Algorithm Design for Connectivity-based Cortex Parcellation

Research Plan

ETH Zurich
Department of Computer Science

PhD Student:	Nico Stephan Gorbach
Supervisor:	Prof. Dr. Joachim M. Buhmann (ETH Zurich)
Co-supervisor:	Prof. Dr. Klaas Enno Stephan (D-ITET, ETH Zurich)
External Collaborator:	Dr. Marc Tittgemeyer (MPI for Neurological Research)

Provisional date of PhD acceptance at ETH Zurich: 22 March 2012
Start date of employment as a research assistant at ETH Zurich: 1 July 2012

1 Summary

White matter connectivity plays a dominant role in brain function and arguably lies at the core of understanding the structure-function relationship in the cerebral cortex. Noninvasive measurement of white matter connectivity in the living tissue became feasible with the advent of diffusion magnetic resonance imaging (dMRI), in particular diffusion tensor imaging (DTI), thereby facilitating an ambitious endeavour to explain brain function in terms of white matter connectivity. In this research we develop a model to explains the structure of diffusion connectomics that is meaningful for functional organization. Our model reveals both connectivity fingerprints and connectional families concurrently by a generative design. In contrast to previous methods, our model is a *problem-centric* description of structure-function relationship in the brain. A further goal is to improve the information content of the diffusion parcellation pipeline with respect to model order selection, fiber tracking and experimental tuning parameters.

2 Introduction

Unravelling the relationship between structure and function in the human brain has been a long-standing ambition in the neuroscience community. The brain is

considered, in concert, as an "ensemble of functionally specialized areas that are coupled" [1] together in a modulatory fashion. The notion that white matter connectivity structure localizes functional specialization in the cerebral cortex forms a basis for understanding the relationship between structure and function. Yet formal proof of this phenomenon remains elusive. However, white matter connections play a dominant role in the functional specialization of a cortical area which suggests that the coherent structure of white matter connections underpins functional localization in the cortex. The framework is known as connectivity-based cortex parcellation and is aimed towards subdividing the cerebral cortex to reveal "connetivity fingerprints" and "connectional families" that support inference of functional organization.

Diffusion magnetic resonance imaging (dMRI) is a noninvasive tool for measuring white matter connectivity in the living tissue which offers enormous potential for elucidating functional organization in the brain [2]. The purpose of this thesis is to develop a model and algorithms that explains the structure of diffusion connectomics meaningful for functional organization. Section 3 describes state-of-the-art including a comparison between inferences on functional organization made by noninvasive techniques with those made by diffusion connectomics. Goals of this thesis are given thereafter followed by a work plan.

3 State-of-the-art

Functional organization is an umbrella term including concepts of functional localization as well as functional integration, connectional families as well as functional systems and stages of information processing. Figure 1 illustrates functional localization of the visual system in the macaque brain obtained by invasive methods which gives some insights into ground truth relevant for connectivity-based cortex parcellation.

3.1 Functional organization inference using diffusion connectomics

Diffusion connectivity-based cortex parcellation is founded upon the premise that functional units are identified by their unique white matter connections, regardless of connection polarity, with the presupposition that their connectivity patterns are also homogenous. The noninvasive and in vivo nature of diffusion connectivity-based cortex parcellation offers enormous potential for the following reasons:

- Arguably its greatest potential is its contribution to multimodal studies that attempt to reveal a more complete picture of brain function than studies that are based on only a single data modality. Brain function is explained by complimentary concepts, namley context-invariant principles such as functional localization measured by dMRI and context-sensitive principles such as the expression of functional units measured by functional MRI (fMRI).

