IN PARTNERSHIP WITH:
Centre CEA-Saclay

Activity Report 2019

Project-Team PARIETAL

Modelling brain structure, function and variability based on high-field MRI data.

IN COLLABORATION WITH: CEA Neurospin

RESEARCH CENTER
Saclay - Île-de-France

THEME
Computational Neuroscience and Medicine

Table of contents

Project-Team PARIETAL

Creation of the Project-Team: 2009 July 01

Keywords:

Computer Science and Digital Science:

A3.3. - Data and knowledge analysis
A3.3.2. - Data mining
A3.3.3. - Big data analysis
A3.4. - Machine learning and statistics
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.4.4. - Optimization and learning
A3.4.5. - Bayesian methods
A3.4.6. - Neural networks
A3.4.7. - Kernel methods
A3.4.8. - Deep learning
A5.3.2. - Sparse modeling and image representation
A5.3.3. - Pattern recognition
A5.9.1. - Sampling, acquisition
A5.9.2. - Estimation, modeling
A5.9.6. - Optimization tools
A6.2.4. - Statistical methods
A6.2.6. - Optimization
A9.2. - Machine learning
A9.3. - Signal analysis

Other Research Topics and Application Domains:

B1.2. - Neuroscience and cognitive science
B1.2.1. - Understanding and simulation of the brain and the nervous system
B1.2.2. - Cognitive science
B2.2.6. - Neurodegenerative diseases
B2.6.1. - Brain imaging

1. Team, Visitors, External Collaborators

Research Scientists

Bertrand Thirion [Team leader, Inria, Senior Researcher, HDR]
Philippe Ciuciu [CEA, Senior Researcher, HDR]
Denis Alexander Engemann [Inria, Starting Research Position]
Alexandre Gramfort [Inria, Senior Researcher, HDR]
Thomas Moreau [Inria, Researcher, from Oct 2019]
Gaël Varoquaux [Inria, Researcher, HDR]
Demian Wassermann [Inria, Researcher]

Faculty Member

Anna Kazeykina [Univ Paris-Saclay, Associate Professor, from Sep 2019]

Post-Doctoral Fellow

Marine Le Morvan [CNRS, Post-Doctoral Fellow, from May 2019]

PhD Students

Pierre Ablin [Inria, PhD Student]
Thomas Bazeille [Inria, PhD Student]
Quentin Bertrand [Inria, PhD Student]
Patricio Cerda Reyes [Inria, PhD Student, until Oct 2019]
Hamza Cherkaoui [CEA, PhD Student]
Jerome-Alexis Chevalier [Inria, PhD Student]
Alexis Cvetkov Iliev [Inria, PhD Student, from Oct 2019]
Kamalaker Reddy Dadi [Inria, PhD Student]
Jérome Dockès [Inria, PhD Student]
Loubna El Gueddari [Ministère de l'Enseignement Supérieur et de la Recherche, PhD Student]
Valentin Iovene [Inria, PhD Student]
Hubert Jacob Banville [Interaxon, PhD Student]
Maeliss Jallais [Inria, PhD Student]
Hicham Janati [Inria, PhD Student]
Antonia Machlouzarides Shalit [Inria, PhD Student]
Lucas Martin [Inria, PhD Student, from Oct 2019]
Mathurin Massias [Inria, PhD Student]
Tuan Binh Nguyen [Inria, PhD Student]
Nicolas Prost [École polytechnique, PhD Student]
Zaccharie Ramzi [CEA, PhD Student, from Feb 2019]
Hugo Richard [Univ Paris-Saclay, PhD Student]
David Sabbagh [INSERM, PhD Student]
Gaston Zanitti [Inria, PhD Student, from Jun 2019]

Technical staff

Alexandre Abadie [Inria, Engineer]
Nicolas Chartier [CEA, Engineer, until Aug 2019]
Kshitij Chawla [Inria, Engineer]
L Emir Omar Chehab [Inria, Engineer, from Dec 2019]
Guillaume Daval Frerot [CEA, Engineer, from Nov 2019]
Jérémie Du Boisberranger [Inria, Engineer]
Merlin Dumeur [École Normale Supérieure de Lyon, Engineer, from Nov 2019]
Guillaume Favelier [Inria, Engineer]
Chaithya Giliyar Radhkrishna [CEA, Engineer, from Jul 2019]
Pierre Glaser [Inria, Engineer, until Sep 2019]
Ana Luisa Grilo Pinho [Inria, Engineer]
Olivier Grisel [Inria, Engineer, until Jan 2019]
Guillaume Lemaitre [Inria, Engineer, until Mar 2019]
Jiaping Liu [Inria, Engineer, from Jul 2019]
Joan Massich Vall [Inria, Engineer, until Oct 2019]
Romuald Menuet [Inria, Engineer, until May 2019]
Thomas Schmitt [Inria, Engineer, from Aug 2019]
Swetha Shankar [Inria, Engineer, from Oct 2019]
Maria Telenczuk [Inria, Engineer, from Feb 2019]
Juan Jesus Torre Tresols [Inria, Engineer]
Joris Van Den Bossche [Inria, Engineer, until Mar 2019]

Interns and Apprentices

Oleh Kozynets [Inria, from Feb 2019 until Aug 2019]
Agnieszka Lorens [Inria, from Apr 2019 until Sep 2019]

Administrative Assistant
Corinne Petitot [Inria, Administrative Assistant]

Visiting Scientist
Huchuan Xia [Univ. Pennsylvania, from Jun 2019 until Sep 2019]

External Collaborators
Danilo Bzdok [Univ. Aachen]
Luigi Gresele [MPI, from Oct 2019]
Joseph Salmon [Institut Telecom ex GET Groupe des Écoles des Télécommunications, HDR]
Darya Chyzhyk [Univ. San Sebastian]
Meyer Scetbon [ENS Paris Saclay]
Elvis Dohmatob [Criteo]
Arthur Mensch [Ecole Normale Supérieure]

2. Overall Objectives

2.1. Overall Objectives

The Parietal team focuses on mathematical methods for modeling and statistical inference based on neuroimaging data, with a particular interest in machine learning techniques and applications of human functional imaging. This general theme splits into four research axes:

- Modeling for neuroimaging population studies,
- Encoding and decoding models for cognitive imaging,
- Statistical and machine learning methods for large-scale data,
- Compressed-sensing for MRI.

Parietal is also strongly involved in open-source software development in scientific Python (machine learning) and for neuroimaging applications.

3. Research Program

3.1. Inverse problems in Neuroimaging

Many problems in neuroimaging can be framed as forward and inverse problems. For instance, brain population imaging is concerned with the *inverse problem* that consists in predicting individual information (behavior, phenotype) from neuroimaging data, while the corresponding *forward problem* boils down to explaining neuroimaging data with the behavioral variables. Solving these problems entails the definition of two terms: a loss that quantifies the goodness of fit of the solution (does the model explain the data well enough?), and a regularization scheme that represents a prior on the expected solution of the problem. These priors can be used to enforce some properties on the solutions, such as sparsity, smoothness or being piece-wise constant.

Let us detail the model used in typical inverse problem: Let $\mathbf{X}$ be a neuroimaging dataset as an $(n_{subjects}, n_{voxels})$ matrix, where $n_{subjects}$ and n_{voxels} are the number of subjects under study, and the image size respectively, $\mathbf{Y}$ a set of values that represent characteristics of interest in the observed population, written as $(n_{subjects}, n_{features})$ matrix, where $n_{features}$ is the number of characteristics that are tested, and $\mathbf{w}$ an array of shape $(n_{voxels}, n_{features})$ that represents a set of pattern-specific maps. In the first place, we may consider the columns $\mathbf{Y}_1, .., \mathbf{Y}_{n_{features}}$ of Y independently, yielding $n_{features}$ problems to be solved in parallel:

$$\mathbf{Y}_i = \mathbf{X}\mathbf{w}_i + \epsilon_i, \forall i \in \{1, .., n_{features}\},$$

where the vector contains $\mathbf{w}_i$ is the i^{th} row of $\mathbf{w}$. As the problem is clearly ill-posed, it is naturally handled in a regularized regression framework:

$$\widehat{w}_i = \text{argmin}_{w_i} \|\mathbf{Y}_i - \mathbf{X}\mathbf{w}_i\|^2 + \Psi(\mathbf{w}_i), \tag{1}$$

where Ψ is an adequate penalization used to regularize the solution:

$$\Psi(\mathbf{w}; \lambda_1, \lambda_2, \eta_1, \eta_2) = \lambda_1 \|\mathbf{w}\|_1 + \lambda_2 \|\mathbf{w}\|_2 + \eta_1 \|\nabla \mathbf{w}\|_{2,1} + \eta_2 \|\nabla \mathbf{w}\|_{2,2} \tag{2}$$

with $\lambda_1,\ \lambda_2,\ \eta_1,\ \eta_2 \geq 0$ (this formulation particularly highlights the fact that convex regularizers are norms or quasi-norms). In general, only one or two of these constraints is considered (hence is enforced with a non-zero coefficient):

- When $\lambda_1 > 0$ only (LASSO), and to some extent, when $\lambda_1, \lambda_2 > 0$ only (elastic net), the optimal solution $\mathbf{w}$ is (possibly very) sparse, but may not exhibit a proper image structure; it does not fit well with the intuitive concept of a brain map.
- Total Variation regularization (see Fig. 1) is obtained for ($\eta_1 > 0$ only), and typically yields a piece-wise constant solution. It can be associated with Lasso to enforce both sparsity and sparse variations.
- Smooth lasso is obtained with ($\eta_2 > 0$ and $\lambda_1 > 0$ only), and yields smooth, compactly supported spatial basis functions.

Note that, while the qualitative aspect of the solutions are very different, the predictive power of these models is often very close.

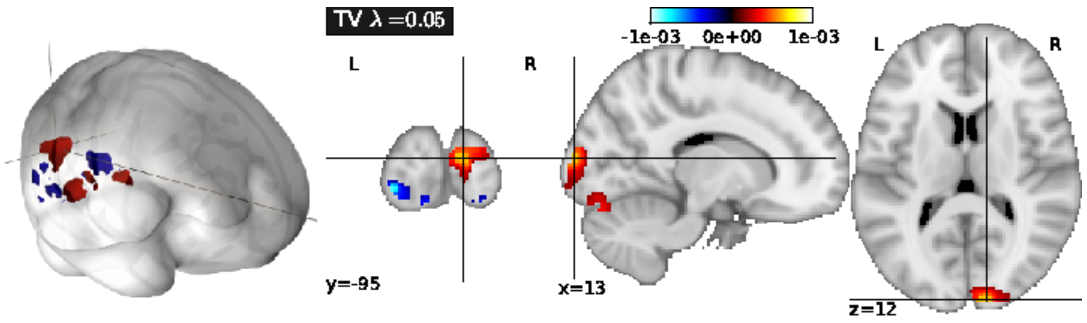

Figure 1. Example of the regularization of a brain map with total variation in an inverse problem. The problem here is to predict the spatial scale of an object presented as a stimulus, given functional neuroimaging data acquired during the presentation of an image. Learning and test are performed across individuals. Unlike other approaches, Total Variation regularization yields a sparse and well-localized solution that also enjoys high predictive accuracy.

The performance of the predictive model can simply be evaluated as the amount of variance in $\mathbf{Y}_i$ fitted by the model, for each $i \in \{1, .., n_{features}\}$. This can be computed through cross-validation, by *learning* $\widehat{\mathbf{w}}_i$ on some part of the dataset, and then estimating $\|\mathbf{Y}_i - \mathbf{X}\widehat{w}_i\|^2$ using the remainder of the dataset.

This framework is easily extended by considering

- *Grouped penalization*, where the penalization explicitly includes a prior clustering of the features, i.e. voxel-related signals, into given groups. This amounts to enforcing structured priors on the solution.
- *Combined penalizations*, i.e. a mixture of simple and group-wise penalizations, that allow some variability to fit the data in different populations of subjects, while keeping some common constraints.
- *Logistic and hinge regression*, where a non-linearity is applied to the linear model so that it yields a probability of classification in a binary classification problem.
- *Robustness to between-subject variability* to avoid the learned model overly reflecting a few outlying particular observations of the training set. Note that noise and deviating assumptions can be present in both $\mathbf{Y}$ and $\mathbf{X}$
- *Multi-task learning*: if several target variables are thought to be related, it might be useful to constrain the estimated parameter vector $\mathbf{w}$ to have a shared support across all these variables.
  For instance, when one of the variables $\mathbf{Y}_i$ is not well fitted by the model, the estimation of other variables $\mathbf{Y}_j, j \neq i$ may provide constraints on the support of $\mathbf{w}_i$ and thus, improve the prediction of $\mathbf{Y}_i$.

$$\mathbf{Y} = \mathbf{X}\mathbf{w} + \epsilon, \tag{3}$$

then

$$\widehat{w} = \text{argmin}_{\mathbf{w}=(\mathbf{w}_i), i=1..n_f} \sum_{i=1}^{n_f} \|\mathbf{Y_i} - \mathbf{X}\mathbf{w_i}\|^2 + \lambda \sum_{j=1}^{n_{voxels}} \sqrt{\sum_{i=1}^{n_f} \mathbf{w_{i,j}^2}} \tag{4}$$

3.2. Multivariate decompositions

Multivariate decompositions provide a way to model complex data such as brain activation images: for instance, one might be interested in extracting an *atlas of brain regions* from a given dataset, such as regions exhibiting similar activity during a protocol, across multiple protocols, or even in the absence of protocol (during resting-state). These data can often be factorized into spatial-temporal components, and thus can be estimated through *regularized Principal Components Analysis* (PCA) algorithms, which share some common steps with regularized regression.
Let $\mathbf{X}$ be a neuroimaging dataset written as an $(n_{subjects}, n_{voxels})$ matrix, after proper centering; the model reads

$$\mathbf{X} = \mathbf{A}\mathbf{D} + \epsilon, \tag{5}$$

where $\mathbf{D}$ represents a set of n_{comp} spatial maps, hence a matrix of shape (n_{comp}, n_{voxels}), and $\mathbf{A}$ the associated subject-wise loadings. While traditional PCA and independent components analysis (ICA) are limited to reconstructing components $\mathbf{D}$ within the space spanned by the column of $\mathbf{X}$, it seems desirable to add some constraints on the rows of $\mathbf{D}$, that represent spatial maps, such as sparsity, and/or smoothness, as it makes the interpretation of these maps clearer in the context of neuroimaging. This yields the following estimation problem:

$$\text{min}_{\mathbf{D},\mathbf{A}} \|\mathbf{X} - \mathbf{A}\mathbf{D}\|^2 + \Psi(\mathbf{D}) \text{ s.t. } \|\mathbf{A}_i\| = 1 \; \forall i \in \{1..n_{features}\}, \tag{6}$$

where $(\mathbf{A}_i),\ i \in \{1..n_{features}\}$ represents the columns of $\mathbf{A}$. Ψ can be chosen such as in Eq. (2) in order to enforce smoothness and/or sparsity constraints.

The problem is not jointly convex in all the variables but each penalization given in Eq (2) yields a convex problem on $\mathbf{D}$ for $\mathbf{A}$ fixed, and conversely. This readily suggests an alternate optimization scheme, where $\mathbf{D}$ and $\mathbf{A}$ are estimated in turn, until convergence to a local optimum of the criterion. As in PCA, the extracted components can be ranked according to the amount of fitted variance. Importantly, also, estimated PCA models can be interpreted as a probabilistic model of the data, assuming a high-dimensional Gaussian distribution (probabilistic PCA).

Ultimately, the main limitations to these algorithms is the cost due to the memory requirements: holding datasets with large dimension and large number of samples (as in recent neuroimaging cohorts) leads to inefficient computation. To solve this issue, online methods are particularly attractive [1].

