

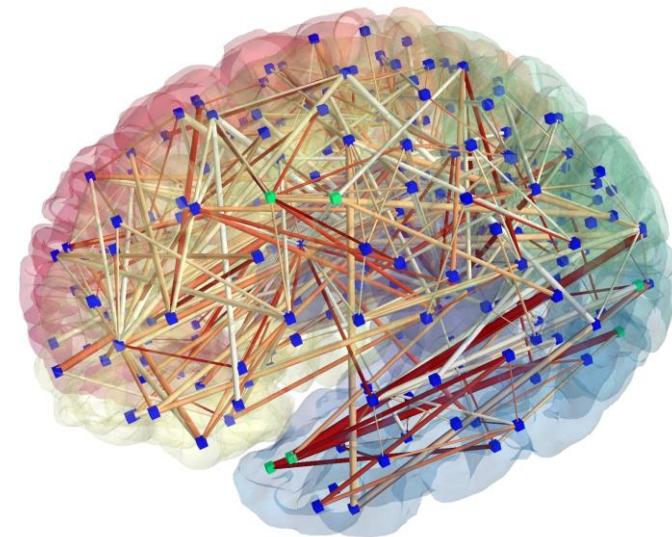
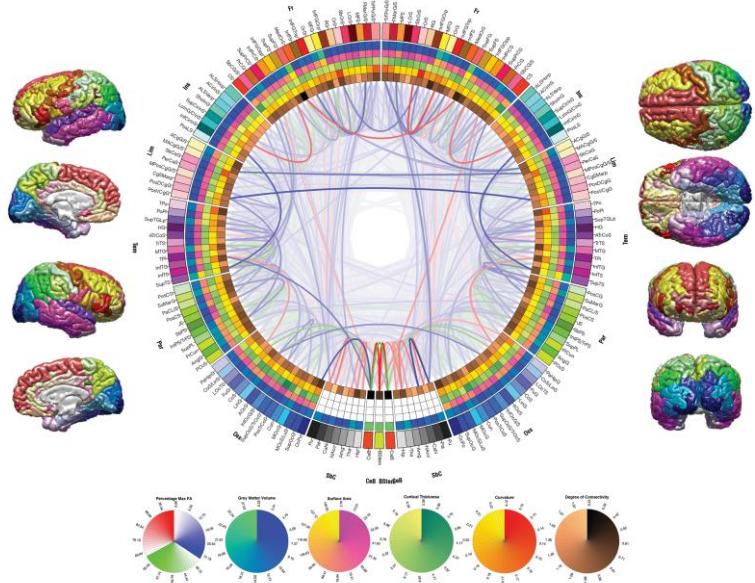
# Bayesian Connectomics

Max Hinne



## The brain as a network

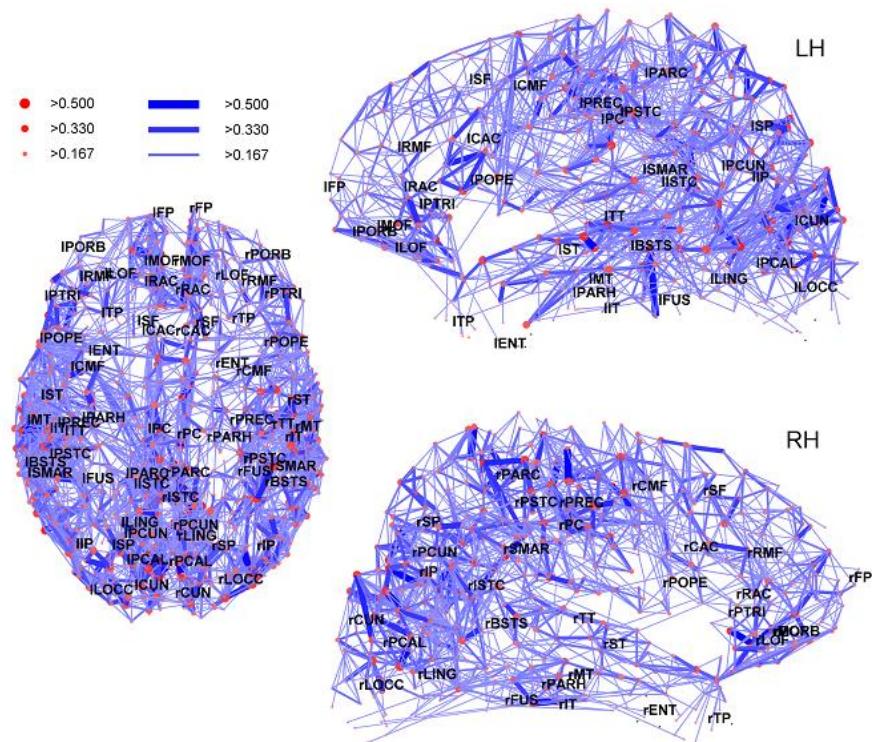
- The human brain shows both functional **segregation** and **integration**.
- Function integration requires coordinated **interaction** between **distributed** neural populations.
- Connectomics refers to the **mapping** and **analysis** of **brain connectivity** at different scales.



# Applications of connectomics

Changes in connectivity have been associated with:

- Aging
  - Alzheimer's Disease
  - ADHD
  - Epilepsy
  - IQ
  - ...





## Flavours of connectomics

Connectivity has a structural, functional and effective component:

- **Structural connectivity:** the white-matter tracts that link neuronal populations
- **Functional connectivity:** the correlations between neuronal time series
- **Effective connectivity:** the causal interactions between neuronal populations

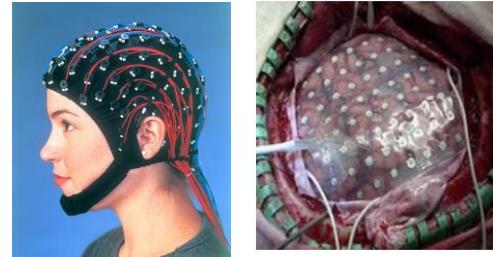
How can we measure connectivity?

# Microscopic & macroscopic approaches

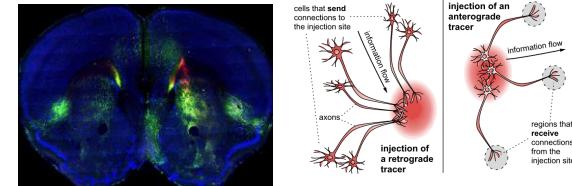
Functional and diffusion MRI



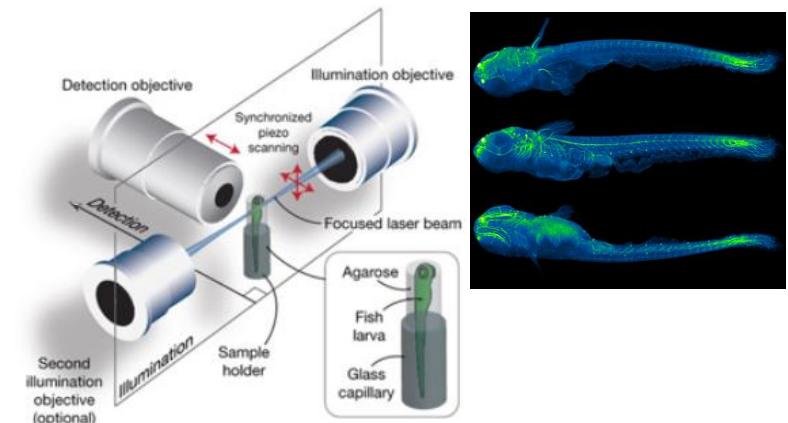
EEG/ECoG



Tracer injections

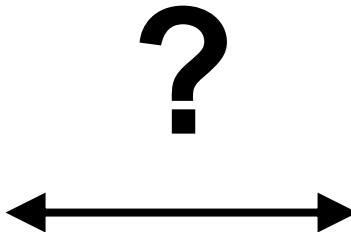
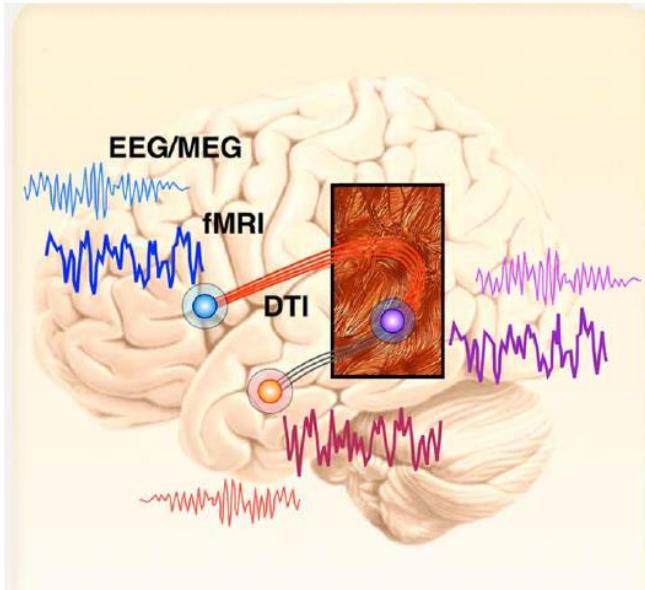


Lightsheet microscopy

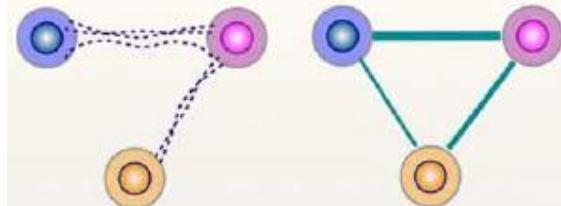




## Connectivity analysis



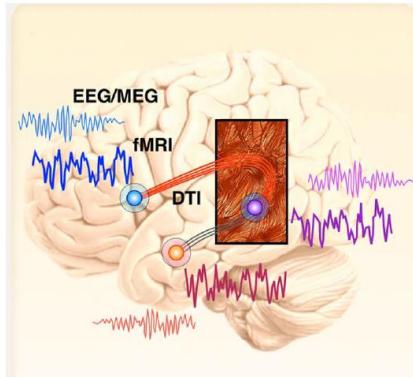
Structural connectivity      Functional connectivity



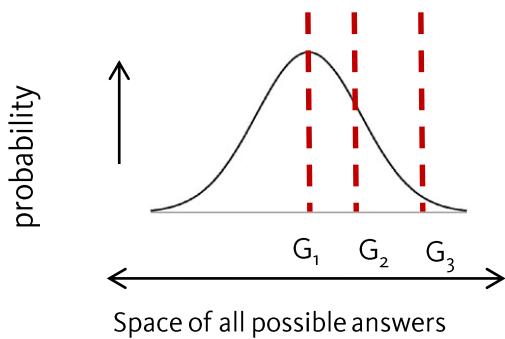
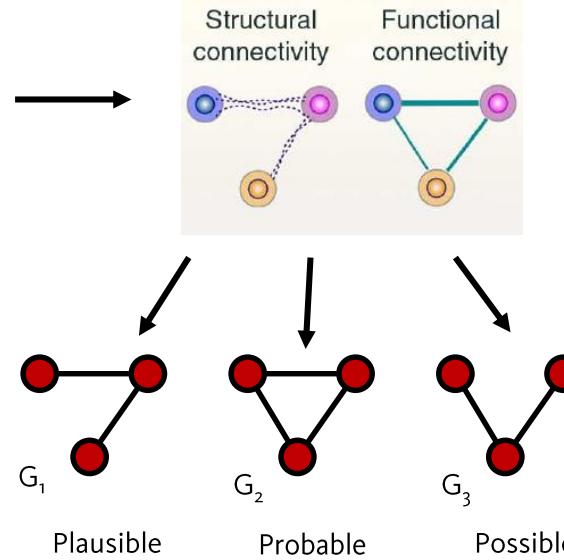
[Park & Friston, Science, 2013]



# Connectivity analysis



[Park & Friston, Science, 2013]



How to express the certainty of our claims?



## Connectivity analysis

- Connectivity analysis is inherently difficult due to the **noisy and incomplete** nature of the observations
- We advocate a Bayesian approach to connectomics: **Bayesian Connectomics** (BaCon for short)
  - Relies on defining **generative models** whose parameters are inferred from data, using various imaging modalities
  - Will be applied in the context of **macroscopic** connectomics:
    - Structural connectivity analysis using diffusion MRI or tract tracing
    - Functional connectivity analysis with functional MRI

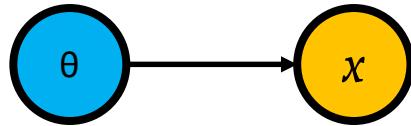
# Generative models

Using Bayes to model data and assumptions



## Generative modeling

- Which **random variables** depend on which other variables?
- Which **probability distribution** reflects this dependency?
- E.g.:

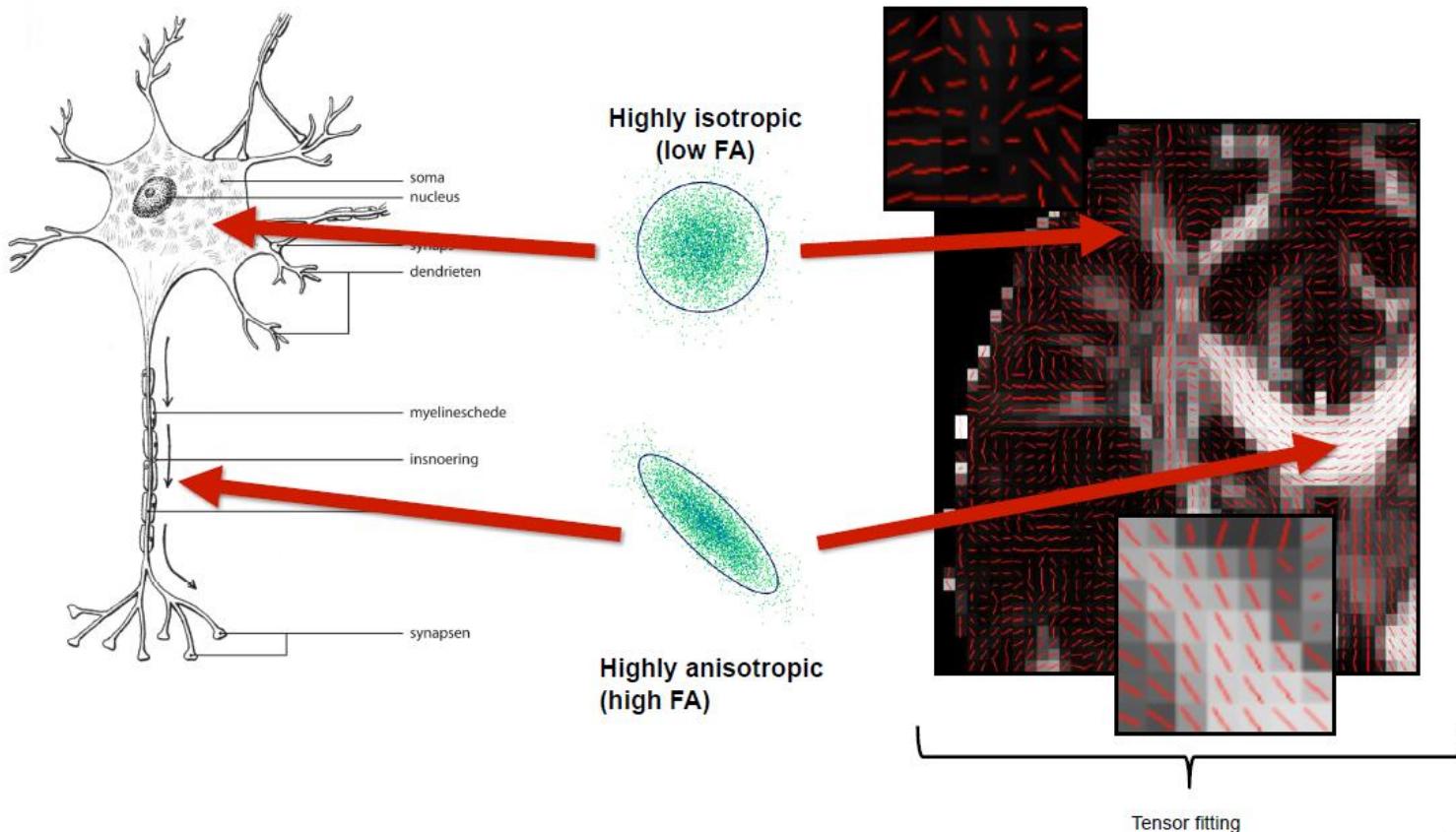


- $P(x|\theta)$  = Bernoulli, or  $P(x|\theta)$  = Gaussian, etc., depending on what the variables represent.
- Use **Bayes' theorem** to compute  $P(\theta|x) = P(x|\theta) P(\theta) / P(x)$  ← Not so easy in practice
- Modeling choices:
  - What is the **likelihood**  $P(x|\theta)$ , i.e. if we know the parameters, what process generates the data  $x$ ?
  - What is the **prior**  $P(\theta)$ , i.e. what parameters do we expect before we make observations?
- Challenge: compute posterior in reasonable time.