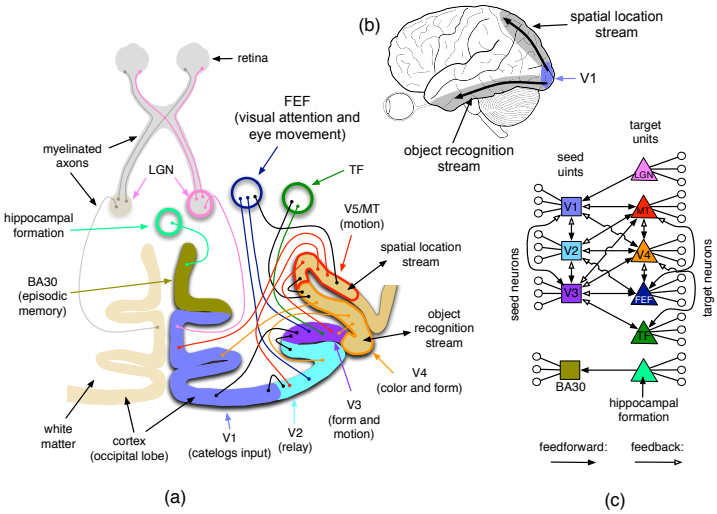

Figure 1: a) The seed region in the occipital lobe of the macaque monkey includes BA30, V1, V2 and V3. Target areas are limited to V4, V5/MT, TF, FEF, LGN and the hippocampal formation for illustration purposes. Only areas V1, V2, V3, V4, V5/MT and LGN were placed in an anatomical frame of reference. Each connection between areas represents an axon bundle. Binary connection information was taken from Hilgetag et al. (2000) [3]. b) Visual information processing can be seperated into two functional systems, the spatial location stream and object recognition stream. c) The block diagram is representative of the same connectome between seed and target functional units with feedforward and feedback connections.

- Population related inferences of functional organization that depend on cytoarchitectonically delineated areas have not been fruitful: "the history of cortical parcellation[1] has produced almost as many different definitions of cortical areas as there were eminent authors" [4]. The differences may be explained by interindividual variability but also by semantic disaccord and by incompatability between different methodologies used for cytoarchitectonic deliniations. Automated diffusion connectivity-based cortex parcellation can process a much larger population sample while conditioning on a single methodology. Due to the high variability of cortical convolutions among individuals such population studies may aid in explaining the relationship between functional localization and cortical morphology.

- Findings by Johansen-Berg et al. (2008) [5] support the feasibility of using diffusion connectivity-based cortex parcellation in guiding surgical interventions, such as deep brain stimulation, for treatment-resistant depression. In general, diffusion connectivity-based cortex parcellation can be useful for improving such therapeutic effects by optimizing the efficacy of stimulating a chosen cortical region within a functional network.

A downside is that the directionality of fibers are unidentifiable by diffusion connectomics. Consequently, diffusion connectomics alone cannot reveal the

[1] using cytoarchitectonic deliniations.

hierarchical information processing structure.

3.1.1 DTI Information Processing Pipeline

Figure 3.1.1 illustrates the pipeline used for connectivity-based cortex parcellation. The diffusion MRI scanner measures the orientation of axon bundles in the brain. Diffusion measurements are captured in a four dimensional matrix with the first three dimensions for spatial location of imaging voxels and the fourth dimension giving the diffusion direction. Image preprocessing includes removing high frequency noise and identifying grey matter, white matter and subcortical structures. The seed region is selected manually on the cortical boundary. The target connectivity regions include the entire cortical boundary and subcortical structures. Fiber tracking is subsequently used to generate n diffusion fibers for each seed voxel in order to empirically determine the connectivity between seed and target voxels. Connectivity measurements are captured in a connectivity matrix that contains connectivity scores between seed and target voxels (i.e. matrix entry i and j quantifies the number of fibers generated between seed voxel i and target voxel j). Finally, a connectivity structure relevant for functional organization is inferred from the connectivity matrix.

