a neurocognitive graph theoretical approach to understanding the relationship between minds and brains

joshua t vogelstein, r jacob vogelstein, carey e priebe

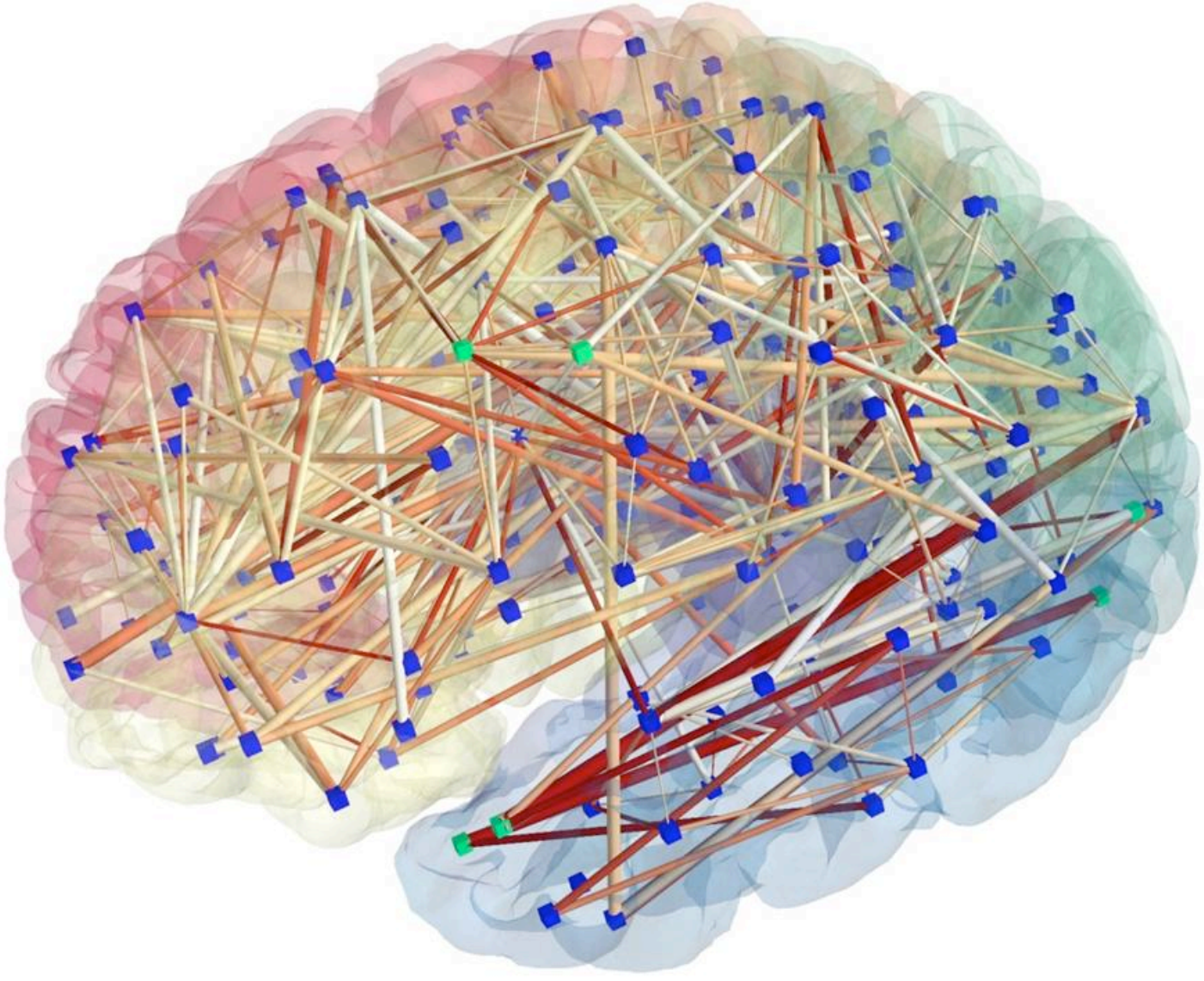
johns hopkins university

march 11th, 2010

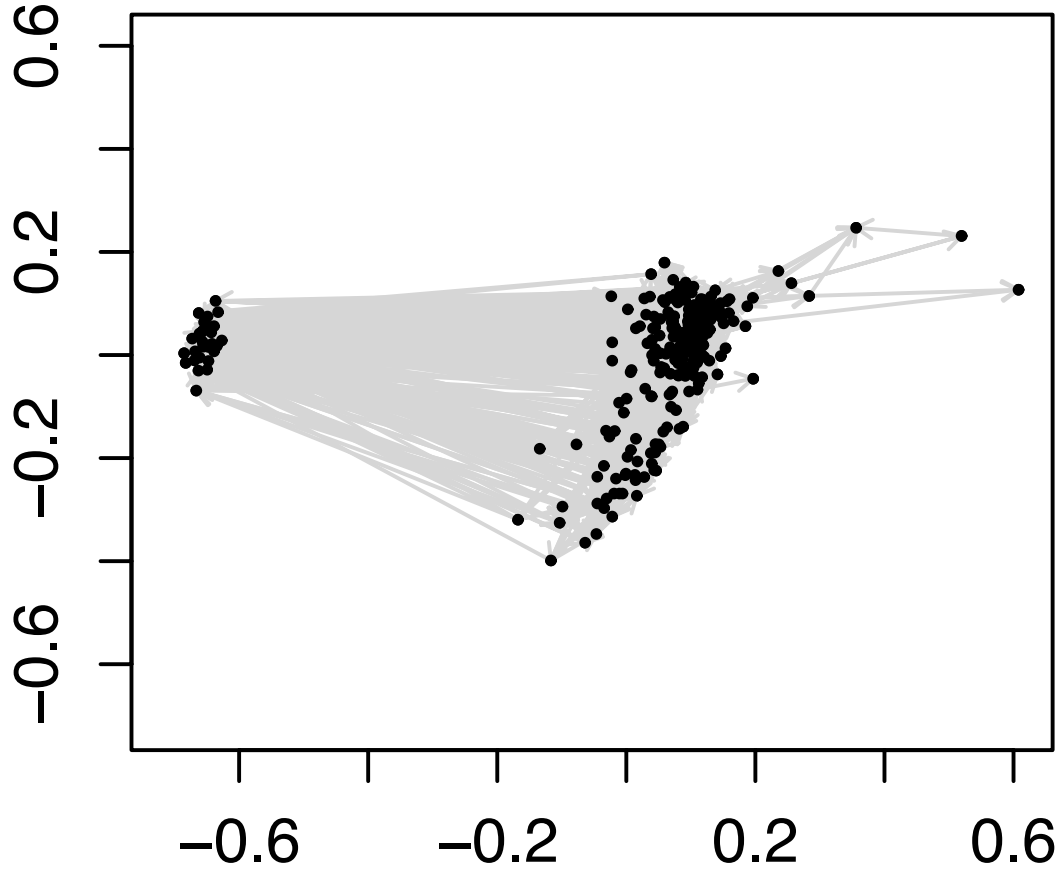
motivation

- connectomes and neural circuits are all the rage
- a natural question is: what do we do with them?
- we aim to construct a rigorous statistical framework that facilitates the elucidation of the causal relationship between minds, M, and brains, B
- this means proposing a model, $P[B, M]$, that characterizes this relationship, and allows us to develop algorithms to perform various inference tasks

human MR connectome



c. elegans electron microscopy (em) chemical connectome



desiderata

model desiderata

- can account for data in a way that is interpretable
- sufficiently generalizable
- latent features are identifiable

inference algorithm desiderata

- universal consistency, meaning that it will converge to the optimal result, for any $P[B, M]$, given infinite data
- rapid convergence rates

a proposal: brains as random graphs

- a random brain-graph, $B = (V, E)$ is defined by a set of vertices V and edges between the vertices E
- a vertex may correspond to any "neural unit", including neurons, voxels, columns, regions
- an edge may correspond to any connection between neural units, including chemical and electrical synapses, functional strength, white matter tracts, etc.
- both vertices and edges may be endowed with (potentially latent) features, including receptive fields, probability of release, cell-type, spatial location, etc.
- observed brain-graphs are random because they are samples from some model