3.3. Covariance estimation

Another important estimation problem stems from the general issue of learning the relationship between sets of variables, in particular their covariance. Covariance learning is essential to model the dependence of these variables when they are used in a multivariate model, for instance to study potential interactions among them and with other variables. Covariance learning is necessary to model latent interactions in high-dimensional observation spaces, e.g. when considering multiple contrasts or functional connectivity data.
The difficulties are two-fold: on the one hand, there is a shortage of data to learn a good covariance model from an individual subject, and on the other hand, subject-to-subject variability poses a serious challenge to the use of multi-subject data. While the covariance structure may vary from population to population, or depending on the input data (activation versus spontaneous activity), assuming some shared structure across problems, such as their sparsity pattern, is important in order to obtain correct estimates from noisy data. Some of the most important models are:

- **Sparse Gaussian graphical models**, as they express meaningful conditional independence relationships between regions, and do improve conditioning/avoid overfit.
- **Decomposable models**, as they enjoy good computational properties and enable intuitive interpretations of the network structure. Whether they can faithfully or not represent brain networks is still an open question.
- **PCA-based regularization of covariance** which is powerful when modes of variation are more important than conditional independence relationships.

Adequate model selection procedures are necessary to achieve the right level of sparsity or regularization in covariance estimation; the natural evaluation metric here is the out-of-sample likelihood of the associated Gaussian model. Another essential remaining issue is to develop an adequate statistical framework to test differences between covariance models in different populations. To do so, we consider different means of parametrizing covariance distributions and how these parametrizations impact the test of statistical differences across individuals.

4. Application Domains

4.1. Cognitive neuroscience

4.1.1. Macroscopic Functional cartography with functional Magnetic Resonance Imaging (fMRI)

The brain as a highly structured organ, with both functional specialization and a complex network organization. While most of the knowledge historically comes from lesion studies and animal electophysiological recordings, the development of non-invasive imaging modalities, such as fMRI, has made it possible to study routinely high-level cognition in humans since the early 90's. This has opened major questions on the interplay between mind and brain , such as: How is the function of cortical territories constrained by anatomy (connectivity) ? How to assess the specificity of brain regions ? How can one characterize reliably inter-subject differences ?

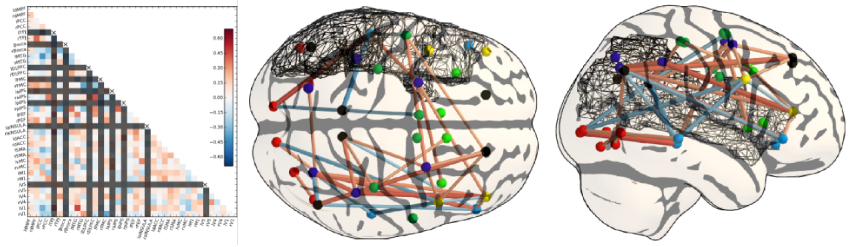

Figure 2. Example of functional connectivity analysis: The correlation matrix describing brain functional connectivity in a post-stroke patient (lesion volume outlined as a mesh) is compared to a group of control subjects. Some edges of the graphical model show a significant difference, but the statistical detection of the difference requires a sophisticated statistical framework for the comparison of graphical models.

4.1.2. Analysis of brain Connectivity

Functional connectivity is defined as the interaction structure that underlies brain function. Since the beginning of fMRI, it has been observed that remote regions sustain high correlation in their spontaneous activity, i.e. in the absence of a driving task. This means that the signals observed during resting-state define a signature of the connectivity of brain regions. The main interest of resting-state fMRI is that it provides easy-to-acquire functional markers that have recently been proved to be very powerful for population studies.

4.1.3. Modeling of brain processes (MEG)

While fMRI has been very useful in defining the function of regions at the mm scale, Magneto-encephalography (MEG) provides the other piece of the puzzle, namely temporal dynamics of brain activity, at the ms scale. MEG is also non-invasive. It makes it possible to keep track of precise schedule of mental operations and their interactions. It also opens the way toward a study of the rhythmic activity of the brain. On the other hand, the localization of brain activity with MEG entails the solution of a hard inverse problem.

4.1.4. Current challenges in human neuroimaging (acquisition+analysis)

Human neuroimaging targets two major goals: *i)* the study of neural responses involved in sensory, motor or cognitive functions, in relation to models from cognitive psychology, i.e. the identification of neurophysiological and neuroanatomical correlates of cognition; *ii)* the identification of markers in brain structure and function of neurological or psychiatric diseases. Both goals have to deal with a tension between

- the search for higher spatial [1] resolution to increase **spatial specificity** of brain signals, and clarify the nature (function and structure) of brain regions. This motivates efforts for high-field imaging and more efficient acquisitions, such as compressed sensing schemes, as well as better source localization methods from M/EEG data.
- the importance of inferring brain features with **population-level** validity, hence, contaminated with high variability within observed cohorts, which blurs the information at the population level and ultimately limits the spatial resolution of these observations.

Importantly, the signal-to-noise ratio (SNR) of the data remains limited due to both resolution improvements [2] and between-subject variability. Altogether, these factors have led to realize that results of neuroimaging studies were **statistically weak**, i.e. plagued with low power and leading to unreliable inference [72], and

[1] and to some extent, temporal, but for the sake of simplicity we focus here on spatial aspects.

[2] The SNR of the acquired signal is proportional to the voxel size, hence an improvement by a factor of 2 in image resolution along each dimension is payed by a factor of 8 in terms of SNR.

particularly so due to the typically number of subjects included in brain imaging studies (20 to 30, this number tends to increase [73]): this is at the core of the *neuroimaging reproducibility crisis*. This crisis is deeply related to a second issue, namely that only few neuroimaging datasets are publicly available, making it impossible to re-assess a posteriori the information conveyed by the data. Fortunately, the situation improves, lead by projects such as NeuroVault or OpenfMRI. A framework for integrating such datasets is however still missing.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- November 2019: Académie des sciences / Dassault System / Inria prize awarded to the developers of scikit-learn: Loic Estève, Alexandre Gramfort, Olivier Grisel, Bertrand Thirion and Gael Varoquaux.
- December 2019: Alexandre Gramfort, Bertrand Thirion and Gael Varoquaux are each awarded a *Chaire IA* following a national call.
- December 2019: Carole Lazarus, former PhD student supervised by Philippe Ciuciu got the Prix de la Chancellerie des Universités de Paris 2019 - Section Sciences.

6. New Software and Platforms

6.1. Mayavi

FUNCTIONAL DESCRIPTION: Mayavi is the most used scientific 3D visualization Python software. Mayavi can be used as a visualization tool, through interactive command line or as a library. It is distributed under Linux through Ubuntu, Debian, Fedora and Mandriva, as well as in PythonXY and EPD Python scientific distributions. Mayavi is used by several software platforms, such as PDE solvers (fipy, sfepy), molecule visualization tools and brain connectivity analysis tools (connectomeViewer).

- Contact: Gaël Varoquaux
- URL: http://mayavi.sourceforge.net/

6.2. Nilearn

NeuroImaging with scikit learn

KEYWORDS: Health - Neuroimaging - Medical imaging

FUNCTIONAL DESCRIPTION: NiLearn is the neuroimaging library that adapts the concepts and tools of scikit-learn to neuroimaging problems. As a pure Python library, it depends on scikit-learn and nibabel, the main Python library for neuroimaging I/O. It is an open-source project, available under BSD license. The two key components of NiLearn are i) the analysis of functional connectivity (spatial decompositions and covariance learning) and ii) the most common tools for multivariate pattern analysis. A great deal of efforts has been put on the efficiency of the procedures both in terms of memory cost and computation time.

- Participants: Alexandre Abraham, Alexandre Gramfort, Bertrand Thirion, Elvis Dohmatob, Fabian Pedregosa Izquierdo, Gaël Varoquaux, Loïc Estève, Michael Eickenberg and Virgile Fritsch
- Contact: Bertrand Thirion
- URL: http://nilearn.github.io/

6.3. Scikit-learn

KEYWORDS: Regession - Clustering - Learning - Classification - Medical imaging

SCIENTIFIC DESCRIPTION: Scikit-learn is a Python module integrating classic machine learning algorithms in the tightly-knit scientific Python world. It aims to provide simple and efficient solutions to learning problems, accessible to everybody and reusable in various contexts: machine-learning as a versatile tool for science and engineering.

FUNCTIONAL DESCRIPTION: Scikit-learn can be used as a middleware for prediction tasks. For example, many web startups adapt Scikitlearn to predict buying behavior of users, provide product recommendations, detect trends or abusive behavior (fraud, spam). Scikit-learn is used to extract the structure of complex data (text, images) and classify such data with techniques relevant to the state of the art.

Easy to use, efficient and accessible to non datascience experts, Scikit-learn is an increasingly popular machine learning library in Python. In a data exploration step, the user can enter a few lines on an interactive (but non-graphical) interface and immediately sees the results of his request. Scikitlearn is a prediction engine . Scikit-learn is developed in open source, and available under the BSD license.

- Participants: Alexandre Gramfort, Bertrand Thirion, Fabian Pedregosa Izquierdo, Gaël Varoquaux, Loïc Estève, Michael Eickenberg and Olivier Grisel
- Partners: CEA - Logilab - Nuxeo - Saint Gobain - Tinyclues - Telecom Paris
- Contact: Olivier Grisel
- URL: http://scikit-learn.org

6.4. MODL

Massive Online Dictionary Learning

KEYWORDS: Pattern discovery - Machine learning

FUNCTIONAL DESCRIPTION: Matrix factorization library, usable on very large datasets, with optional sparse and positive factors.

- Participants: Arthur Mensch, Gaël Varoquaux, Bertrand Thirion and Julien Mairal
- Contact: Arthur Mensch
- Publications: Subsampled online matrix factorization with convergence guarantees - hal-01431618v3
- URL: http://github.com/arthurmensch/modl

6.5. MNE

MNE-Python

KEYWORDS: Neurosciences - EEG - MEG - Signal processing - Machine learning

FUNCTIONAL DESCRIPTION: Open-source Python software for exploring, visualizing, and analyzing human neurophysiological data: MEG, EEG, sEEG, ECoG, and more.

RELEASE FUNCTIONAL DESCRIPTION: http://martinos.org/mne/stable/whats_new.html

- Partners: HARVARD Medical School - New York University - University of Washington - CEA - Aalto university - Telecom Paris - Boston University - UC Berkeley
- Contact: Alexandre Gramfort
- URL: http://martinos.org/mne/

6.6. Dmipy

Diffusion MRI Multi-Compartment Modeling and Microstructure Recovery Made Easy

KEYWORDS: Diffusion MRI - Multi-Compartment Modeling - Microstructure Recovery

FUNCTIONAL DESCRIPTION: Non-invasive estimation of brain microstructure features using diffusion MRI (dMRI) – known as Microstructure Imaging – has become an increasingly diverse and complicated field over the last decades. Multi-compartment (MC)-models, representing the measured diffusion signal as a linear combination of signal models of distinct tissue types, have been developed in many forms to estimate these features. However, a generalized implementation of MC-modeling as a whole, providing deeper insights in its capabilities, remains missing. To address this fact, we present Diffusion Microstructure Imaging in Python (Dmipy), an open-source toolbox implementing PGSE-based MC-modeling in its most general form. Dmipy allows on-the-fly implementation, signal modeling, and optimization of any user-defined MC-model, for any PGSE acquisition scheme. Dmipy follows a "building block"-based philosophy to Microstructure Imaging, meaning MC-models are modularly constructed to include any number and type of tissue models, allowing simultaneous representation of a tissue's diffusivity, orientation, volume fractions, axon orientation dispersion, and axon diameter distribution. In particular, Dmipy is geared toward facilitating reproducible, reliable MC-modeling pipelines, often allowing the whole process from model construction to parameter map recovery in fewer than 10 lines of code. To demonstrate Dmipy's ease of use and potential, we implement a wide range of well-known MC-models, including IVIM, AxCaliber, NODDI(x), Bingham-NODDI, the spherical mean-based SMT and MC-MDI, and spherical convolution-based single- and multi-tissue CSD. By allowing parameter cascading between MC-models, Dmipy also facilitates implementation of advanced approaches like CSD with voxel-varying kernels and single-shell 3-tissue CSD. By providing a well-tested, user-friendly toolbox that simplifies the interaction with the otherwise complicated field of dMRI-based Microstructure Imaging, Dmipy contributes to more reproducible, high-quality research.

- Authors: Rutger Fick, Demian Wassermann and Rachid Deriche
- Contact: Rachid Deriche

6.7. PySAP

Python Sparse data Analysis Package

KEYWORDS: Image reconstruction - Image compression

FUNCTIONAL DESCRIPTION: The PySAP (Python Sparse data Analysis Package, https://github.com/CEA-COSMIC/pysap) open-source image processing software package has been developed for the 3 years between the Compressed Sensing group at Iniria-CEA Parietal team led by Philippe Ciuciu and the CosmoStat team (CEA/IRFU) led by Jean-Luc Statck. It has been developed for the COmpressed Sensing for Magnetic resonance Imaging and Cosmology (COSMIC) project. This package provides a set of flexible tools that can be applied to a variety of compressed sensing and image reconstruction problems in various research domains. In particular, PySAP offers fast wavelet transforms and a range of integrated optimisation algorithms. It also offers a variety of plugins for specific application domains: on top of Pysap-MRI and PySAP-astro plugins, several complementary modules are now in development for electron tomography and electron microscopy for CEA colleagues. In October 2019, PySAP has been released on PyPi (https://pypi.org/project/python-pySAP/, currently version 0.0.3) and in conda (https://anaconda.org/agrigis/python-pysap).

The Pysap-MRI has been advertised through a specific abstract accepted to the next workshop of ISMRM on Data Sampling & Image Reconstruction in late January 2020. It will be presented during a power pitch session together wih an hands-on demo session using JuPyter notebooks.

- Partner: CEA
- Contact: Philippe Ciuciu

7. New Results

7.1. The visual word form area (VWFA) is part of both language and attention circuitry

While predominant models of visual word form area (VWFA) function argue for its specific role in decoding written language, other accounts propose a more general role of VWFA in complex visual processing. However, a comprehensive examination of structural and functional VWFA circuits and their relationship to behavior has been missing. Here, using high-resolution multimodal imaging data from a large Human Connectome Project cohort (N = 313), we demonstrate robust patterns of VWFA connectivity with both canonical language and attentional networks. Brain-behavior relationships revealed a striking pattern of double dissociation: structural connectivity of VWFA with lateral temporal language network predicted language, but not visuo-spatial attention abilities, while VWFA connectivity with dorsal fronto-parietal attention network predicted visuo-spatial attention, but not language abilities. Our findings support a multiplex model of VWFA function characterized by distinct circuits for integrating language and attention, and point to connectivity-constrained cognition as a key principle of human brain organization.

