Effects of higher-order statistics on activity in V1

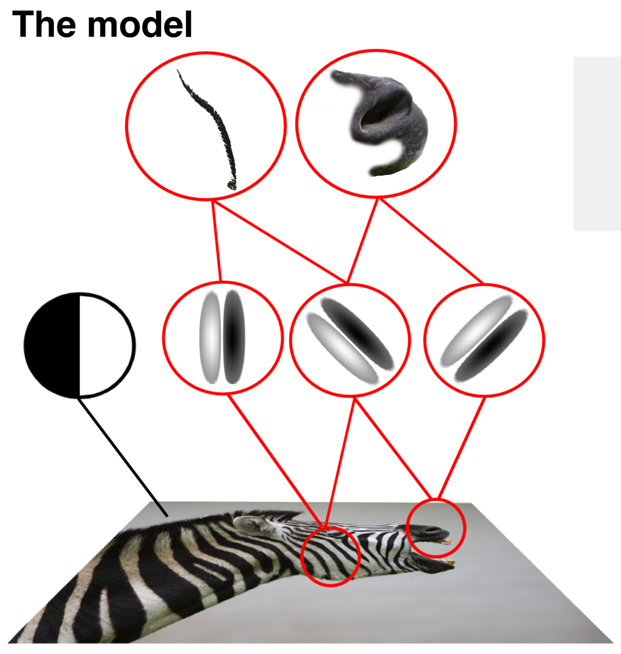
Mihály Bányai, Gergő Orbán

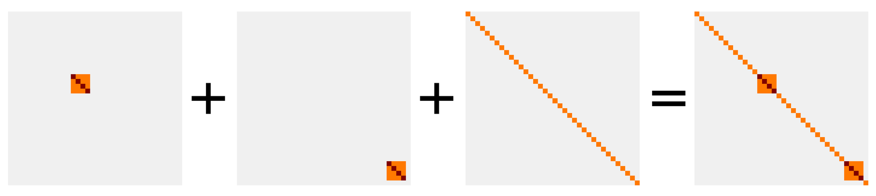
Computational Systems Neuroscience Lab
Wigner Research Centre for Physics, Budapest

Introduction

- The visual system is representing a hierarchical **generative model** of the environment.
- V1 simple cell responses are organised by **latent variables** representing higher level statistics of sensory input.
- The latent structure determining the covariance structure of V1 simple cells corresponds to **Gestalt principles**.
- Full **Bayesian inference** is assumed in the model, posteriors are represented by stochastic **samples.**

The model





$$p(v \mid g) = \mathcal{N}(v; 0, \sum_{j=1}^{K} g_j C_j)$$

$$p(x \mid v, z) = \mathcal{N}(x; zAv, \sigma_x I)$$

- Gestalts are learned from data as covariance components $C_{1...K}$
- Coefficient variables of covariance components, g, are generated from Gamma priors
- The global contrast variable, z, is also generated from a Gamma prior
- Linear filter set A transforms V1-level variables, v to pixels, x, adding independent observation noise

Sampling model posteriors

- Gibbs sampling
- Conditional v activations can be sampled directly from a Gaussian conditional

$$p(v \mid x, g, z) = \mathcal{N}\left(v; \frac{z}{\sigma_x} C_{v|x,g,z} A^T x, C_{v|x,g,z}\right), \quad C_{v|x,g,z} = \left(\frac{z^2}{\sigma_x} A^T A + \left(\sum_{j=1}^{K} g_j C_j\right)^{-1}\right)^{-1}$$

- Conditional g and z activations can be sampled by MCMC schemes

$$\log p(g \mid x, v, z) \sim -\frac{1}{2}\left[\log\left(\det\left(\sum_{k=1}^{K} g_j C_j\right)\right) + v^T \left(\sum_{k=1}^{K} g_j C_j\right)^{-1} v\right] + \log p(g)$$

$$\log p(z \mid x, v, g) \sim -\frac{1}{2}\left[D_x \log(\sigma_x) + \frac{1}{\sigma_x}(x - zAv)^T (x - zAv)\right] + \log p(z)$$