# Bayesian structural connectivity analysis

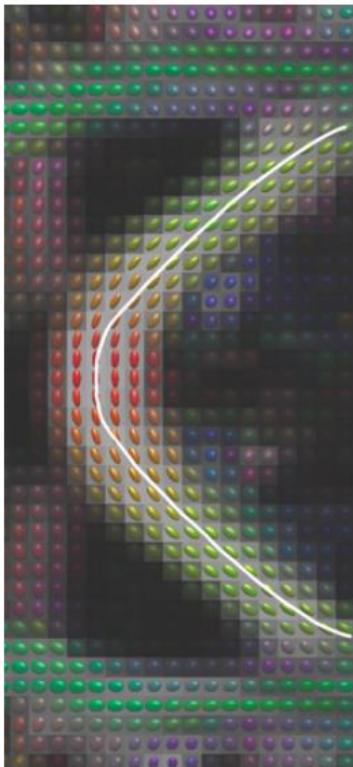
What is the anatomical connectivity between brain regions?

## Diffusion imaging



# Tractography

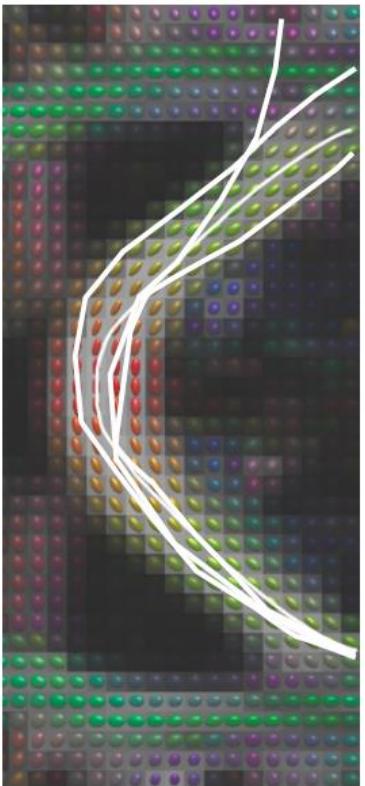
deterministic streamlining



- for each voxel, draw a streamline in the principal diffusion direction
- this gives a whole-brain structural connectivity estimate

# Tractography

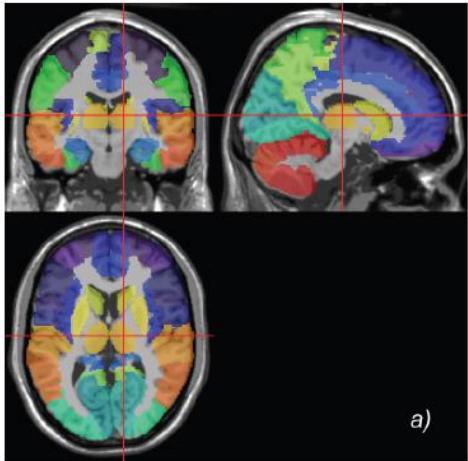
probabilistic streamlining



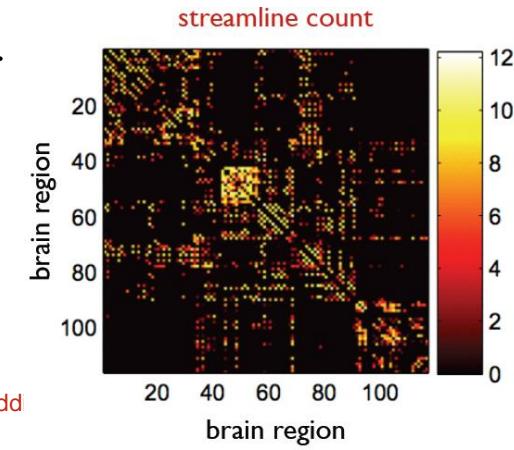
- for each voxel, draw N streamlines according to the local pdf on fiber direction
- this does not give a whole-brain structural connectivity estimate



## Streamlines between brain regions

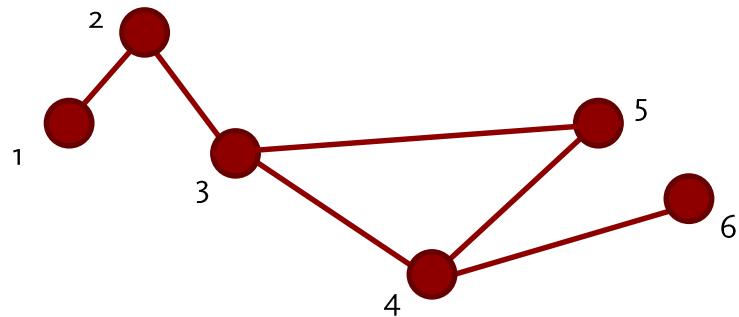


- Each node is **the center of mass** of one region.
- Streamlines connection region A to region B are **aggregated**.
- Intra-region streamlines are discarded.



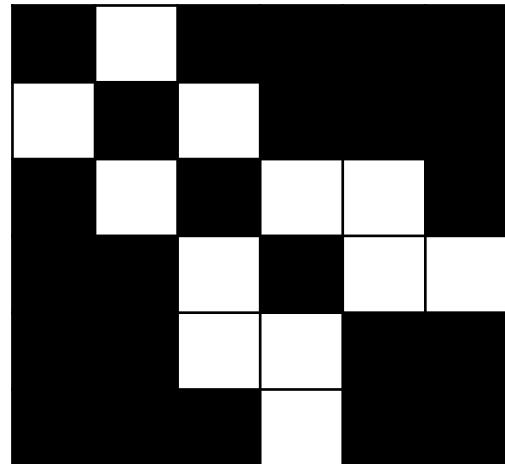
## Representing a brain network

- A graph is a mathematical representation of sets of objects ('nodes' or 'vertices') and their relations ('edges', 'connections' or 'links').
- Formally, we write:  $G = (V, E)$ , with  $E \subseteq V \times V$

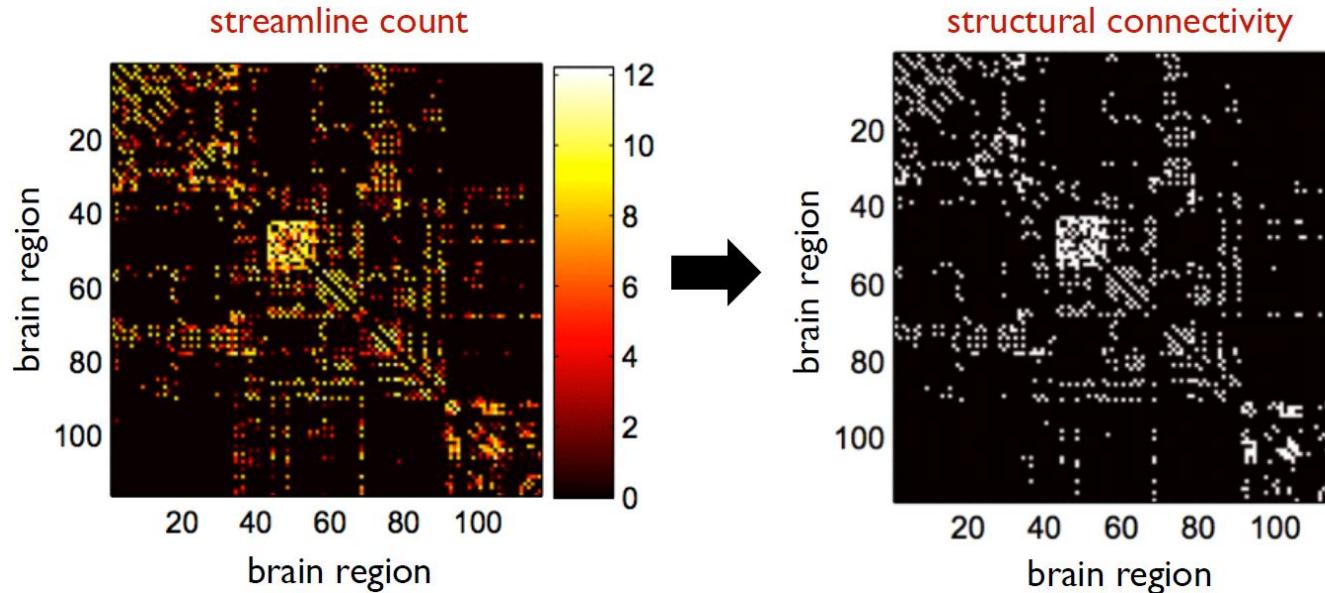


- Vertices =  $\{1, 2, 3, 4, 5, 6\}$
- Edges =  $\{\{1,2\}, \{2,3\}, \{3,4\}, \{3,5\}, \{4,5\}, \{4,6\}\}$

- Equivalent representation: adjacency matrix  $G_{ij} = 1 \leftrightarrow (i, j) \in E$



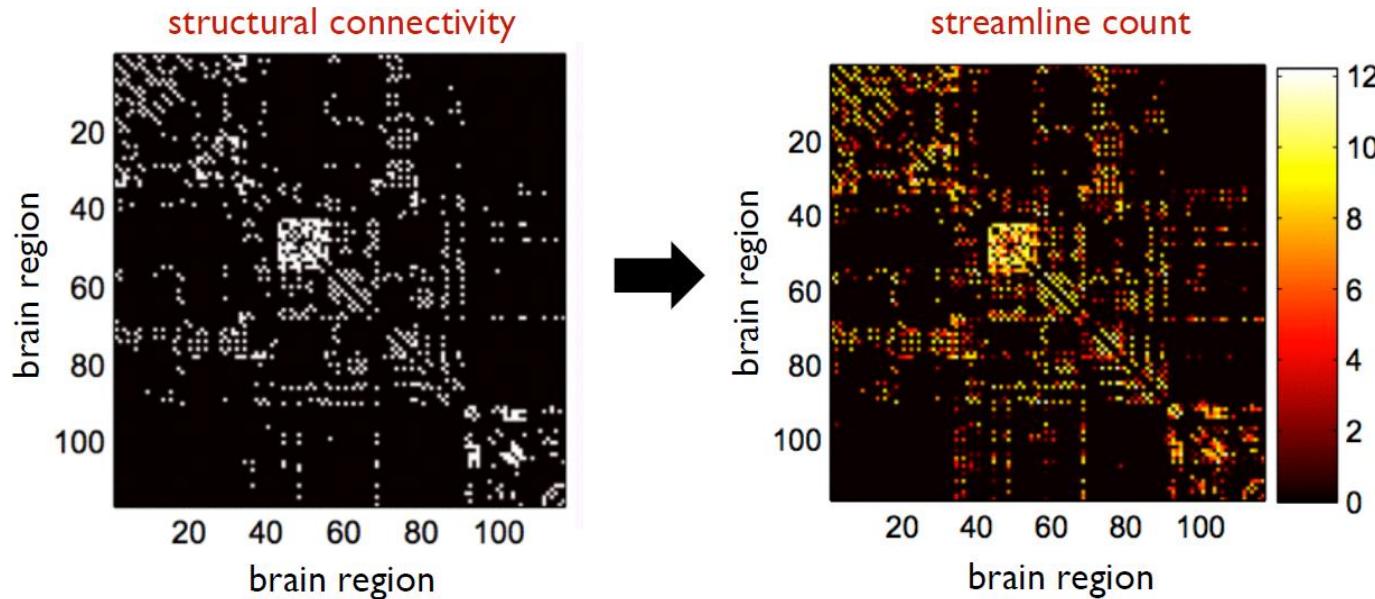
## Goal: estimate (structural) connectivity from streamline counts



'Naive' approach: **threshold** count matrix → high counts become an edge.

- What threshold to choose?
- (Normalized) streamline counts are not edge probabilities.
  - They are conditioned on the (streamline) starting node.

## Bayesian structural connectivity estimation



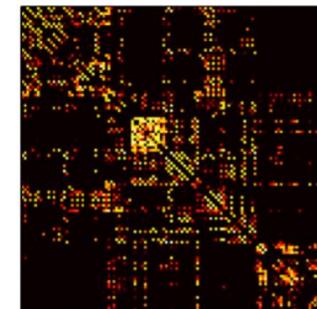
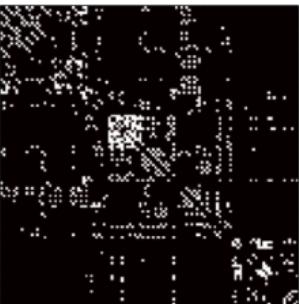
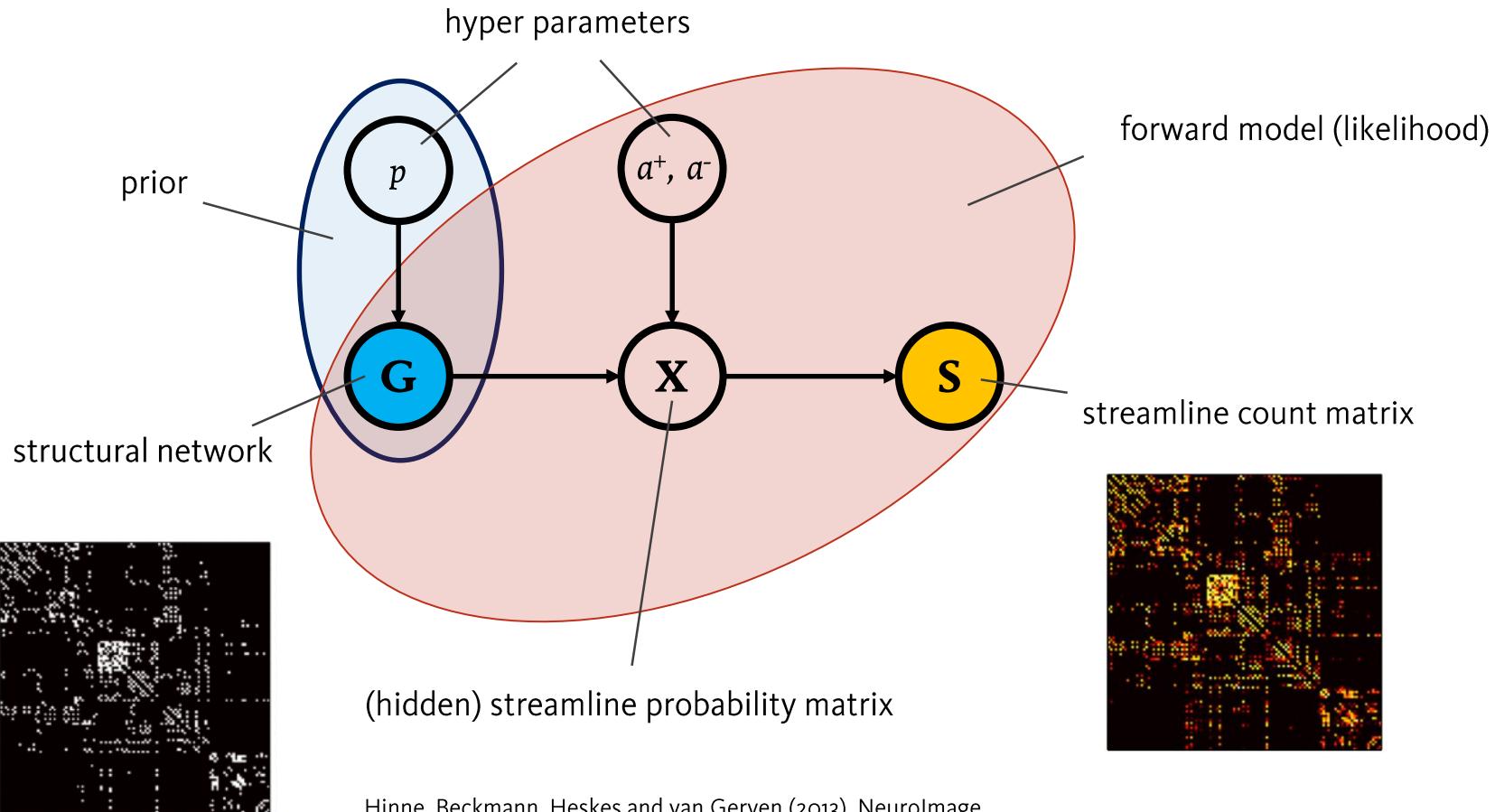
1. Model how a network **generates** the observed counts.
2. Use **Bayes' rule** to learn the network, given the observations.



## Bayesian structural connectivity estimation

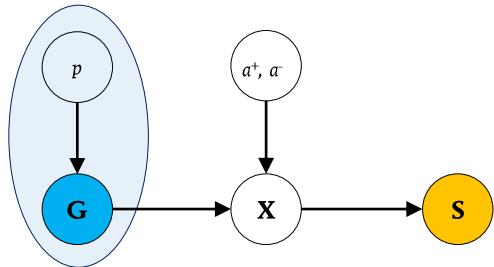
- We want to take into account the **uncertainty** in probabilistic streamlining.
- We'd rather avoid arbitrary thresholds, which we don't know how to pick.
- We want to incorporate **prior knowledge**:
  - Graph-theoretical properties (e.g. degree distribution, small-worldness),
  - Anatomical constraints (i.e. minimize total edge length),
  - Known connections from other studies (i.e., edges found in tracer studies),
  - Data from other subjects/modalities.
- A **Bayesian approach** gives all of this (and more).