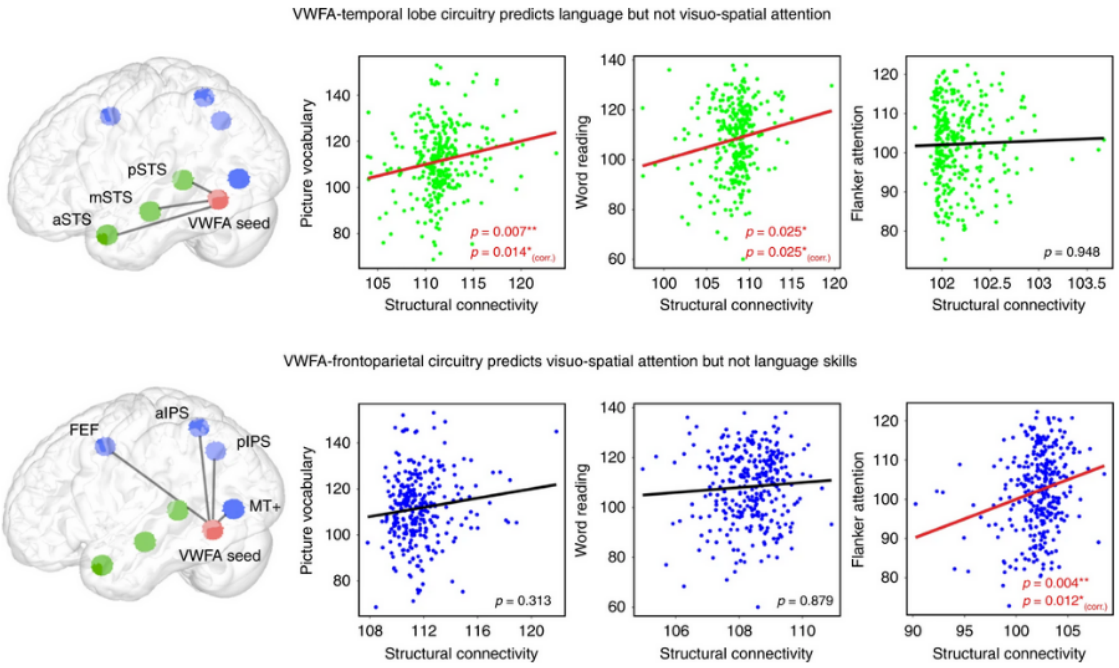

Figure 3. Structural connectivity of VWFA with STS nodes (anterior, middle, and posterior STS) was significantly correlated with individuals' performance on picture vocabulary and word reading tasks, but not on the Flanker visuo-spatial attention task. b Structural connectivity of VWFA with fronto-parietal attention network nodes (FEF, aIPS, pIPS, and MT+) was significantly correlated with individuals' performance on the Flanker attention task but not on the word reading or picture vocabulary tasks. Y-axis is age-adjusted performance scores on cognitive tasks and x-axis presents the predicted performance scores on cognitive tasks from the probability of structural connectivity of VWFA to either the language or the attention network ROIs.

More information can be found in [7].

7.2. SPARKLING: variable-density k-space filling curves for accelerated T2☆ -weighted MRI

Funding information Purpose: To present a new optimization-driven design of optimal k-space trajectories in the context of compressed sensing: Spreading Projection Algorithm for Rapid K-space sampLING

(SPARKLING). Theory: The SPARKLING algorithm is a versatile method inspired from stippling techniques that automatically generates optimized sampling patterns compatible with MR hardware constraints on maximum gradient amplitude and slew rate. These non-Cartesian sampling curves are designed to comply with key criteria for optimal sampling: a controlled distribution of samples (e.g., variable density) and a locally uniform k-space coverage. Methods: Ex vivo and in vivo prospective $T2^{\star}$-weighted acquisitions were performed on a 7 Tesla scanner using the SPARKLING tra-jectories for various setups and target densities. Our method was compared to radial and variable-density spiral trajectories for high resolution imaging. Results: Combining sampling efficiency with compressed sensing, the proposed sampling patterns allowed up to 20-fold reductions in MR scan time (compared to fully-sampled Cartesian acquisitions) for two-dimensional $T2^{\star}$-weighted imaging without deterioration of image quality, as demonstrated by our experimental results at 7 Tesla on in vivo human brains for a high in-plane resolution of 390 um. In comparison to existing non-Cartesian sampling strategies, the proposed technique also yielded superior image quality. Conclusion: The proposed optimization-driven design of k-space trajectories is a versatile framework that is able to enhance MR sampling performance in the context of compressed sensing.

More information can be found in [14].

7.3. Benchmarking functional connectome-based predictive models for resting-state fMRI

Functional connectomes reveal biomarkers of individual psychological or clinical traits. However, there is great variability in the analytic pipelines typically used to derive them from rest-fMRI cohorts. Here, we consider a specific type of studies, using predictive models on the edge weights of functional connectomes, for which we highlight the best modeling choices. We systematically study the prediction performances of models in 6 different cohorts and a total of 2 000 individuals, encompassing neuro-degenerative (Alzheimer's, Post-traumatic stress disorder), neuro-psychiatric (Schizophrenia, Autism), drug impact (Cannabis use) clinical settings and psychological trait (fluid intelligence). The typical prediction procedure from rest-fMRI consists of three main steps: defining brain regions, representing the interactions, and supervised learning. For each step we benchmark typical choices: 8 different ways of defining regions –either pre-defined or generated from the rest-fMRI data– 3 measures to build functional connectomes from the extracted time-series, and 10 classification models to compare functional interactions across subjects. Our benchmarks summarize more than 240 different pipelines and outline modeling choices that show consistent prediction performances in spite of variations in the populations and sites. We find that regions defined from functional data work best; that it is beneficial to capture between-region interactions with tangent-based parametrization of covariances, a midway between correlations and partial correlation; and that simple linear predictors such as a logistic regression give the best predictions. Our work is a step forward to establishing reproducible imaging-based biomarkers for clinical settings.

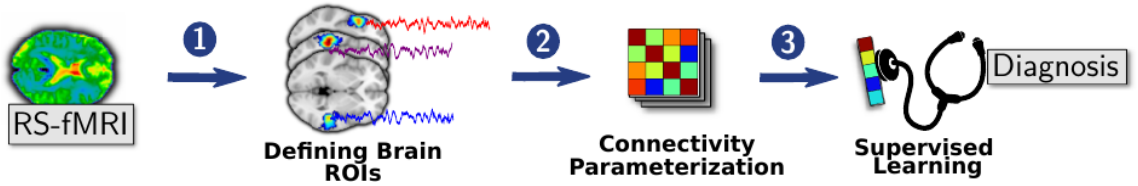

Figure 4. Functional connectome-prediction pipeline with three main steps: 1) definition of brain regions (ROIs) from rest-fMRI images or us- ing already defined reference atlases, 2) quantifying functional interactions from time series signals extracted from these ROIs and 3) comparisons of functional interactions across subjects using supervised learning.

More information can be found in [8].

7.4. Population shrinkage of covariance (PoSCE) for better individual brain functional-connectivity estimation

Estimating covariances from functional Magnetic Resonance Imaging at rest (r-fMRI) can quantify interactions between brain regions. Also known as brain functional connectivity, it reflects inter-subject variations in behavior and cognition, and characterizes neuropathologies. Yet, with noisy and short time-series, as in r-fMRI, covariance estimation is challenging and calls for penalization, as with shrinkage approaches. We introduce population shrinkage of covariance estimator (PoSCE) : a covariance estimator that integrates prior knowledge of covariance distribution over a large population, leading to a non-isotropic shrinkage. The shrinkage is tailored to the Riemannian geometry of symmetric positive definite matrices. It is coupled with a probabilistic modeling of the individual and population covariance distributions. Experiments on two large r-fMRI datasets (HCP n=815, Cam-CAN n=626) show that PoSCE has a better bias-variance trade-off than existing covariance estimates: this estimator relates better functional-connectivity measures to cognition while capturing well intra-subject functional connectivity.

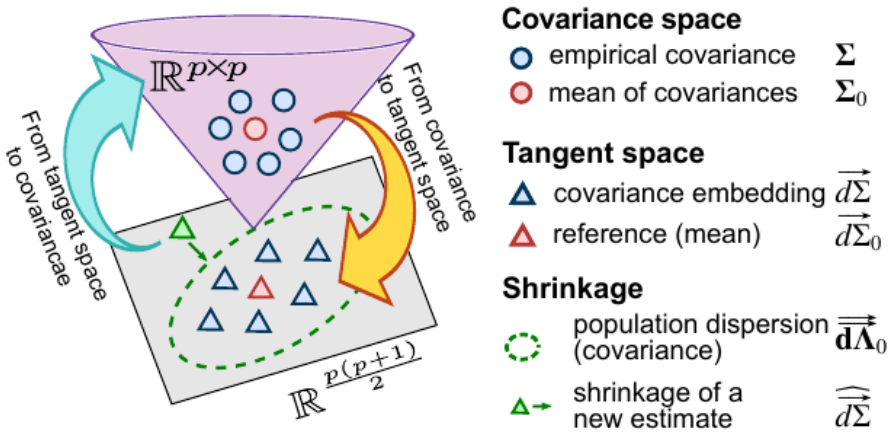

Figure 5. Tangent embedding and population prior modeling. Σ_0 is the mean covariance from a train set of covariances. It is the reference point in the tangent space. The population prior is defined as a Gaussian multivariate dis- tribution centered on $d\Sigma_0$. Λ_0 is the covariance dispersion over the population. The arrows depict the mapping between the non-Euclidean covariance space and the tangent space.

More information can be found in [20].

7.5. Feature Grouping as a Stochastic Regularizer for High-Dimensional Structured Data

In many applications where collecting data is expensive , for example neuroscience or medical imaging, the sample size is typically small compared to the feature dimension. It is challenging in this setting to train expressive, non-linear models without overfitting. These datasets call for intelligent regularization that exploits known structure, such as correlations between the features arising from the measurement device. However, existing structured regularizers need specially crafted solvers, which are difficult to apply to complex models. We propose a new regularizer specifically designed to leverage structure in the data in a way that can be applied efficiently to complex models. Our approach relies on feature grouping, using a fast clustering algorithm inside

a stochas-tic gradient descent loop: given a family of feature groupings that capture feature covariations, we randomly select these groups at each iteration. We show that this approach amounts to enforcing a denoising regularizer on the solution. The method is easy to implement in many model archi-tectures, such as fully connected neural networks, and has a linear computational cost. We apply this regularizer to a real-world fMRI dataset and the Olivetti Faces datasets. Experiments on both datasets demonstrate that the proposed approach produces models that generalize better than those trained with conventional regularizers, and also improves convergence speed.

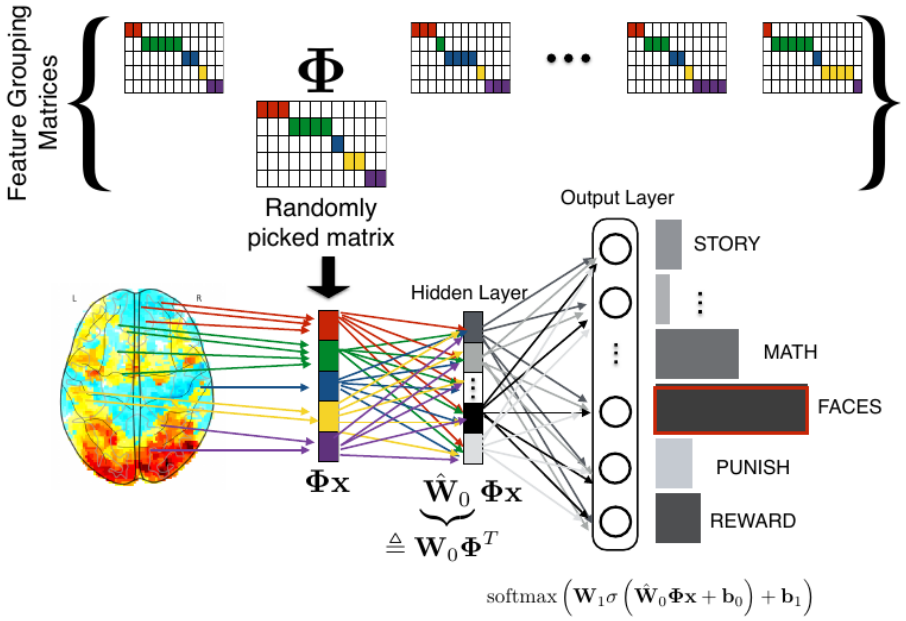

Figure 6. Illustration of the proposed approach: Forward propa- gation of a neural network with a single hidden layer using feature grouping during training. The parameters of the neural network to be estimated are $\mathbf{W}_0$, $\mathbf{b}_0$, $\mathbf{W}_1$, $\mathbf{b}1$. *A bank of feature grouping matrices are pre-generated where each matrix is calculated from a sub-sample of the training test. At each SGD iteration, a feature grouping matrix is sampled from the bank of pre-generated matri- ces. The gradient is then computed with respect to* $\widehat{\mathbf{W}}_0$ *to update* $\mathbf{W}_0$ *in backpropagation.*

More information can be found in [25].

7.6. Manifold-regression to predict from MEG/EEG brain signals without source modeling

Magnetoencephalography and electroencephalography (M/EEG) can reveal neuronal dynamics non-invasively in real-time and are therefore appreciated methods in medicine and neuroscience. Recent advances in modeling brain-behavior relationships have highlighted the effectiveness of Riemannian geometry for summarizing the spatially correlated time-series from M/EEG in terms of their covariance. However , after artefact-suppression, M/EEG data is often rank deficient which limits the application of Riemannian concepts. In this article, we focus on the task of regression with rank-reduced covariance matrices. We study two Riemannian approaches that vectorize the M/EEG covariance between-sensors through projection into a tangent space. The Wasserstein distance readily applies to rank-reduced data but lacks affine-invariance. This can be overcome by finding a common sub-space in which the covariance matrices are full rank, enabling the affine-invariant geometric

distance. We investigated the implications of these two approaches in synthetic generative models, which allowed us to control estimation bias of a linear model for prediction. We show that Wasserstein and geometric distances allow perfect out-of-sample prediction on the generative models. We then evaluated the methods on real data with regard to their effectiveness in predicting age from M/EEG covariance matrices. The findings suggest that the data-driven Riemannian methods outperform different sensor-space estimators and that they get close to the performance of biophysics-driven source-localization model that requires MRI acquisitions and tedious data processing. Our study suggests that the proposed Riemannian methods can serve as fundamental building-blocks for automated large-scale analysis of M/EEG.

Figure 7. Age prediction on Cam-CAN MEG dataset for different methods, ordered by out-of-sample MAE. The y-axis depicts the projection method, with identity denoting the absence of projection. Colors indicate the subsequent embedding. The biophysics-driven MNE method (blue) performs best. The Riemannian methods (orange) follow closely and their performance depends little on the projection method. The non-Riemannian methods log-diag (green) perform worse, although the supervised projection clearly helps.

More information can be found in [47].

7.7. Stochastic algorithms with descent guarantees for ICA

Independent component analysis (ICA) is a widespread data exploration technique, where observed signals are modeled as linear mixtures of independent components. From a machine learning point of view, it amounts to a matrix factorization problem with a statistical independence criterion. Infomax is one of the most used ICA algorithms. It is based on a loss function which is a non-convex log-likelihood. We develop a new majorization-minimization framework adapted to this loss function. We derive an online algorithm for the streaming setting, and an incremental algorithm for the finite sum setting, with the following benefits. First, unlike most algorithms found in the literature, the proposed methods do not rely on any critical hyper-parameter like a step size, nor do they require a line-search technique. Second, the algorithm for the finite sum setting, although

stochas-tic, guarantees a decrease of the loss function at each iteration. Experiments demonstrate progress on the state-of-the-art for large scale datasets, without the necessity for any manual parameter tuning.

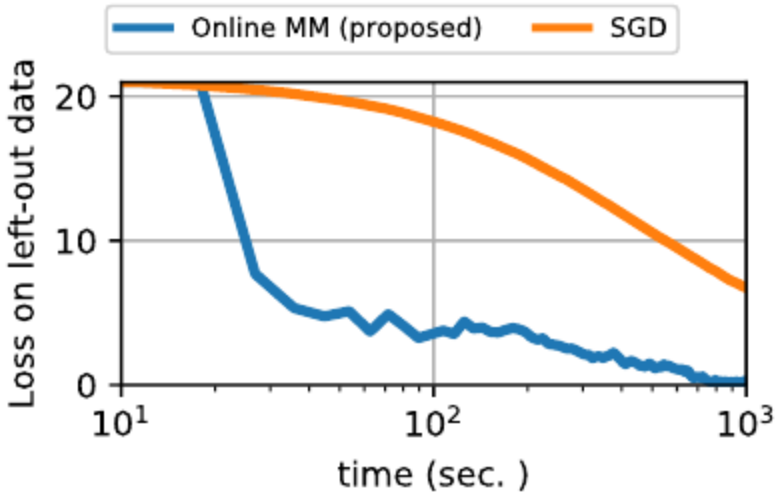

Figure 8. Online algorithms applied on a 32 GB real dataset with p = 100 and n = 4×10^7 . Time is in logarithmic scale. Values of the loss on left out data greater than its initial value are truncated.

