Neuroscience*, 9(5), 690–696.

Niell, C. M., & Stryker, M. P. (2008). Highly selective receptive fields in mouse visual cortex. *Journal of Neuroscience*, 28(30), 7520–7536.

Ohki, K., Chung, S., Ch'ng, Y. H., Kara, P., & Reid, R. C. (2005). Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature*, 433(7026), 597–603.

Touboul, J., & Ermentrout, G. B. (2011). Finite-size and correlation effects in mean-field neural populations. *Physical Review E*, 84(3), 031910.

Wilson, H. R., & Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12(1), 1–24.

Correspondence to:
Carolina E. Reinhardt, Faculty of Psychology & Neuroscience, Maastricht University

ce.reinhardt@student.maastrichtuniversity.nl

Maastricht University

P.O. Box 616
6200 MD Maastricht, The Netherlands

GitHub repository:
github.com/C-Reinhardt/MFM_DecodingComparison