## Generative model for structural connectivity



Hinne, Beckmann, Heskes and van Gerven (2013), NeuroImage.

## Prior on connectivity



$$\text{Bernoulli}(x|p) = p^x(1-p)^{1-x}$$

- A priori we assume **independent edges**.
- Each edge has a **Bernoulli distribution** as prior, i.e.:

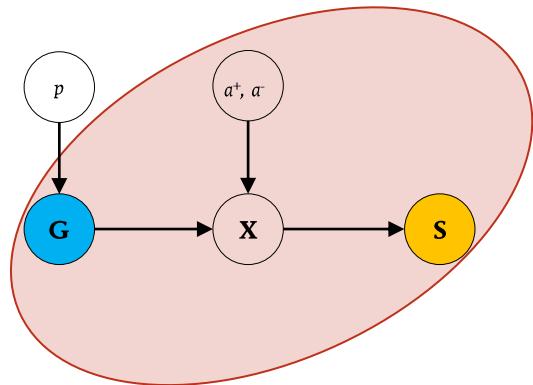
$$\begin{aligned} P(\mathbf{G}|p) &= \prod_{i<j} p(g_{ij}|p) \\ &= \prod_{i<j} p^{g_{ij}}(1-p)^{1-g_{ij}} \end{aligned}$$

- $i < j$  because we assume the graph is **symmetric**.



(weighted) coin flip for each edge

## Forward model: how does a graph ‘generate’ data



- We assume again independence between regions, given  $\mathbf{G}$ , for the streamline count vectors.
- Hence, the likelihood term **factorizes over regions** as:

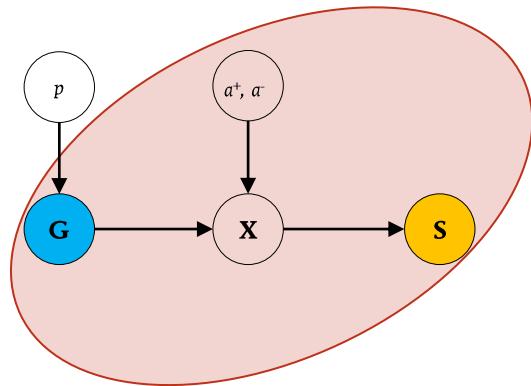
$$P(\mathbf{S}|\mathbf{G}, \alpha^+, \alpha^-) = \prod_i p(\mathbf{s}_i|\mathbf{g}_i, \alpha^+, \alpha^-)$$

where

$$p(\mathbf{s}_i|\mathbf{g}_i, \alpha^+, \alpha^-) = \int p(\mathbf{s}_i|\mathbf{x}_i)p(\mathbf{x}_i|\mathbf{g}_i, \alpha^+, \alpha^-) d\mathbf{x}_i$$

- We **integrate out** the latent streamline probability vectors.
- Can be done analytically if we pick the right distributions for  $p( )$ .

## Forward model: how does a graph ‘generate’ data



- For the streamline probability vectors  $\mathbf{x}_i$  we need a distribution over probability vectors, so that  $\sum_j x_{ij} = 1$  and  $0 \leq x_{ij} \leq 1$ . We assume:

$$p(\mathbf{x}_i | \mathbf{g}_i, a^+, a^-) = \text{Dirichlet}(\mathbf{x}_i | \mathbf{b}_i)$$

with  $\mathbf{b}_i = a^+ \mathbf{g}_i + a^- (1 - \mathbf{g}_i)$ .

This is where the connections and hyperparameters are combined to inform the streamline probabilities.

- Next step: use the probability vector  $\mathbf{x}_i$  to draw streamline counts:

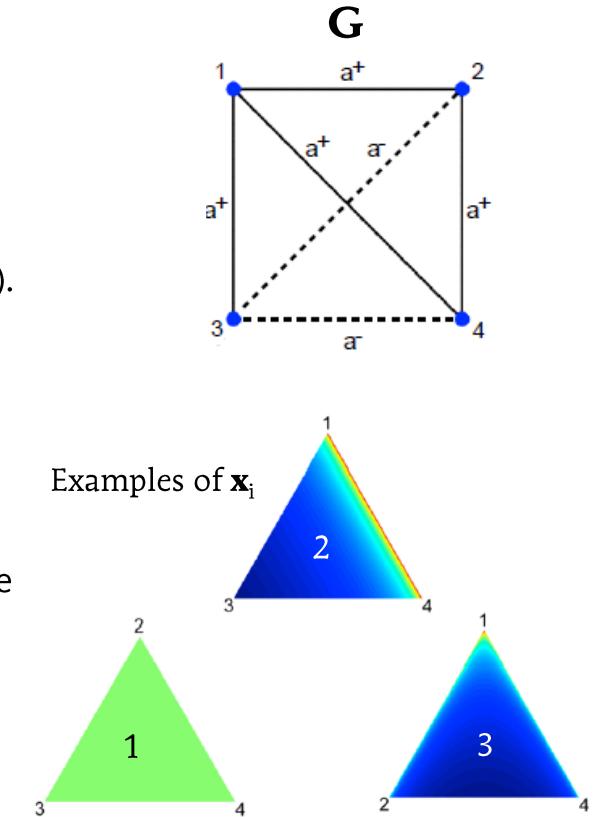
$$p(\mathbf{s}_i | \mathbf{x}_i, S_i) = \text{Multinomial}(\mathbf{n}_i | \mathbf{x}_i, S_i)$$

with  $S_i$  the total number of streamlines to draw for region i.



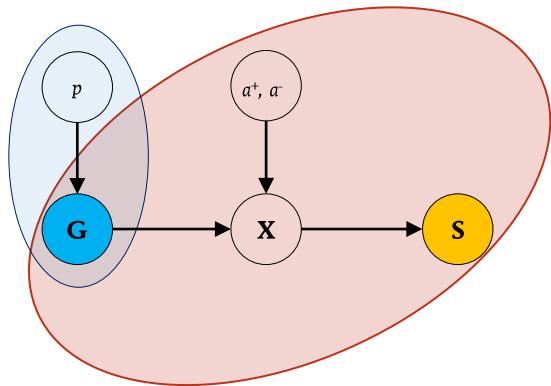
## A note on Bernoulli's, Dirichlet's, Multinomials and more

- Distribution choices seem arbitrary, but they follow from the data type.
  - **Bernoulli** is appropriate for events with a binary outcome, e.g. success or no success: a weighted coinflip, with a weight parameter.
  - **Binomial** is the extension of the Bernoulli that is about the number of successes: how many successes out of N trials (same weight parameter).
  - **Categorical** is another extension of the Bernoulli, which has M outcomes: e.g. a 6-sided dice. Parameter is a loading vector with probabilities for each outcome.
  - **Multinomial** is to Categorical as Binomial is to Bernoulli: how many successes for each face of the dice. It has the same loading vector as the Categorical.
  - **Dirichlet** is a probability distribution over loading vectors: which faces have a *high probability of having a high probability*. Parameter is a vector of concentration parameters.





## The generative model



- Putting it all together:
- For each vector  $i$ , we integrate out the latent variables  $\mathbf{x}_i$ :

$$\begin{aligned}
 p(\mathbf{s}_i | \mathbf{g}_i, a^+, a^-) &= \int p(\mathbf{s}_i | \mathbf{x}_i) p(\mathbf{x}_i | \mathbf{g}_i, a^+, a^-) d\mathbf{x}_i \\
 &= \frac{S_i!}{\prod_j s_{ij}} \frac{\Gamma(\sum_j b_{ij})}{\Gamma(\sum_j (b_{ij} + s_{ij}))} \prod_j \frac{\Gamma(b_{ij} + s_{ij})}{\Gamma(b_{ij})}
 \end{aligned}$$

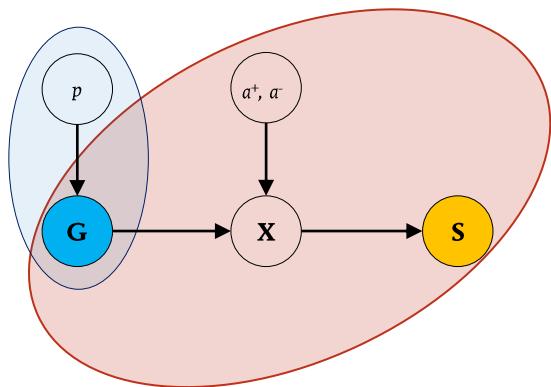
with  $\mathbf{b}_i = a^+ \mathbf{g}_i + a^- (1 - \mathbf{g}_i)$  and  $S_i = \sum_j s_{ij}$ .

- And for the prior we had

$$P(\mathbf{G}|p) = \prod_{i < j} p^{g_{ij}} (1-p)^{1-g_{ij}}$$



## The generative model



- Bayes rule time!

$$P(\mathbf{G}|\mathbf{S}, \theta) = \frac{P(\mathbf{S}|\mathbf{G}, \theta)P(\mathbf{G}|\theta)}{P(\mathbf{S}|\theta)} \mathbf{1}_{\mathbf{G} \in \mathcal{M}}$$

With  $\theta = (a^+, a^-, p)$ , to keep it readable and  $\mathbf{1}_{\mathbf{G} \in \mathcal{M}}$  is 1 if  $\mathbf{G}$  is symmetric and without loops, and 0 otherwise.

- So far so good, but we **cannot actually compute**  $P(\mathbf{G}|\mathbf{S}, \theta)$ ; for  $K$  regions there are  $2^{K(K-1)/2}$  possible networks!
  - 10 regions =  $3.52 \times 10^{13}$  networks
  - This problem is **intractable**, so we need **approximate inference** to explore the networks with reasonable probability.
  - Who cares about  $P(\mathbf{G}|\mathbf{S}, \theta) \approx 0$  anyway!



## Markov chain Monte Carlo

- We use MCMC to collect samples from all possible matrices, that are supported by the data and our prior beliefs.
- Starting from some matrix  $\mathbf{G}^t$ , we propose a new matrix  $\mathbf{G}^*$  by flipping one element. We compute:

$$\alpha = \frac{P(\mathbf{S}|\mathbf{G}^*, \theta)P(\mathbf{G}^*|\theta)}{P(\mathbf{S}|\mathbf{G}^t, \theta)P(\mathbf{G}^t|\theta)}$$

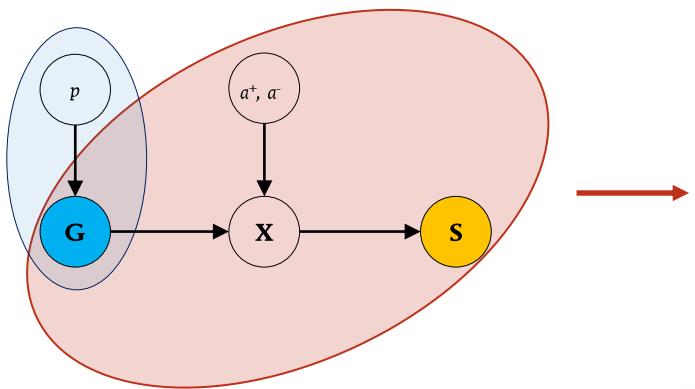
Note that we don't need  $P(\mathbf{S}|\theta)$ ; it is the same for  $\mathbf{G}^t$  and  $\mathbf{G}^*$ .



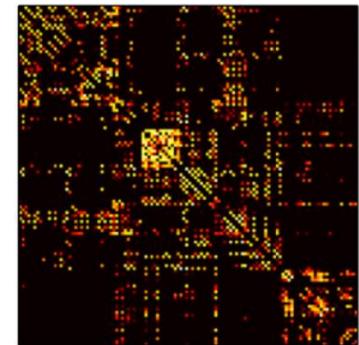
- With probability  $\alpha$  we accept the proposed sample;  $\mathbf{G}^{t+1} = \mathbf{G}^*$ . Otherwise, we remain where we are;  $\mathbf{G}^{t+1} = \mathbf{G}^t$ .
- As  $t \rightarrow \infty$ , the collection of samples goes to  $P(\mathbf{G}|\mathbf{S}, \theta)$ .



# Markov chain Monte Carlo in action

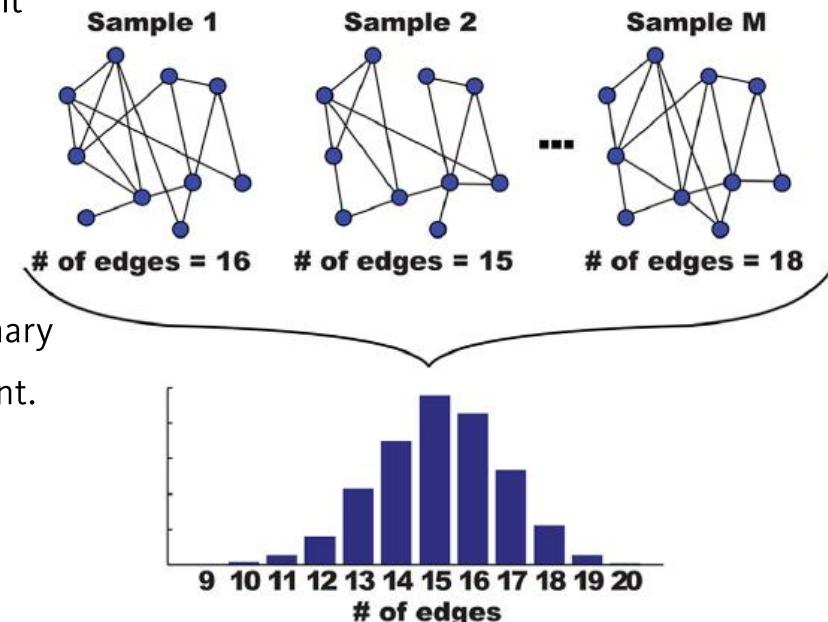
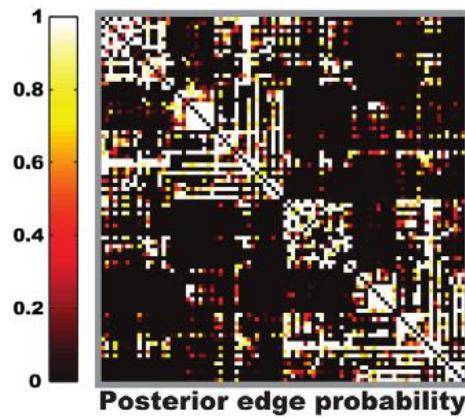


MCMC



## Posterior distributions instead of point estimates

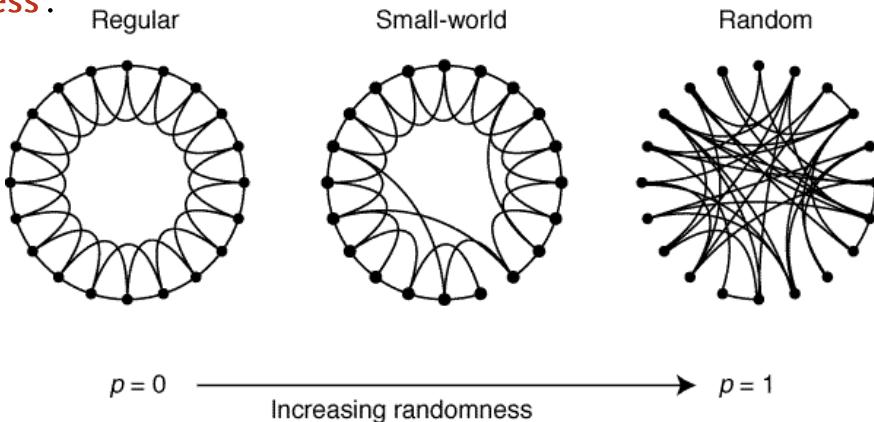
- We have a(n approximate) posterior distribution, not a point estimate.
- For each sample, we can compute any **graph-theoretical** quantity.
  - This implies a **distribution over this quantity** as well.
- The posterior expectation  $\mathbb{E}(\mathbf{G}|\mathbf{S}, \theta) = \frac{1}{T} \sum_{t=1}^T \mathbf{G}^t$  is a summary statistic; it shows the probability for each edge to be present.



Janssen, Hinne, Heskes, van Gerven (2014), Front. Comp. Neuro.

