Post-doc proposal : Approximate Message Passing for the efficient identification of cognitive networks (BrainAMP)

Description of the project

Context and Motivation In many fields of physics or life sciences, improving the resolution of observations (signal, images, spectra) is a major endeavor, as it is a pre-requisite toward more accurate information. Data acquisition devices benefit thus of hardware improvement to increase the resolution of the observed phenomena, leading to ever larger datasets. From a statistical perspective, these datasets have a high dimensionality and the signals show some prominent structures that require adequate modeling. While the **dimensionality has increased**, the **number of samples available is sometimes limited**, due to physical or financial limits. This becomes a problem when these data are processed with estimators that have a large sample complexity, such as many multivariate estimators (classifiers, regression models, covariance estimators, structure learning). In that case it is very useful to rely on **structured priors**, so that the resulting models reflect the state of knowledge on the phenomena of interest. Well-chosen priors improve the accuracy of the models and decrease the sample complexity of the estimators.

The study of the human brain activity through functional Magnetic Resonance Imaging belongs among these problems. The number of features per image reaches 10^5 to 10^6 —thanks to the rise of high-field MRI acquisitions that cross the mm scale— yet the number of observations is limited by the duration of scanning sessions and the number of subjects that can be included in studies. The key challenge addressed here is to **set up a novel generation of efficient techniques to enforce structured priors on high-resolution MRI datasets** to improve statistical analysis.

State-of-the art and positioning of BrainAMP partners The Parietal team has set up analytic tools for predictive modeling on these images, using the framework of convex M-estimators [12, 6, 2, 1, 4], obtaining state-of-the-art and nearly computationally optimal solutions. An example is given in Fig. 1.

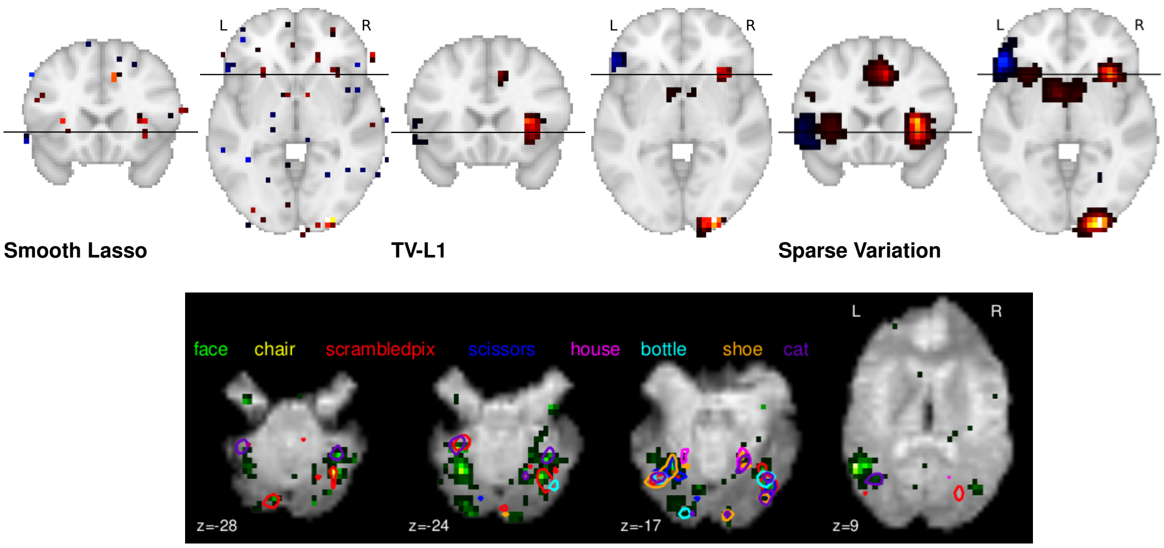

FIGURE 1 – (Top) Weight vectors from estimating gain from brain functional activity on a gambling task [14]. Prediction target is the gain proposed in a series of gambles proposed to the subjects. This inter-subject analysis shows broad regions of activation. Sophisticated estimators, such as Sparse variation and TV-l1, yield a clean map and obtain high accuracy. Parameter setting for these estimators is also easier than the cheaper Smooth Lasso estimator. See [4] for details. (Bottom) These effective estimators can be used to single out specialized regions in the ventral occipital cortex from [7], providing an elegant solution to fundamental brain mapping problems.

In this field, Parietal is a major player when it comes to applying structured penalties to multivariate estimators of brain activity. Yet an essential problem remains computation time, because of the non-smooth optimization steps involved when imposing these penalties. Another major open question is to gain statistical control (e.g. confidence intervals, p-values) on the solutions of this problem : for instance, if a certain pattern of activity predicts an autism diagnosis, one would like to test whether a given brain region is significantly loaded by this pattern.