More information can be found in [39].

7.8. Comparing distributions: $l1$ geometry improves kernel two-sample testing

Are two sets of observations drawn from the same distribution? This problem is a two-sample test. Kernel methods lead to many appealing properties. Indeed state-of-the-art approaches use the L2 distance between kernel-based distribution representatives to derive their test statistics. Here, we show that Lp distances (with $p \geq 1$) between these distribution representatives give metrics on the space of distributions that are well-behaved to detect differences between distributions as they metrize the weak convergence. Moreover, for analytic kernels, we show that the L1 geometry gives improved testing power for scalable computational procedures. Specifically, we derive a finite dimensional approximation of the metric given as the ℓ_1 norm of a vector which captures differences of expectations of analytic functions evaluated at spatial locations or frequencies (i.e, features). The features can be chosen to maximize the differences of the distributions and give interpretable indications of how they differs. Using an ℓ_1 norm gives better detection because differences between representatives are dense as we use analytic kernels (non-zero almost everywhere). The tests are consistent, while much faster than state-of-the-art quadratic-time kernel-based tests. Experiments on artificial and real-world problems demonstrate improved power/time tradeoff than the state of the art, based on ℓ_2 norms, and in some cases, better outright power than even the most expensive quadratic-time tests. More information can be found in [37].

7.9. Wasserstein regularization for sparse multi-task regression

We focus in this work on high-dimensional regression problems where each regressor can be associated to a location in a physical space, or more generally a generic geometric space. Such problems often employ sparse priors, which promote models using a small subset of regressors. To increase statistical power, the so-called multi-task techniques were proposed, which consist in the simultaneous estimation of several related models.

Combined with sparsity assumptions, it lead to models enforcing the active regressors to be shared across models, thanks to, for instance L1 / Lq norms. We argue in this paper that these techniques fail to leverage the spatial information associated to regressors. Indeed, while sparse priors enforce that only a small subset of variables is used, the assumption that these regressors overlap across all tasks is overly simplistic given the spatial variability observed in real data. In this paper, we propose a convex reg-ularizer for multi-task regression that encodes a more flexible geometry. Our regularizer is based on unbalanced optimal transport (OT) theory, and can take into account a prior geometric knowledge on the regressor variables, without necessarily requiring overlapping supports. We derive an efficient algorithm based on a regularized formulation of OT, which iterates through applications of Sinkhorn's algorithm along with coordinate descent iterations. The performance of our model is demonstrated on regular grids with both synthetic and real datasets as well as complex triangulated geometries of the cortex with an application in neuroimaging.

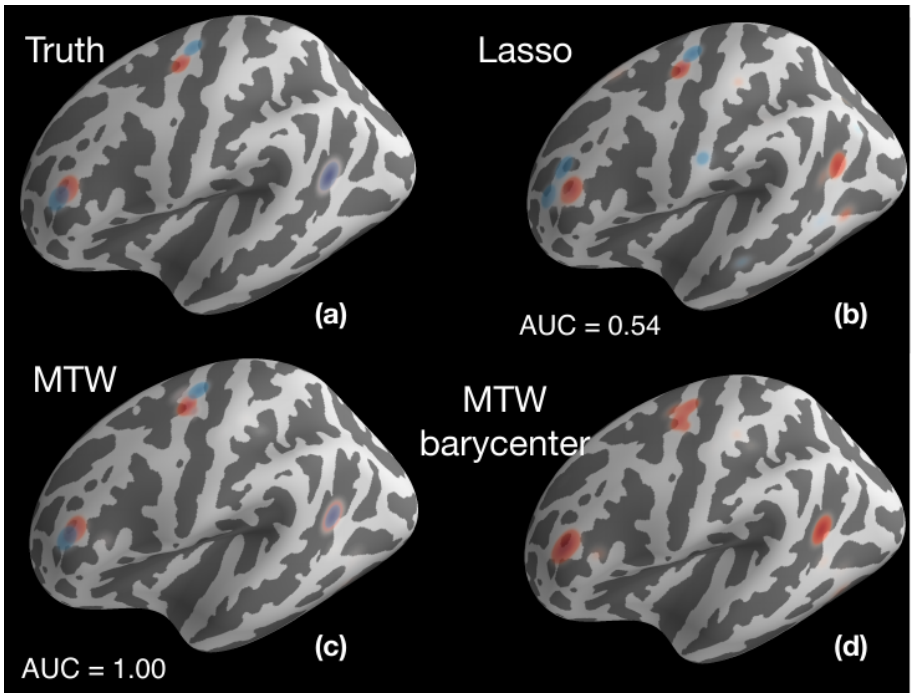

Figure 9. Each color corresponds to one of the two subjects (except for (d)). (a): True sources: one common feature (right side of the displayed hemisphere) and two non-overlapping sources. (b, c): Sources estimated by (b) Lasso and (c) MTW with the highest AUC score. (d) Shows the barycenter textthetā associated with MTW model. In this figure, the displayed activations were smoothed for the sake of visibility.

More information can be found in [34].

7.10. Ensemble of Clustered Knockoffs for robust multivariate inference on MRI data

Continuous improvement in medical imaging techniques allows the acquisition of higher-resolution images. When these are used in a predictive setting, a greater number of explanatory variables are potentially related to the dependent variable (the response). Meanwhile, the number of acquisitions per experiment remains limited. In such high dimension/small sample size setting, it is desirable to find the explanatory variables

that are truly related to the response while controlling the rate of false discoveries. To achieve this goal, novel multivariate inference procedures, such as knockoff inference, have been proposed recently. However, they require the feature covariance to be well-defined, which is impossible in high-dimensional settings. In this paper, we propose a new algorithm, called Ensemble of Clustered Knockoffs, that allows to select explanatory variables while controlling the false discovery rate (FDR), up to a prescribed spatial tolerance. The core idea is that knockoff-based inference can be applied on groups (clusters) of voxels, which drastically reduces the problem's dimension; an ensembling step then removes the dependence on a fixed clustering and stabilizes the results. We benchmark this algorithm and other FDR-controlling methods on brain imaging datasets and observe empirical gains in sensitivity, while the false discovery rate is controlled at the nominal level.

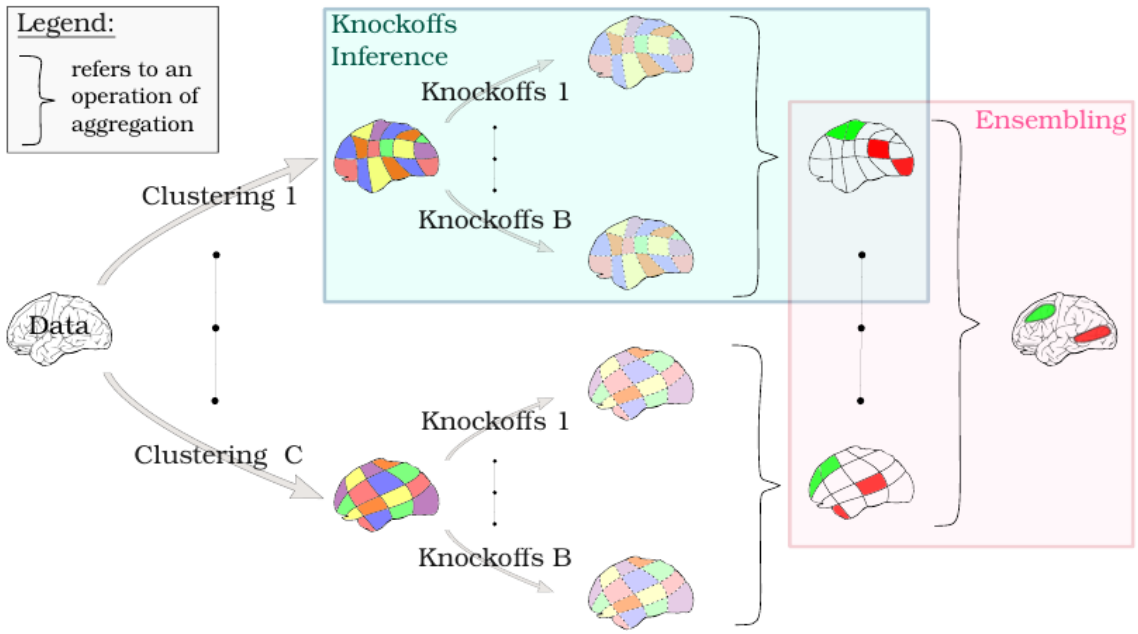

Figure 10. Representation of the ECKO algorithm. To create a stable inference result, we introduce ensembling steps both within each cluster level and at the voxel- level, across clusterings.

More information can be found in [46].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

In 2019, a CIFRE PhD thesis was launched with Siemens-Healthineers France. This contract supports the PhD thesis of Guillaume Daval-Frérot.

8.2. Scikit-learn Consortium

Scikit-learn is a machine-learning library in Python. It is the engine that powers many applications of artifical intelligence and data science.

Scikit-learn is used on a regular basis by more than half a million people in the world, with applications ranging from medical imaging to product recommendation.

Scikit-learn is an open-source software, under BSD license that facilitates commercial usage. It is developed by a world-wide community, gathering many different expertise on statistics, algorithms and software production.

The quality of scikit-learn, its algorithms, its interfaces, its documentation, are universally acclaimed. Its development follows a strict process to ensure this quality.

The goal of the foundation is to enable maintaining scikit-learn's high standards addressing new challenges.

The foundation employs central contributors to the project, to support scikit-learn's community and to develop new ambitious features. The priorities of the foundation are set jointly by the community and its sponsors.

More information can be found here http://scikit-learn.fondation-inria.fr/home.

The consortium is supported by 8 companies and has an annual budget of about half a million euros.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. Inserm-Inria project

This project is funded by the joint Inserm and Inria program 'médecine numérique' and is conducted in collaborations with our clinical partners from the Lariboisière hospital, Inserm uni U942 BioCANVAS (Biomarkers in Cardio-Neuro-VAScular diseases). It supports the PhD thesis of David Sabbagh.

Participants:

- Denis Engemann [coordinator, co-advisor]
- Alexandre Gramfort [thesis director, co-advisor]
- Etienne Gayat [clinical collaborator, co-advisor]
- Fabrice Vallée [clinical collaborator]
- David Sabbagh [PhD Student]

Post-operative delirium (POD) is a potential complication of anesthesia during surgery. It is often associated with adverse outcomes and is aggravated by aging. In elderly patients, post-operative complications have been estimated to incur tens of million US dollars of costs each year in the United States by prolonging hospitalization and potentially affecting health prognosis. Recent studies suggest that POD can already be prevented by improving electrophysiological monitoring of anesthesia depth and individual dosage of anesthetic agents. Doing so probably minimizes the time patients spend in a coma-like state that manifests itself in isoelectric burst suppression, an electroencephalogram (EEG) pattern characterized by alternation between quiescence and high-amplitude bursts, and causally linked to POD. However, such an enterprise, currently, depends on the trained clinical electrophysiologist and guidance by commercially provided EEG indices of states of consciousness. One such metric is the bispectral index (BIS), which, like other related metrics, does not explicitly take into account baseline changes related to normative aging and may therefore be biased when used naively.

While electrophysiological signatures of aging (e.g. drop in Alpha and Gamma band power), states of consciousness (e.g. drop in Theta band long-range connectivity) and drug response (e.g. anteriorization of alpha band power in propofol anesthesia) have been separately investigated in the past years, their common denominators are not known. It is therefore difficult to detect individual risk, choose the optimal dosage, and automate anesthesia monitoring readily for any patient in any hospital.

The goal of this research project is to build statistical models that enable prediction of burst suppression and subsequent POD by exploiting diverse EEG-signatures of states of consciousness in the context of aging. We approach this challenge by recasting it as a problem of learning brain-age from the point of view of electrophysiology of consciousness.

9.1.2. CoSmic project

Participants: Philippe Ciuciu [Correspondant], Nicolas Chartier, Loubna El Gueddari, Zaccharie Ramzi, Chaithya Giliyar Radhkrishna.

This project is funded by CEA DRF-Impulsion.

the DRF-impulsion CEA program which has been transformed into a CEA PTC program for 2 years (2018-2020), in collaboration with Pierre Kestener, La Maison de la Simulation (CEA/CNRS).

Compressed Sensing is a recent theory in maths that allows the perfect recovery of signals or images from compressive acquisition scenarios. This approach has been popularized in MRI over the last decade as well as in astrophysics (noticeably in radio-astronomy). So far, both of these fields have developed skills in CS separately. The aim of the COSMIC project is to foster collaborations between CEA experts in MRI (Parietal team within NeuroSpin) and in astrophysics (CosmoStat lab within the Astrophysics Department). These interactions will allow us to share different expertise in order to improve image quality, either in MRI or in radio-astronomy (thanks to the interferometry principle). In this field, given the data delivered by radio-telescopes, the goal consists in extracting high temporal resolution information in order to study fast transient events.

9.1.3. Metacog

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Jérome Dockès.

This project is funded by Digiteo.

This is a Digicosme project (2016-2019) and a collaboration with Fabian Suchanek (Telecom Paritech).

Understanding how cognition emerges from the billions of neurons that constitute the human brain is a major open problem in science that could bridge natural science –biology– to humanities –psychology. Psychology studies performed on humans with functional Magnetic Resonance Imaging (fMRI) can be used to probe the full repertoire of high-level cognitive functions. While analyzing the resulting image data for a given experiment is a relatively well-mastered process, the challenges in comparing data across multiple datasets poses serious limitation to the field. Indeed, such comparisons require to pool together brain images acquired under different settings and assess the effect of different *experimental conditions* that correspond to psychological effects studied by neuroscientists.

Such meta-analyses are now becoming possible thanks to the development of public data resources –OpenfMRI http://openfmri.org and NeuroVault http://neurovault.org. As many others, researchers of the Parietal team understand these data sources well and contribute to them. However, in such open-ended context, the description of experiments in terms of cognitive concepts is very difficult: there is no universal definition of cognitive terms that could be employed consistently by neuroscientists. Hence meta-analytic studies loose power and specificity. On the other hand, http://brainspell.org provide a set of curated annotation, albeit on much less data, that can serve as a seed or a ground truth to define a consensual ontology of cognitive concepts. Relating these terms to brain activity poses another challenge, of statistical nature, as brain patterns form high-dimensional data in perspective with the scarcity and the noise of the data.

The purpose of this project is to learn a semantic structure in cognitive terms from their occurrence in brain activation. This structure will simplify massive multi-label statistical-learning problems that arise in brain mapping by providing compact representations of cognitive concepts while capturing the imprecision on the definition these concepts.

9.1.4. HidimStat

Participants: Bertrand Thirion [Correspondant], Jerome-Alexis Chevalier, Joseph Salmon.

This project is funded by Digiteo.

This is a Digicosme project (2017-2020) and a collaboration with Joseph Salmon (Telecom Paritech).

The HiDimStat project aims at handling uncertainty in the challenging context of high dimensional regression problem. Though sparse models have been popularized in the last twenty years in contexts where many features can explain a phenomenon, it remains a burning issue to attribute confidence to the predictive models that they produce. Such a question is hard both from the statistical modeling point of view, and from a computation perspective. Indeed, in practical settings, the amount of features at stake (possibly up to several millions in high resolution brain imaging) limit the application of current methods and require new algorithms to achieve computational efficiency. We plan to leverage recent developments in sparse convex solvers as well as more efficient reformulations of testing and confidence interval estimates to provide several communities with practical software handling uncertainty quantification. Specific validation experiments will be performed in the field of brain imaging.

9.1.5. Template estimation for arbitrary alignments: application to brain imaging.

Participants: Bertrand Thirion [Correspondant], Thomas Bazeille.

This project is funded by Digiteo.