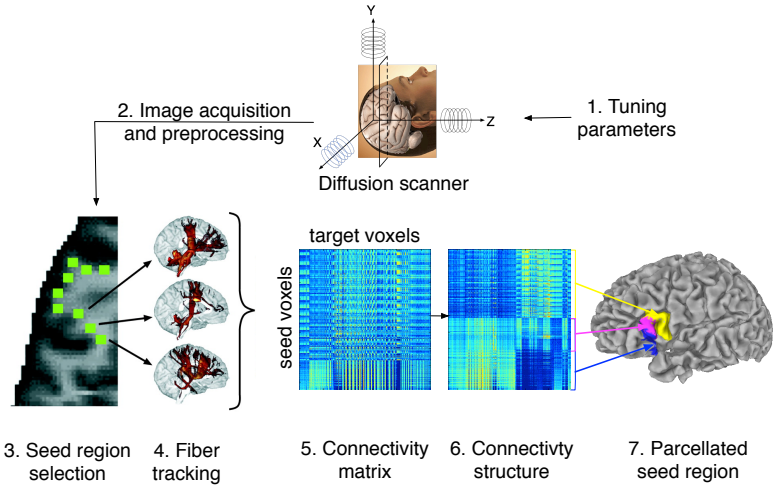

Figure 2: Diffusion tensor imaging is typically used for diffusion connectivity-based cortex parcellation and conforms to a standard pipeline producing a connectivity matrix. From a machine learning perspective, the crucial degrees of freedom are the tuning parameters the image preprocessing, fiber tracking and, most importantly, inferring the connectivity structure. More precisely, 6 illustrates a minimalistic projection of the connectivity structure onto the connectivity matrix.

3.1.2 Fiber tracking

Several probabilistic fiber tracking algorithms are described in literature that attempt to compensate for partial volume affects as well as improve its robustness against noise in the dMRI. Such tracking algorithms are single-directional or

multi-directional and are computed step-wise using local diffusion measurements or by a global measure-of-fit. In contrast to single-directional, multi-directional fiber tracking is capable of resolving two or more fiber trajectories within an imaging voxel but are unable to determine the number of trajectories. Global measure-of-fit fiber tracking is proposed to be less affected by noise than their local step-wise counterparts. A recent review on fiber tracking is provided by Bastiani et al. (2012) [6].

3.1.3 Clustering

Diffusion connectivity-based cortex parcellation has been modelled as single-assignment clustering of only seed voxels by minimizing centroid-based cost functions ranging from K-means [7], affinity propagation [8, 9] as well as the negative Gaussian log-likelihood [10] and a noncentroid-based cost function given by the information bottleneck method [11]. Each of the aforementioned methods have their own inference method; expectation-maximization for k-means clustering, minimizing a complexity term given by the mutual information between objects and clusters for the information bottleneck method and Gibbs sampling by utilizing a conjugate prior for the Gaussian likelihood function. The latter used a hierarchical Dirichlet model for multi-subject parcellation. My previous model was the first study to infer connectional families from diffusion connectomics [9].

One may, however, argue that inferences made by the aforementioned methods is not meaningful without extrinsic validation because the loss functions are not chosen based on the problem category; that is, previous methods have not adopted a problem-centric perspective for diffusion connectivity-based cortex parcellation [12]. Furthermore, meaningful inference by intrinsic validation to select the preferred loss function (i.e. model type) has not yet been implemented either with exception to my previous work [9, 11].

3.2 Equivalence with other Problems

Role mining, formulated by Frank et al. (2009) [13], belongs to the same category of problems involving matrix factorization. The true power of matrix factorization in the generalized clustering setting is its flexible, customizable design which enables it to be applied to a large category of problems. Similar to the goals of this thesis, Frank et al. (2009) [13] customized matrix factorization to incorporate overlapping clusters in order to provide a more suitable structural desciption of role mining. However, model design in role mining is considerably easier for the following reasons:

- Role mining is a well defined problem for which design decisions can be easily justified.
- The binary nature of role mining is less complex than the binomial nature of the connectivity matrix.

- The data structure is not obscured by phenomenal effects such as partial voluming in dMRI.
- Role mining typically does not condition on previous data processing with many degrees of freedom.
- The performance of a role mining model can be measured more easily.

4 Goals of Thesis

The primary goal of this thesis is to adopt a problem-centric perspective for diffusion connectivity-based cortex parcellation. To this end, we build upon my previous work [8, 9, 11] to provide a much richer inference of functional organization using a more suitable description of connectivity fingerprints and connectional families, quantified by diffusion connectomics. We therefore customize matrix factorization to include the following novel characteristics:

- *Biclustering*: The nature of connectivity in the functional network necessitates an analysis between regions which, in machine learning terms, translates to identifying seed and target clusters. Connectivity stengths between seed and target clusters reveal the distinguishability of functional units.
- *Connectivity strength quantified as a natural number*: Given a hard clustering framework, the connectivity strength is given by the number of diffusion tractograms connecting seed and target regions. We assume that axons do not branch in the white matter.
- *Overlapping clusters*: We attempt to compensate for partial volume effects by considering overlapping clusters.
- *Approximate hierarchy*: In contrast to the perfect hierarchy modelled previously [9], connectional families are described by an approximate topological seed hierarchy which is characterized by fuzzy assignments of objects to connectivity paths in the hierarchy. In the interest of pragmatic design, our model approximates the fuzzy hierarchy by hard multi-assignments of objects to connectivity paths in the hierarchy.

