Olfactory inference using correlated priors with sister cells

Sina Tootoonian
Schaefer Lab

Neuroscience Interest Group
2025/05/28

Section 1

Overview

Outline

- Olfactory bulb, sister cells

Outline

- Olfactory bulb, sister cells
- Top-down understanding of neural circuits

Outline

- Olfactory bulb, sister cells
- Top-down understanding of neural circuits
- Olfactory inference

Outline

- Olfactory bulb, sister cells
- Top-down understanding of neural circuits
- Olfactory inference
- Circuits for inference

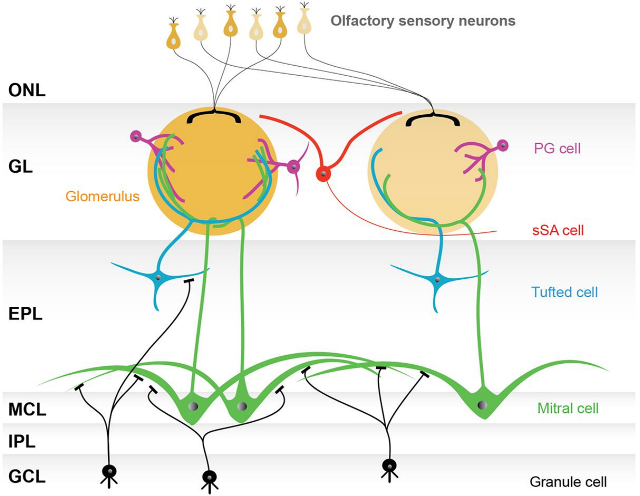
Outline

- Olfactory bulb, sister cells
- Top-down understanding of neural circuits
- Olfactory inference
- Circuits for inference
- Incorporating sister cells

Outline

- Olfactory bulb, sister cells
- Top-down understanding of neural circuits
- Olfactory inference
- Circuits for inference
- Incorporating sister cells
- Preprint at http://arxiv.org/abs/2505.14579

The olfactory bulb



Nagayama et al. 2014

Two approaches to understanding neural circuits

- Bottom-up

Two approaches to understanding neural circuits

- Bottom-up
- Top-down

Two approaches to understanding neural circuits

- Bottom-up
- Top-down
  - Marr's levels:

Two approaches to understanding neural circuits

- Bottom-up
- Top-down
  - Marr's levels:
    1. What is the computation?

Two approaches to understanding neural circuits

- Bottom-up
- Top-down
  - Marr's levels:
    1. What is the computation?
    2. What is the algorithm?

Two approaches to understanding neural circuits

- Bottom-up
- Top-down
  - Marr's levels:
    1. What is the computation?
    2. What is the algorithm?
    3. What is the implementation?

The computation: Olfactory inference

- The environment contains odours $\mathbf{x}$ animals care about.

The computation: Olfactory inference

- The environment contains odours $\mathbf{x}$ animals care about.
- These have a *prior* probability $p(\mathbf{x})$.

The computation: Olfactory inference

- The environment contains odours $\mathbf{x}$ animals care about.
- These have a *prior* probability $p(\mathbf{x})$.
- Odours produce receptor responses $\mathbf{r}$ according to $p(\mathbf{r}|\mathbf{x})$.

The computation: Olfactory inference

- The environment contains odours $\mathbf{x}$ animals care about.
- These have a *prior* probability $p(\mathbf{x})$.
- Odours produce receptor responses $\mathbf{r}$ according to $p(\mathbf{r}|\mathbf{x})$.
- Combining these gives a *posterior* probability $p(\mathbf{x}|\mathbf{r})$