In the recent years, the nature of scientific inference has shifted quite substantially from model-based to predictive approaches, thanks to the generalization of powerful machine learning techniques. While this has certainly improved scientific standards, this has also obscured the objects and concepts on which inference is drawn. For instance, it is now possible –based on some initial data– to predict individual brain activity topographies, yet the very notion of a standard brain template has become increasingly elusive. Given the importance of establishing models for the progress of knowledge, we revisit the problem of model inference on data with high variance. Specifically, in a context where almost arbitrary transformation can successfully warp observations to each other with high accuracy, what is the common definition of a population model underlying all these observations? What is the working definition of a template ? We plan to leverage recent developments on optimal transport and multivariate analysis to build working definition of templates; we will use them in a brain imaging context to build a novel generation of brain templates.

9.1.6. CDS2

Participants: Alexandre Gramfort [Correspondant], Gaël Varoquaux, Maria Telenczuk, Jiaping Liu.

CDS2 is an "Strategic research initiative" of the Paris Saclay University Idex https://www.datascience-paris-saclay.fr/. Although it groups together many partners of the Paris Saclay ecosystem, Parietal has been deeply involved in the project. It currently funds 2 engineers: Maria Telenczuk and Jiaping (Lucy) Liu.

9.2. National Initiatives

9.2.1. ANR

9.2.1.1. Neuroref: Mathematical Models of Anatomy / Neuroanatomy / Diffusion MRI

Participants: Demian Wassermann [Correspondant], Antonia Machlouzarides Shalit, Valentin Iovene.

While mild traumatic brain injury (mTBI) has become the focus of many neuroimaging studies, the understanding of mTBI, particularly in patients who evince no radiological evidence of injury and yet experience clinical and cognitive symptoms, has remained a complex challenge. Sophisticated imaging tools are needed to delineate the kind of subtle brain injury that is extant in these patients, as existing tools are often ill-suited for the diagnosis of mTBI. For example, conventional magnetic resonance imaging (MRI) studies have focused on seeking a spatially consistent pattern of abnormal signal using statistical analyses that compare average differences between groups, i.e., separating mTBI from healthy controls. While these methods are successful in many diseases, they are not as useful in mTBI, where brain injuries are spatially heterogeneous.

The goal of this proposal is to develop a robust framework to perform subject-specific neuroimaging analyses of Diffusion MRI (dMRI), as this modality has shown excellent sensitivity to brain injuries and can locate subtle brain abnormalities that are not detected using routine clinical neuroradiological readings. New algorithms will be developed to create Individualized Brain Abnormality (IBA) maps that will have a number of clinical and research applications. In this proposal, this technology will be used to analyze a previously acquired dataset from the INTRuST Clinical Consortium, a multi-center effort to study subjects with Post-Traumatic Stress Disorder (PTSD) and mTBI. Neuroimaging abnormality measures will be linked to clinical and neuropsychological assessments. This technique will allow us to tease apart neuroimaging differences between PTSD and mTBI and to establish baseline relationships between neuroimaging markers, and clinical and cognitive measures.

9.2.1.2. DirtyData: Data integration and cleaning for statistical analysis

Participants: Gaël Varoquaux [Correspondant], Patricio Cerda Reyes, Pierre Glaser.

Machine learning has inspired new markets and applications by extracting new insights from complex and noisy data. However, to perform such analyses, the most costly step is often to prepare the data. It entails correcting errors and inconsistencies as well as transforming the data into a single matrix-shaped table that comprises all interesting descriptors for all observations to study. Indeed, the data often results from merging multiple sources of informations with different conventions. Different data tables may come without names on the columns, with missing data, or with input errors such as typos. As a result, the data cannot be automatically shaped into a matrix for statistical analysis.

This proposal aims to drastically reduce the cost of data preparation by integrating it directly into the statistical analysis. Our key insight is that machine learning itself deals well with noise and errors. Hence, we aim to develop the methodology to do statistical analysis directly on the original dirty data. For this, the operations currently done to clean data before the analysis must be adapted to a statistical framework that captures errors and inconsistencies. Our research agenda is inspired from the data-integration state of the art in database research combined with statistical modeling and regularization from machine learning.

Data integrating and cleaning is traditionally performed in databases by finding fuzzy matches or overlaps and applying transformation rules and joins. To incorporate it in the statistical analysis, an thus propagate uncertainties, we want to revisit those logical and set operations with statistical-learning tools. A challenge is to turn the entities present in the data into representations well-suited for statistical learning that are robust to potential errors but do not wash out uncertainty.

Prior art developed in databases is mostly based on first-order logic and sets. Our project strives to capture errors in the input of the entries. Hence we formulate operations in terms of similarities. We address typing entries, deduplication -finding different forms of the same entity- building joins across dirty tables, and correcting errors and missing data.

Our goal is that these steps should be generic enough to digest directly dirty data without user-defined rules. Indeed, they never try to build a fully clean view of the data, which is something very hard, but rather include in the statistical analysis errors and ambiguities in the data.

The methods developed will be empirically evaluated on a variety of dataset, including the French public-data repository, http://www.data.gouv.fr. The consortium comprises a company specialized in data integration, Data Publica, that guides business strategies by cross-analyzing public data with market-specific data.

9.2.1.3. FastBig Project

Participants: Bertrand Thirion [Correspondant], Jerome-Alexis Chevalier, Tuan Binh Nguyen.

In many scientific applications, increasingly-large datasets are being acquired to describe more accurately biological or physical phenomena. While the dimensionality of the resulting measures has increased, the number of samples available is often limited, due to physical or financial limits. This results in impressive amounts of complex data observed in small batches of samples.

A question that arises is then : what features in the data are really informative about some outcome of interest ? This amounts to inferring the relationships between these variables and the outcome, conditionally to all other variables. Providing statistical guarantees on these associations is needed in many fields of data science, where competing models require rigorous statistical assessment. Yet reaching such guarantees is very hard.

FAST-BIG aims at developing theoretical results and practical estimation procedures that render statistical inference feasible in such hard cases. We will develop the corresponding software and assess novel inference schemes on two applications : genomics and brain imaging.

9.2.1.4. MultiFracs project

Participant: Philippe Ciuciu [Correspondant].

The scale-free concept formalizes the intuition that, in many systems, the analysis of temporal dynamics cannot be grounded on specific and characteristic time scales. The scale-free paradigm has permitted the relevant analysis of numerous applications, very different in nature, ranging from natural phenomena (hydrodynamic turbulence, geophysics, body rhythms, brain activity,...) to human activities (Internet traffic, population, finance, art,...).

Yet, most successes of scale-free analysis were obtained in contexts where data are univariate, homogeneous along time (a single stationary time series), and well-characterized by simple-shape local singularities. For such situations, scale-free dynamics translate into global or local power laws, which significantly eases practical analyses. Numerous recent real-world applications (macroscopic spontaneous brain dynamics, the central application in this project, being one paradigm example), however, naturally entail large multivariate data (many signals), whose properties vary along time (non-stationarity) and across components (non-homogeneity), with potentially complex temporal dynamics, thus intricate local singular behaviors.

These three issues call into question the intuitive and founding identification of scale-free to power laws, and thus make uneasy multivariate scale-free and multifractal analyses, precluding the use of univariate methodologies. This explains why the concept of scale-free dynamics is barely used and with limited successes in such settings and highlights the overriding need for a systematic methodological study of multivariate scale-free and multifractal dynamics. The Core Theme of MULTIFRACS consists in laying the theoretical foundations of a practical robust statistical signal processing framework for multivariate non homogeneous scale-free and multifractal analyses, suited to varied types of rich singularities, as well as in performing accurate analyses of scale-free dynamics in spontaneous and task-related macroscopic brain activity, to assess their natures, functional roles and relevance, and their relations to behavioral performance in a timing estimation task using multimodal functional imaging techniques.

This overarching objective is organized into 4 Challenges:

1. Multivariate scale-free and multifractal analysis,
2. Second generation of local singularity indices,
3. Scale-free dynamics, non-stationarity and non-homogeneity,
4. Multivariate scale-free temporal dynamics analysis in macroscopic brain activity.

9.2.1.5. DARLING: Distributed adaptation and learning over graph signals

Participant: Philippe Ciuciu [Correspondant].

The project will be starting in 2020 with a post-doc to be hired probably in 2021.

The DARLING project will aim to propose new adaptive learning methods, distributed and collaborative on large dynamic graphs in order to extract structured information of the data flows generated and/or transiting at the nodes of these graphs. In order to obtain performance guarantees, these methods will be systematically accompanied by an in-depth study of random matrix theory. This powerful tool , never exploited so far in this context although perfectly suited for inference on random graphs, will thereby provide even avenues for improvement. Finally, in addition to their evaluation on public data sets, the methods will be compared with each other using two advanced imaging techniques in which two of the partners are involved: radio astronomy with the giant SKA instrument (Obs. Côte d'Azur) and magnetoencephalographic brain imaging (Inria Parietal at NeuroSpin, CEA Saclay). These involve the processing of time series on graphs while operating at extreme observation scales.

9.2.1.6. meegBIDS.fr: Standardization, sharing and analysis of MEEG data simplified by BIDS

Participant: Alexandre Gramfort [Correspondant].

The project accepted by ANR in 2019 will be starting in 2020 with an engineer to be hired in 2020. This project is in collaboration with the MEG groups at CEA NeuroSpin and the Brain and Spine Institute (ICM) in Paris.

The neuroimaging community recently started an international effort to standardize the sharing of data recorded with magnetoencephalography (MEG) and with electroencephalography (EEG). This format, known as the Brain Imaging Data Structure (BIDS), now needs a wider adoption, notably in the French neuroimaging community, along with the development of dedicated software tools that operate seamlessly on BIDS formatted datasets. The meegBIDS.fr project has three aims: 1) accelerate the research cycles by allowing analysis software tools to work with BIDS formated data, 2) simplify data sharing with high quality standards thanks to automated validation tools, 3) train French neuroscientists to leverage existing public BIDS MEG/EEG datasets and to share their own data with little efforts.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. VirtualBrainCloud

Title:

Programm: H2020 FET Open

Duration: 01/01/2019 - 31/12/2022

Coordinator: Petra Ritter

Inria contact: Bertrand Thirion

Summary:

The central goal of this project is the development of a cloud-based platform for biomedical research and clinical decision-making that helps to improve early patient-specific diagnosis and treatment of NDD and has substantial potential for significant positive socioeconomic impact.

The platform integrates several aims that revolve around early diagnosis, prognosis, and personalized treatment of neurodegenerative diseases (NDD) like Alzheimer's disease (AD) and Parkinson's disease (PD). It is becoming increasingly clear that meeting this objective requires a multifactorial approach that takes into account individual genetic, metabolic and environmental aspects, and that integrates them with the understanding of the biophysical processes underlying NDD.

More information can be found here https://virtualbraincloud-2020.eu/tvb-cloud-main.html.

9.3.1.2. Neurolang

Title: Accelerating Neuroscience Research by Unifying Knowledge Representation and Analysis Through a Domain Specific Language

Programm: ERC Starting researcher

Duration: 01/03/2018 - 28/02/2023

Coordinator: Demian Wassermann

Inria contact: Demian Wassermann

Summary:

Neuroscience is at an inflection point. The 150-year old cortical specialization paradigm, in which cortical brain areas have a distinct set of functions, is experiencing an unprecedented momentum with over 1000 articles being published every year. However, this paradigm is reaching its limits. Recent studies show that current approaches to atlas brain areas, like relative location, cellular

population type, or connectivity, are not enough on their own to characterize a cortical area and its function unequivocally. This hinders the reproducibility and advancement of neuroscience.

Neuroscience is thus in dire need of a universal standard to specify neuroanatomy and function: a novel formal language allowing neuroscientists to simultaneously specify tissue characteristics, relative location, known function and connectional topology for the unequivocal identification of a given brain region.

The vision of NeuroLang is that a unified formal language for neuroanatomy will boost our understanding of the brain. By defining brain regions, networks, and cognitive tasks through a set of formal criteria, researchers will be able to synthesize and integrate data within and across diverse studies. NeuroLang will accelerate the development of neuroscience by providing a way to evaluate anatomical specificity, test current theories, and develop new hypotheses.

NeuroLang will lead to a new generation of computational tools for neuroscience research. In doing so, we will be shedding a novel light onto neurological research and possibly disease treatment and palliative care. Our project complements current developments in large multimodal studies across different databases. This project will bring the power of Domain Specific Languages to neuroscience research, driving the field towards a new paradigm articulating classical neuroanatomy with current statistical and machine learning-based approaches.

9.3.1.3. SLAB (698)

Title: Signal processing and Learning Applied to Brain data

Programm: ERC Starting researcher

Duration: 01/04/2017 - 31/08/2021

Coordinator: Alexandre Gramfort

Partner: LTCI , Telecom ParisTech (France)

Inria contact: Alexandre Gramfort

Summary:

Understanding how the brain works in healthy and pathological conditions is considered as one of the challenges for the 21st century. After the first electroencephalography (EEG) measurements in 1929, the 90's was the birth of modern functional brain imaging with the first functional MRI and full head magnetoencephalography (MEG) system. In the last twenty years, imaging has revolutionized clinical and cognitive neuroscience.

After pioneering works in physics and engineering, the field of neuroscience has to face two major challenges. The size of the datasets keeps growing. The answers to neuroscience questions are limited by the complexity of the signals observed: non-stationarity, high noise levels, heterogeneity of sensors, lack of accurate models. SLAB will provide the next generation of models and algorithms for mining electrophysiology signals which offer unique ways to image the brain at a millisecond time scale.

SLAB will develop dedicated machine learning and signal processing methods and favor the emergence of new challenges for these fields. SLAB focuses on five objectives: 1) source localization with M/EEG for brain imaging at high temporal resolution 2) representation learning to boost statistical power and reduce acquisition costs 3) fusion of heterogeneous sensors 4) modeling of non-stationary spectral interactions to identify functional coupling between neural ensembles 5) development of fast algorithms easy to use by non-experts.

SLAB aims to strengthen mathematical and computational foundations of brain data analysis. The methods developed will have applications across fields (computational biology, astronomy, econometrics). Yet, the primary impact of SLAB will be on neuroscience. The tools and high quality open software produced in SLAB will facilitate the analysis of electrophysiology data, offering new perspectives to understand how the brain works at a mesoscale, and for clinical applications (epilepsy, autism, tremor, sleep disorders).

9.3.1.4. HBP SGA2

Title: Interactive Computing E-Infrastructure for the Human Brain Project

Programm: FET Flagship

Duration: 01/04/2018 - 31/03/2020

Coordinator: Katrin Amunts

Partners: see https://www.humanbrainproject.eu/en/open-ethical-engaged/contributors/partners/

Inria contact: Bertrand Thirion

Summary:

The HBP Flagship was launched by the European Commission's Future and Emerging Technologies (FET) scheme in October 2013, and is scheduled to run for ten years. The Flagships, represent a new partnering model for visionary, long-term European cooperative research in the European Research Area, demonstrating the potential for common research efforts. The HBP has the following main objectives:

- Create and operate a European scientific Research Infrastructure for brain research, cognitive neuroscience, and other brain-inspired sciences
- Gather, organise and disseminate data describing the brain and its diseases
- Simulate the brain
- Build multi-scale scaffold theory and models for the brain
- Develop brain-inspired computing, data analytics and robotics
- Ensure that the HBP's work is undertaken responsibly and that it benefits society.

More information on the HBP's Flagship Objectives is available in the Framework Partnership Agreement.

The timeline of the Project is split into multiple phases, each of which will be covered by a separate funding agreement. The current phase is Specific Grant Agreement Two (SGA2), which spans the two-year period from April 2018–April 2020. The HBP is funded via several sources. Total funding is planned to be in the region of EUR 1 billion; around one half of which will be provided by the European Union, and the other by Member States and private funding sources. The European Union contributed EUR 54 million to the Project in the Ramp-Up Phase (October 2013 to March 2016), EUR 89 million for the second phase (SGA1), and EUR 88 million for the current phase (SGA2). The FET Flagships Staff Working Document provides further information on how Flagships are funded.