4.1 Further goals: quantifying the information content of the pipeline

Sources of noise in dMRI are multifarious ranging from thermal (Johnson) noise in the MRI electronic circuitry to motion artifacts [14] and physiological noise. In DTI noise is analytically expressed as a Rician distribution with residual variance. Such noise determines the information content of our model which can be used to remove degrees of freedom in the pipeline. More precisely, we

wish to select the optimal number of clusters, rank fiber tracking algorithms[2] and optimize tuning parameters such as the b-value.

Once more, intrinsic validation builds upon my previous work [9, 11] to select the model order. The channel capacity between two sampled connectivity structures quantifies the information content of diffusion parcellation pipeline which is given by a tradeoff of two measures of model uncertainty [15]. The first uncertainty is the entropy of the structure which is a measure of the structure information content in the absence of noise and can be computed in the same way as described by Buhmann (2010) [15]. The second uncertainty is the conditional entropy of the model which is a measure of the cardinality of statistically indistinguishable structures (i.e. the cardinality of the approximation set). Given that we have an analytical definition of noise in the data space we have two options for computing the cardinality of the approximation set for a given model complexity:

1. The cardinality of the approximation set can be estimated by analytically expressing the propagation of noise from the diffusion space (i.e. data space) to the connectivity structure space (i.e. solution space).

2. Two diffusion image datasets of the same brain can be acquired independently in order to sufficiently (hopefully) capture noise in the connectivity structure space. The novel coding scenario in approximation set coding can be used to characterize noise in the solution space and therefore the cardinality of the approximation set.

Approximation set coding, seems favourable since an analytical expression of noise propagation in the pipeline will likely be subject to heuristics. Note that approximation set coding suffers from the downside that it conditions on knowledge of deterministic, bijective correspondence between objects across sampled datasets; that is, for rich learning such correspondence information should ideally be available extrinsic to the data analysis. Fortunately, diffusion connectivity-based cortex parcellation caters for extrinsic, nearly deterministic and nearly bijective correspondence by providing another category of seed voxel (i.e. objects) description, namely their spatial location on the cortical surface.

4.2 Progress to date

4.2.1 Cortex parcellation by matrix factorization: formal definition

The key idea of using matrix factorization for connectivity-based cortex parcellation is to superimpose a connectivity structure onto the connectivity matrix shown in figure 2 which is reminiscent of the generative design of role mining by Frank et al. (2009) [13]. Any deviation from the connectivity structure is interpreted as noise. The nature of the connectivity matrix X is characterized by

[2]Note that the most elegant solution would be to condition our model directly on the diffusion images instead of the connectivity matrix thereby removing degrees of freedom in fiber tracking. This, however, seems to be a too ambitious endeavour.

natural numbers that quantify the strength of structural connectivity between seed voxels (i.e. objects associated with rows of the matrix) and target voxels (i.e. object descriptions associated with columns of the matrix). Matrix X is factorized into matrices Z, V, W and Y by the following multiplication rule[3]:

$$X = Z \times V \times W \times Y, \tag{1}$$

where W is a matrix containing natural numbers that quantify the strength of connectivity between seed and target clusters. W is further decomposed into a sum over Boolean matrices, $W = \sum_t^n G^{(t)}$, where t indexes a particular diffusion fiber and n is the number of diffusion fibers generated per seed voxel. The Boolean nature of matrices Z, V and Y fascilitate a straightforward interpretation of hard clustering. For example, diffusion fiber t connects seed voxel i and j by a path in the connectivity structure indexed by k_1, k_2 and l if $z_{ik_1} = 1$ ($z_{ik_1} \in Z$), $v_{k_1,k_2} = 1$ ($v_{ik_2} \in V$), $g^{(t)}_{k_2l} = 1$ ($g^{(t)}_{k_2l} \in G^{(t)}$) and $y_{lj} = 1$ ($y_{lj} \in Y$). More precisely, the bitstring $[z_{i1}\ v_{11}\ g_{11}\ y_{1j}, \ldots, z_{ik_1}\ v_{k_1k_2}\ g_{k_2l}\ y_{lj}, \ldots]$ of length $k_1 \times k_2 \times l$ identifies the connectivity structure between seed voxel i and target voxel j. Given that i and j connect along m paths, the number of possible connectivity structures between i and j is determined by the number of combinations of the bitstring, $\binom{k_1 \times k_2 \times l}{m}$.