The computation: Olfactory inference

- The environment contains odours $\mathbf{x}$ animals care about.
- These have a *prior* probability $p(\mathbf{x})$.
- Odours produce receptor responses $\mathbf{r}$ according to $p(\mathbf{r}|\mathbf{x})$.
- Combining these gives a *posterior* probability $p(\mathbf{x}|\mathbf{r})$
- Computation: Find the odour with highest posterior probability

$$\hat{\mathbf{x}} = \max_{\mathbf{x}} p(\mathbf{x}|\mathbf{r}).$$

The algorithm: Hill climbing

- Our computation:

$$\hat{\mathbf{x}} = \max_{\mathbf{x}} p(\mathbf{x}|\mathbf{r}) = \max_{\mathbf{x}} \log p(\mathbf{x}|\mathbf{r})$$

The algorithm: Hill climbing

- Our computation:

$$\hat{\mathbf{x}} = \max_{\mathbf{x}} p(\mathbf{x}|\mathbf{r}) = \max_{\mathbf{x}} \log p(\mathbf{x}|\mathbf{r})$$

- Improve a guess $\hat{\mathbf{x}}$ by moving up the probability hill

$$\frac{d\hat{\mathbf{x}}}{dt} \propto \nabla \log p(\hat{\mathbf{x}}|\mathbf{r})$$

The algorithm: Hill climbing

- Our computation:

$$\hat{\mathbf{x}} = \max_{\mathbf{x}} p(\mathbf{x}|\mathbf{r}) = \max_{\mathbf{x}} \log p(\mathbf{x}|\mathbf{r})$$

- Improve a guess $\hat{\mathbf{x}}$ by moving up the probability hill

$$\frac{d\hat{\mathbf{x}}}{dt} \propto \nabla \log p(\hat{\mathbf{x}}|\mathbf{r})$$

- Use Bayes' rule to relate the posterior to the prior and likelihood:

$$\log p(\mathbf{x}|\mathbf{r}) = \log p(\mathbf{x}) + \log p(\mathbf{r}|\mathbf{x}).$$

The algorithm: Hill climbing

- Our computation:

$$\hat{\mathbf{x}} = \max_{\mathbf{x}} p(\mathbf{x}|\mathbf{r}) = \max_{\mathbf{x}} \log p(\mathbf{x}|\mathbf{r})$$

- Improve a guess $\hat{\mathbf{x}}$ by moving up the probability hill

$$\frac{d\hat{\mathbf{x}}}{dt} \propto \nabla \log p(\hat{\mathbf{x}}|\mathbf{r})$$

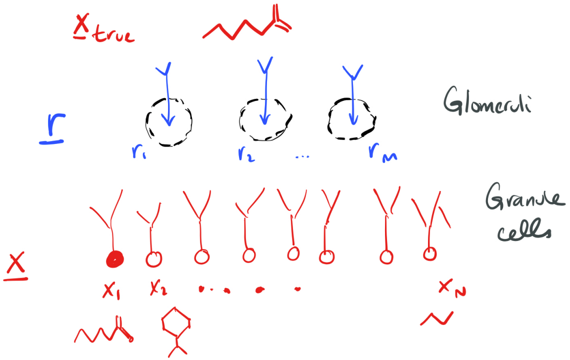
- Use Bayes' rule to relate the posterior to the prior and likelihood:

$$\log p(\mathbf{x}|\mathbf{r}) = \log p(\mathbf{x}) + \log p(\mathbf{r}|\mathbf{x}).$$

- Two contributions to the dynamics

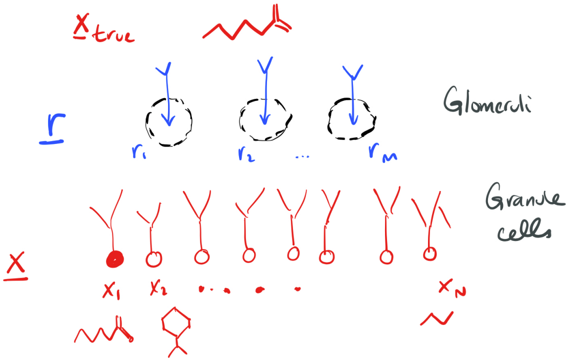
$$\frac{d\hat{\mathbf{x}}}{dt} \propto \nabla \log p(\hat{\mathbf{x}}) + \nabla \log p(\mathbf{r}|\hat{\mathbf{x}}).$$

Implementation: Representations



- Receptors $\mathbf{r}$: glomeruli.