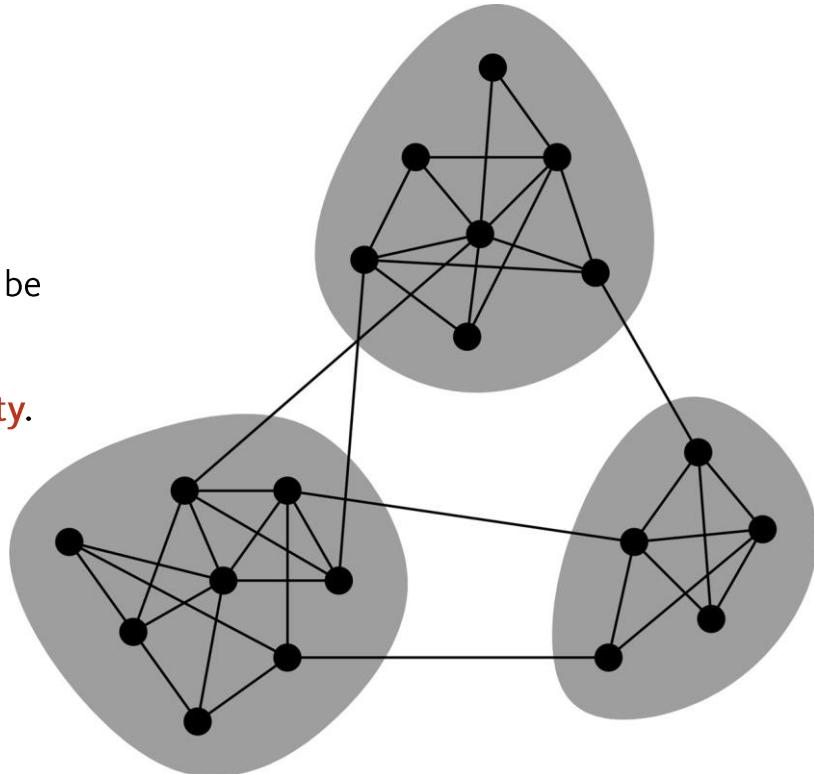
## Network theory: Small-world phenomenon

- Seminal paper by Watts & Strogatz described the network of a small roundworm, *C. Elegans*, and showed that it had
  - **High clustering:**
    - Many connected nodes have the same neighbours.
  - **Short path length**
    - It takes only few ‘hops’ to reach any node in the network.
- These properties together are known as '**small-worldness**'.
- Brain networks are often small-world.



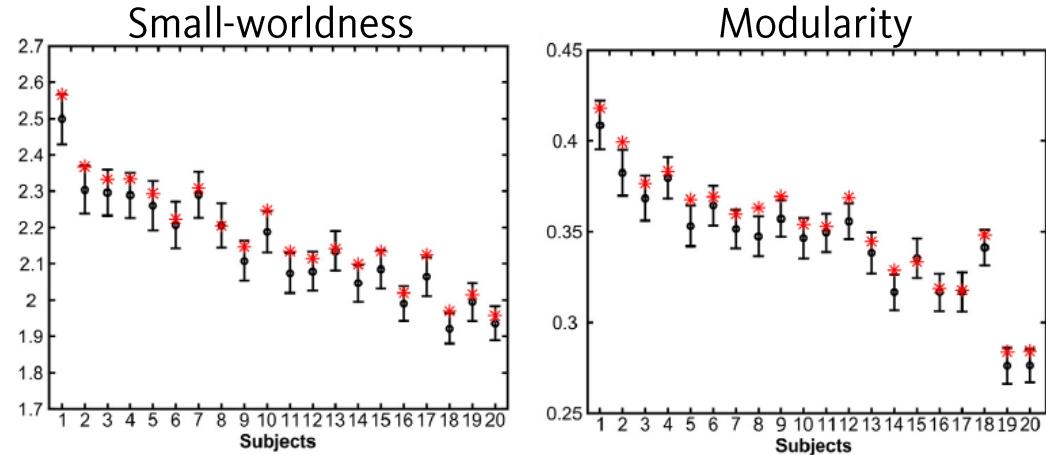
## Network theory: Modularity

- Networks can be subdivided into **clusters** (a.k.a. modules, communities);
- A common way to define a cluster is to say that
  - A cluster has **many internal connections**, but
  - **Few connections to other clusters.**
- **Modularity** is a score that indicates how well a network can be clustered.
- Brain networks are often clustered by **optimizing modularity**.



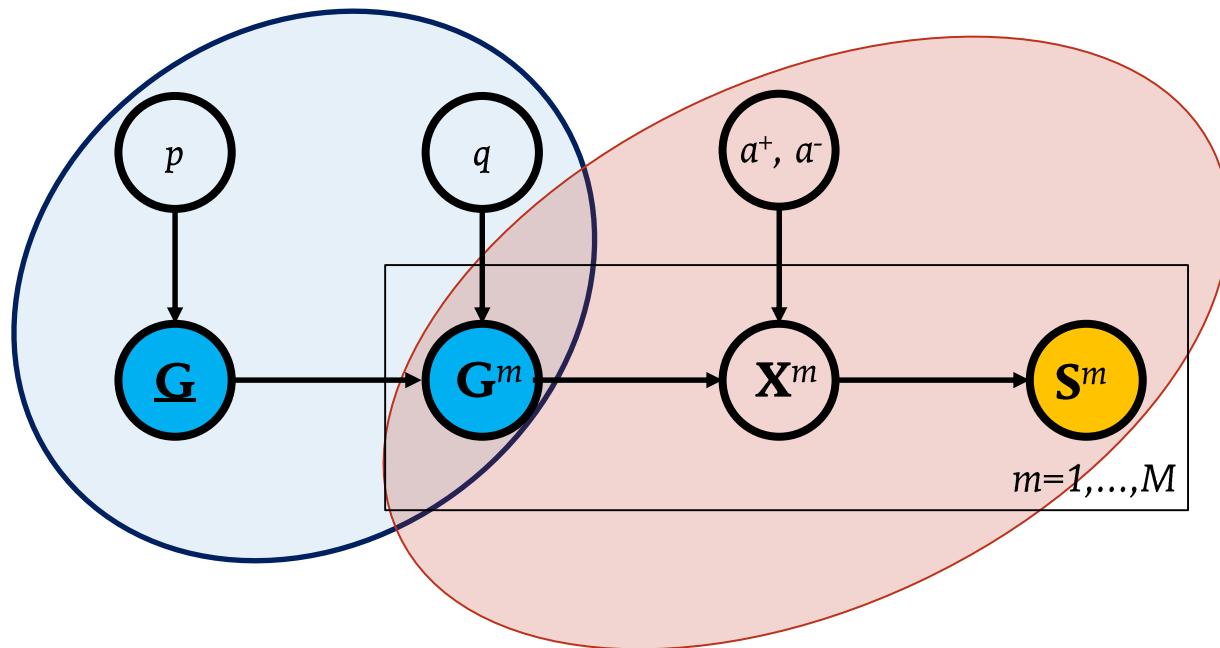
## Posterior distributions instead of point estimates

- Posterior distribution of network measure  $f$ ;
  - Compute  $f$  for each sample.
  - Compare with thresholded network.
    - Threshold set to have network density match the mean density of the samples.



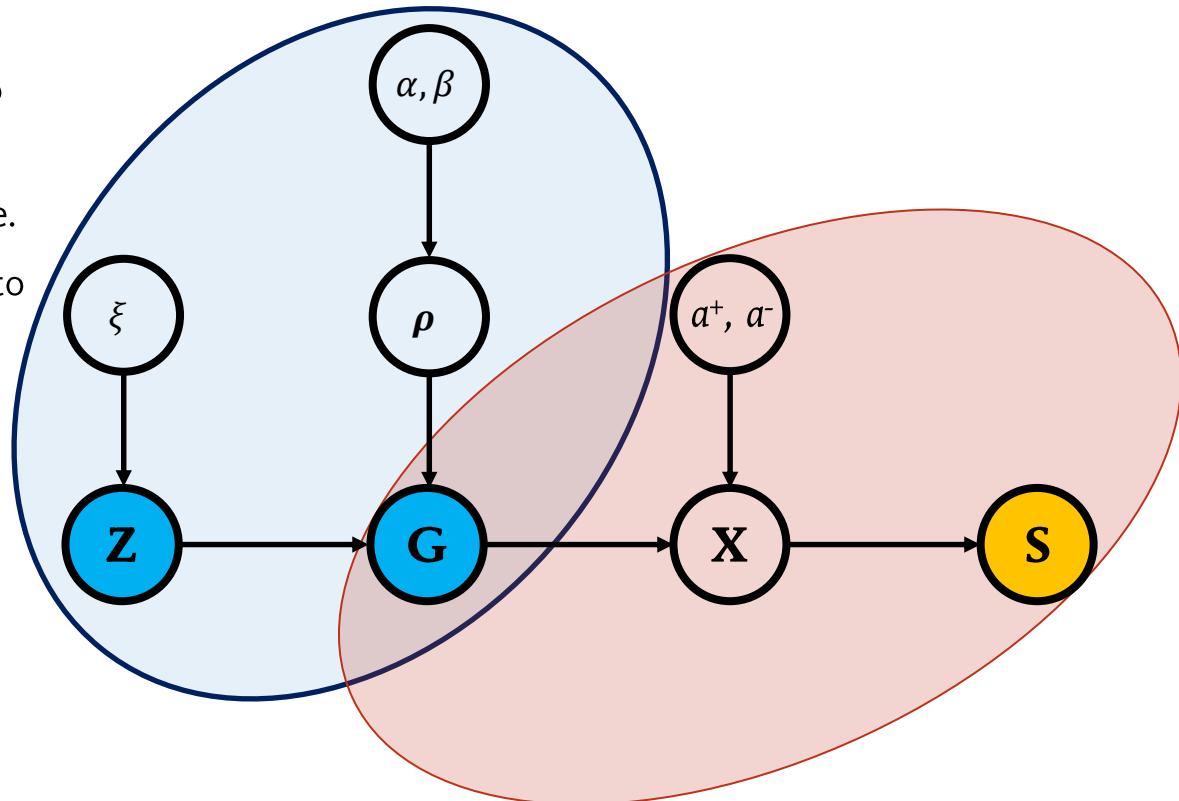
## Model extensions

- Hierarchical model: Assume each subject's brain network  $G^m$  is an instantiation of an ur-network  $\underline{G}$ . Connections are inherited from  $\underline{G}$  with probability  $q$ .
- Consequence:
  - Subjects inform each other.



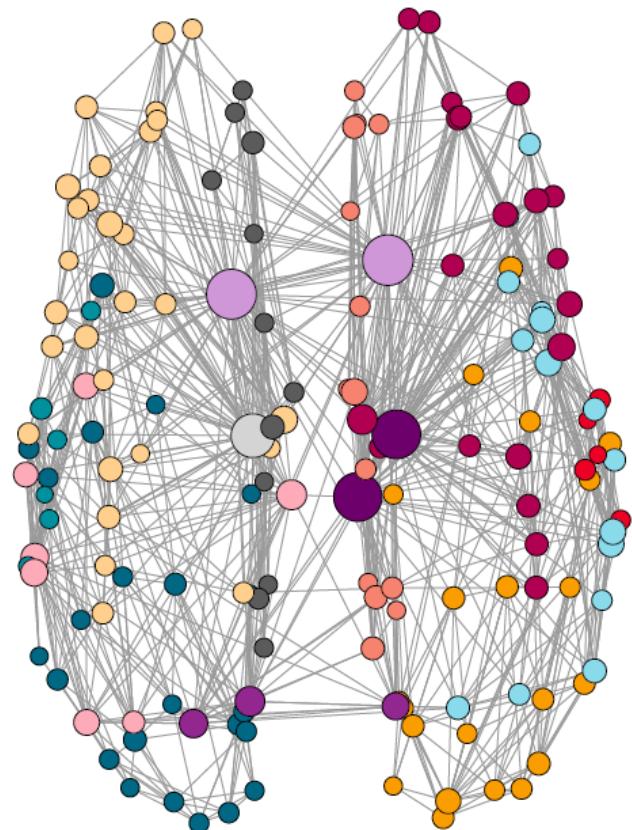
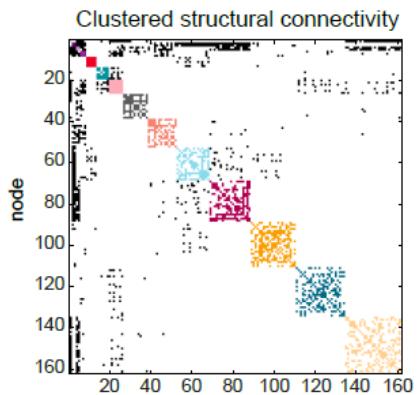
## Model extensions

- Clustering: replace coin-flip prior with a clustering model.
- Connections from cluster a to cluster b all have the same probability.,,
- Nodes within a cluster are exchangeable.
- Learn the latent clusters  $Z$  in addition to connectivity  $G$ , i.e.:  $P(G, Z|\theta)$
- Models can be combined ad infinitum.



## Model extensions

- Clustering: replace coin-flip prior with a clustering model.
- Connections from cluster a to cluster b all have the same probability.,,
- Nodes within a cluster are exchangable.
- Learn the latent clusters  $Z$  in addition to connectivity  $G$ , i.e.:  $P(G, Z|\theta)$
- Models can be combined ad infinitum.

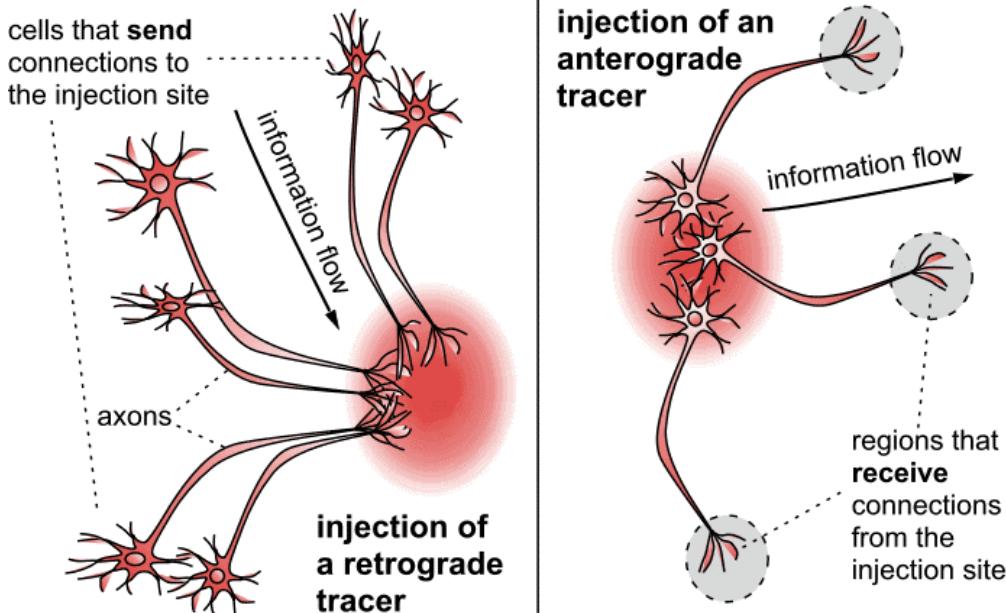


# Predicting structural connectivity

Spatial embedding of network structures

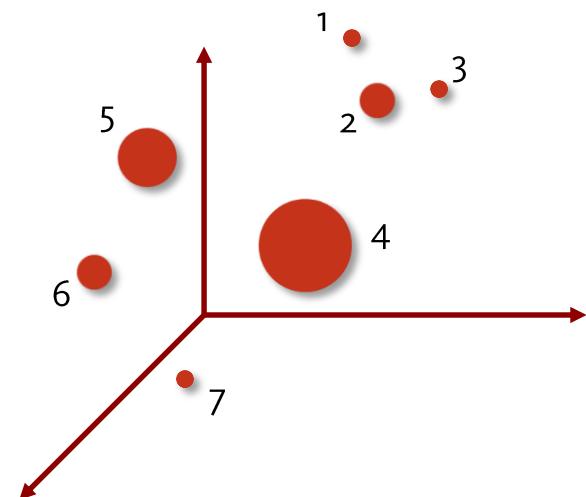
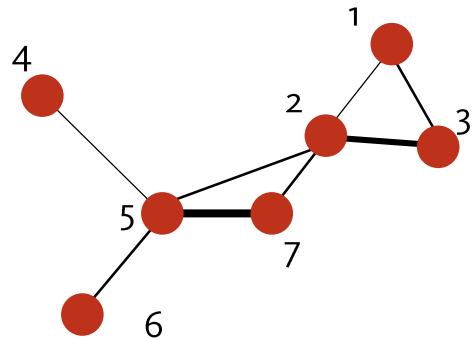
## Different modalities for structural connectivity

- Downsides of diffusion MRI
  - **Indirect measurement.**
  - Lot of noise/uncertainty.
  - Difficulty with long-range connections.
- **Tracer studies** however
  - Very accurate.
  - Have directionality.
- But:
  - Requires animal sacrifice.
  - Labour intensive.