The IPhT partner, on the other hand, is an expert on the so-called **Approximate Message Passing** (AMP) techniques [13, 9, 10, 11]. AMP is a class of algorithms that derive from a probabilistic setting of the inference problem where m data samples are presented, and each data sample $\mathbf{F}_\mu \in \mathbb{R}^n$, $\mu = 1, \ldots, m$ is associated with an observation $y_\mu \in \mathbb{R}$ (e.g. $\mathbf{F}_\mu$ is a brain image and y_μ an associated label). The goal is to estimate a latent sparse vector $\mathbf{x} \in \mathbb{R}^n$ such that $z_\mu = \sum_{i=1}^n \mathbf{F}_{\mu i}\mathbf{x}_i$ explains well the observations y_μ. AMP estimates marginals (and their variances) of a posterior distribution of $\mathbf{x}$ given $\mathbf{F}$ and y, it is able to deal with a rather generic class of priors on the vector $\mathbf{x}$ and "output channels" $P_{\text{out}}(y_\mu|z_\mu)$. AMP is a message passing algorithm, related to loopy belief propagation ; as such it is a distributed iterative procedure making it computationally very efficient. There are well studied strategies that ensure that this iterative procedure converges ; they have been shown to work well for problems such as the present one, as studied in [17]. In a range of statistical models, AMP works optimally in the low-sample regime, where it outperforms (both in speed and precision) more traditional sparse estimation methods [8]. AMP methods have been shown to perform well in compressed sensing, where they achieve information theoretically optimal performance [9, 3], or in tomographic reconstruction, both problems formally similar to the one posed by Parietal researchers [5]. AMP methods have several very appealing features, one of them is the possibility to obtain approximate confidence intervals on the solution (indeed it outputs the estimators as well as variances of the latent variables). Another promising feature is a natural way to tune hyper-parameters of the model by imposing consistency of expectations, analogously to expectation maximization.

Objectives, Planning and deliverable We thus propose to join forces between the two groups to design **a new generation of inverse problem solvers** that can take into account the complex structure of brain images and provide guarantees in the low-sample regime where they are used. To this end, we will first adapt AMP tools to the neuroimaging setting, using first standard sparsity priors (a.k.a. ℓ_0 norm, Gauss-Bernoulli etc.) on the estimated model. In this we will follow the approach described in [17]. We will then consider more complex structured priors, that control the variation of the learned image patterns in space [16, 15]. This is related to the well known *analysis sparsity* framework : the signals of interest should have a sparse representation in a well-chosen signal basis.

Crucial gains are expected from the use of the variational strategy for parameter setting (EM type of algorithm) that comes naturally with AMP approaches [10]. An important task will be to benchmark the resulting estimator against alternatives, in terms of prediction accuracy, support recovery and computation time. We expect different regimes, depending on the number of features or samples in the dataset and the noise level in the data. We will also examine the statistical guarantees provided by the AMP approach : are the confidence intervals returned by the estimators reliable, compared to e.g. bootstrapped estimators, so that non-expert neuroscientists can rely on them ?

The project will deliver a **technical publication** and a **software implementation** in Python, which, after a thorough assessment (code quality control, efficiency, simplicity of the API, documentation), will be included in the Nilearn open source library. In parallel, these contributions will serve as a basis for large-scale analyses carried out on human cognition based on the data accumulated at Neurospin ; in the longer term they will also be considered for other problems (image reconstruction, compressed sensing). The efficiency of the estimator will be crucial for the feasibility of the approach on large-scale data (existing datasets now scale in tens of Gigabytes, and keep increasing).

Work environment The Parietal team has set up analytic tools for predictive modeling on these images, using the framework of convex M-estimators [12], with important validation experiments [6] and technical improvement to reduce computation time [2, 1] or finesse the model [4]. These tools are disseminated the Nilearn open-source software (`http://nilearn.github.io`), which has an increasing impact in neuroimaging.

The IPhT partner is an expert on the so-called Approximate Message Passing techniques [13], which are an alternative solution for more generic class of estimators with less formal guarantees, that are more efficient in many settings. For instance, AMP methods have been shown to perform well in problems formally similar to the one posed by Parietal researchers [5, 9].

Further collaborations This BrainAMP proposal is related to the COSMIC project, that proposes enhancements for compressed sensing for MRI and links the parietal team with the CosmoStat laboratory. The two projects share in common the interest in sparse methods for brain imaging, that rely on the statistical structure of images to optimize acquisition or statistical analysis related to these images. As the two projects will run in parallel, they will share intellectual and software contributions.

Duration of the project : 10 months. Further extensions are under active consideration.

Required skills

— Prior experience in machine learning, high-dimensional estimation problems.
— If possible, some knowledge on graphical models and related inference techniques.
— Knowledge of statistical inference.
— Knowledge of Python language.
— Interest in Open Software development.