In a *perfect hierarchy* (i.e. characterized by only single-assignments of objects to clusters) the set consisting of x_{ij} diffusion fibers between i and j connect along *only one unique path* (i.e. $m = 1$) in the connectivity structure. The likelihood of counting x_{ij} fibers is binomially distributed since each corresponding diffusion fiber has the same probability of successfully connecting i and j and the same probability of failure at each trial[4]:

$$p(x_{ij}|\mathbf{z}_{i*}, \mathbf{y}_{*j}) = \binom{n}{x_{ij}} p\{u_{ij} = 1|\mathbf{z}_{i*}, \mathbf{y}_{*j}\}^{x_{ij}} \left(p\{u_{ij} = 0|\mathbf{z}_{i*}, \mathbf{y}_{*j}\}\right)^{(n-x_{ij})}, \tag{2}$$

$$p\{u_{ij} = 1|\mathbf{z}_{i*}, \mathbf{y}_{*j}\} = \prod_{k_1k_2l} \left(1 - \alpha_{k_1k_2}\lambda_{k_2l}\right)^{z_{ik_1}y_{lj}}, \tag{3}$$

where u_{ij} is the connectivity bit whos probability is conditioned on the clustering (and therefore functional localization) of seed and target voxels on the first hierarchical level. Parameters $\alpha_{k_1k_2}$ and λ_{k_2l} substitute $p(v_{k_1k_2})$ and $p(g_{k_2l})$, respectively.

The *approximate hierarchy*, considered in this thesis, is characterized by multi-assignments of objects to clusters which implies that there are multiple paths in the connectivity structure that connect seed voxel i and target voxel j. In contrast to the perfect hierarchy, the fiber count x_{ij} is therefore divided among the different paths in the connectivity structure. The likelihood of counting x_{ij} fibers is thus multinomially distributed among the different connectivity paths[5].

[3]The seed hierarchy is restricted to two levels for simplification purposes but can easily be extended to more than two levels using the same methodology

[4]The derivation is lengthy and therefore not shown in this proposal.

[5]The formula for an approximate hierarchy is complex and therefore not shown in this proposal.

4.2.2 Inference and Learning

Clustering assignments z_{ik} and y_{lj} and parameters $\alpha_{k_1 k_2}$ and $\lambda_{k_2 l}$ are optimized by altering between an estimation step and a maximization step (i.e. EM algorithm). Hard assignments z_{ik} and y_{lj} are inferred by the estimation step which determines the posterior distribution over assignment sets, given the parameters. In the maximization step, the parameters $\alpha_{k_1 k_2}$ and $\lambda_{k_2 l}$ are learned by minimizing the free energy given by:

$$F := -\frac{1}{\beta} \log \mathcal{Z}, \tag{4}$$

where β is the inverse computational temperature and $\mathcal{Z}$ is the partition function given by the sum over Boltzmann weights, $\exp(-\beta R(X, Z, Y))$. The loss function, $R(X, Z, Y)$, depends on the connectivity structure and the connectivity matrix and is given by the negative log-likelihood:

$$R = -\log p(X|Z, Y), \tag{5}$$

$$p(X|Z, Y) = \prod_i \prod_j p(x_{ij}|\mathbf{z}_{i*}, \mathbf{y}_{*j}). \tag{6}$$

A means to compute the partition function is to estimate the loss function as a sum over potentials which in this case is the cost of assigning row i and column j to the seed hierarchical path k_1 and k_2 and target cluster l, respectively. Such potentials are not readily available from equation 5 and will be estimated using mean field approximation.