9.4. International Initiatives

9.4.1. Inria International Labs

Inria@SiliconValley
Associate Team involved in the International Lab:

9.4.1.1. Meta&Co

Title: Meta-Analysis of Neuro-Cognitive Associations

International Partner (Institution - Laboratory - Researcher):

Stanford (United States) - Psychology department. - Russel Poldrack

Start year: 2018

See also: http://team.inria.fr/parietal

Cognitive science and psychiatry describe mental operations: cognition, emotion, perception and their dysfunction. Cognitive neuroimaging bridge these mental concepts to their implementation in the brain, neural firing and wiring, by relying on functional brain imaging. Yet aggregating results from experiments probing brain activity into a consistent description faces the roadblock that cognitive concepts and brain pathologies are ill-defined. Separation between them is often blurry. In addition, these concepts and subdivisions may not correspond to actual brain structures or systems. To tackle this challenge, we propose to adapt data-mining techniques used to learn relationships in computational linguistics. Natural language processing uses distributional semantics to build semantic relationships and ontologies. New models are needed to learn relationships from heterogeneous signals: functional magnetic resonance images (fMRI), on the one hand, combined with related psychology and neuroimaging annotations or publications, on the other hand. Such a joint effort will rely on large publicly-available fMRI databases shared by Podrack Lab, as well as literature mining.

Inria@SiliconValley

Associate Team involved in the International Lab:

9.4.1.2. LargeSmallBrainNets

Title: Characterizing Large and Small-scale Brain Networks in Typical Populations Using Novel Computational Methods for dMRI and fMRI-based Connectivity and Microstructure

International Partner (Institution - Laboratory - Researcher):

Stanford (United States) - Stanford Cognitive and Systems Neuroscience Laboratory - Vinod Menon

Start year: 2019

See also: http://pages.saclay.inria.fr/demian.wassermann/largesmallbrainnets/

In the past two decades, brain imaging of neurotypical individuals and clinical populations has primarily focused on localization of function and structures in the brain, revealing activation in specific brain regions during performance of cognitive tasks through modalities such as functional MRI. In parallel, technologies to identify white matter structures have been developed using diffusion MRI. Lately, interest has shifted towards developing a deeper understanding of the brain's macroscopic and microscopic architectures and their influence on cognitive and affective information processing. Using for this resting state fMRI and diffusion MRI to build the functional and structural networks of the human brain.

The human brain is a complex patchwork of interconnected regions, and graph-theoretical approaches have become increasingly useful for understanding how functionally connected systems engender, and constrain, cognitive functions. The functional nodes of the human brain, i.e. cortical regions, and their structural inter-connectivity, collectively the brain's macrostructure or "connectome", are, however, poorly understood. Quantifying in vivo how these nodes' microstructure, specifically cellular composition or cytoarchitecture, influences the cognitive tasks in which these are involved is fundamental problem in understanding the connectome. Furthermore, the coupling between within and across-subject contributions to the connectome and cognitive differences hampers the identification and understanding of the link between brain structure and function, and human cognition.

Critically, there is a dearth of computational methods for reliably identifying functional nodes of the brain, their micro and macrostructure in vivo, and separating the population and subject-specific effects. Devising and validating methods for investigating the human connectome has therefore taken added significance.

The first major goal of this project is to develop and validate appropriate sophisticated computational and mathematical tools relate the brain's macrostructure with its function. Specifically, we will focus on being able to separate population and subject-specific contributions within these models using state-of-the-art human brain imaging techniques and open-source data from the Human Connectome

Project (HCP) and the Adolescent Brain Cognitive Development study (ABCD). To this end, we will first develop and validate novel computational tools for (1) formulating and fitting large scale random effect models on graphs derived from functional and structural connectivity and (2) implement techniques enabling us to impose different regularization schemes based on sparsity and multicollinearity of the model parameters.

The second major goal of this project is characterizing the cytoarchitecture of the nodes, i.e. cortical regions, at the microscopic level and their relationship with the brain's hemodynamical function and cognition. For this, we will (1) identify cortical areas with specific cytoarchitecture in the human cortex and use them to develop diffusion MRI-based models, (2) validate these models with numerical simulations of the dMRI signal and animal models, and (3) establish the relationship between cytoarchitecture and hemodynamical function measured from fMRI and cognition. For this we will leverage multi-shell high-angular diffusion MRI from public databases such as HCP and ABCD.

Finally, we will use to use our newly developed computational tools to characterize normal structural and functional brain networks in neurotypical adults. Due to the complementarity of the cognitive science and imaging techniques expertise the synergy between the two laboratories of this associate team will allow us to reveal in unprecedented detail the structural and functional connectivity of the human brain and its relation to cognition.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- Aapo Havärinen has been with parietal since April 2019 for a one-year visit, funded by the Dataia convergence institute.
- Luigi Gresele (MPI Tübingen) has been visiting the team in November-December 2019.
- Cedric Xu (UPenn) visited the team for three months in June-August 2019.
- James Cole (UCL) visited the team in December 2019.
- Yu Zhang (BIC, Montreal) visited the team during one week in September 2019
- Bemsibom Toh (HWU, Edinburgh) visited the team during two months in September-November 2019

9.5.2. Visits to International Teams

9.5.2.1. Sabbatical programme

Gael Varoquaux is spending one year in Montreal (September 2019-September 2020), hosted at MILA and Montreal Neuroimaging Institute at McGill Unviversity.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

Bertrand Thirion co-organized the "Imagerie médicale et intelligence artificielle " colloque at Collège de France on April, 23rd, 2019.

Thomas Moreau co-organized the Palaisien seminar: a monthly scientific seminar in statistics and statistical learning bringing together researchers in this field working on the Saclay Plateau. The seminar started in November and brings together about fifty researchers (http://palaisien.herokuapp.com).

10.1.1.2. Member of the Organizing Committees

Demian Wassermann: MICCAI 2019, ISMRM 2019

Gaël Varoquaux: PyParis 2019

10.1.2. Scientific Events: Selection

10.1.2.1. Member of the Conference Program Committees

- Philippe Ciuciu: Chairman of scientific oral sessions in ISBI 2019 (Venice, Italy), program commitee of EUSIPCO 2019.
- Gaël Varoquaux: program committee of NIPS, ICML, ICLR, senior program committee IJCAI.
- Alexandre Gramfort: program committee of NIPS, ICML, ICLR.
- Demian Wassermann: ISMRM 2019
- Thomas Moreau: JDSE 2019

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Philippe Ciuciu: Senior Area Editor of IEEE OJSP
- Alexandre Gramfort: Editor, NeuroImage, Journal of Machine Learning Research (JMLR), Frontiers in Brain Imaging Methods
- Bertrand Thirion: Editor, Frontiers in Brain Imaging Methods
- Gael Varoquaux: Editor, elife

10.1.3.2. Reviewer - Reviewing Activities

- Philippe Ciuciu is reviewer for Nature Communication, Biological Pysichiatry, Plos Computational Biology, Neuroimage, Scientific Reports, Journal of Neuroscience, IEEE TSP/TIP/TMI/TCI/SPL, Frontiers in Neuroscience, MRM, SIAM Imaging Science
- Gaël Varoquaux: Nature Methods, JMLR, PLOS Bio, NeuroImage, NeuroImage Clinical, IEEE TMI, Annals of Applied Statistics, Biological Psychiatry, MedIA, Science
- Alexandre Gramfort: JMLR, PLOS Computational Biology, NeuroImage, IEEE TBME, IEEE TMI, IEEE TSP, MedIA, NIPS, ICML, ICLR, ICASSP, Scientific Reports, Frontiers in Brain Imaging Methods, Journal of Neuroscience Methods
- Bertrand Thirion: Nature communications, Neuroimage, Medical Image Analysis, IEEE TMI, PNAS, PLOS Comp Bio, Brain Structure and Function, NIPS, ICML, IPMI, AISTATS
- Denis Engemann: Brain, PLOS Biology, PLOS Computational Biology, Psychological Medicine, Scientific Reports, Neuroimage, Neuroimage Clinical, Human Brain Mapping, Journal of Machine Learning Research, Brain Topography, Brain Connectivity, Journal of Alzheimer's Disease, Neuroscience of Consciousness, PLOS ONE, Frontiers in Neuroscience, Journal of Computational Neuroscience, Psychiatry and Clinical Neurosciences, Sensors
- Demian Wassermann: NeuroImage, MRM, JMRI, Brain Structure and Function, Cortex, MedIA.
- Thomas Moreau: ICML, NeurIPS, EUSIPCO, JDSE, AISTAT, JMLR, TSP, LSP, JMIV
- Olivier Grisel: JMLR

10.1.4. Invited Talks

- Philippe Ciuciu, Ecole Centrale Nantes, Nov. 2019
- Philippe Ciuciu, I3S/CNRS (Sophia-Antipolis), Oct. 2019
- Philippe Ciuciu, Wavelets & Sparsity XVIII, San Diego, Aug. 2019
- Philippe Ciuciu, Collège de France, Apr. 2019
- Denis Engemann, Max Planck Institute AES, Frankfurt, Germany, Jan. 2019

- Denis Engemann, University of Birmingham, Birmingham, UK, Apr. 2019
- Denis Engemann, University of Glasgow, Glasgow, UK, Jun. 2019
- Denis Engemann, Max Planck Institute CBS, Leipzig, Germany, Nov. 2019
- Alexandre Gramfort at IHP, Paris, France, April. 2019
- Alexandre Gramfort at Univ Genoa, Italy, Sept. 2019
- Alexandre Gramfort at Univ Jyvaskyla, Finland, May. 2019
- Alexandre Gramfort at Institut Pasteur, Paris, France, June 2019
- Alexandre Gramfort at National Institute of Health, Washington, USA, Oct. 2019
- Bertrand Thirion at Donders Institute, Netherlands Feb. 2019
- Bertrand Thirion at MRC CBU, Cambridge, UK, Nov. 2019
- Bertrand Thirion at INT, Marseille, Nov. 2019
- Bertrand Thirion at GFAIH, Paris, Oct. 2019
- Bertrand Thirion at Université Paris-Sud, faculté de médecine, Nov. 2019
- Bertrand Thirion at Alan turing Institute, London, June 2019
- Gael Varoquaux at MNI (Montréal), Nov. 2019
- Gael Varoquaux at SparkAISummit, Amsterdam, Oct. 2019
- Gael Varoquaux at PyData Amsterdam, Oct. 2019
- Gael Varoquaux at Mila (Montréal), Sept. 2019
- Gael Varoquaux at Etalab (Paris), June 2019
- Gael Varoquaux at MGH (Harvard), June 2019
- Gael Varoquaux at Yale, June 2019
- Demian Wassermann at NYU, February 2019
- Demian Wassermann at ERC, June 2019
- Demian Wassermann at BWH (Harvard), June 2019
- Demian Wassermann at Stanford, October 2019
- Thomas Moreau **ISMRM, Educational Session, Montreal, Canada, 2019:** "Best Practices & Pitfalls in Applying Machine Learning to Magnetic Resonance Imaging".
- Thomas Moreau **Télécom ParisTech, SMILE seminar, Paris, France, 2019:** "Multivariate Convolutional Sparse Coding for Electromagnetic Brain Signals".
- Thomas Moreau **ENS Paris-Saclay, Centre de Mathématiques et de leurs Applications, MLMDA seminar, Cachan, France, 2019:** "Multivariate Convolutional Sparse Coding for Electromagnetic Brain Signals".
- Thomas Moreau **Telecom ParisTech, Laboratoire Traitement et Communication de l'Information, seminar, Paris, France, 2019:** "Using the Dictionary Structure for efficient Convolutional Dictionary Learning".
- Thomas Moreau **Centre pour l'Énergie Atomique Saclay, Cosmostat, annual seminar day, Saclay, France, 2019:** "Learning Recurring Patterns in Large Signals with Convolutional Dictionary Learning".

10.1.5. Leadership within the Scientific Community

- Demian Wassermann: Organising committee of the BrainHack community — Paris Chapter
- Demian Wassermann: ISMRM and MICCAI action organization towards reducing gender and minority biases.

10.1.6. Scientific Expertise

- Philippe Ciuciu has been reviewer for NWO, The Netherlands (Veni/Vidi/Vici call)
- Philippe Ciuciu has been hired as H2020 expert evaluator for the H2020-WIDESPREAD-2020-5 Twinning RCA programme since Dec 2019.
- Denis Engemann has been reviewer for the Swiss National Science Foundation in 2019
- Alexandre Gramfort has been member of an NSF panel in the USA in January 2019.
- Demian Wassermann FET-OPEN ERCA Action, and DataIA
- Gaël Varoquaux was expert for DataIA

10.1.7. Research Administration

- In Nov 2018, Philippe Ciuciu Elected Vice-Chair of the SAT Biomedical Image & Signal Analytics (EURASIP technical committee) for 2019-2020.
- In Dec 2019, Philippe Ciuciu has been renewed as the IEEE Signal Processing Society Representative for the 2020 IEEE ISBI conference.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master: Alexandre Gramfort, Optimization for Data Science, 20h, Msc 2 Data Science Master Ecole Polytechnique, France
Master: Alexandre Gramfort, DataCamp, 20h, Msc 2 Data Science Master Ecole Polytechnique, France
Master: Alexandre Gramfort, Source Imaging with EEG and MEG, 7h, Msc 2 in Biomedical Imaging at Télécom Paristech
Doctoral School: Alexandre Gramfort, Practical machine learning, 6h, The mathematics of imaging: the CIRM pre-school, Marseille, France
Master: Bertrand Thirion, Functional neuroimaging and BCI, 12h, Master MVA, ENS Paris-Saclay, France
Master: Bertrand Thirion, Kshitih Chawla, Thomas Bazeille: tutorial on Nilearn, 8h, INT, Marseille.
Master: Philippe Ciuciu, fMRI course: From acquisition to data Analysis, 6h, Msc 2 in Biomedical Imaging, Université Paris-Sud
Bachelor: Demian Wassermann, CSE201 class, 15h, C++ programming, Ecole Polytechnique
Master: Demian Wassermann, 7h Biomedical Engineering, Msc 2 Biomedical Engineering
Extension: Demian Wassermann, Data Science, 20h, Ecole Polytechnique
Master: Gaël Varoquaux, Machine learning in Python, 3h, ENSAE
Master: Gaël Varoquaux, Functional brain connectivity, 7h, Msc 2 in Biomedical Imaging at Télécom Paristech
Doctoral school: Gaël Varoquaux, representation learning with limited sample, 4h30, Deep Learning Summer School, Warsaw
Doctoral school: Gaël Varoquaux, machine learning for neuroimaging, 3h, Unique days, Montréal
Doctoral school: Gaël Varoquaux, machine learning with scikit-learn, 4h, IBPRIA, Madrid
Master: Thomas Moreau, **Formation continue Telecom ParisTech (7h)** TP pour Machine learning and Python.
Master: Thomas Moreau, **Master Spécialisé courses, Telecom ParisTech (20h)** Assistant TP pour le cours de machine learning avancé.
Master: Thomas Moreau, **Exectutive Master Statistique et Big Data, Université Paris-Dauphine (9h)** Cours Python pour le machine learning.
Master: Thomas Moreau, **Master Data Science, Université Paris-Saclay (25h)** Cours machine learning associé à un challenge.
Master: Olivier Grisel, **Master Data Science, Université Paris-Saclay (40h)** Cours de Deep Learning.

10.2.2. Supervision

Demian Wassermann has defended his habilitation thesis in May 2019.