why random graph models

- most of probability and statistics deals with random variables, random vectors, or random point-processes (e.g., spike trains)
- random graph theory provides a rigorous statistical framework for performing inference on objects that are characterized by graphs
- our perspective is that each observed neural circuit or connectome is a sample from a random brain-graph model
- given such models, we can infer all sorts of fun stuff

the formal setting

- let b be a randomly sampledbrain-graph, $B = (V, E)$
- let m be a randomly sampled mental property, M (e.g., intelligence)
- given a novel brain-graph, b, compute the maximum a posteriori associated mental property, m^*:

$$m^* = \underset{m}{\text{argmax}}\, P[m|b] \tag{1}$$

  where $P[m|b]$ is the posterior
- the optimal classifier g^* provides this:

$$m^* = g^*(b) \tag{2}$$

the formal approach

- since neither the model $P[B, M]$ nor the optimal classifier g^* are known, one must be estimated from the data
- let $(b_1, m_1), \ldots, (b_n, m_n) \overset{iid}{\sim} P[B, M]$ be observed pairs sampled iid from the model
- given these pairs, we can build a classifier to estimate the optimal one, by minimizing the misclassification rate:

$$\widehat{g} = \underset{g}{\operatorname{argmin}}\, P[g(b) \neq m | (b_1, m_1), \ldots, (b_n, m_n)] \tag{3}$$

- then, for a new observed brain, b, we use our estimated classifier to predict the most likely mental property, m:

$$\widehat{m} = \widehat{g}(b; (b_1, m_1), \ldots, (b_n, m_n)) \tag{4}$$

bayes plug-in classifier is universally consistent

- for each b_i, estimate the posterior using the Bayes plug-in,

$$\widehat{P}[M = m|B = b] = \frac{1}{n}\sum_{i=1}^{n} \delta\{m_i = m|B = b_i\} \qquad (5)$$

  where n is the number of training samples, and $\delta\{\cdot\} = 1$ when the argument is true, and zero otherwise
- note that if $b \neq b_i$ for any b_i, then the prior is optimal
- so this would work in the limit, but converges very slowly [2]

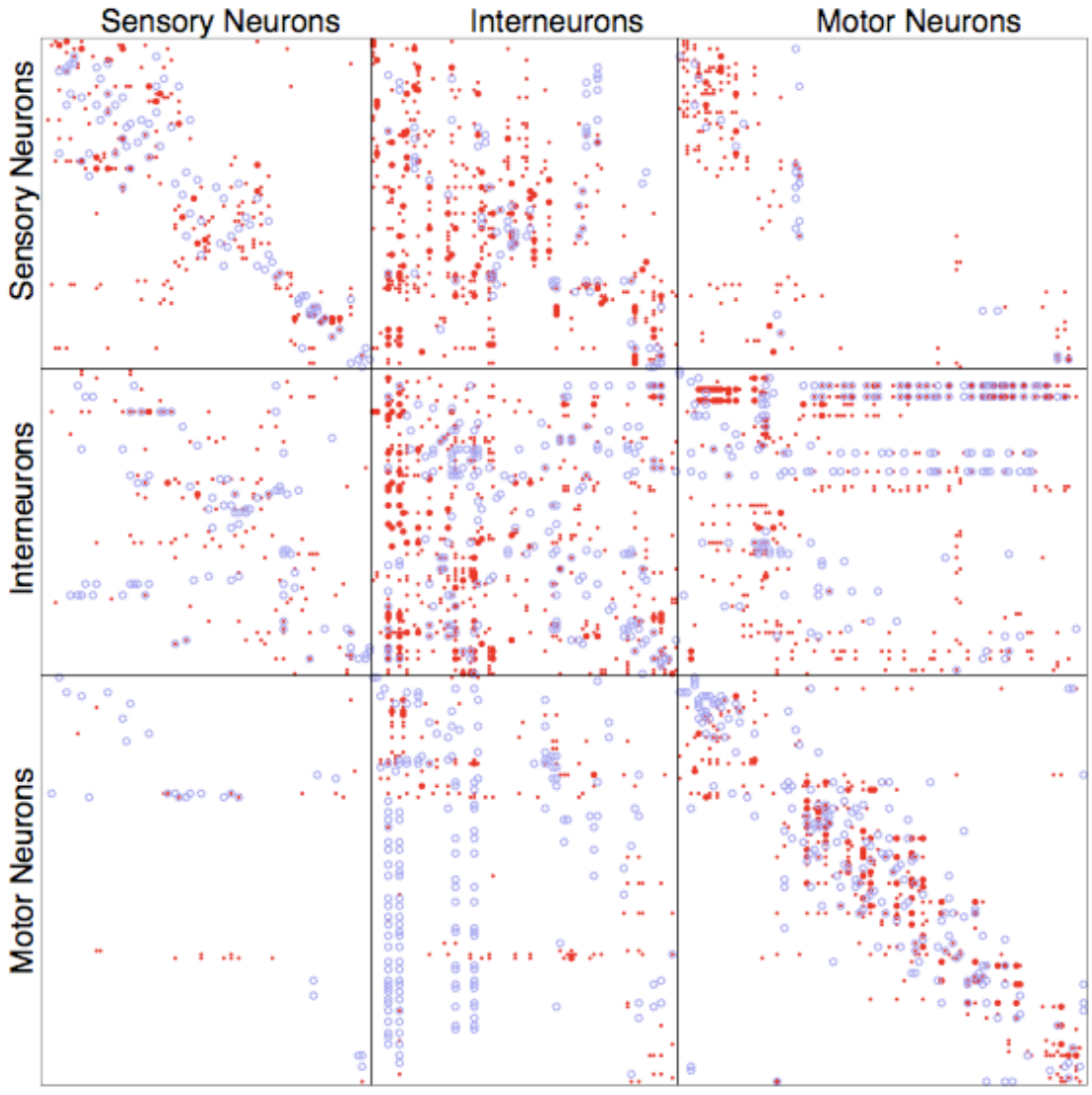
k_n nearest neighbor is universally consistent with faster convergence rates (often)