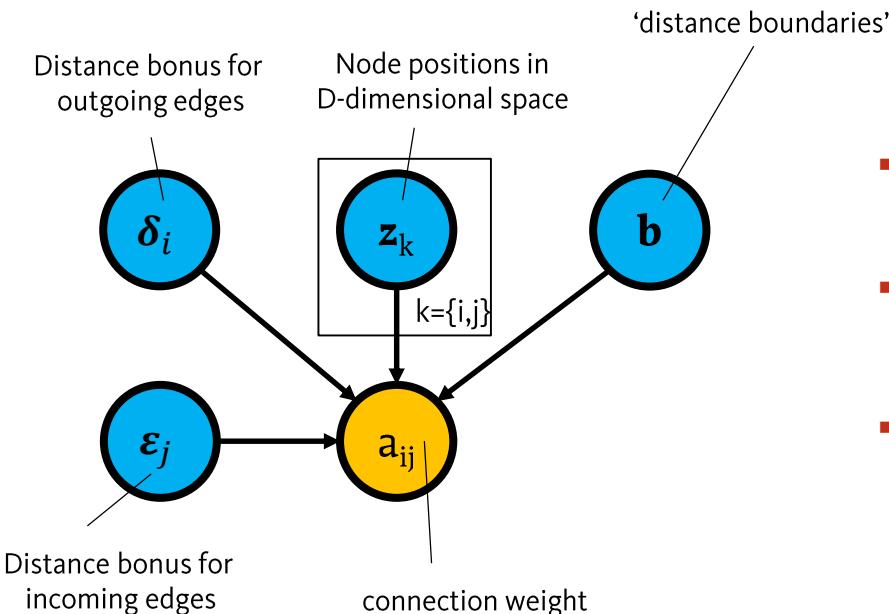


## Link prediction

- Predicting connections avoids new data collection
- Idea:
  - Connections depend on latent factors
  - Learn factors from available data
  - Use factors to predict (absence/presence of) new connections
- Here: **latent space models**

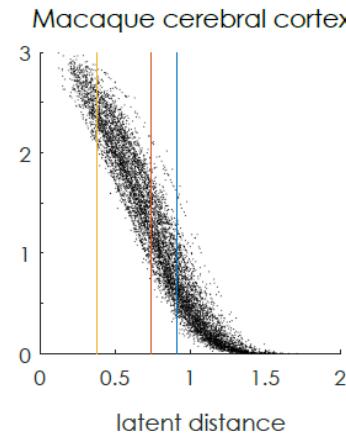
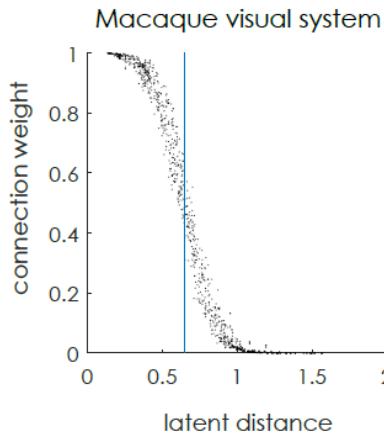


# Ordinal latent space model

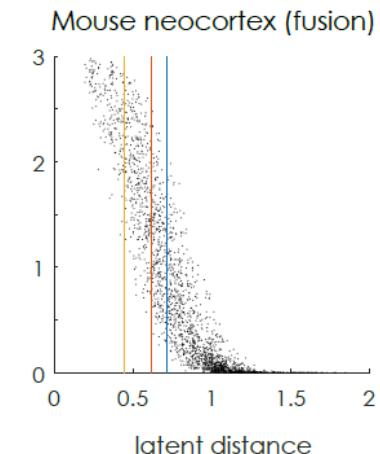
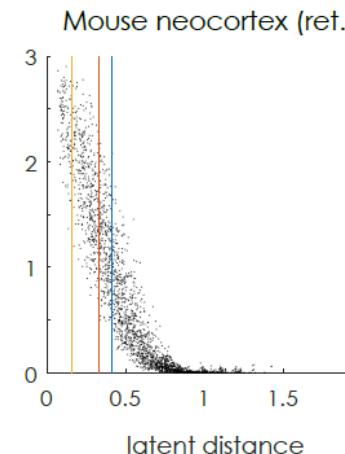
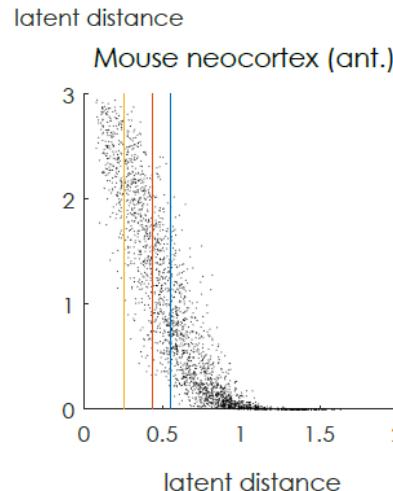


- Data: matrix  $A$  with elements *absent*, *weak*, *moderate* or *strong* (connection weights)
- Weak/uninformative priors not shown.
- Posterior  $P(\mathbf{A}, \mathbf{Z}, \mathbf{b}, \boldsymbol{\delta}, \boldsymbol{\epsilon}, |\theta)$ , with  $\theta$  the hyper parameters not shown here, is approximated using the Stan software package (Hamiltonian Monte Carlo).
- Optimal latent dimensionality  $D$  is determined using cross-validation.

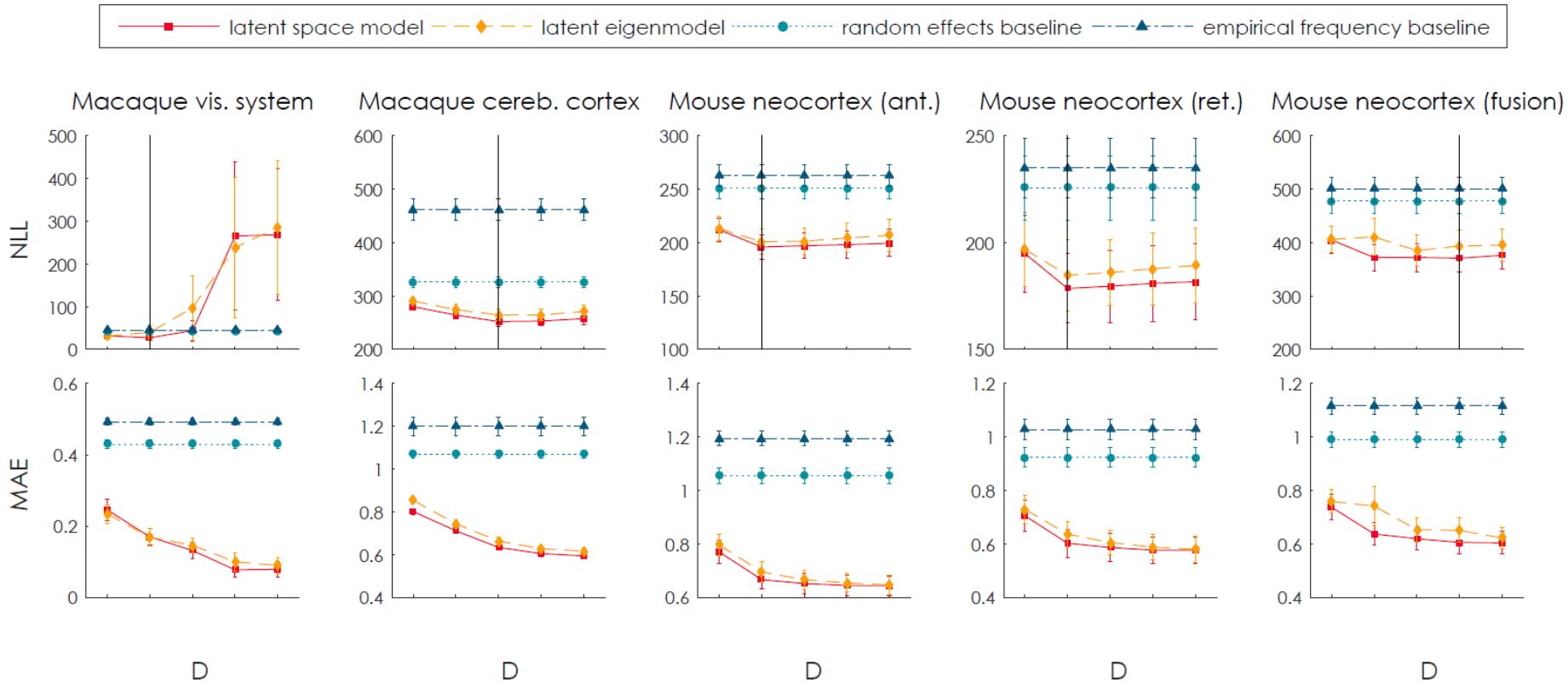
## Learning latent embedding



- Connection weight as function of (latent) distance, for optimal number of dimensions,  $D = (2, 3, 2, 2, 4)$ .
- Vertical variance due to  $\delta$  and  $\varepsilon$ .
- Boundaries  $b$  indicated.

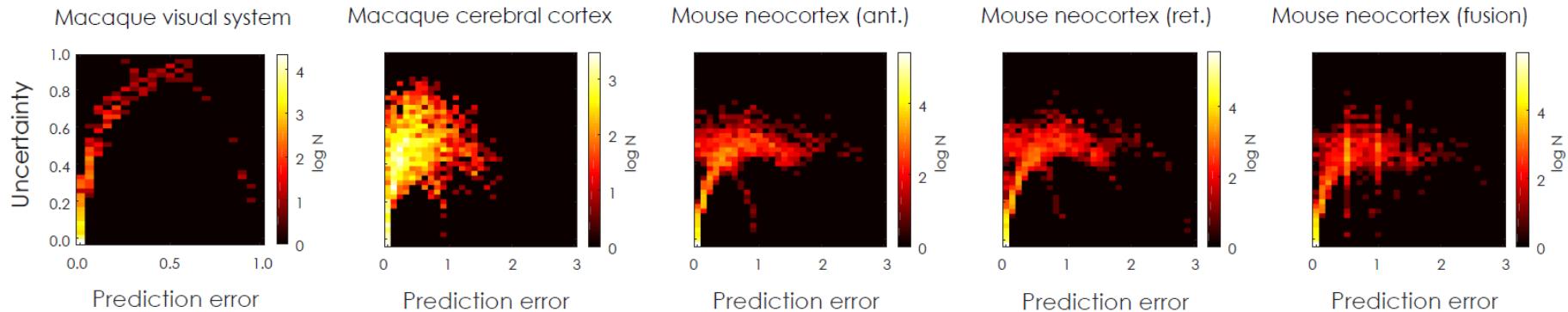


# Experiments on macaque and mouse tracer data



## Experiments on macaque and mouse tracer data

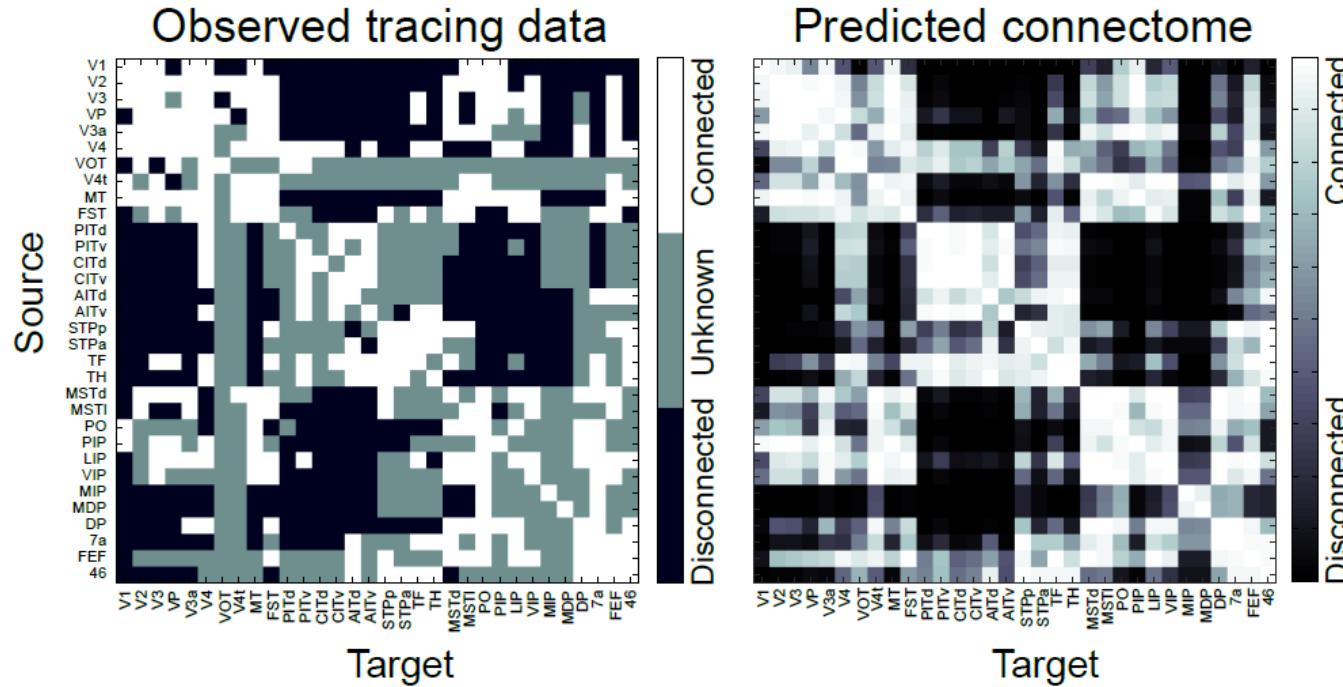
- Uncertainty in prediction reflects accuracy (in cross-validation).
- Low uncertainty goes hand-in-hand with low error.





## Predicted connectome for the macaque visual system

- Using D=2 latent dimensions.

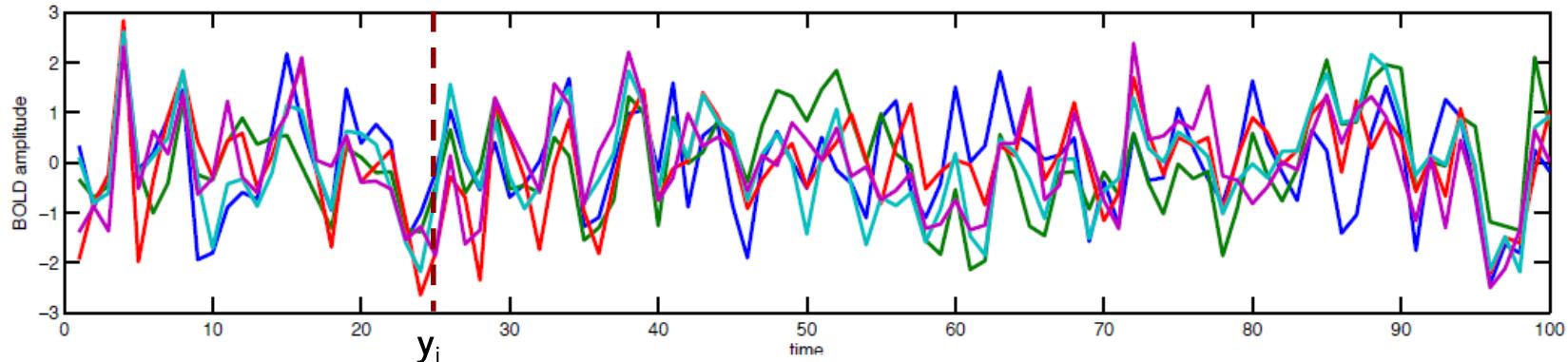


# Bayesian functional connectivity analysis

How do different brain regions communicate?

## Functional connectivity

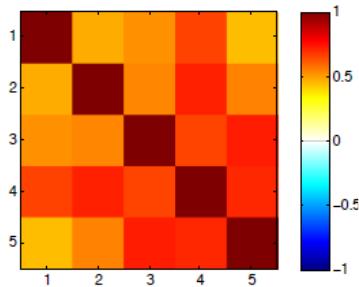
- Functional connectivity refers to **correlated activity** in spatially segregated brain regions.
- Consider the observed **BOLD response time series** in multiple brain regions:



- Let the  $p$ -dimensional vector  $\mathbf{y}_t$  denote observed BOLD response for  $p$  brain regions at time  $i$ .
- Let  $\mathbf{Y} \in \mathbb{R}^{p \times n}$  be the full data set, with  $n$  observations (time points) for  $p$  regions.

## Functional connectivity: covariance

- The easiest way to express functional connectivity is by computing the **covariance** between the time series of pairs of regions.
  - High covariance = strong connectivity.
- Covariances collected in a matrix give the covariance matrix  $\Sigma$ :

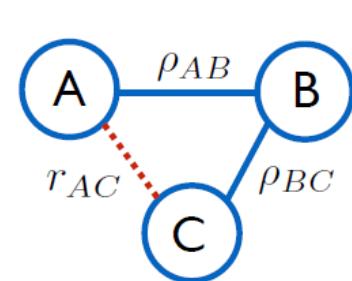


- If the data is standardized (0-mean and unit variance), the covariance matrix is equal to the **correlation** matrix.