Contacts

— Bertrand Thirion, Neurospin, Parietal team, bertrand.thirion@inria.fr
— Lenka Zdeborová, Institut de Physique Théorique, lenka.zdeborova@cea.f

How to apply The position must start in 2016, ideally November 1^{st}, 2016. Applicants should send there cv and a short statement of interest to Bertrand Thirion and Lenka Zdeborova starting from July and certainly before August 10th, 2016.

Références

[1] Elvis Dohmatob, Michael Eickenberg, Bertrand Thirion, and Gaël Varoquaux. Speeding-up model-selection in GraphNet via early-stopping and univariate feature-screening. In *PRNI*, Stanford, United States, June 2015.

[2] Elvis Dohmatob, Alexandre Gramfort, Bertrand Thirion, and Gaël Varoquaux. Benchmarking solvers for TV-l1 least-squares and logistic regression in brain imaging. In *Pattern Recoginition in Neuroimaging (PRNI)*, Tübingen, Allemagne, April 2014. IEEE.

[3] David L Donoho, Adel Javanmard, and Alessandro Montanari. Information-theoretically optimal compressed sensing via spatial coupling and approximate message passing. *Information Theory, IEEE Transactions on*, 59(11) :7434–7464, 2013.

[4] Michael Eickenberg, Elvis Dohmatob, Bertrand Thirion, and Gaël Varoquaux. Total Variation meets Sparsity : statistical learning with segmenting penalties. In *Medical Image Computing and Computer Aided Intervention (MICCAI)*, Proceedings of MICCAI 2015, München, Germany, October 2015.

[5] Emmanuelle Gouillart, Florent Krzakala, Marc Mezard, and Lenka Zdeborová. Belief-propagation reconstruction for discrete tomography. *Inverse Problems*, 29(3) :035003, 2013.

[6] Alexandre Gramfort, Bertrand Thirion, and Gaël Varoquaux. Identifying predictive regions from fMRI with TV-L1 prior. In *Pattern Recognition in Neuroimaging (PRNI)*, Philadelphia, États-Unis, June 2013. IEEE. ANR grant BrainPedia, ANR-10-JCJC 1408-01, FMJH Program Gaspard Monge in optimization and operation research with support from EDF.

[7] James V. Haxby, Ida M. Gobbini, Maura L. Furey, Alumit Ishai, Jennifer L. Schouten, and Pietro Pietrini. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293 :2425, 2001.

[8] Florent Krzakala, Andre Manoel, Eric W Tramel, and Lenka Zdeborova. Variational free energies for compressed sensing. In *Information Theory (ISIT), 2014 IEEE International Symposium on*, pages 1499–1503. IEEE, 2014.

[9] Florent Krzakala, Marc Mézard, François Sausset, YF Sun, and Lenka Zdeborová. Statistical-physics-based reconstruction in compressed sensing. *Physical Review X*, 2(2) :021005, 2012.

[10] Florent Krzakala, Marc Mézard, Francois Sausset, Yifan Sun, and Lenka Zdeborová. Probabilistic reconstruction in compressed sensing : algorithms, phase diagrams, and threshold achieving matrices. *Journal of Statistical Mechanics : Theory and Experiment*, 2012(08) :P08009, 2012.

[11] Andre Manoel, Florent Krzakala, Eric Tramel, and Lenka Zdeborovà. Swept approximate message passing for sparse estimation. In *Proceedings of the 32nd International Conference on Machine Learning (ICML-15)*, pages 1123–1132, 2015.

[12] Vincent Michel, Alexandre Gramfort, Gaël Varoquaux, Evelyn Eger, and Bertrand Thirion. Total variation regularization for fMRI-based prediction of behaviour. *IEEE Transactions on Medical Imaging*, 30(7) :1328 – 1340, February 2011.

[13] Sundeep Rangan. Generalized approximate message passing for estimation with random linear mixing. In *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, pages 2168–2172. IEEE, 2011.

[14] Sabrina M Tom, Craig R Fox, Christopher Trepel, and Russell A Poldrack. The neural basis of loss aversion in decision-making under risk. *Science*, 315(5811) :515–8, Jan 2007.

[15] Eric W Tramel, Angélique Drémeau, and Florent Krzakala. Approximate message passing with restricted boltzmann machine priors. *arXiv preprint arXiv :1502.06470*, 2015.

[16] Xing Wang and Jie Liang. Approximate message passing-based compressed sensing reconstruction with generalized elastic net prior. *Image Commun.*, 37(C) :19–33, September 2015.

[17] Justin Ziniel, Philip Schniter, and Per Sederberg. Binary linear classification and feature selection via generalized approximate message passing. In *Information Sciences and Systems (CISS), 2014 48th Annual Conference on*, pages 1–6. IEEE, 2014.