- k_n-nearest neighbor (knn) algorithm is known to be universally consistent (u.c.) for vectors [2]
- we prove that knn is u.c. for graphs [3]
- proof of faster convergence rates under certain very lax conditions is in preparation
- the 1-nearest neighbor algorithm for graphs proceeds as follows:
  - compute the distance between b and each b_i, $d_i = d(b, b_i)$
  - let $j = \text{argmin}_i\, d_i$, and $m = m_j$
- for $k > 1$, find k smallest d_i's, and let m be whichever class is the plurality/majority
- we let $d(x, y) = (\sum_{ij}(x_{ij} - y_{ij})^2)^{1/2}$ be the Froebenius norm

application to em data from c. elegans

- the c. elegans connectome has been determined for approximately 1 worm [1]
- the wild type c. elegans connectome is believed to be relatively stereotyped [1]
- odor evoked behavior is believed to be determined by synapses between a small number of neurons [4]
- thus, we generated a simulation containing two distinct populations of c. elegans
- the intention is to be able to classify a new c. elegans into either wild type, $m = 0$ or odor evoked behavior impaired, $m = 1$

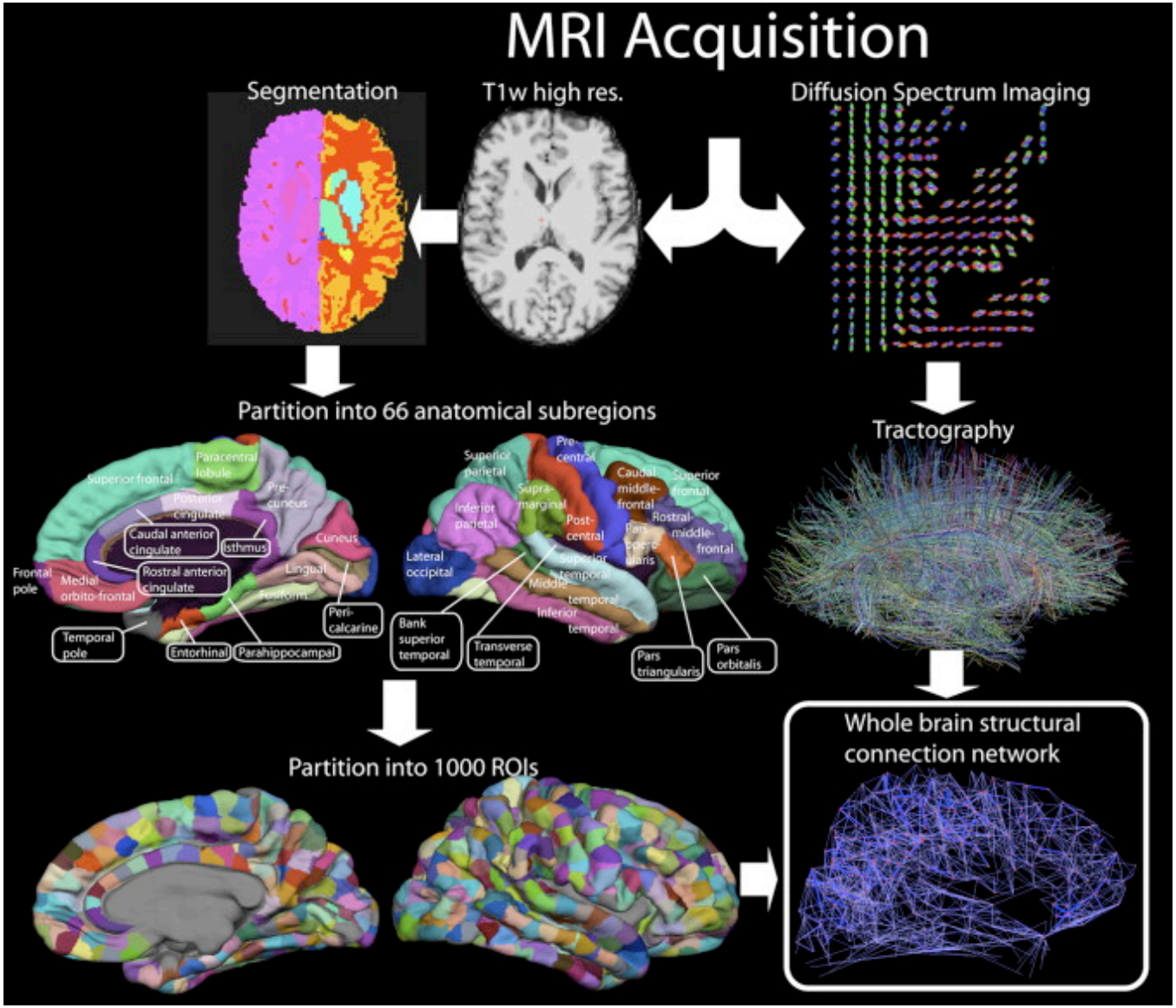
c. elegans electron microscopy (em) connectome [1]