## Functional connectivity: precision & partial correlation

- Computing/ the (empirical) covariance is easy;  $\widehat{\Sigma} = \frac{1}{n} \mathbf{Y}^T \mathbf{Y}$ .
- But covariance has an **inherent problem**; it captures **indirect** connectivity too!
  - If A-B and B-C, then also A-C.
  - We cannot distinguish direct from indirect connections by covariance only.
- Alternative: use the **inverse covariance** (precision) matrix  $\mathbf{K} = \Sigma^{-1}$ .
- Precision captures only (sort of) **direct connections**.
  - **Partial correlations** can be computed from the precision:  $\rho_{ij} = -\frac{k_{ij}}{\sqrt{k_{ii}k_{jj}}}$
- The precision matrix has an interesting property:
  - **Zero elements** of the precision matrix reflect that the corresponding brain regions are **conditionally independent**, given the other regions.





## Functional connectivity: sparse partial correlation matrix

- Zeros in the partial correlation matrix tell us which regions are **conditionally independent**.
- We obtain the precision matrix by computing the **maximum likelihood estimate** (MLE):

$$\hat{\mathbf{K}} = \arg \max_{\mathbf{K} \in \mathcal{P}^+} [\log|\mathbf{K}| - \langle \hat{\Sigma}, \mathbf{K} \rangle] = \hat{\Sigma}^{-1}$$

- In practice, random fluctuations due to finite data will make par. corr. elements **non-zero**.
- With a graphical-lasso **prior** on the magnitude of the elements of  $\mathbf{K}$  we compute instead:

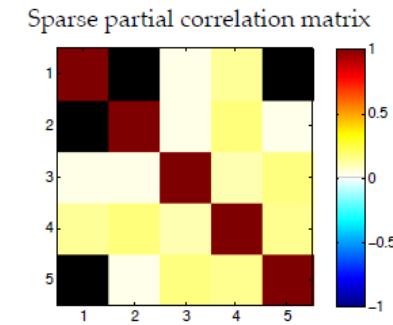
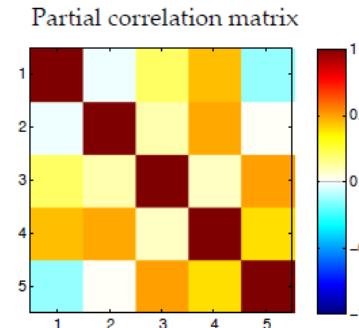
$$\hat{\mathbf{K}} = \arg \max_{\mathbf{K} \in \mathcal{P}^+} [\log|\mathbf{K}| - \langle \hat{\Sigma}, \mathbf{K} \rangle - \lambda \|\mathbf{K}\|_1]$$

Likelihood term (as before)

Shrinkage term (prior)

$$\langle \mathbf{A}, \mathbf{B} \rangle = \text{tr}(\mathbf{A}^T \mathbf{B})$$

$P^+$  is the space of pos. def. matrices





## Bayesian functional connectivity analysis

- The graphical lasso introduces a **bias** towards smaller partial correlations.
- Point estimates **ignore uncertainty**.

What could a **Bayesian approach** to functional connectivity analysis look like?

- We have a **multivariate Gaussian likelihood** term:

$$P(\mathbf{Y}|\mathbf{K}) = \prod_{i=1}^n P(\mathbf{y}_i|\mathbf{K}) = \prod_{i=1}^n \text{Normal}(\mathbf{y}_i|0, K) \propto |\mathbf{K}|^{n/2} \exp\left[-\frac{1}{2}\langle \mathbf{K}, \mathbf{Y}^T \mathbf{Y} \rangle\right]$$

- We need a **prior** on the precision matrix  $\mathbf{K}$ .
- A default prior on precision matrices is known as the **Wishart distribution**.
  - It doesn't encourage zero partial correlations either.
  - We wish to explicitly model the conditional independence structure of the precision matrix, cf. the graphical lasso!



## Bayesian functional connectivity analysis

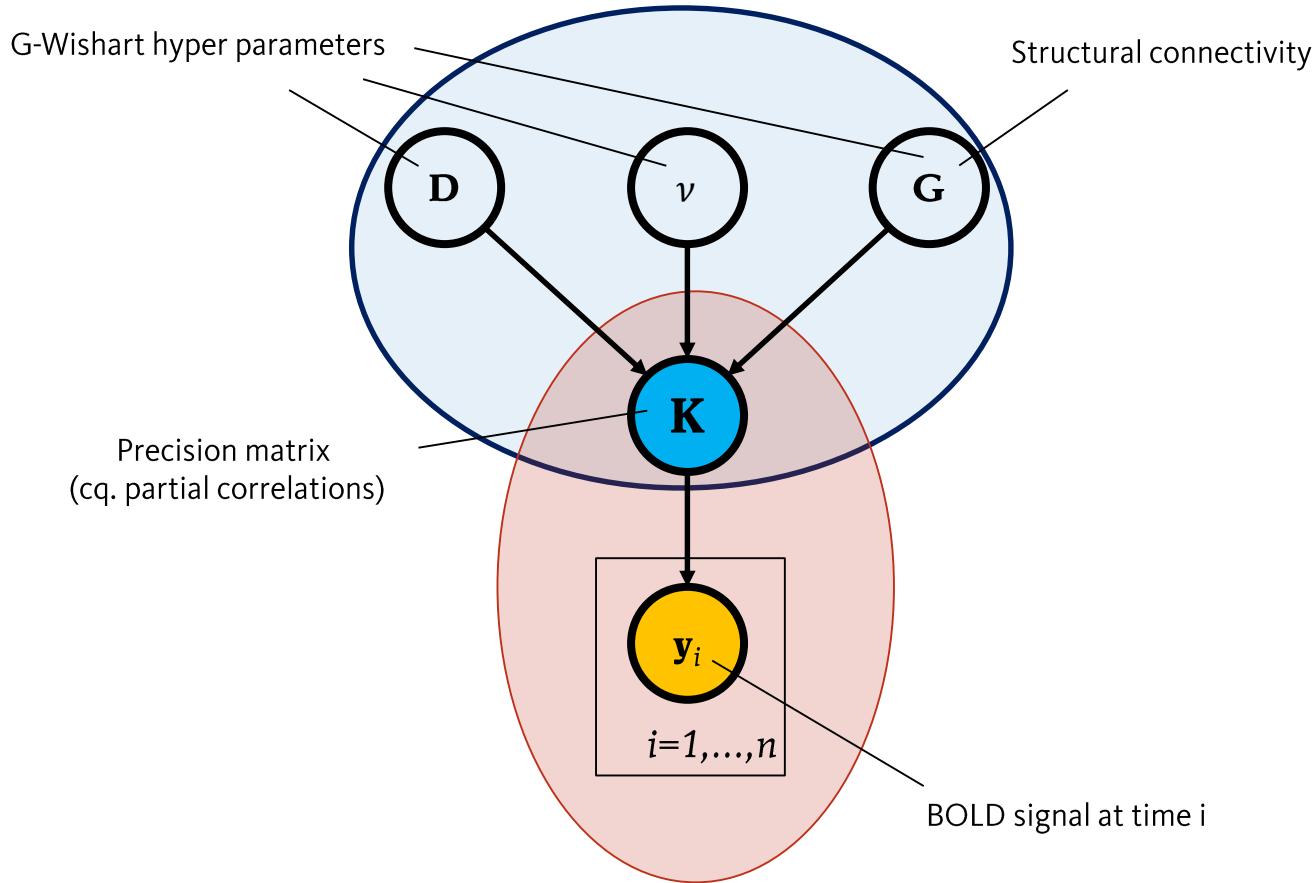
- Conditional independence structure is represented as a graph  $G \rightarrow$  **independence = no edge**
- $G$ -Wishart prior on precision matrix  $K$

$$P(K|G, \nu, D) = \frac{|K|^{\frac{\nu-2}{2}}}{Z_G(\nu, D)} \exp\left[-\frac{1}{2}\langle D, K \rangle\right] 1_{K \in \mathcal{P}^+(G)}$$

with hyper parameters  $\nu$  and  $D$ .

- Key idea: **structural connectivity** serves as an estimate for  $G$ .
  - Rationale: two regions can only have functional communication if they're anatomically connected.
  - Partial correlations are **not shrunk**; they are either zero or they may have any weight.
- Sampling from the posterior  $P(K|G, \nu, D, Y)$  is straightforward (Lenkoski, 2013).

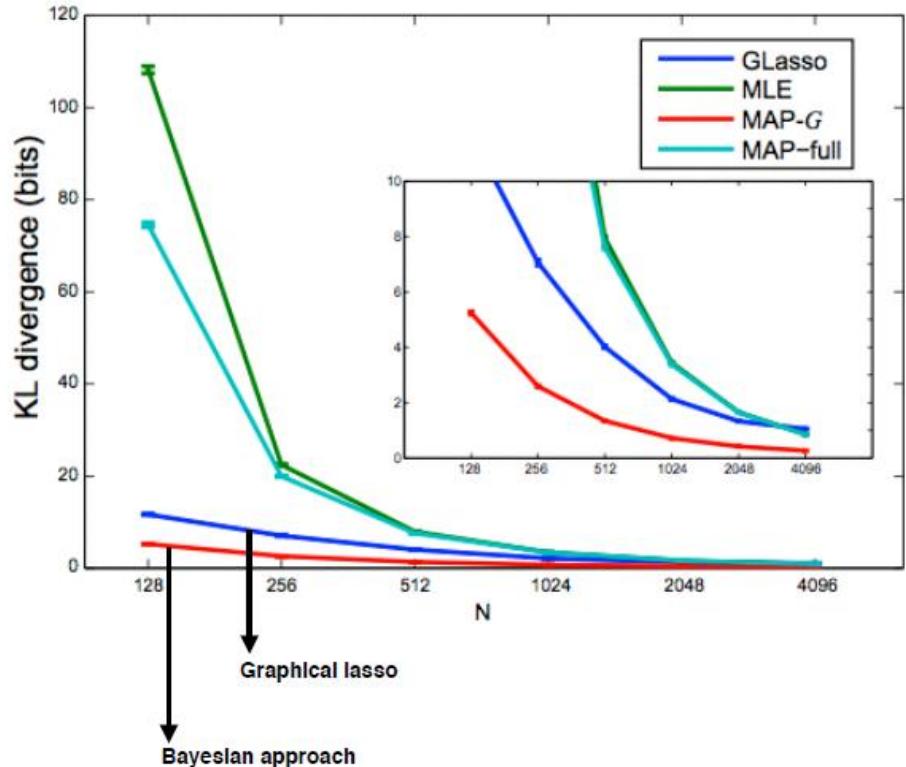
## Bayesian functional connectivity: generative model





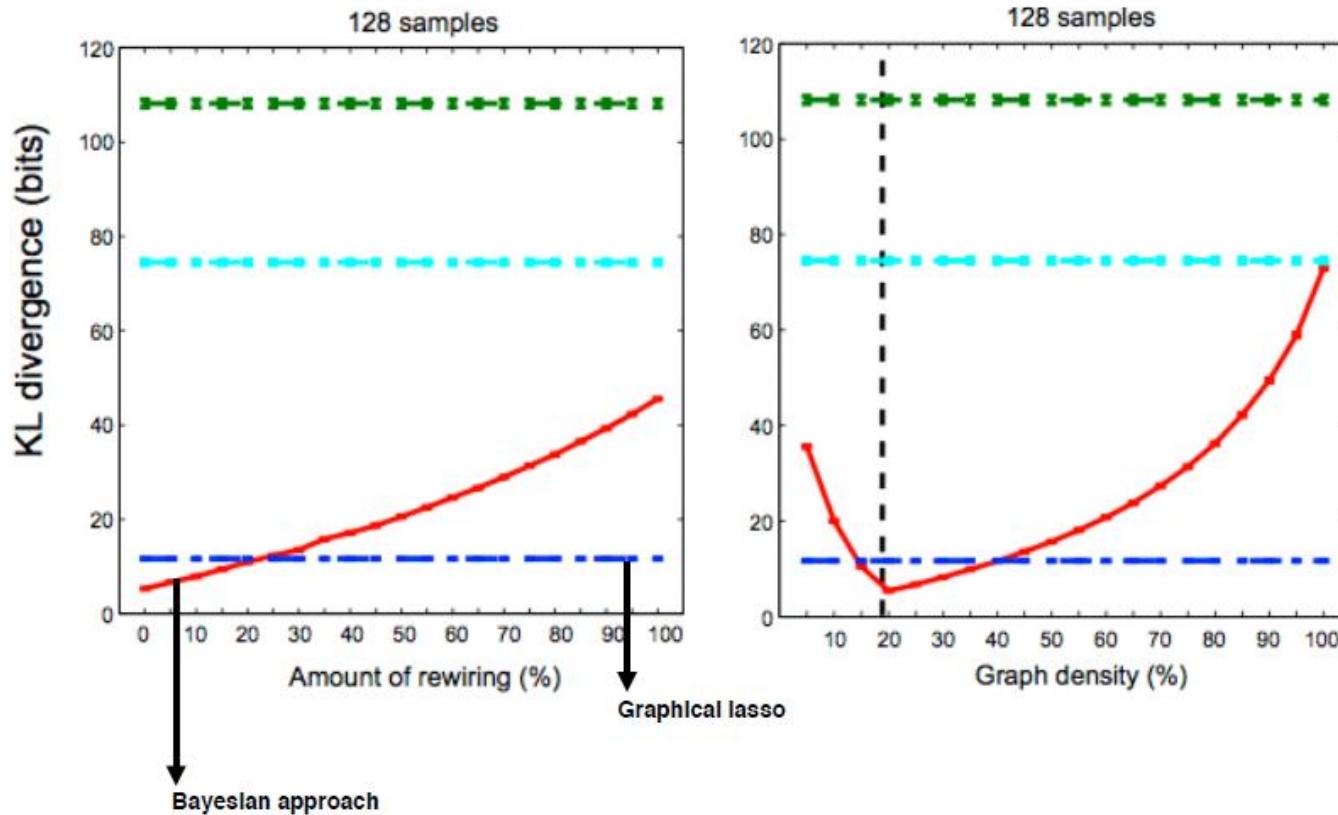
## Simulation results

- Distance to ground-truth precision matrix (assuming known  $G$ )



## Simulation results

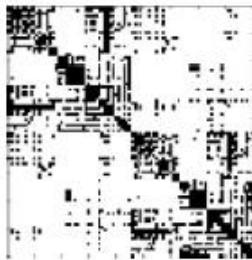
- Bayesian approach is somewhat robust to errors in  $G$ .