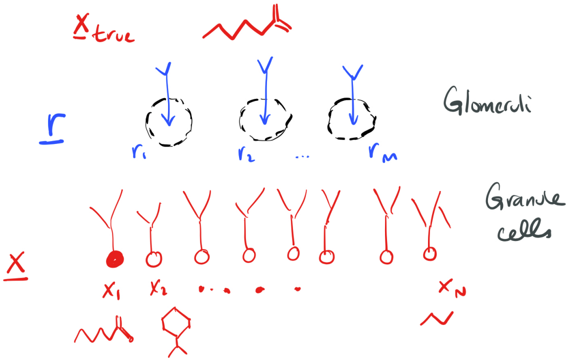
Implementation: Representations



- Receptors $\mathbf{r}$: glomeruli.
- Odours = Vectors of molecular concentrations:

$$\mathbf{x} = \{x_1, x_2, \ldots, x_N\}.$$

Implementation: Representations



- Receptors $\mathbf{r}$: glomeruli.
- Odours = Vectors of molecular concentrations:

$$\mathbf{x} = \{x_1, x_2, \ldots, x_N\}.$$

- Assign one molecular species to one granule cell.

Specifying a prior

- How to specify $p(\mathbf{x}) = p(x_1, x_2, \ldots, x_N)$?

Specifying a prior

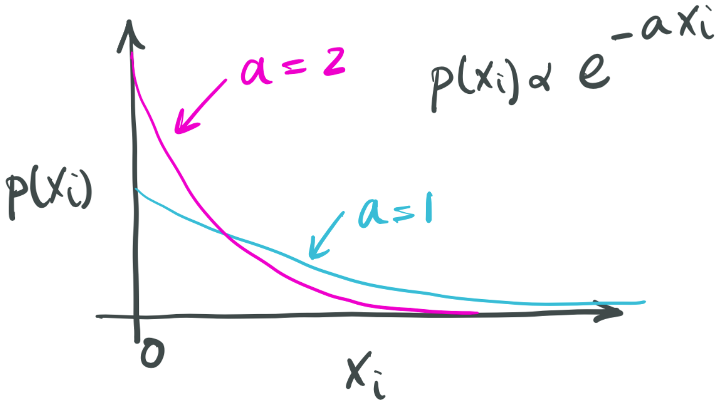
- How to specify $p(\mathbf{x}) = p(x_1, x_2, \ldots, x_N)$?
- Natural odours are *sparse*: most likely concentration is zero.

Specifying a prior

- How to specify $p(\mathbf{x}) = p(x_1, x_2, \ldots, x_N)$?
- Natural odours are *sparse*: most likely concentration is zero.
- Suggests e.g.

$$p(x_i) \propto e^{-ax_i}.$$

Specifying a prior



Specifying a prior

- How to get $p(x_1, \ldots, x_N)$ from $p(x_i)$?

Specifying a prior

- How to get $p(x_1, \ldots, x_N)$ from $p(x_i)$?
- Assume independence!

$$p(x_1, \ldots, x_N) = p(x_1)p(x_2)\ldots$$

Specifying a prior

- How to get $p(x_1, \ldots, x_N)$ from $p(x_i)$?
- Assume independence!

$$p(x_1, \ldots, x_N) = p(x_1)p(x_2)\ldots$$

- Plays nicely with dynamics

$$\frac{dx_i}{dt} \propto \frac{\partial \log p(\mathbf{x})}{\partial x_i} + \cdots = -a + \ldots.$$

Correlated priors

- Independence is nice, but natural odour components are correlated.