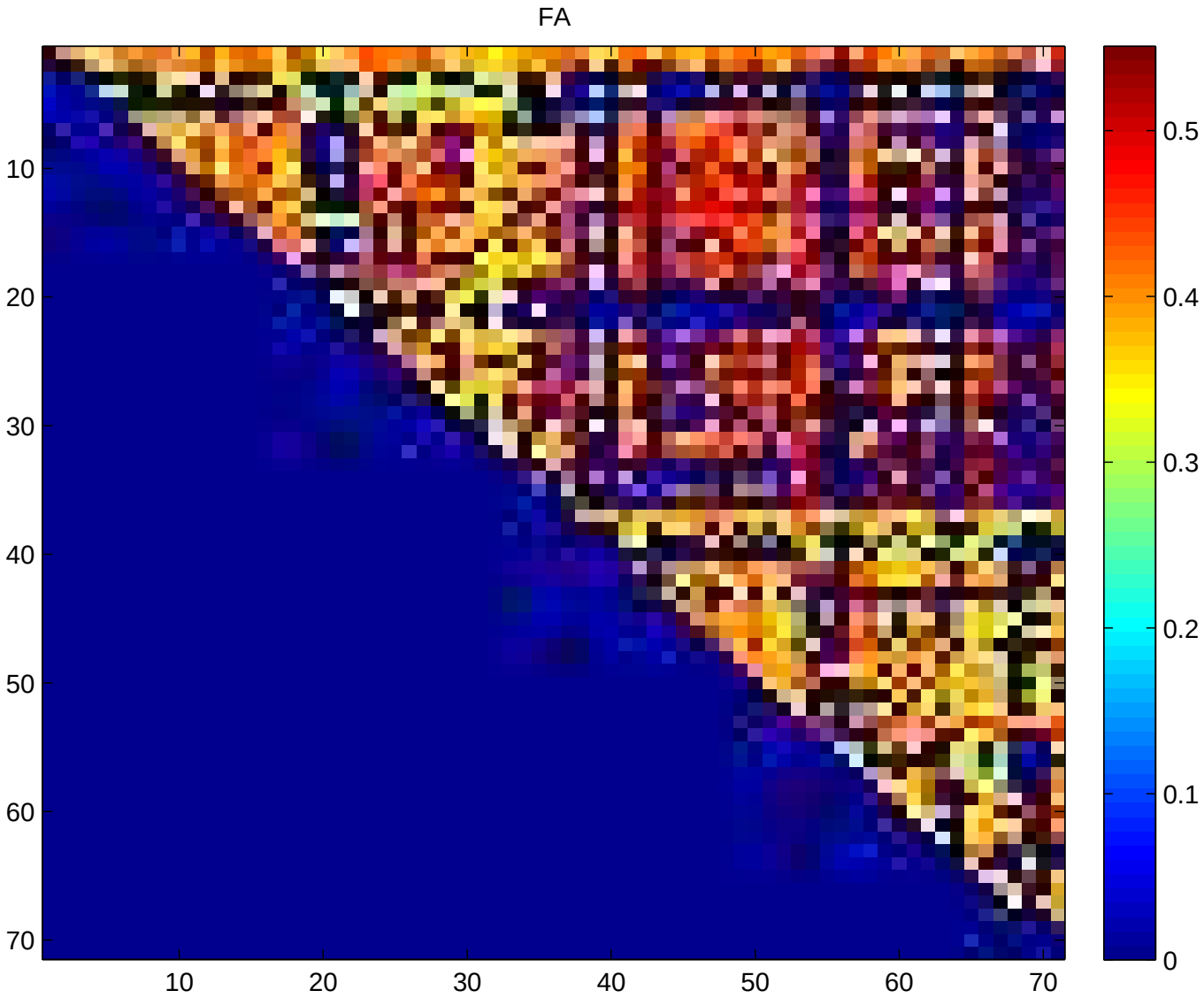
application to MR data from humans

- diffusion MRI can be used to infer brain-graphs [5]
- various cognitive features have been linked to connectivity [5]
- we utilized custom software to infer labeled brain-graphs from $n = 1$ human
- we obtain mean FA for fibers connecting 70 labeled anatomical cortical regions [6]
- based on this, we generated a simulation containing two distinct populations of humans