Figure 3 illustrates preliminary results for single assignment biclustering. Further analysis included inferring overlapping clusters and revealing hierarchi-

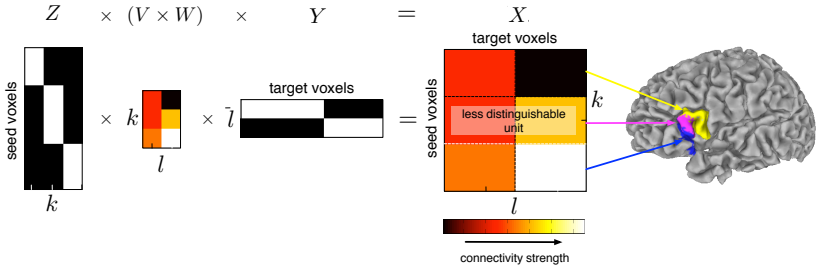

Figure 3: The connectivity matrix is factorized into matrices Z, $(V \times W)$ and Y. Functional localization in the cortical mantle is inferred from seed clusters (row clusters). Relations between seed clusters (row clusters) and target clusters (column clusters) are quantified by natural numbers and reveal the degree of distinction between functional units.

cal structure. Moreover, the advantage of biclustering will subsequently be fully realized by visualizing [16] the connectivity between seed and target clusters in the connectivity structure.

4.2.3 Intrinsic validation

Figure 4 demonstrates intrinsic validation using approximation set coding to determine the model order for pairwise clustering of the anterior prefrontal cortex (aPFC). Pairwise clustering (i.e. hard single assignment clustering of only seed voxels) was used as a preliminary loss function together with constant shift embedding to reduce the data size while optimally preserving clustering structure [17]. Noise limits the information content of pairwise clustering and therefore the number of effective clusters.

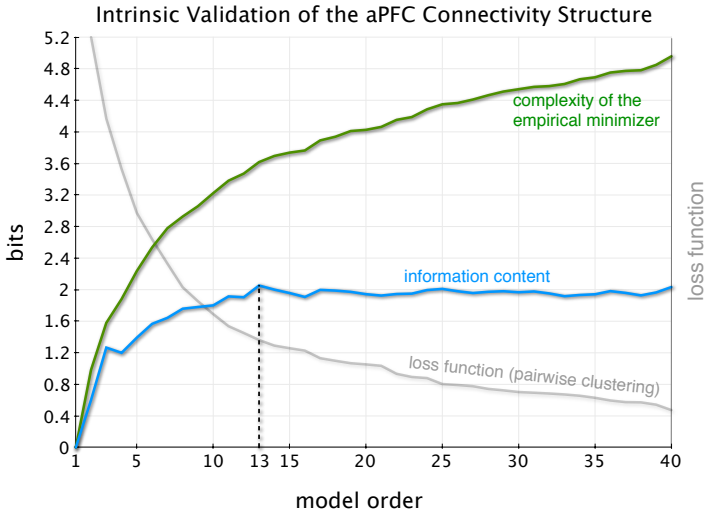

Figure 4: Approximation set coding fascilitates intrinsic validation of diffusion connectivity-based cortex parcellation by quantifying the information content of the pairwise clustering loss function with respect noise.

5 Work Plan: Work Packages

High angular resolution diffusion imaging (HARDI) data is provided by the Max Planck Institute for Neurological Research for multiple subjects under the supervision of Dr. Marc Tittgemeyer. The major work packages of this thesis are:

- a) Customizing matrix factorization for diffusion connectivity-based cortex parcellation and implementation.
- b) Quantify the information content of the pipeline to:
  1. determine the model order.
  2. rank diffusion fiber tracking algorithms.

3. rank different b-values used for diffusion image acquisition (optional).

- c) Thesis write-up.

Figure 5 shows the projected completion dates of my work packages.

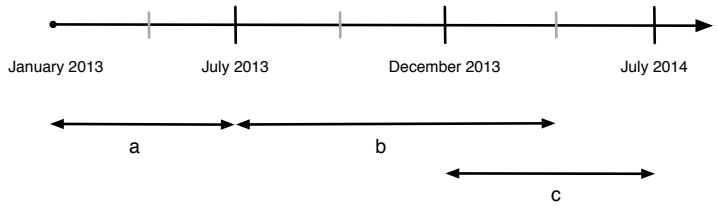

Figure 5: Projected completion dates of work packages.

Note that work package (b) builds upon my previous work to select the model order using approximation set coding [11]. An ongoing project with the Max Planck Institute for Neurological Research is to provide an interim automated parcellation algorithm and assist my colleagues at the MPI insitute in using it to parcellate the anterior prefrontal cortex (aPFC) of, at least 8 subjects. The automated parcellation algorithm is also required to include intrinsic validation by approximation set coding.