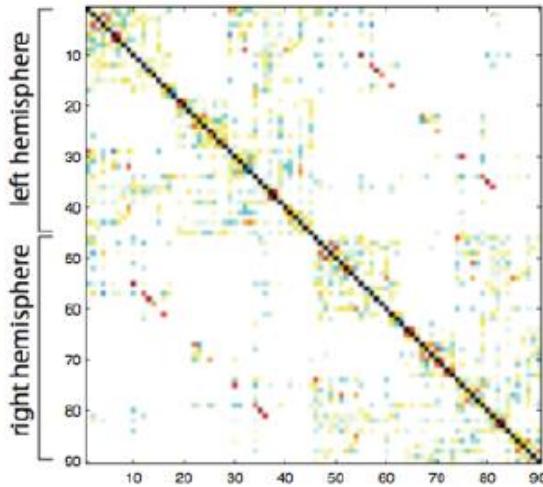
## Empirical results: partial correlations for 90 brain regions

structural graph

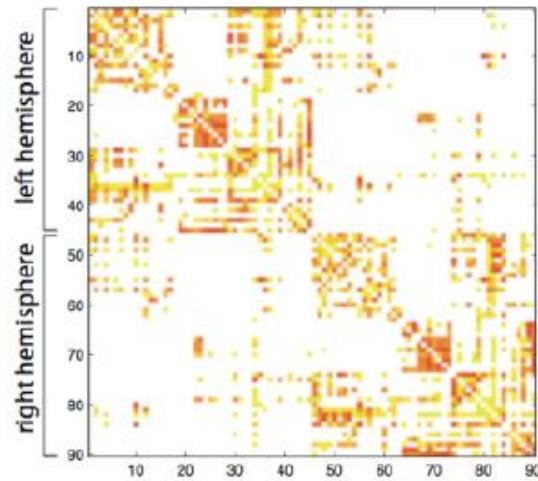


Hinne, Ambrogioni, Janssen, Heskes, van Gerven (2013), NeuroImage

S1 mean partial correlation

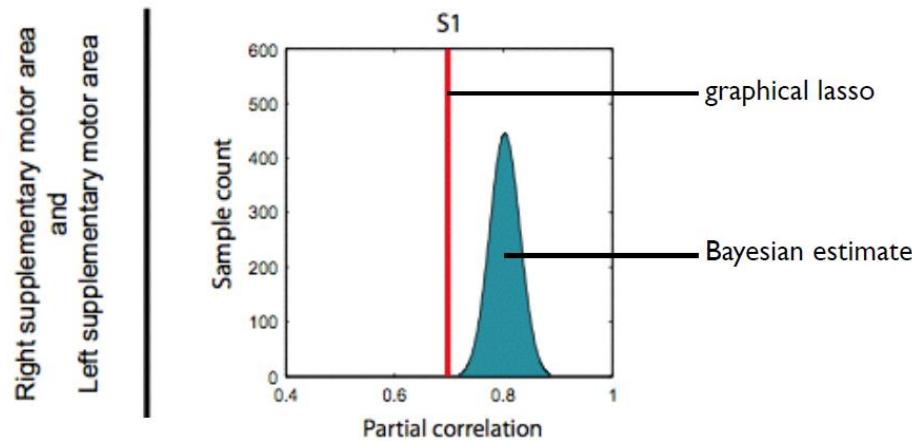


S1 SD partial correlation



## Empirical results: partial correlations for 90 brain regions

- Highest partial correlations for:
  - **Functional homologues** in left and right hemisphere
  - Pallidum—putamen
  - Intra-hemispheric frontal areas
- Bayesian approach gives **credible intervals** on partial correlations

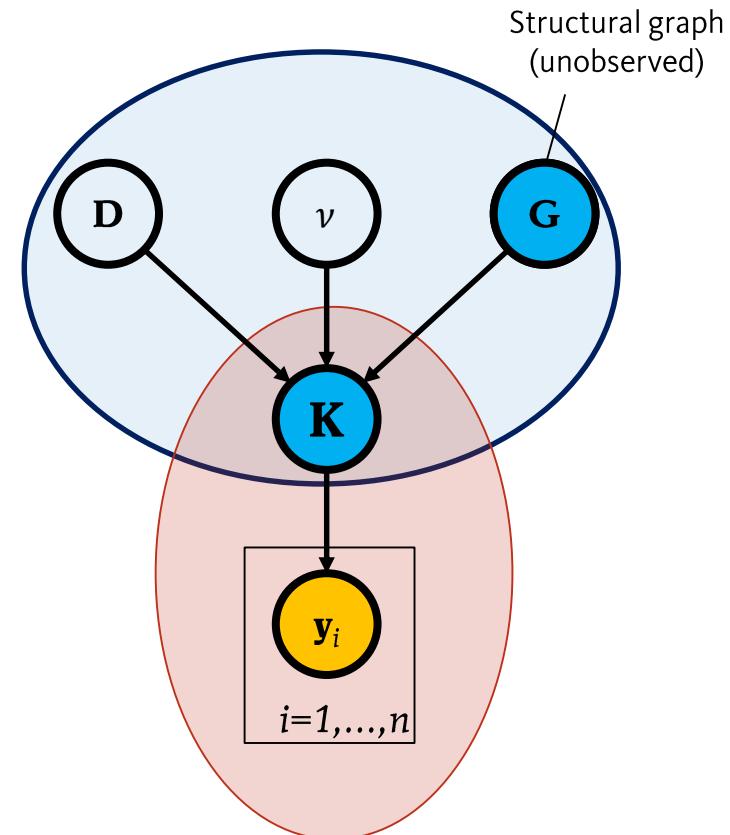


- Graphical lasso is biased towards small partial correlations.

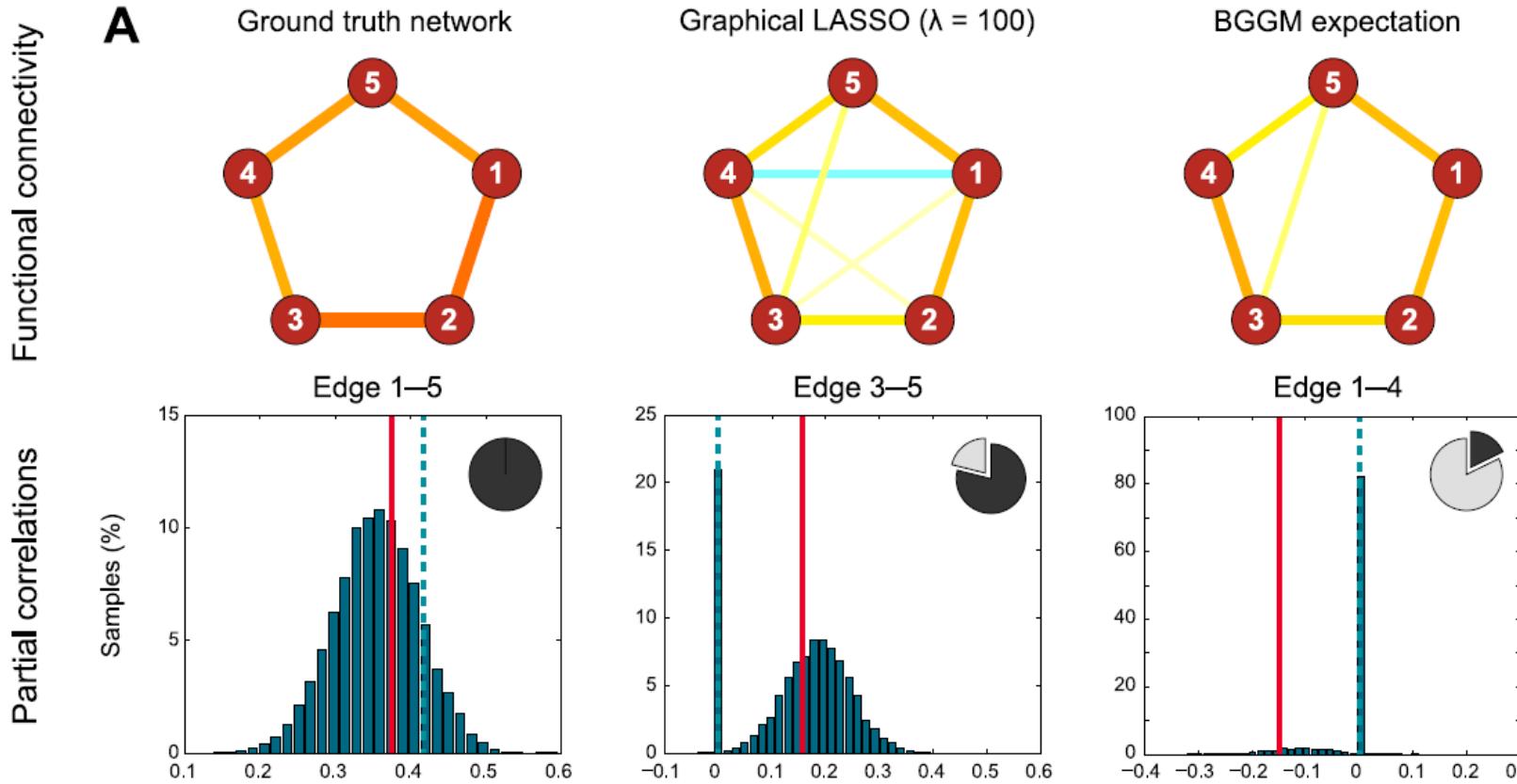


# Joint inference of functional and structural connectivity

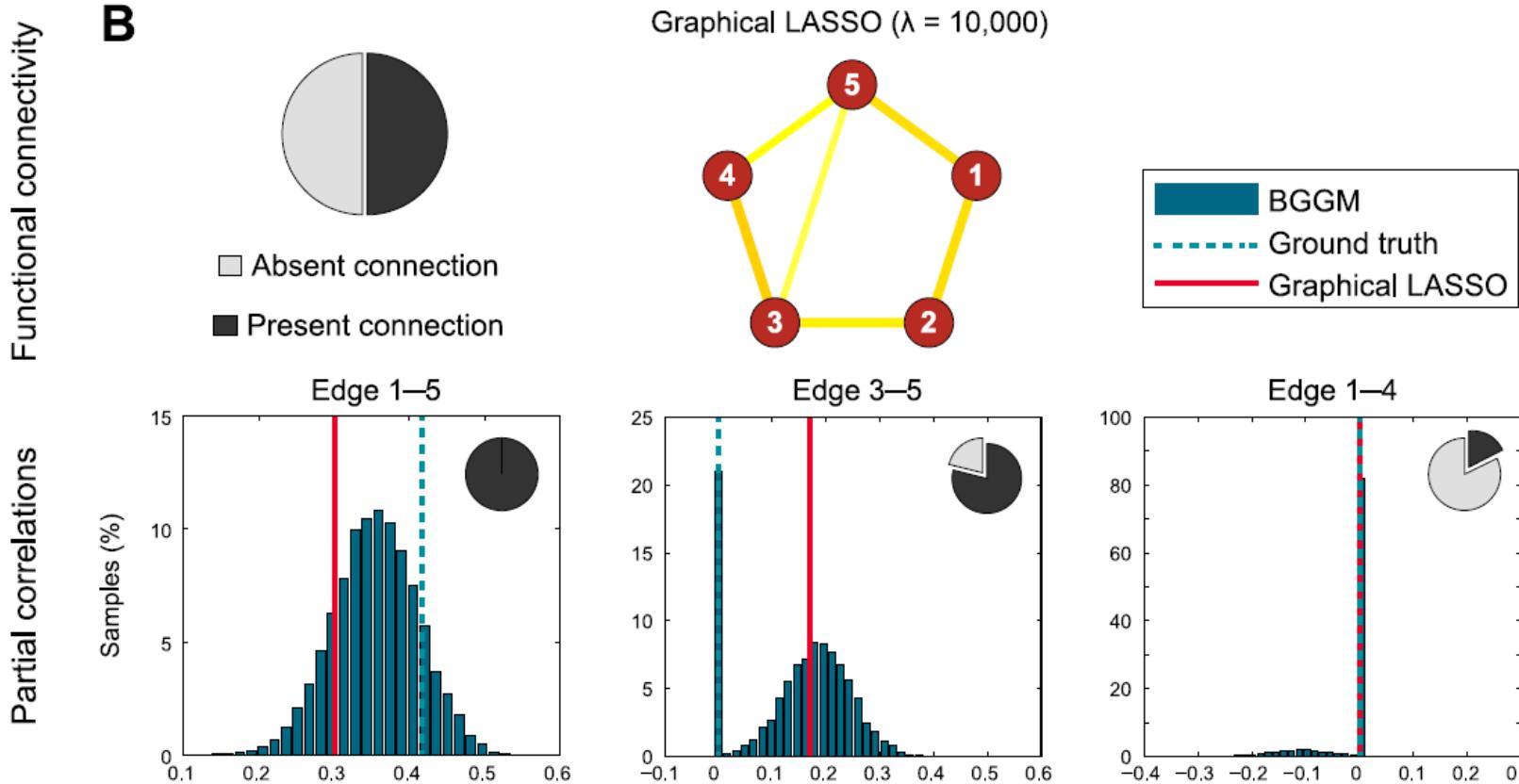
- We assumed  $\mathbf{G}$  was given.
  - Structural connectivity estimate was a useful starting point.
- Can we jointly estimate both **functional and structural** connectivity? E.g. can we obtain  $P(\mathbf{G}, \mathbf{K}|\mathbf{Y}, \mathbf{D}, \nu)$ ?
- **Extremely hard** inference problem
  - Even computing acceptance proposals is intractable!
- Requires development of more efficient MCMC samplers.
- Recent work allows inference for small-scale problems.



## Further comparison with graphical lasso



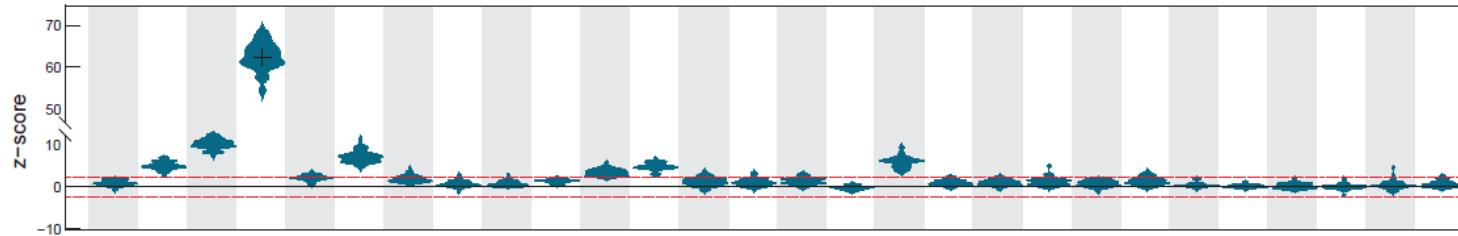
## Further comparison with graphical lasso





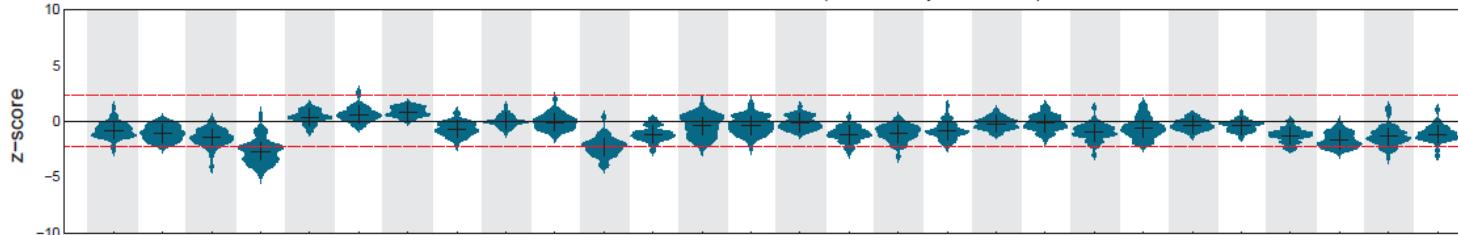
## Further comparison with graphical lasso

Total distribution of error z-score



28 simulations with different properties (#nodes, noise levels, etc.)

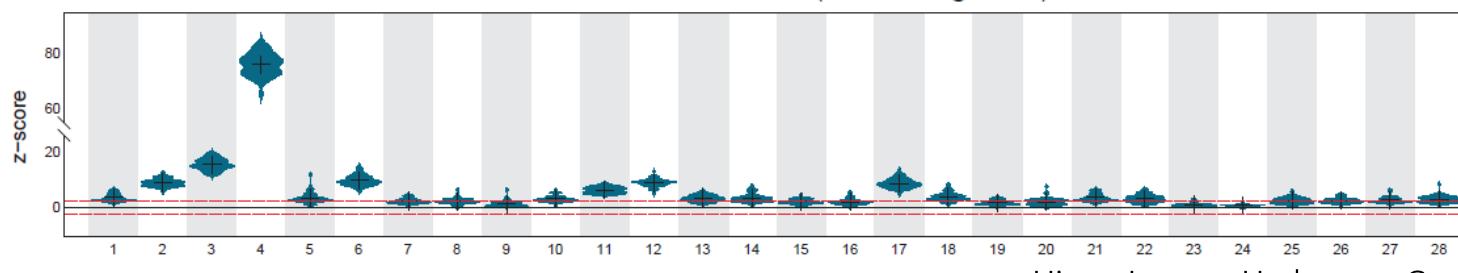
Distribution of error z-score (for true positives)



Y-axis: z-score of G-lasso error vs errors of posterior distribution of our Bayesian approach.

High z-score: G-lasso performs worse.