from MR data to brain-graph



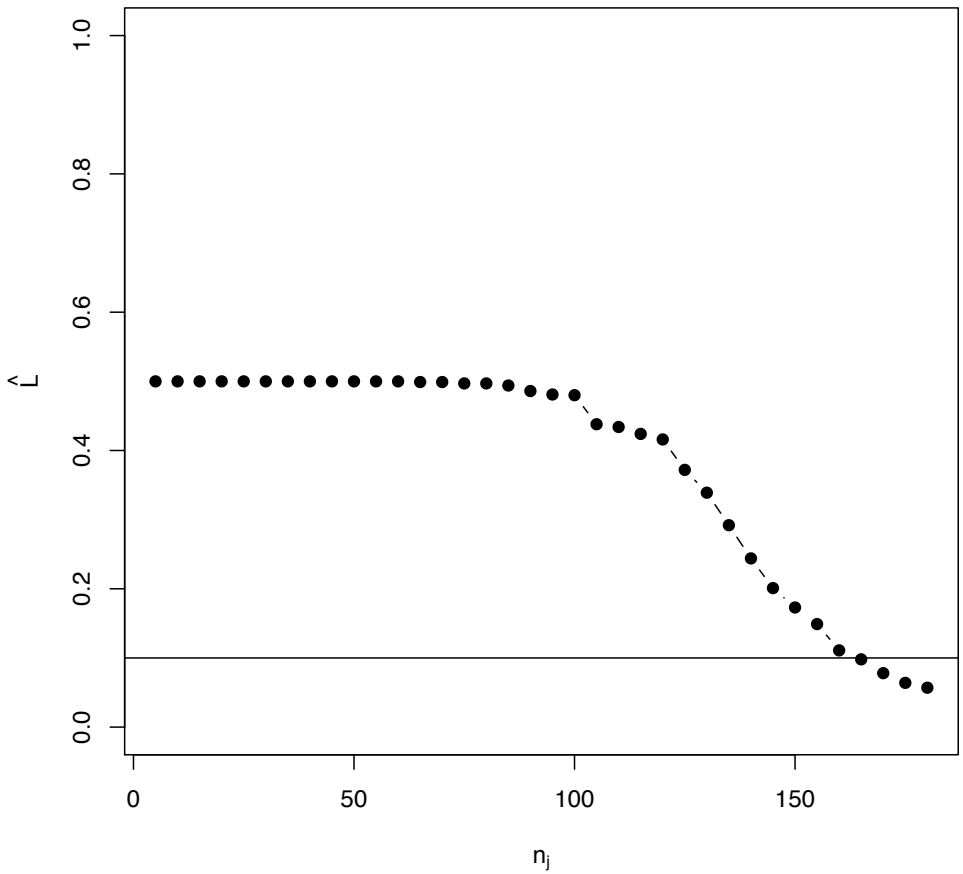
human diffusion MRI connectome



simulation details

- let A_{ij} indicate the number of chemical synapses or white matter tracts between neurons or neuroanatomical regions i and j
- let $0 < \eta \ll 1$ indicate a noise parameter
- let $\mathcal{E}$ be the set of edges that differ between the two classes
- let n be the number of training samples
- sample $n_0 \sim \mathcal{U}(0, \ldots, n)$, and $n_1 = n - n_0$
- let $C_{ij} \overset{iid}{\sim} \mathcal{U}(-lb, ub)$ for $(i, j) \in \mathcal{E}$, and $C_{ij} = 0$ for $(i, j) \notin \mathcal{E}$
- sample n_0 adjacency matrices from Poisson$(A_{ij} + \eta)$ to obtain our simulated wild type population
- sample n_1 adjacency matrices from Poisson$(A_{ij} + C_{ij})$ to obtain our simulated impaired population