Correlated priors

- Independence is nice, but natural odour components are correlated.
- Using that information would improve inference.

Correlated priors

- Independence is nice, but natural odour components are correlated.
- Using that information would improve inference.
- Easy to incorporate into the prior, e.g.

$$p(\mathbf{x}) \propto e^{-C_{ij}x_i x_j}.$$

Correlated priors

- Independence is nice, but natural odour components are correlated.
- Using that information would improve inference.
- Easy to incorporate into the prior, e.g.

$$p(\mathbf{x}) \propto e^{-C_{ij}x_i x_j}.$$

- Introduces interactions into the dynamics

$$\frac{dx_i}{dt} \propto \frac{\partial \log p(\mathbf{x})}{\partial x_i} + \cdots = -a - C_{ij}x_j + \ldots.$$

Evidence from receptor activations

- Assume receptors are linearly excited by odours:

$$r_i = \sum_j A_{ij} x_j + \text{noise}.$$

Evidence from receptor activations

- Assume receptors are linearly excited by odours:

$$r_i = \sum_j A_{ij} x_j + \text{noise.}$$

- A_{ij} is the **affinity** of receptor i for molecule j.

Evidence from receptor activations

- Assume receptors are linearly excited by odours:

$$r_i = \sum_j A_{ij} x_j + \text{noise.}$$

- A_{ij} is the **affinity** of receptor i for molecule j.
- Then error when guessing odour $\mathbf{x}$ caused the response r_i, is

$$(r_i - \sum_j A_{ij} x_j)^2$$

Evidence from receptor activations

- Assume receptors are linearly excited by odours:

$$r_i = \sum_j A_{ij} x_j + \text{noise.}$$

- A_{ij} is the **affinity** of receptor i for molecule j.
- Then error when guessing odour $\mathbf{x}$ caused the response r_i, is

$$(r_i - \sum_j A_{ij} x_j)^2$$

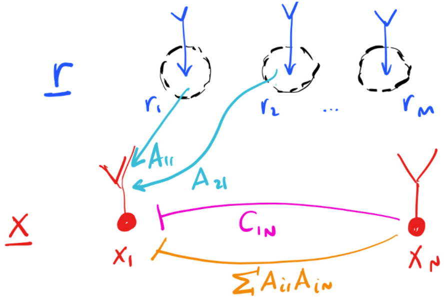
- We sum the evidence from all the receptors

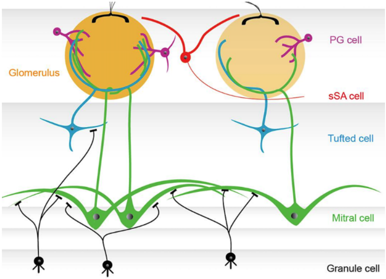
$$\log p(\mathbf{r}|\mathbf{x}) = -\sum_i (r_i - \sum_j A_{ij} x_j)^2.$$

Inference circuit, v. 1

- We have all the pieces we need to specify the dynamics:

$$\frac{dx_i}{dt} \propto \underbrace{-a - \sum_j C_{ij} x_j}_{\partial\text{prior}/\partial x_i} + \underbrace{\sum_k A_{ki} r_k - \sum_j \left[\sum_k A_{ki} A_{kj} \right] x_j}_{\partial\text{evidence}/\partial x_i} .$$



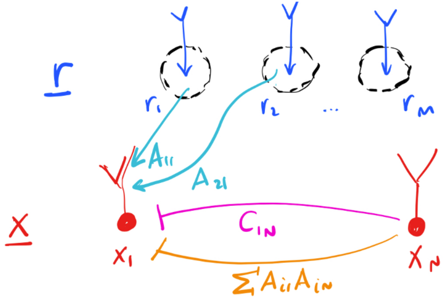


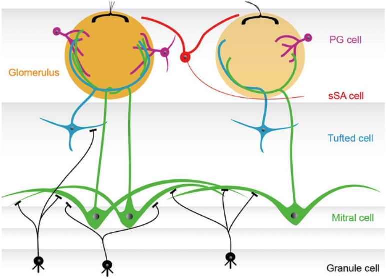
Inference circuit, v. 1

- We have all the pieces we need to specify the dynamics:

$$\frac{dx_i}{dt} \propto \underbrace{-a - \sum_j C_{ij} x_j}_{\partial \text{prior}/\partial x_i} \underbrace{+ \sum_k A_{ki} r_k - \sum_j \left[\sum_k A_{ki} A_{kj} \right] x_j}_{\partial \text{evidence}/\partial x_i} .$$

- But it doesn't look like the bulb...





Inference circuit, v. 2

- Drop the cross term in the prior

$$\frac{dx_i}{dt} = -a + \cancel{\sum_j C_{ij} x_j} + \sum_k A_{ki}(r_k - \sum_j A_{kj} x_j)$$

Inference circuit, v. 2

- Drop the cross term in the prior

$$\frac{dx_i}{dt} = -a + \cancel{\sum_j C_{ij} x_j} + \sum_k A_{ki}(r_k - \sum_j A_{kj} x_j)$$

- Add a population of 'mitral cells' to encode glomerular residuals:

$$\frac{dm_k}{dt} = -m_k + r_k - \sum_j A_{kj} x_j.$$

- Drop the cross term in the prior

$$\frac{dx_i}{dt} = -a + \cancel{\sum_j C_{ij} x_j} + \sum_k A_{ki}(r_k - \sum_j A_{kj} x_j)$$

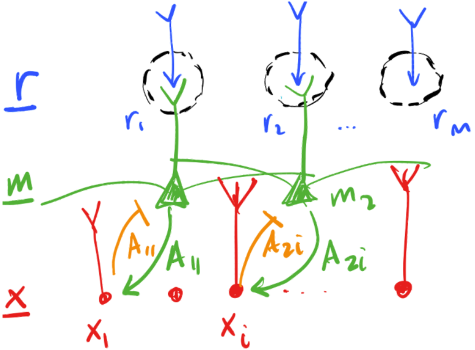
- Add a population of 'mitral cells' to encode glomerular residuals:

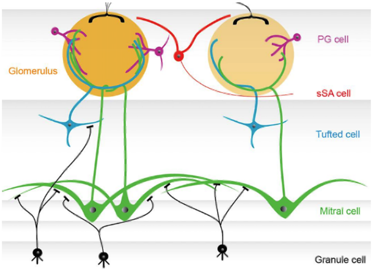
$$\frac{dm_k}{dt} = -m_k + r_k - \sum_j A_{kj} x_j.$$

- Granule cells now only interact with mitral cells:

$$\frac{dx_j}{dt} = -a + \sum_k A_{kj} m_k.$$

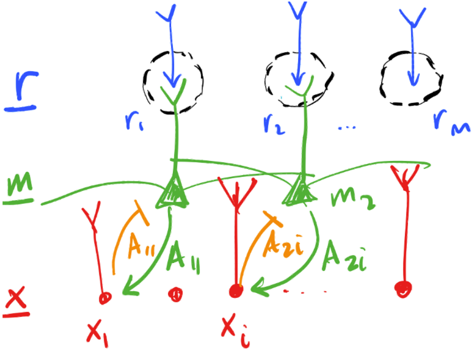
Mission accomplished?

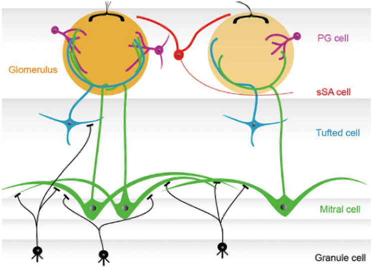




- Looks like the bulb!