Learning the parameters

- Iterative stochastic generalised expectation maximisation
- Collecting L samples with Gibbs sampling as E-step
- Reparametrise with Cholesky components to ensure validity of covariance matrices

$$C_v = \sum_{k=1}^{K} g_k U_k^T U_k$$

- Gradient of the complete-data log-likelihood over N observations

$$\frac{\partial \mathcal{L}}{\partial [U_k]_{i,j}} = \sum_{l=1}^{NL} \mathrm{Tr}\left[\frac{\partial \log p(x^l, v^l, g^l \mid U_{1...K})}{\partial C_v^l} \frac{\partial C_v^l}{\partial [U_k]_{i,j}}\right]$$

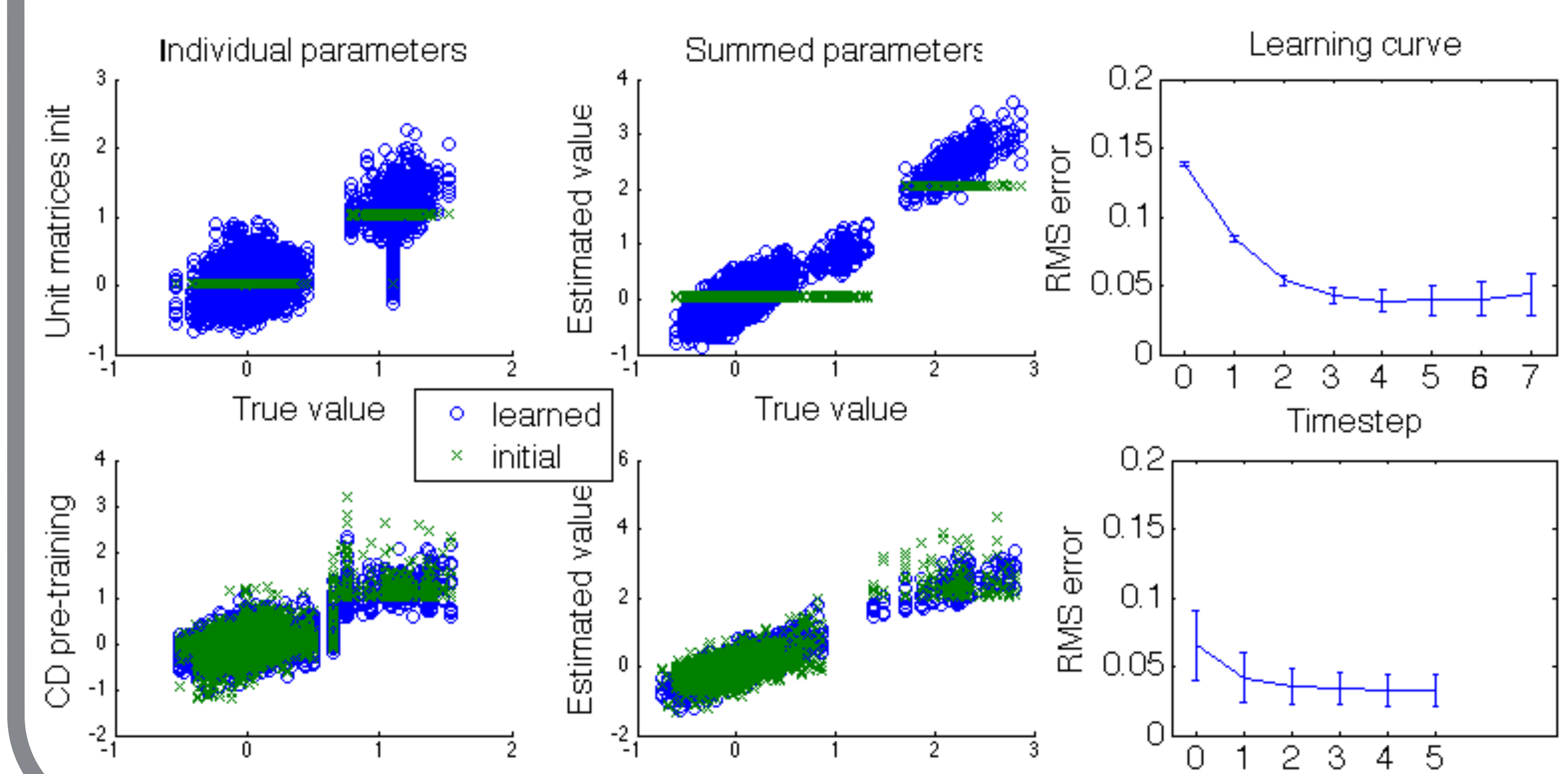
- Gradient descent as generalised M-step