References

[1] K. Friston, "Functional integration and inference in the brain," *Progress Neurobiology*, vol. 590, pp. 1–31, 2002.

[2] S. Jbabdi and H. Johansen-Berg, "Tractography: Where do we go from here," *Brain Connectivity*, vol. 1, pp. 169–183, 2011.

[3] C. Hilgetag, G. Burns, M. O'Neill, J. Scannell and M. Young, "Anatomical connectivity defines the organization of clusters of cortical areas in the macaque monkey and the cat." *Phil Trans Royal Soc B*, vol. 355, pp. 91–110, 2000a.

[4] R. Kötter, K. Stephan, N. Palomero-Gallagher, S. Geyer, A. Schleicher and K. Zilles, "Multimodal characterization of cortical areas by multivariate analyses of receptor binding and connectivity data," *Anat*, vol. 204, pp. 333–350, 2001.

[5] H. Johansen-Berg, D. Gutman, T. Behrens, P. Mathews, M. Rushworth, E. Katz, A. Lozano and H. Mayberg, "Anatomical connectivity of the subgenual cingulate region targeted with deep brain stimulation for treatment-resistant depression," *Cerebral Cortex*, vol. 18, pp. 1374–1383, 2008.

[6] M. Bastiani, N. J. Shah, R. Goebel and A. Roebroek, "Human cortical connectome reconstractions from diffusion weighted mri: The effect of tractography algorithm," *Neuroimage*, vol. 62, pp. 1732–1749, 2012.

[7] A. Anwander, M. Tittgemeyer, D. von Cramon, A. Friederici and T. Knösche, "Connectivity-based parcellation of broca's area," *Cerebral Cortex*, vol. 17, pp. 826–825, 2007.

[8] N. S. Gorbach, C. Melzer, C. Schütte, K. Amunts, T. Douglas and M. Tittgemeyer, "Hierarchical clustering for connectivity-based cortex parcellation," in *16th Annual Meeting of the Organization for Human Brain Mapping*, 2010.

[9] N. S. Gorbach, C. Schütte, C. Melzer, M. Goldau, O. Sujazow, J. Jitsev, T. Douglas and M. Tittgemeyer, "Hierarchical information-based clustering for connectivity-based cortex parcellation," *Frontiers in Neuroinformatics*, vol. 5, pp. 1–13, 2011.

[10] S. Jbabdi, M. W. Woolrich and T. E. J. Behrens, "Multiple-subjects connectivity-based parcellation using hierarchical dirichlet process mixture models." *NeuroImage*, vol. 44, no. 2, pp. 373–384, 2009.

[11] N. S. Gorbach, S. Siep, J. Jitsev, C. Melzer and M. Tittgemeyer, "Information-theoretic connectivity-based cortex parcellation," in *Lecture Notes in Computer Science*, 2012.

[12] U. von Luxburg, R. C. Williamson and I. Guyon, "Clustering: Science or art?" *Journal of Machine Learning Research*, vol. 27, pp. 65–79, 2012.

[13] M. Frank, J. M. Buhmann and D. Basin, "A probabilistic approach to role mining." in *16th ACM Conference on Computer and Communications Security (CCS2009)*, 2009.

[14] D. K. Jones, *Diffusion MRI: Theory, Methods and Applications*, D. K. Jones, Ed. Oxford University Press, 2011.

[15] J. Buhmann, "Information theoretic model validation for clustering," *International Symposium on Information Theory*, 2010.

[16] M. Goldau, A. Wiebel, N. Gorbach, C.Melzer, M. Hlawitschka, G. Scheuermann and M. Tittgemeyer, "Fiber stippling: An illustrative rendering for probabilistic diffusion tractography," in *1st IEEE Symposium on Biological Data Visualization BioVis*, to appear.

[17] V. Roth, J. Laub, M. Kawanabe and J. M. B, "Optimal cluster preserving embedding of nonmetric proximity data," *IEEE Transactions on Pattern A*, vol. 25, pp. 1540–1551, 2003.

Prof. Dr. Joachim M. Buhmann **Nico Stephan Gorbach**

Zürich, January 17, 2014