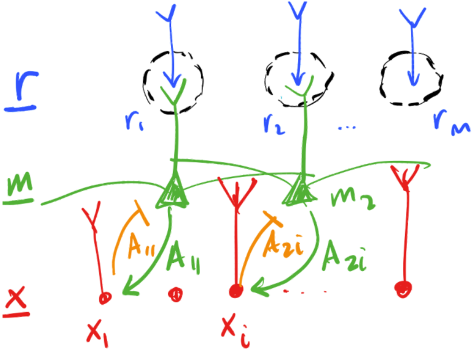
Mission accomplished?

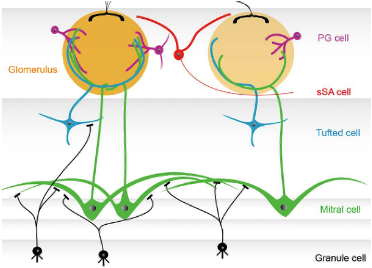




- Looks like the bulb!
- MC-to-GC weights A_{ij} store affinities.

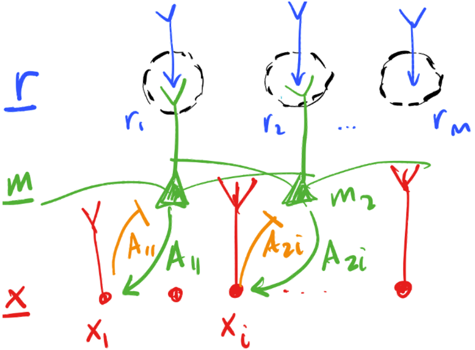
Mission accomplished?

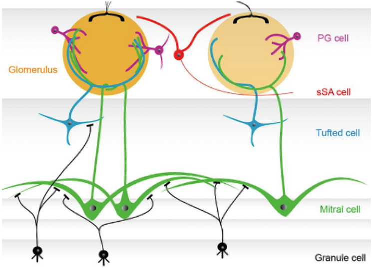




- Looks like the bulb!
- MC-to-GC weights A_{ij} store affinities.
- But: using an independent prior

Mission accomplished?



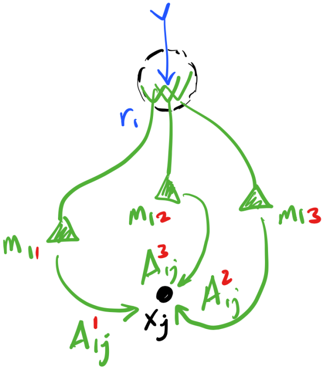


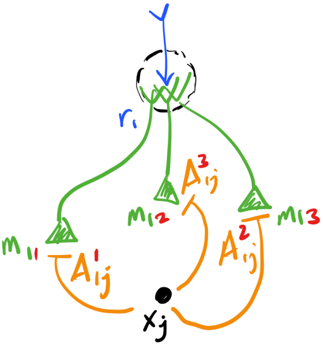
- Looks like the bulb!
- MC-to-GC weights A_{ij} store affinities.
- But: using an independent prior
- No sister cells. . .

To include sister cells, just add 's' to the dynamics

$$\frac{dm_{is}}{dt} = -m_{is} + r_i - \sum_j A^s_{ij} x_j.$$

$$\frac{dx_j}{dt} = -a + \sum_{i,s} A^s_{ij} m_{is}.$$





The objective of the dynamics

- We *derived* the old dynamics to maximize the posterior.

$$L(\mathbf{x}) = \underbrace{-\sum_j a x_j}_{\text{log prior}} \underbrace{-\sum_i (r_i - A_{ij} x_j)^2}_{\text{log evidence}} .$$

The objective of the dynamics

- We *derived* the old dynamics to maximize the posterior.

$$L(\mathbf{x}) = \underbrace{-\sum_j a x_j}_{\text{log prior}} \underbrace{-\sum_i (r_i - A_{ij} x_j)^2}_{\text{log evidence}}.$$

- We derived the sister dynamics heuristically - no guarantee that they maximize anything.