$$[U_k]_{i,j}^{new} = [U_k]_{i,j}^{old} + \epsilon \frac{\partial \mathcal{L}}{\partial [U_k]_{i,j}}$$

Pre-training with contrastive divergence

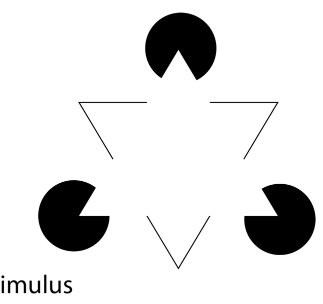
- Constructing a restricted Boltzmann machine of the two layers of hidden variables
- Sample v for each observation, and run CD1 between v and g layers
- Construct covariance components from learned weights between v and each g_k

$$[C_k]_{i,j} = W_{i,k} W_{j,k}$$



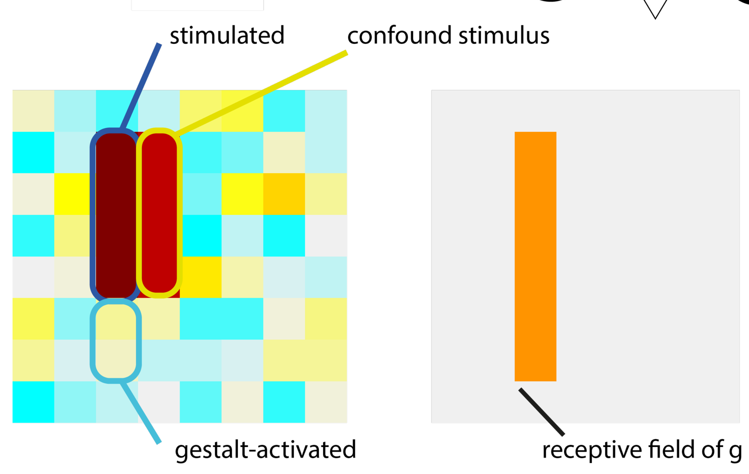
Response to illusory contours

- We propose that IC responses are elicited by top-down Gestalt effects (g activations)
- The IC can be regarded as the non-stimulus-activated part of the receptive field of a covariance component partially activated by the stimulus