main result: knn classifier converges on c. elegans data

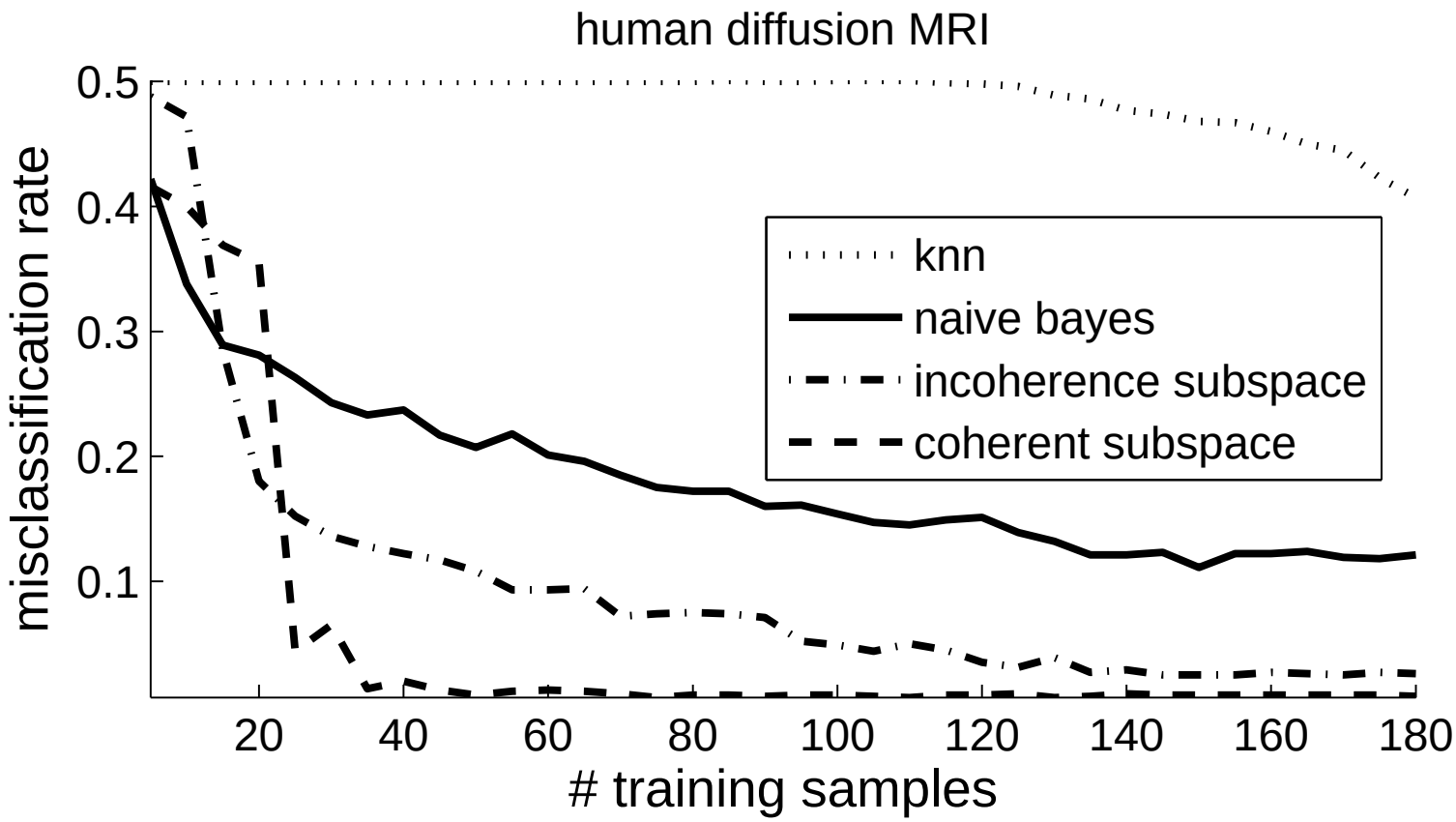


$\widehat{L}$ is the misclassification rate

making convergence rates faster

- if simplifying assumptions about $P[B, M]$ can be made without introducing much bias, but substantially reducing variance, then algorithms optimal given these assumptions will tend to converge faster [1]
- the edge independent approximation is poor, but often leads to improved results (called naïve bayes)
- $\mathcal{E}$ is known, then one can define d_i as only the distance between the subgraphs induced by $\mathcal{E}$
- if $\mathcal{E}$ is unknown, it can be estimated from the data (incoherent subspace)
- if $\mathcal{E}$ is unknown, but it is known that all the edges within it share a common set of vertices, this knowledge can be utilized (coherent subspace)

secondary result: simplifying assumptions yield faster convergence rates



discussion

- unraveling the mind-brain relationship is a central tenet of contemporary neuroscience
- we introduce a coherent framework for quantifying this relationship utilizing the concept of random brain-graphs
- we prove the existence of a universally consistent classifier
- faster convergence rates come from simplifying assumptions that don't introduce much bias
- results on real data could lead to prognostics or other therapeutics
- brain-graph model fitting, model selection, etc., all seamlessly integrate with this approach
- applications abound

bibliography & acknowledgements

bibliography

1. Varshney LR, et al. *Structural Properties of the Caenorhabditis elegans Neuronal Network.* `http://arxiv.org/abs/0907.2373`
2. Devroye L, Györfi L, Lugosi, G. *A Probabilistic Theory of Pattern Recognition.* Springer, 1996.
3. Vogelstein JT, Vogelstein RJ, Priebe CE. *Are mental properties supervenient on brain properties?* `http://arxiv.org/abs/0912.1672`
4. Chalasani SH, et al. *Dissecting a circuit for olfactory behaviour in caenorhabditis elegans.* Nature, 2007.
5. Bullmore E, Sporns O. *Complex brain networks: graph theoretical analysis of structural and functional systems.* Nat Rev Neurosci, 2009.
6. Desikan Rs, et al. *An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest.* NeuroImage, 2006.

acknowledgements

- This work was supported in part by the NSA Research Program in Applied Neuroscience