The objective of the dynamics

- We *derived* the old dynamics to maximize the posterior.

$$L(\mathbf{x}) = \underbrace{-\sum_j a x_j}_{\text{log prior}} \underbrace{-\sum_i (r_i - A_{ij} x_j)^2}_{\text{log evidence}} .$$

- We derived the sister dynamics heuristically - no guarantee that they maximize anything.
- But in fact, they maximize

$$L(\mathbf{x}) = \underbrace{-\sum_j a x_j}_{\text{log prior}} \underbrace{-\sum_{i,s} (r_i - A^{s}_{ij} x_j)^2}_{\text{log evidence?}} .$$

The objective of the dynamics

- We *derived* the old dynamics to maximize the posterior.

$$L(\mathbf{x}) = \underbrace{-\sum_j a x_j}_{\text{log prior}} \underbrace{-\sum_i (r_i - A_{ij} x_j)^2}_{\text{log evidence}} .$$

- We derived the sister dynamics heuristically - no guarantee that they maximize anything.
- But in fact, they maximize

$$L(\mathbf{x}) = \underbrace{-\sum_j a x_j}_{\text{log prior}} \underbrace{-\sum_{i,s} (r_i - A^{s}_{ij} x_j)^2}_{\text{log evidence?}} .$$

- How to interpret this?

Intrepreting the new objective

- The original objective maximized a log posterior:

$$\log p(\mathbf{x}|\mathbf{r}) = \underbrace{-\sum_j a x_j}_{\text{log prior}} \underbrace{-\sum_i (r_i - A_{ij} x_j)^2}_{\text{log evidence}} .$$

Intrepreting the new objective

- The original objective maximized a log posterior:

$$\log p(\mathbf{x}|\mathbf{r}) = \underbrace{-\sum_j a x_j}_{\text{log prior}} \underbrace{-\sum_i (r_i - A_{ij}x_j)^2}_{\text{log evidence}}.$$

- To interpret the new objective...

$$L(\mathbf{x}) = -\sum_j a x_j - \sum_{i,s} (r_i - A^s_{ij} x_j)^2.$$

Intrepreting the new objective

- The original objective maximized a log posterior:

$$\log p(\mathbf{x}|\mathbf{r}) = \underbrace{-\sum_j a x_j}_{\text{log prior}} \underbrace{-\sum_i (r_i - A_{ij} x_j)^2}_{\text{log evidence}} .$$

- To interpret the new objective...

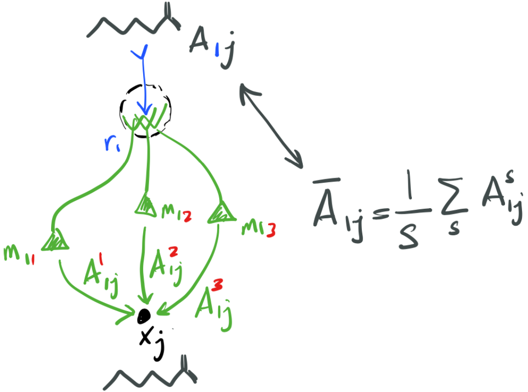
$$L(\mathbf{x}) = -\sum_j a x_j - \sum_{i,s} (r_i - A^s_{ij} x_j)^2 .$$

- ...we express it in the same form as the old

$$L(\mathbf{x}) = \underbrace{-\sum_j a x_j - \sum_{j,k} C_{jk} x_j x_k}_{\text{log prior}} \underbrace{-\sum_i S_i (r_i - \sum_j \overline{A_{ij}} x_j)^2}_{\text{log evidence}} .$$