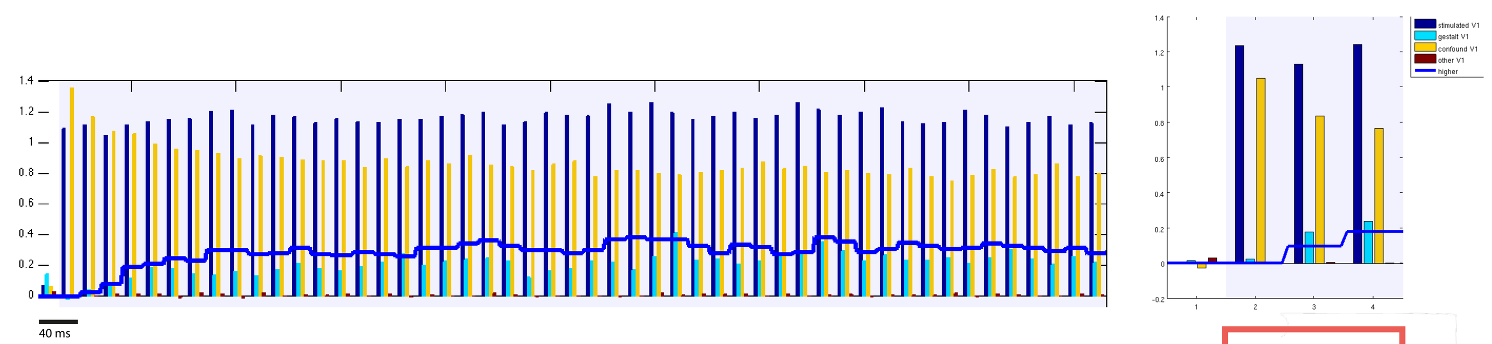
Stimulus

- To compare model responses to experimental results with IC a test stimulus is constructed
- Left: mean stimulus in the v space
- Right: receptive field of one of the covariance component in the v space

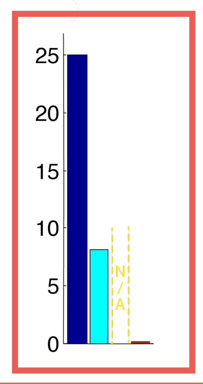


Firing rates

- Alternating sampling from v and g conditioned on the observation, starting with v
- Non-stimulus-evoked activity in v follows the activation of higher-level areas

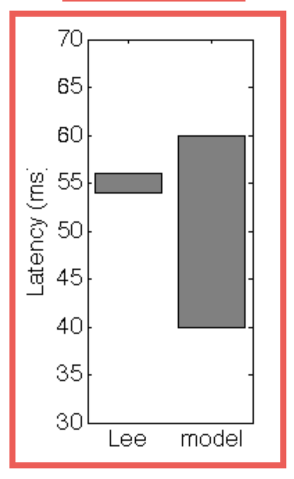


- The model reproduces the magnitude ratio of response rates to real contours, ICs and background
- Mean spike count in V1 as a response to different stimuli in receptive field from **Lee & Nguyen, PNAS, 2001.**



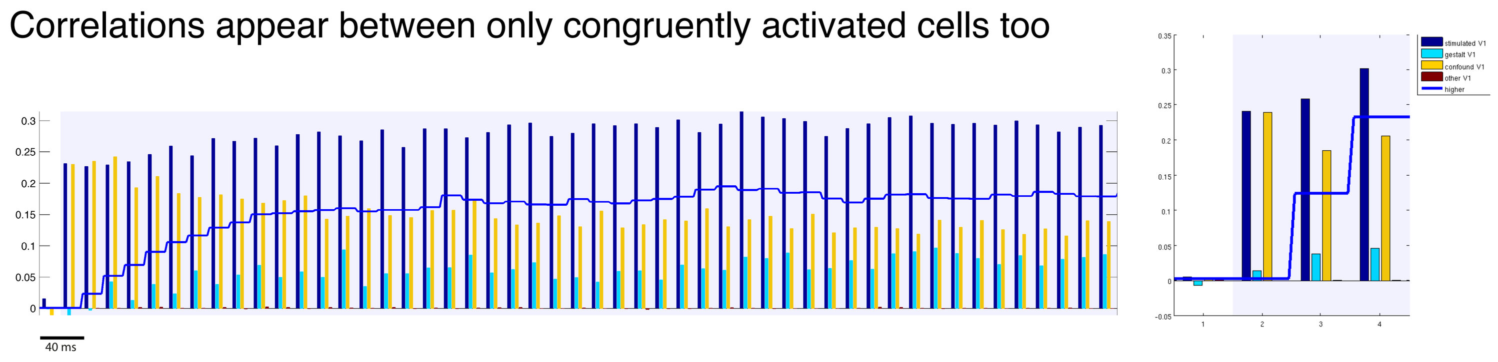
Latencies

- Lee & Nguyen (2001) report that the response to ICs arrives with an additional latency of 55 ms compared to real contours
- In the model, rates of Gestalt-activated v units increase 2-3 Gibbs steps later
- Membrane potential autocorrelation typically diminishes after 20 ms, taking this as the duration of a sampling step, the model predicts 40-60 ms latency