The following PhD defenses have taken place

Guillermo Gallardo, "Inferring and comparing structural parcellations of the human brain using diffusion MRI", defended Jan 2019, co-directed by Demian Wassermann

Patricio Cerda, "encoding high-cardinality string categories", defended Nov 2019, directed by Gaël Varoquaux

Jérome Dockes, "large-scale predictive neuroimaging meta-analyses" defended Nov 2019, directed by Gaël Varoquaux & Fabian Suchanek (Telecom)

Pierre Ablin, defended Nov 2019, directed by Alexandre Gramfort

Mathurin Massias defended Dec 2019, directed by Alexandre Gramfort

Loubna El Guedari, defended Dec 2019, directed by Philippe Ciuciu

Ongoing PhD theses

Kamalakar Daddi, "predictive population analysis for mental health", defense planned Oct 2020, directed by Gaël Varoquaux, Denis Engemmann & Bertrand Thirion

Alexis Cvetkov Iliev, "embeddings of database entries", defense planned Oct 2022, directed by Gaël Varoquaux & Alexandre Allauzen (ESPCI)

Hicham Janati, 01/10/2017, coadvised by Alexandre Gramfort and Marco Cuturi

Quentin Bertrand, 01/10/2018, coadvised by Alexandre Gramfort and Joseph Salmon

Hubert Banville, 01/10/2018, coadvised by Alexandre Gramfort and Denis Engemann

David Sabbagh, 01/10/2018, coadvised by Alexandre Gramfort and Denis Engemann

Jérôme-Alexis Chevalier, 01/10/2017, coadvised by Bertrand Thirion and Joseph Salmon

Hugo Richard, 01/10/2018, coadvised by Bertrand Thirion and Olivier Grisel

Thomas Bazeille, 01/10/2018, advised by Bertrand Thirion

Tuan Binh Nguyen, 01/10/2018, coadvised by Bertrand thirion and Sylvain Arlot

Valentin Iovene, 01/10/2018, advised by Demian Wassermann

Antonia Machlouzarides Shalit, 01/10/2018, coadvised by Demian Wassermann and Bertrand Thirion

Hamza Cherkaoui, 01/10/2017, co-advised by Philippe Ciuciu and Thomas Moreau

Zaccharie Ramzi, 01/02/2019, co-advised by Philippe Ciuciu and Jean-Luc Starck

Anaïïs Artiges, 01/10/2019, co-advised by Philippe Ciuciu and Cyril Poupon

Guillaume Daval-Frérot, 01/11/2019, advised by Philippe Ciuciu

Maeliss Jallais, 01/10/2018, advised by Demian Wassermann

Lucas Martin, 01/10/2019, advised by Bertrand Thirion and Julie Josse.

Chengran Fan, 01/10/2019, co-advised by Demian Wassermann and Jing-Rebecca Li

Gaston Zanitti, 01/11/2019, advised by Demian Wassermann

10.2.3. Juries

Bertrand Thirion has been involved in the following committees:

- PhD - Florian Tilquin / University of Strasbourg
- PhD - Han Bossier / University of Ghent
- Habilitation - Laurent Oudre / Université Paris XIII

Philippe Ciuciu has been involved in the following committees:

- PhD - Marica Pesce (reviewer) / Heriot-Watta University, Edinburgh
- PhD - Christian El Hajj (reviewer) / Ecole Centrale Nantes.

Denis Engemann has been involved in the following committees:

- PhD - Laura Sophie Imperatori, IMT School for Advanced Studies Lucca, Lucca, Italy
- MD - Cyril Touchard, Universite´ Paris Diderot (Paris 7)

Gael Varoquaux has been involved in the following committees:

- PhD - Beyrem KHALFAOUI / Mines ParisTech

Alexandre Gramfort has been involved in the following committees:

- PhD - Benjamin Donnot / LRI, Université Paris-Sud
- PhD - Belhal Karimi / CMAP, Polytechnique
- PhD -Pedro Rodriges (reviewer) / GIPSA Lab, Grenoble
- PhD - Kostiantyn Maksymenko (reviewer) / Inria, Sophia
- MD - Cyril Touchard, Universite´ Paris Diderot (Paris 7)

Demian Wassermann has been involved in the following committees:

- PhD-midterm - Katrin Rojkova / Université Paris-Sorbonne
- Habilitation - Daniel Margulies / Université Paris-Sorbonne

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

- Philippe Ciuciu has been member of the Inria Saclay scientific commission since 2016
- Denis Engemann has been managing the Twitter accounts of Parietal and the MNE software since 2017
- Alexandre Gramfort is in charge of the Paris-Saclay Center for Data Science
- Alexandre is a Member of the technical development committee of Inria Saclay since Dec 2018
- Bertrand Thirion, head of the Dataia Convergence Institute since January 1st, 2019
- Gaël Varoquaux was Member of the technical development committee of Inria Saclay until Dec 2018
- Gaël Varoquaux is member of the doctoral monitoring committee of Inria Saclay
- Gaël Varoquaux is member of the steering committee of the Paris-Saclay Center for Data Science
- Thomas Moreau: Co-organisation of the weekly Parietal team seminar bringing together the 45 members of the team for short presentations and a longer scientific presentation.
- Gaël Varoquaux is in charge of the scikit-learn consortium

10.3.2. Internal action

- Guillaume Lemaître, Marie Telenczuk, et Lucie Liu have taught a 1-day training on data-science with Python in Nov 2019.

10.3.3. Creation of media or tools for science outreach

- Guillaume Lemaître, Gael Varoquaux, Olivier Grisel started developing new teaching material for MOOC of Scikit-learn to be hosted in 2020 on the FUN platform in collaboration with Loic Esteve and the team members of the Inria Learning Lab).

11. Bibliography

Major publications by the team in recent years

[1] A. MENSCH, J. MAIRAL, B. THIRION, G. VAROQUAUX. *Stochastic Subsampling for Factorizing Huge Matrices*, in "IEEE Transactions on Signal Processing", January 2018, vol. 66, n^{o} 1, pp. 113-128 [*DOI :* 10.1109/TSP.2017.2752697], https://hal.archives-ouvertes.fr/hal-01431618

Publications of the year

Doctoral Dissertations and Habilitation Theses

[2] M. MASSIAS. *Sparse high dimensional regression in the presence of colored heteroscedastic noise: application to M/EEG source imaging*, Telecom Paristech, December 2019, https://tel.archives-ouvertes.fr/tel-02401628

Articles in International Peer-Reviewed Journals

[3] D. BZDOK, D. FLORIS, A. MARQUAND. *Analyzing Brain Networks in Population Neuroscience: A Case for the Bayesian Philosophy*, in "Philosophical Transactions of the Royal Society of London. B (1887–1895)", January 2020, https://hal.archives-ouvertes.fr/hal-02447507

[4] D. BZDOK, J. P. A. IOANNIDIS. *Exploration, inference and prediction in neuroscience and biomedicine*, in "Trends in Neurosciences", March 2019, https://arxiv.org/abs/1903.10310 [*DOI :* 10.1016/J.TINS.2019.02.001], https://hal.archives-ouvertes.fr/hal-02044120

[5] D. BZDOK, T. E. NICHOLS, S. SMITH. *Towards Algorithmic Analytics for Large-scale Datasets*, in "Nature Machine Intelligence", July 2019 [*DOI :* 10.1038/S42256-019-0069-5], https://hal.archives-ouvertes.fr/hal-02178410

[6] S. CHAMBON, V. THOREY, P. J. ARNAL, E. J. M. MIGNOT, A. GRAMFORT. *DOSED: A deep learning approach to detect multiple sleep micro-events in EEG signal*, in "Journal of Neuroscience Methods", June 2019, vol. 321, pp. 64-78, https://arxiv.org/abs/1812.04079 [*DOI :* 10.1016/J.JNEUMETH.2019.03.017], https://hal.inria.fr/hal-02121090

[7] L. CHEN, D. WASSERMANN, D. ABRAMS, J. KOCHALKA, G. GALLARDO-DIEZ, V. MENON. *The visual word form area (VWFA) is part of both language and attention circuitry*, in "Nature Communications", December 2019, vol. 10, n° 1, Lang Chen, Demian Wassermann, and Daniel Abrams contributed equally [*DOI :* 10.1038/S41467-019-13634-Z], https://hal.inria.fr/hal-02401938

[8] K. DADI, M. RAHIM, A. ABRAHAM, D. CHYZHYK, M. MILHAM, B. THIRION, G. VAROQUAUX. *Benchmarking functional connectome-based predictive models for resting-state fMRI*, in "NeuroImage", May 2019, n° 192, pp. 115-134 [*DOI :* 10.1016/J.NEUROIMAGE.2019.02.062], https://hal.inria.fr/hal-01824205

[9] R. H. FICK, D. WASSERMANN, R. DERICHE. *The Dmipy Toolbox: Diffusion MRI Multi-Compartment Modeling and Microstructure Recovery Made Easy*, in "Frontiers in Neuroinformatics", October 2019, vol. 13 [*DOI :* 10.3389/FNINF.2019.00064], https://hal.archives-ouvertes.fr/hal-02400877

[10] P. FILIPIAK, R. H. FICK, A. PETIET, M. SANTIN, A.-C. PHILIPPE, S. LEHÉRICY, P. CIUCIU, R. DERICHE, D. WASSERMANN. *Reducing the number of samples in spatiotemporal dMRI acquisition design*, in "Magnetic Resonance in Medicine", 2019 [*DOI :* 10.1002/MRM.27601], https://hal.archives-ouvertes.fr/hal-01928734

[11] T. HE, R. KONG, A. J. HOLMES, M. Q. NGUYEN, M. SABUNCU, S. B. EICKHOFF, D. BZDOK, J. FENG, B. T. YEO. *Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics*, in "NeuroImage", October 2019, 116276 p. [*DOI :* 10.1016/J.NEUROIMAGE.2019.116276], https://hal.archives-ouvertes.fr/hal-02314718

[12] M. JORDING, D. ENGEMANN, H. ECKERT, G. BENTE, K. VOGELEY. *Distinguishing Social From Private Intentions Through the Passive Observation of Gaze Cues*, in "Frontiers in Human Neuroscience", December 2019, vol. 13 [*DOI :* 10.3389/FNHUM.2019.00442], https://hal.archives-ouvertes.fr/hal-02416981

[13] T. M. KARRER, D. S. BASSETT, B. DERNTL, O. GRUBER, A. ALEMAN, R. JARDRI, A. R. LAIRD, P. FOX, S. B. EICKHOFF, O. GRISEL, G. VAROQUAUX, B. THIRION, D. BZDOK. *Brain-based ranking of cognitive domains to predict schizophrenia*, in "Human Brain Mapping", July 2019 [*DOI :* 10.1002/HBM.24716], https://hal.archives-ouvertes.fr/hal-02180311

[14] C. LAZARUS, P. WEISS, N. CHAUFFERT, F. MAUCONDUIT, L. EL GUEDDARI, C. DESTRIEUX, I. ZEMMOURA, A. VIGNAUD, P. CIUCIU. *SPARKLING: variable-density k-space filling curves for accelerated T 2 * -weighted MRI*, in "Magnetic Resonance in Medicine", June 2019, vol. 81, n^{o} 6, pp. 3643-3661 [*DOI :* 10.1002/MRM.27678], https://hal.inria.fr/hal-02361265

[15] J. LEBENBERG, J.-F. MANGIN, B. THIRION, C. POUPON, L. HERTZ-PANNIER, F. LEROY, P. ADIBPOUR, G. DEHAENE-LAMBERTZ, J. DUBOIS. *Mapping the asynchrony of cortical maturation in the infant brain: a MRI multi-parametric clustering approach*, in "NeuroImage", January 2019, vol. 185, pp. 641-653 [*DOI :* 10.1016/J.NEUROIMAGE.2018.07.022], https://hal.archives-ouvertes.fr/hal-01966812

[16] J. LI, T. BOLT, D. BZDOK, J. S. NOMI, B. T. YEO, R. NATHAN SPRENG, L. Q. UDDIN. *Topography and Behavioral Relevance of the Global Signal in the Human Brain*, in "Scientific Reports", October 2019, vol. 9, n^{o} 1 [*DOI :* 10.1038/S41598-019-50750-8], https://hal.archives-ouvertes.fr/hal-02305188

[17] G. NGUYEN, N. AUNAI, D. FONTAINE, E. LE PENNEC, J. VAN DEN BOSSCHE, A. JEANDET, B. BAKKALI, L. VIGNOLI, B. REGALDO-SAINT BLANCARD. *Automatic Detection of Interplanetary Coronal Mass Ejections from In Situ Data: A Deep Learning Approach*, in "The Astrophysical journal letters", April 2019, vol. 874, n^{o} 2, 145 p. , https://arxiv.org/abs/1903.10780 [*DOI :* 10.3847/1538-4357/AB0D24], https://hal.sorbonne-universite.fr/hal-02103805

[18] V.-D. NGUYEN, M. DE LEONI, T. DANCHEVA, J. JANSSON, J. HOFFMAN, D. WASSERMANN, J.-R. LI. *Portable simulation framework for diffusion MRI*, in "Journal of Magnetic Resonance", December 2019, vol. 309, 106611 p. [*DOI :* 10.1016/J.JMR.2019.106611], https://hal.archives-ouvertes.fr/hal-02431598

[19] L. J. O'DONNELL, A. DADUCCI, D. WASSERMANN, C. LENGLET. *Advances in computational and statistical diffusion MRI*, in "NMR in Biomedicine", 2019, vol. 32, n^{o} 4, e3805 p. [*DOI :* 10.1002/NBM.3805], https://hal.inria.fr/hal-02432249

[20] M. RAHIM, B. THIRION, G. VAROQUAUX. *Population shrinkage of covariance (PoSCE) for better individual brain functional-connectivity estimation*, in "Medical Image Analysis", March 2019 [*DOI :* 10.1016/J.MEDIA.2019.03.001], https://hal.inria.fr/hal-02068389

[21] S. SAREMI, A. HYVÄRINEN. *Neural Empirical Bayes*, in "Journal of Machine Learning Research", 2019, vol. 20, pp. 1 - 23, https://hal.inria.fr/hal-02419496

[22] R. TOMI-TRICOT, V. GRAS, B. THIRION, F. MAUCONDUIT, N. BOULANT, H. CHERKAOUI, P. ZERBIB, A. VIGNAUD, A. LUCIANI, A. AMADON. *SmartPulse, a Machine Learning Approach for Calibration-Free Dynamic RF Shimming: Preliminary Study in a Clinical Environment*, in "Magnetic Resonance in Medicine", June 2019, vol. 82, pp. 2016–2031 [*DOI :* 10.1002/MRM.27870], https://hal-cea.archives-ouvertes.fr/cea-02141266

[23] G. VAROQUAUX, R. POLDRACK. *Predictive models avoid excessive reductionism in cognitive neuroimaging*, in "Current Opinion in Neurobiology", April 2019, vol. 55, forthcoming [*DOI :* 10.1016/J.CONB.2018.11.002], https://hal.archives-ouvertes.fr/hal-01856412

International Conferences with Proceedings

[24] P. ABLIN, D. FAGOT, H. WENDT, A. GRAMFORT, C. FÉVOTTE. *A Quasi-Newton Algorithm on the Orthogonal Manifold for NMF with Transform Learning*, in "IEEE-ICASSP 2019 - International Conference on Acoustics, Speech and Signal Processing", Brighton, United Kingdom, Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 2019, https://hal.archives-ouvertes.fr/hal-02346829

[25] S. AYDÖRE, B. THIRION, G. VAROQUAUX. *Feature Grouping as a Stochastic Regularizer for High-Dimensional Structured Data*, in "NeurIPS 2019 - 33th Annual Conference on Neural Information Processing Systems", Vancouver, Canada, December 2019, https://hal.inria.fr/hal-02318458

[26] T. BIN MASOOD, J. BUDIN, M. FALK, G. FAVELIER, C. GARTH, C. GUEUNET, P. GUILLOU, L. HOFMANN, P. HRISTOV, A. KAMAKSHIDASAN, C. KAPPE, P. KLACANSKY, P. LAURIN, J. A. LEVINE, J. LUKASCZYK, D. SAKURAI, M. SOLER, P. STENETEG, J. TIERNY, W. USHER, J. VIDAL, M. WOZNIAK. *An Overview of the Topology ToolKit*, in "TopoInVis 2019 - Topological Methods in Data Analysis and Visualization", Nykoping, Sweden, June 2019, https://hal.archives-ouvertes.fr/hal-02159838