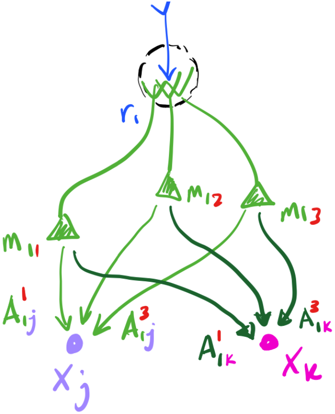
Mean connectivity encodes affinity

$$L(\mathbf{x}) = -\sum_j a x_j - \sum_{j,k} C_{jk} x_j x_k - \sum_i S_i (r_i - \sum_j \overline{A_{ij}} x_j)^2.$$



Covariance of connectivity encodes correlated priors

$$L(\mathbf{x}) = -\sum_j a x_j - \sum_{j,k} C_{jk} x_j x_k - \sum_i S_i (r_i - \sum_j \overline{A_{ij}} x_j)^2.$$



$$C^1_{jk} = \mathrm{Cov}_s(A^s_{1j}, A^s_{1k})$$

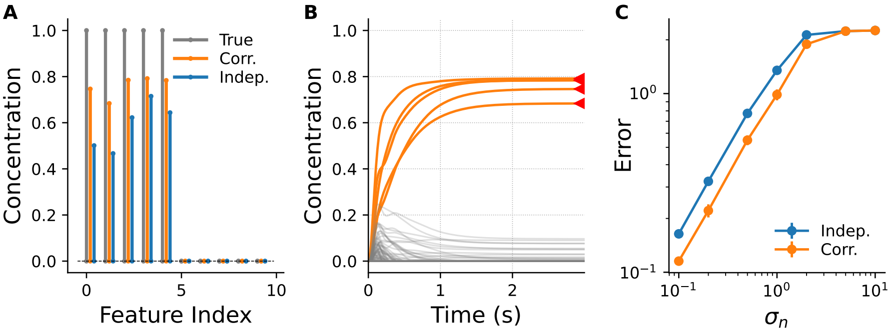
$$C_{jk} = \sum_i S_i C^i_{jk}$$

How many odours can we encode correlations for?

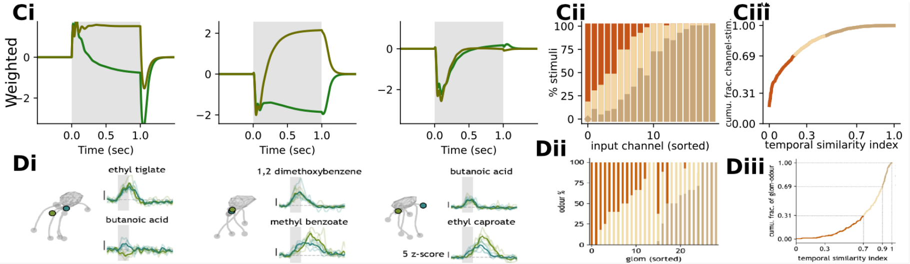
Connectivity solutions exist whenever

$$n \leq \underbrace{S}_{\text{total \# sisters}} - \underbrace{M}_{\text{\# glomeruli}}.$$

Simulations: Using correlated priors helps



Simulations: Comparison to experimental data



Data from Zhang et al. 2025

Conclusions

- Natural odour components are correlated.

Conclusions

- Natural odour components are correlated.
- Optimal inference should use correlated priors.

Conclusions

- Natural odour components are correlated.
- Optimal inference should use correlated priors.
- Naive implementation requires GCs to talk to each other (they don't).

Conclusions

- Natural odour components are correlated.
- Optimal inference should use correlated priors.
- Naive implementation requires GCs to talk to each other (they don't).
- Sisters cells can encode correlated priors in their connectivity.

Conclusions

- Natural odour components are correlated.
- Optimal inference should use correlated priors.
- Naive implementation requires GCs to talk to each other (they don't).
- Sisters cells can encode correlated priors in their connectivity.
- Details: `https://arxiv.org/abs/2505.14579`

Section 2

Thank you Schaefer Lab,
esp. Yuxin and Carles,
and you for your attention!