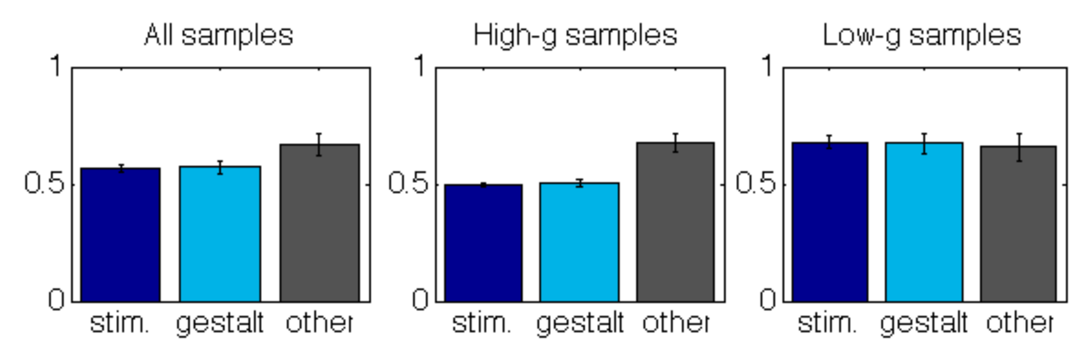
Predicted correlations

- The model predicts that noise correlations increase with g activations
- Correlations appear between only congruently activated cells too

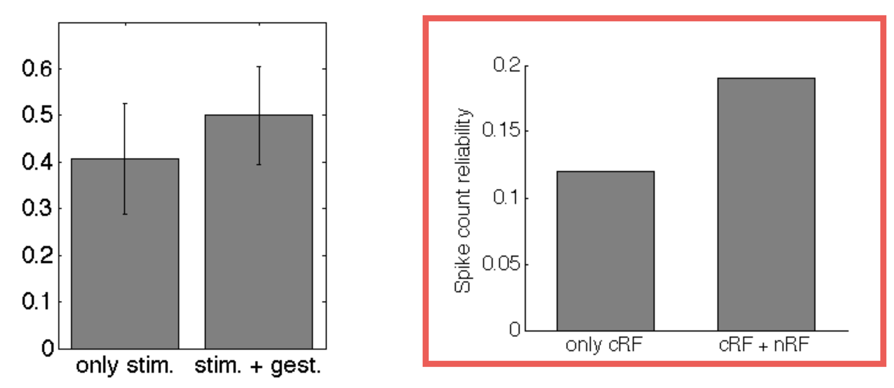


Variances

- Within-trial variance of stimulated and Gestalt-activated v units decrease in the model as the g activation of the covariance component that has them in its receptive field increases



- **Haider et al. Neuron, 2010.** report that the spike train reliability of V1 cells increase with non-classical RF activation when added to classical RF stimuli.



Conclusions

- Gestalt principles can be formalised in a probabilistic generative model
- Inference on the model can be realised as Gibbs sampling
- Firing rates during sampling reproduces the temporal order and the magnitude ratio of activation in cortical areas responding to illusory contours
- The model predicts diminishing variance in cells with Gestalt-activated receptive fields
- Top-down effects may contribute to non-classical receptive field activations
- The model can be trained on a set of stimuli with an EM algorithm
- Contrastive divergence can approximate model parameters and speed up learning

Acknowledgement

This work was supported by an MTA "Lendület" Fellowship to G.O.