[27] H. CHERKAOUI, T. MOREAU, A. HALIMI, P. CIUCIU. *Sparsity-based blind deconvolution of neural activation signal in fMRI*, in "IEEE-ICASSP 2019 - International Conference on Acoustics, Speech and Signal Processing", Brighton, United Kingdom, May 2019, https://hal.inria.fr/hal-02085810

[28] O. D. DOMINGUES, P. CIUCIU, D. LA ROCCA, P. ABRY, H. WENDT. *Multifractal analysis for cumulant-based epileptic seizure detection in eeg time series*, in "ISBI 2019 - IEEE International Symposium on Biomedical Imaging", Venise, Italy, April 2019, https://hal.inria.fr/hal-02108099

[29] L. EL GUEDDARI, P. CIUCIU, E. CHOUZENOUX, A. VIGNAUD, J.-C. PESQUET. *Calibrationless oscar-based image reconstruction in compressed sensing parallel MRI*, in "ISBI 2019 - IEEE International Symposium on Biomedical Imaging", Venise, Italy, April 2019, https://hal.inria.fr/hal-02101262

[30] P. FILIPIAK, R. FICK, A. PETIET, M. SANTIN, A.-C. PHILIPPE, S. LEHÉRICY, R. DERICHE, D. WASSERMANN. *Coarse-Grained Spatiotemporal Acquisition Design for Diffusion MRI*, in "ISBI 2019 - IEEE International Symposium on Biomedical Imaging", Venice, Italy, April 2019, https://hal.inria.fr/hal-01973588

[31] L. E. GUEDDARI, E. CHOUZENOUX, A. VIGNAUD, J.-C. PESQUET, P. CIUCIU. *Online MR image reconstruction for compressed sensing acquisition in T2* imaging*, in "SPIE Conference - Wavelets and Sparsity XVIII", San Diego, United States, August 2019, https://hal.inria.fr/hal-02265538

[32] A. HALIMI, P. CIUCIU, A. MCCARTHY, S. MCLAUGHLIN, G. S. BULLER. *Fast adaptive scene sampling for single-photon 3D lidar images*, in "IEEE CAMSAP 2019 - International Workshop on Computational Advances in Multi-Sensor Adaptive Processing", Le Gosier (Guadeloupe), France, December 2019, https://hal.inria.fr/hal-02298998

[33] H. JANATI, T. BAZEILLE, B. THIRION, M. CUTURI, A. GRAMFORT. *Group level MEG/EEG source imaging via optimal transport: minimum Wasserstein estimates*, in "IPMI 2019 - 26th international conference on

Information Processing in Medical Imaging", Hong Kong, Hong Kong SAR China, SPRINGER (editor), June 2019, vol. Lecture Notes in Computer Science, https://arxiv.org/abs/1902.04812 , https://hal.inria.fr/hal-02013889

[34] H. JANATI, M. CUTURI, A. GRAMFORT. *Wasserstein regularization for sparse multi-task regression*, in "AISTATS 2019 - 22nd International Conference on Artificial Intelligence and Statistics", Naha, Japan, April 2019, vol. 89, https://hal.inria.fr/hal-02304176

[35] Z. RAMZI, P. CIUCIU, J.-L. STARCK. *Benchmarking proximal methods acceleration enhancements for CS-acquired MR image analysis reconstruction*, in "SPARS 2019 - Signal Processing with Adaptive Sparse Structured Representations Workshop", Toulouse, France, July 2019, https://hal.inria.fr/hal-02298569

[36] Z. RAMZI, P. CIUCIU, J.-L. STARCK. *Benchmarking Deep Nets MRI Reconstruction Models on the FastMRI Publicly Available Dataset*, in "ISBI 2020 - International Symposium on Biomedical Imaging", Iowa City, United States, April 2020, https://hal.inria.fr/hal-02436223

[37] M. SCETBON, G. VAROQUAUX. *Comparing distributions: $l1$ geometry improves kernel two-sample testing*, in "NeurIPS 2019 - 33th Conference on Neural Information Processing Systems", Vancouver, Canada, 2019, https://hal.inria.fr/hal-02292545

Conferences without Proceedings

[38] P. ABLIN, D. FAGOT, H. WENDT, A. GRAMFORT, C. FÉVOTTE. *A quasi-Newton algorithm on the orthogonal manifold for NMF with transform learning*, in "IEEE-ICASSP 2019 - International Conference on Acoustics, Speech and Signal Processing", Brighton, United Kingdom, May 2019, https://arxiv.org/abs/1811.02225 , https://hal.archives-ouvertes.fr/hal-01912918

[39] P. ABLIN, A. GRAMFORT, J.-F. CARDOSO, F. BACH. *Stochastic algorithms with descent guarantees for ICA*, in "AISTATS 2019 - 22nd International Conference on Artificial Intelligence and Statistics", Naha, Japan, April 2019, https://hal.archives-ouvertes.fr/hal-02372092

[40] T. BAZEILLE, H. RICHARD, H. JANATI, B. THIRION. *Local Optimal Transport for Functional Brain Template Estimation*, in "IPMI 2019 - 26th International Conference on Information Processing in Medical Imaging", Hong Kong, China, June 2019 [*DOI :* 10.1007/978-3-030-20351-1_18], https://hal.archives-ouvertes.fr/hal-02278663

[41] H. CHERKAOUI, T. MOREAU, A. HALIMI, P. CIUCIU. *fMRI BOLD signal decomposition using a multivariate low-rank model*, in "Eusipco 2019 - 27th European Signal Processing Conference", Corunna, Spain, September 2019, https://hal.archives-ouvertes.fr/hal-02163497

[42] L. EL GUEDDARI, P. CIUCIU, E. CHOUZENOUX, A. VIGNAUD, J.-C. PESQUET. *Online compressed sensing MR image reconstruction for high resolution T2* imaging*, in "ISMRM 2019 - 27th Annual Meeting and Exhibition", Montréal, Canada, May 2019, https://hal.archives-ouvertes.fr/hal-02314904

[43] L. EL GUEDDARI, P. CIUCIU, E. CHOUZENOUX, A. VIGNAUD, J.-C. PESQUET. *OSCAR-based reconstruction for compressed sensing and parallel MR imaging*, in "ISMRM 2019 - 27th Annual Meeting and Exhibition", Montréal, Canada, May 2019, https://hal.archives-ouvertes.fr/hal-02314911

[44] C. LAZARUS, P. WEISS, F. MAUCONDUIT, A. VIGNAUD, P. CIUCIU. *3D SPARKLING for accelerated ex vivo T2*-weighted MRI with compressed sensing*, in "ISMRM 2019 - 27th Annual Meeting & Exhibition", Montréal, France, May 2019, https://hal.archives-ouvertes.fr/hal-02374538

[45] M. MASSIAS, S. VAITER, A. GRAMFORT, J. SALMON. *Exploiting regularity in sparse Generalized Linear Models*, in "SPARS 2019 - Signal Processing with Adaptive Sparse Structured Representations", Toulouse, France, July 2019, https://hal-univ-bourgogne.archives-ouvertes.fr/hal-02288859

[46] T.-B. NGUYEN, J.-A. CHEVALIER, B. THIRION. *ECKO: Ensemble of Clustered Knockoffs for robust multivariate inference on MRI data*, in "IPMI 2019 - International Conference on Information Processing in Medical Imaging", Hong Kong, Hong Kong SAR China, June 2019, https://hal.archives-ouvertes.fr/hal-02076510

[47] D. SABBAGH, P. ABLIN, G. VAROQUAUX, A. GRAMFORT, D. ENGEMANN. *Manifold-regression to predict from MEG/EEG brain signals without source modeling*, in "NeurIPS 2019 - 33th Annual Conference on Neural Information Processing Systems", Vancouver, Canada, December 2019, https://hal.archives-ouvertes.fr/hal-02147708

Scientific Books (or Scientific Book chapters)

[48] L. ESCH, C. DINH, E. LARSON, D. ENGEMANN, M. JAS, S. KHAN, A. GRAMFORT, M. HÄMÄLÄINEN. *MNE: Software for Acquiring, Processing, and Visualizing MEG/EEG Data*, in "Magnetoencephalography", Springer International Publishing, October 2019, pp. 355-371 [*DOI :* 10.1007/978-3-030-00087-5_59], https://hal.inria.fr/hal-02369299

Research Reports

[49] P. CIUCIU, A. KAZEYKINA. *Anisotropic compressed sensing for non-Cartesian MRI acquisitions*, CEA Paris Saclay ; Inria ; Université Paris Sud, Université Paris Saclay, October 2019, https://arxiv.org/abs/1910.14513 , https://hal.archives-ouvertes.fr/hal-02339689

Other Publications

[50] S. ABBOUD, D. ENGEMANN, L. COHEN. *Semantic coding in the occipital cortex of early blind individuals*, February 2019, working paper or preprint [*DOI :* 10.1101/539437], https://hal.archives-ouvertes.fr/hal-02018272

[51] P. ABLIN, T. MOREAU, M. MASSIAS, A. GRAMFORT. *Learning step sizes for unfolded sparse coding*, May 2019, working paper or preprint, https://hal.inria.fr/hal-02140383

[52] F. BAILLIN, A. LEFEBVRE, A. PEDOUX, Y. BEAUXIS, D. ENGEMANN, A. MARUANI, F. AMSELLEM, T. BOURGERON, R. DELORME, G. DUMAS. *Interactive Psychometrics for Autism with the Human Dynamic Clamp: Interpersonal Synchrony from Sensory-motor to Socio-cognitive Domains*, January 2020, Literature reference for the preprint first posted on medRxiv: https://www.medrxiv.org/content/10.1101/19013771v1 [*DOI :* 10.1101/19013771], https://hal.inria.fr/hal-02396923

[53] H. BANVILLE, I. ALBUQUERQUE, A. HYVÄRINEN, G. MOFFAT, D. ENGEMANN, A. GRAMFORT. *Self-supervised representation learning from electroencephalography signals*, November 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02361350

[54] Q. BERTRAND, M. MASSIAS, A. GRAMFORT, J. SALMON. *Handling correlated and repeated measurements with the smoothed Multivariate square-root Lasso*, September 2019, https://arxiv.org/abs/1902.02509 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02010014

[55] P. CERDA, G. VAROQUAUX. *Encoding high-cardinality string categorical variables*, July 2019, https://arxiv.org/abs/1907.01860 - working paper or preprint, https://hal.inria.fr/hal-02171256

[56] L. EL GUEDDARI, E. CHOUZENOUX, A. VIGNAUD, P. CIUCIU. *Calibration-less parallel imaging compressed sensing reconstruction based on OSCAR regularization*, September 2019, working paper or preprint, https://hal.inria.fr/hal-02292372

[57] D. ENGEMANN, O. KOZYNETS, D. SABBAGH, G. LEMAÎTRE, G. VAROQUAUX, F. LIEM, A. GRAMFORT. *Combining electrophysiology with MRI enhances learning of surrogate-biomarkers*, December 2019, working paper or preprint [*DOI :* 10.1101/856336], https://hal.archives-ouvertes.fr/hal-02395406

[58] B. HERMANN, F. RAIMONDO, L. HIRSCH, Y. HUANG, M. VALENTE, P. PÉREZ, D. ENGEMANN, F. FAUGERAS, N. WEISS, S. DEMERET, B. ROHAUT, L. PARRA, J. D. SITT, L. NACCACHE. *Combined behavioral and electrophysiological evidence for a direct cortical effect of prefrontal tDCS on disorders of consciousness*, April 2019, working paper or preprint [*DOI :* 10.1101/612309], https://hal.archives-ouvertes.fr/hal-02107872

[59] A. JAISWAL, J. NENONEN, M. STENROOS, A. GRAMFORT, S. S. DALAL, B. U. WESTNER, V. LITVAK, J. C. MOSHER, J.-M. SCHOFFELEN, C. WITTON, R. OOSTENVELD, L. PARKKONEN. *Comparison of beamformer implementations for MEG source localization*, November 2019, working paper or preprint [*DOI :* 10.1101/795799], https://hal.inria.fr/hal-02369296

[60] H. JANATI, T. BAZEILLE, B. THIRION, M. CUTURI, A. GRAMFORT. *Multi-subject MEG/EEG source imaging with sparse multi-task regression*, October 2019, working paper or preprint, https://hal.inria.fr/hal-02304194

[61] H. JANATI, M. CUTURI, A. GRAMFORT. *Spatio-Temporal Alignments: Optimal transport through space and time*, October 2019, working paper or preprint, https://hal.inria.fr/hal-02309340

[62] J. JOSSE, N. PROST, E. SCORNET, G. VAROQUAUX. *On the consistency of supervised learning with missing values*, March 2019, https://arxiv.org/abs/1902.06931 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02024202

[63] D. LA ROCCA, P. CIUCIU, D. ENGEMANN, V. VAN WASSENHOVE. *Emergence of β and γ networks following multisensory training*, February 2019, working paper or preprint [*DOI :* 10.1101/560235], https://hal.archives-ouvertes.fr/hal-02052443

[64] C. LAZARUS, P. WEISS, L. E. GUEDDARI, F. MAUCONDUIT, A. VIGNAUD, P. CIUCIU. *3D SPARKLING trajectories for high-resolution T2*-weighted Magnetic Resonance imaging*, March 2019, working paper or preprint, https://hal.inria.fr/hal-02067080

[65] A. MACHLOUZARIDES-SHALIT, V. IOVENE, N. MAKRIS, D. WASSERMANN. *A Novel Sulcal Hierarchy Based on Manually Labelled Sulci*, June 2019, Organization for Human Brain Mapping, Poster, https://hal.inria.fr/hal-02347268

[66] M. MASSIAS, Q. BERTRAND, A. GRAMFORT, J. SALMON. *Support recovery and sup-norm convergence rates for sparse pivotal estimation*, January 2020, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02444978

[67] M. MASSIAS, S. VAITER, A. GRAMFORT, J. SALMON. *Dual Extrapolation for Sparse Generalized Linear Models*, August 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02263500

[68] V. MENON, G. GUILLERMO, M. PINSK, V.-D. NGUYEN, J.-R. LI, W. CAI, D. WASSERMANN. *Quantitative modeling links in vivo microstructural and macrofunctional organization of human and macaque insular cortex, and predicts cognitive control abilities*, January 2020, working paper or preprint [*DOI :* 10.1101/662601], https://hal.inria.fr/hal-02434382

[69] T. MOREAU, A. GRAMFORT. *DiCoDiLe: Distributed Convolutional Dictionary Learning*, November 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02371715

[70] D. SABBAGH, P. ABLIN, G. VAROQUAUX, A. GRAMFORT, D. ENGEMANN. *Predictive regression modeling with MEG/EEG: from source power to signals and cognitive states*, November 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02367411

[71] M.-A. SCHULZ, B. T. YEO, J. T. VOGELSTEIN, J. MOURAO-MIRANADA, J. N. KATHER, K. KORDING, B. A. RICHARDS, D. BZDOK. *Deep learning for brains?: Different linear and nonlinear scaling in UK Biobank brain images vs. machine-learning datasets*, September 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02276649

References in notes

[72] K. S. BUTTON, J. P. IOANNIDIS, C. MOKRYSZ, B. A. NOSEK, J. FLINT, E. S. ROBINSON, M. R. MUNAFÒ. *Power failure: why small sample size undermines the reliability of neuroscience*, in "Nature Reviews Neuroscience", 2013, vol. 14, n^{o} 5, pp. 365–376

[73] R. A. POLDRACK, C. I. BAKER, J. DURNEZ, K. J. GORGOLEWSKI, P. M. MATTHEWS, M. R. MUNAFÒ, T. E. NICHOLS, J.-B. POLINE, E. VUL, T. YARKONI. *Scanning the horizon: towards transparent and reproducible neuroimaging research*, in "Nature Reviews Neuroscience", 2017, vol. 18, n^{o} 2, pp. 115–126