Distribution of error z-score (for true negatives)

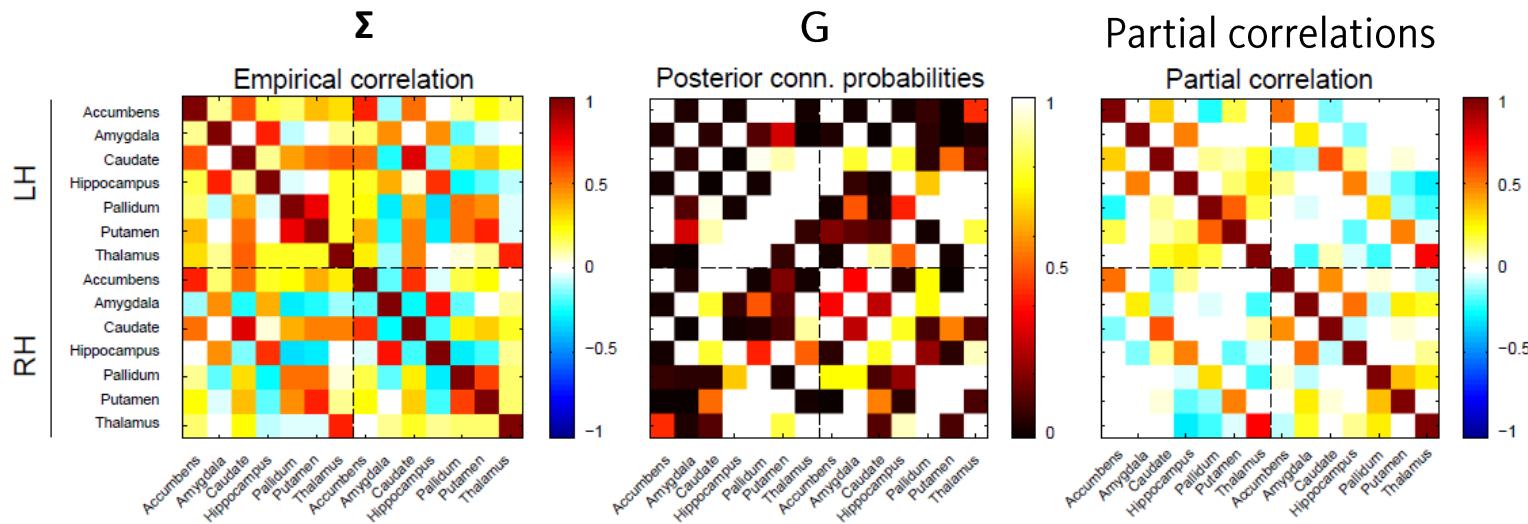
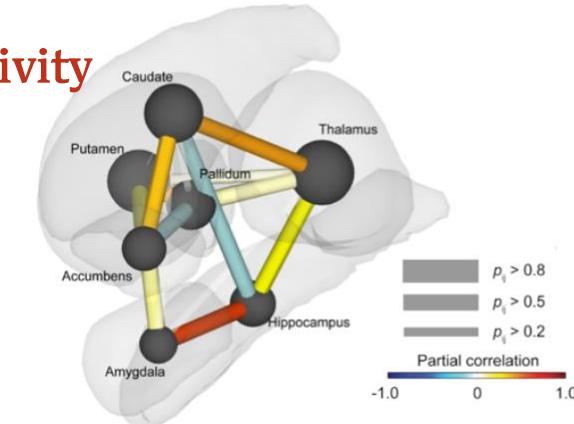


Simulation #

Hinne, Janssen, Heskes, van Gerven (2015), PLoS Comp. Bio.

# Joint inference of functional and structural connectivity

- Functional and structural connectivity using only functional data.
- 14 subcortical areas

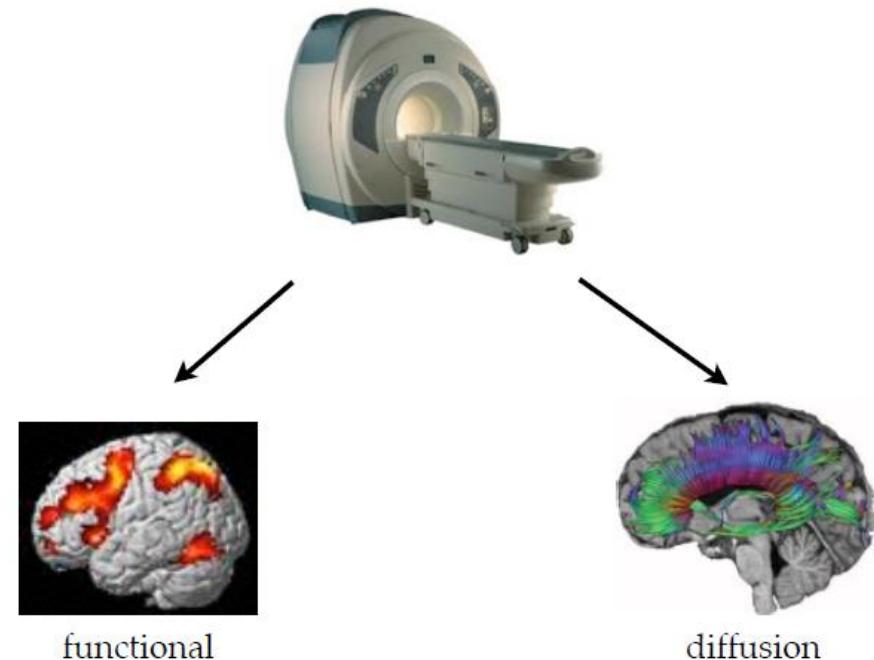


# Bayesian data fusion

How do we combine different sources of information?

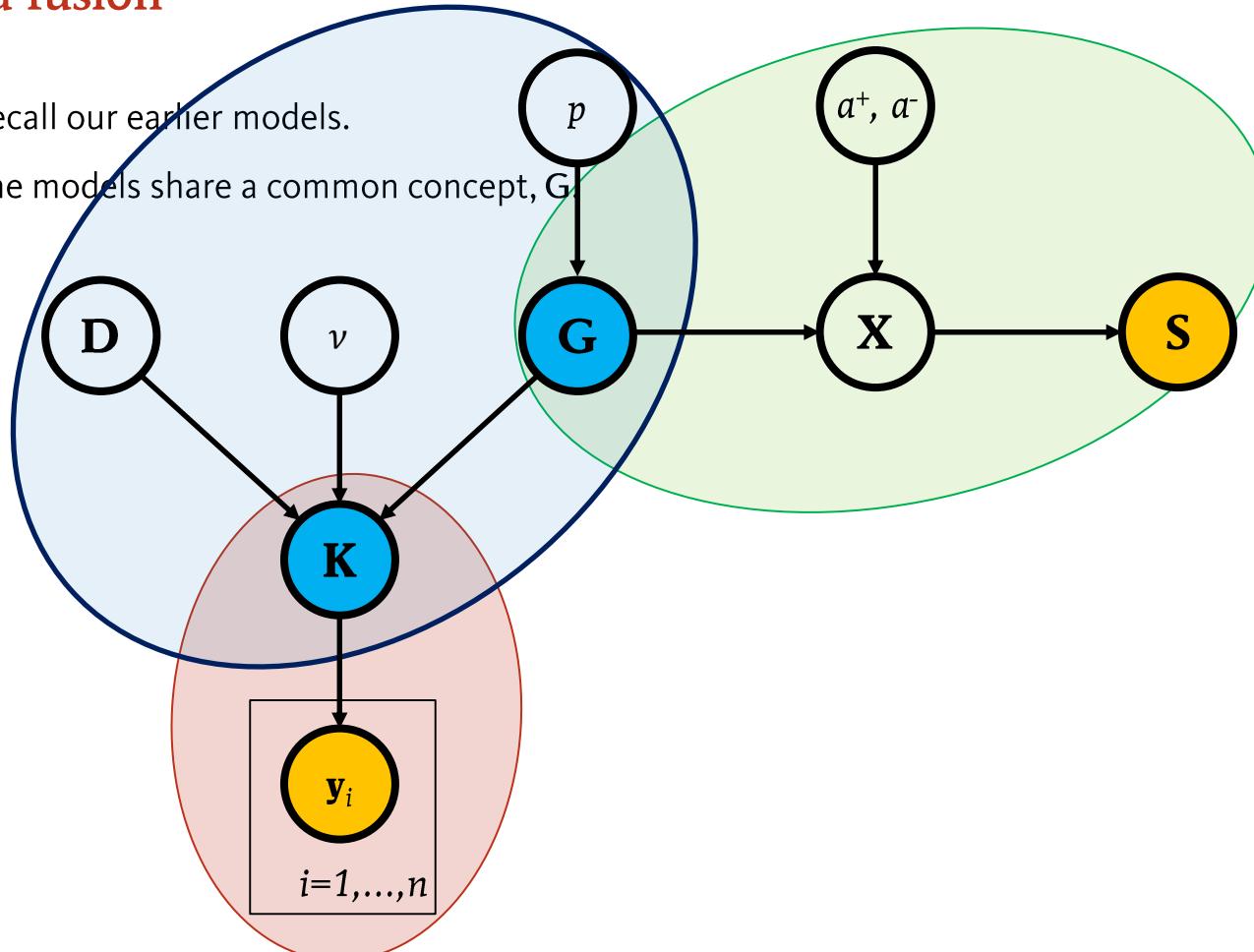
## Data fusion

- Suppose we collect both resting-state fMRI and diffusion imaging data.
- How can we **integrate data** from multiple sources to get more reliable estimates of connectivity?
- This is a **data fusion** problem.



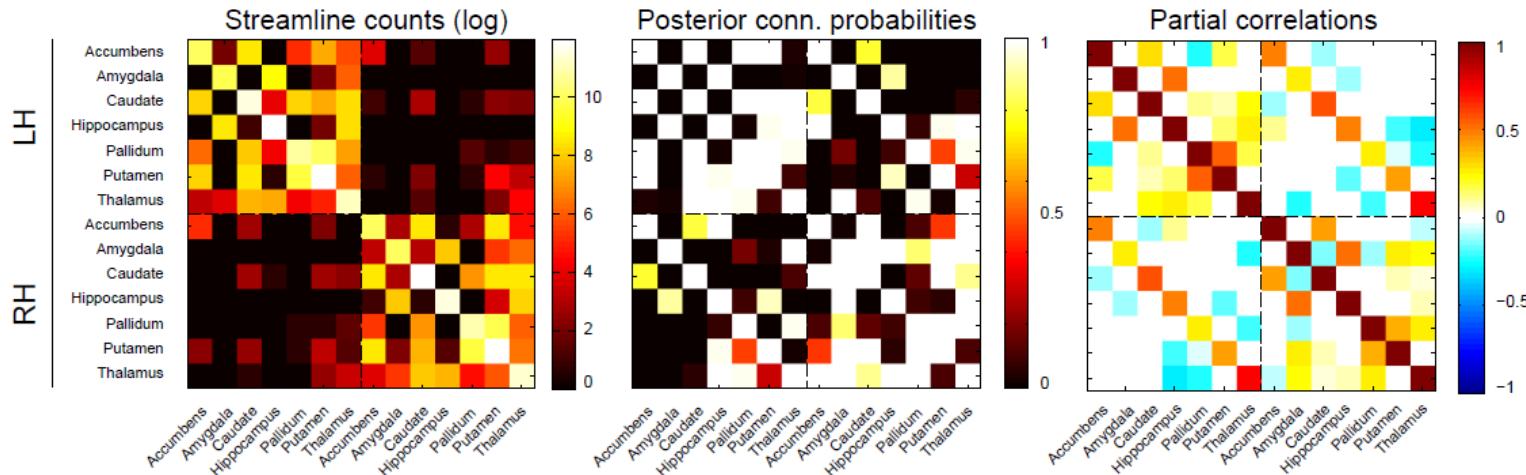
## Data fusion

- Recall our earlier models.
- The models share a common concept,  $G$ .



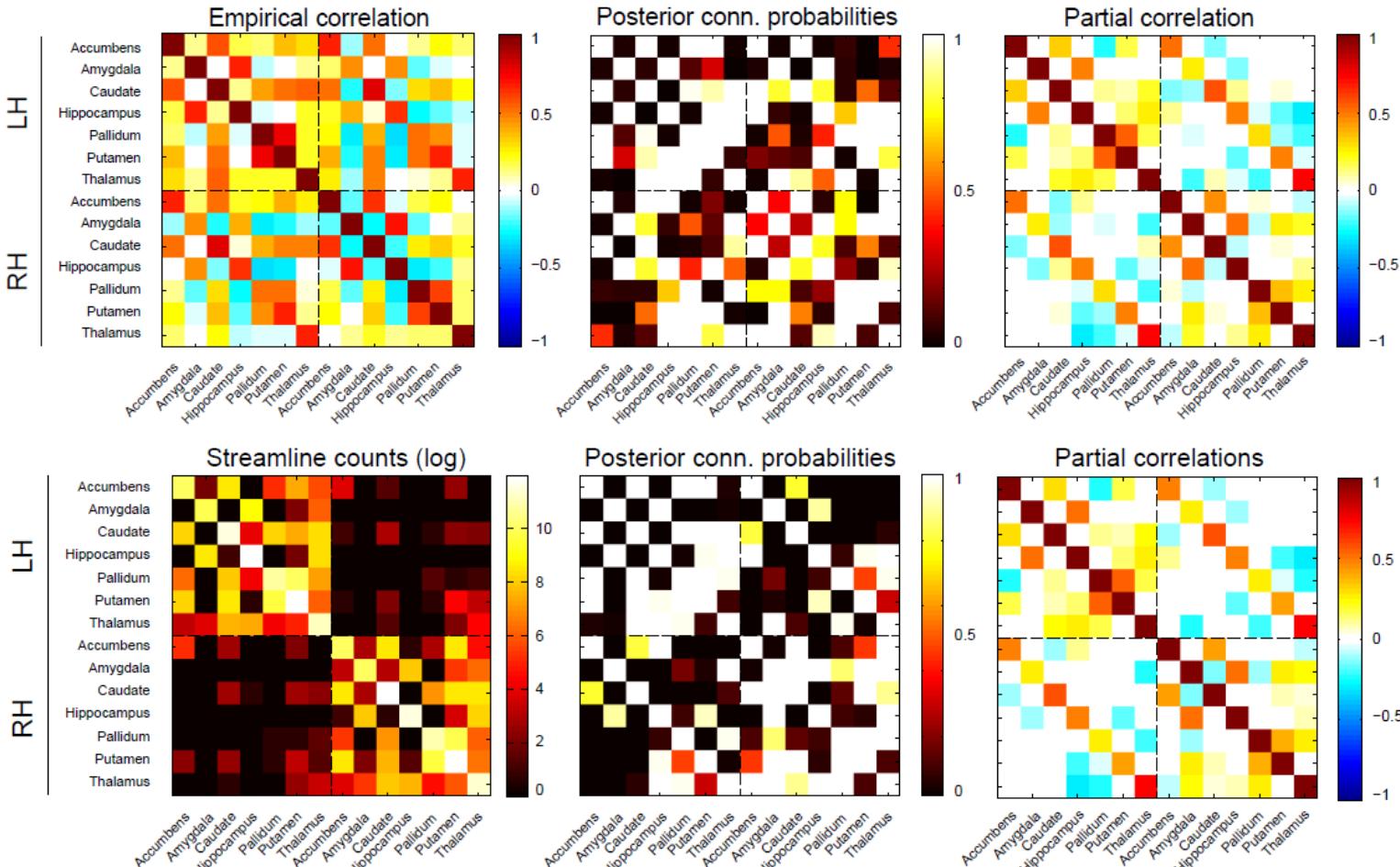
## Data fusion

- Streamlines **regularize** partial correlations.
- Conditional (in)dependencies **inform** structural connections.

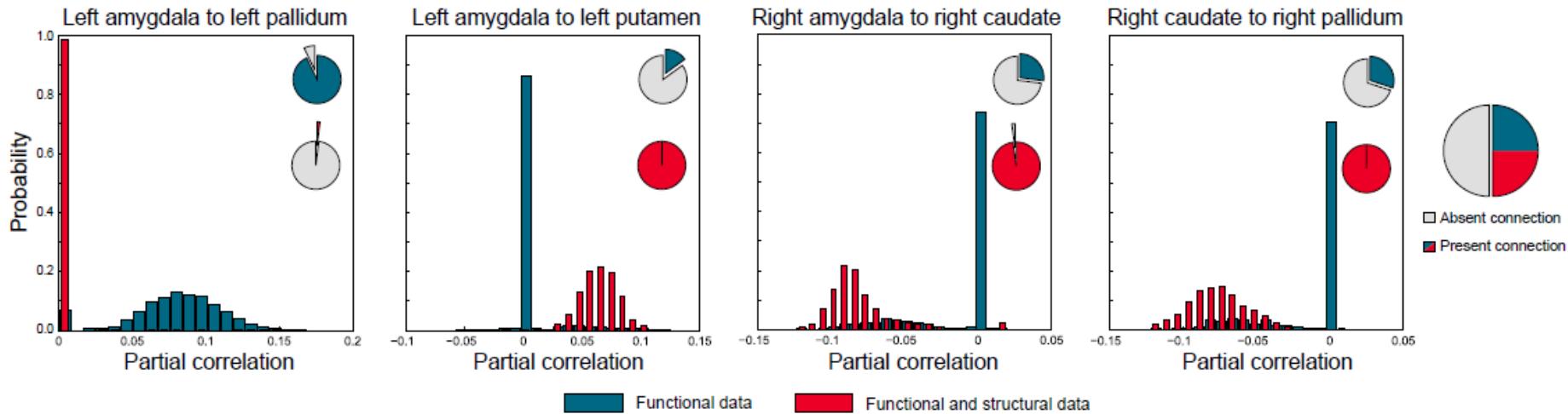




# Data fusion



## Data fusion



- Streamlining subcortical connectivity is prone to errors.
- Proof-of-concept shows value/applicability of data fusion.
- Bayesian approach allows **incorporating all available observations**.



## Conclusions & future directions/challenges

### Fundamental advantages

- Bayesian generative modeling gives an principled way of learning connectivity from neuroimaging data.
- Posterior distributions enable expression of uncertainty.
- Generative models can be combined in a plug-and-play fashion.

### Practical challenges

- Large and complex models are difficult to sample.
- Application on large graphs requires development of more efficient approximate inference.

### What lies ahead

- More realistic models.
- Dynamics of connectivity.
- Large-scale networks of microscopic connectivity.



Answers

Questions