Accepted Manuscript

Improved statistical evaluation of group differences in connectomes by screening-filtering strategy with application to study maturation of brain connections between childhood and adolescence

Djalel-Eddine Meskaldji, Lana Vasung, David Romascano, Jean-Philippe Thiran, Patric Hagmann, Stephan Morgenthaler, Dimitri Van De Ville

Please cite this article as: Meskaldji, Djalel-Eddine, Vasung, Lana, Romascano, David, Thiran, Jean-Philippe, Hagmann, Patric, Morgenthaler, Stephan, Van De Ville, Dimitri, Improved statistical evaluation of group differences in connectomes by screening-filtering strategy with application to study maturation of brain connections between childhood and adolescence, *NeuroImage* (2014), doi: 10.1016/j.neuroimage.2014.11.059

Improved statistical evaluation of group differences in connectomes by screening-filtering strategy with application to study maturation of brain connections between childhood and adolescence

Djalel-Eddine Meskaldji[a,b,*], Lana Vasung[c], David Romascano[d], Jean-Philippe Thiran[d,e], Patric Hagmann[e], Stephan Morgenthaler[f], Dimitri Van De Ville[a,b]

[a]*Institute of Bioengineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland*
[b]*Department of radiology and medical informatics, University of Geneva, Geneva, Switzerland*
[c]*Division of Development and Growth. Department of Pediatrics. University of Geneva, Geneva, Switzerland*
[d]*LTS5, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland*
[e]*Department of Radiology, University Hospital Center and University of Lausanne (CHUV-UNIL), Lausanne, Switzerland*
[f]*Applied statistics, Institute of mathematics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland*

Abstract

Detecting local differences between groups of connectomes is a great challenge in neuroimaging, because the large number of tests that have to be performed and the impact on multiplicity correction. Any available information should be exploited to increase the power of detecting true between-group effects. We present an adaptive strategy that exploits the data structure and the prior information concerning positive dependence between nodes and connections, without relying on strong assumptions. As a first step, we decompose the brain network, i.e., the connectome, into subnetworks and we apply a screening at the subnetwork level. The subnetworks are defined either according to prior knowledge or by applying a data driven algorithm. Given the results of the screening step, a filtering is performed to seek real differences at the node/connection level. The proposed strategy could be used to strongly control either the family-wise error rate or the false discovery rate. We show by means of different simulations the benefit of the proposed strategy, and we present a real application of comparing connectomes of preschool children and adolescents.

Keywords: Brain connectivity, complex networks, multiple testing, family-wise error rate (FWER), false discovery rate FDR, fronto-limbic circuitry, adolescence.

Introduction

The study of brain connectivity has become an important aspect of neuroscience as it can help to understand brain organization and function (Fornito et al., 2013; Sporns, 2011). Moreover, the metrics of brain connectivity, assessed through neuroimaging methods, have been recognized as an important marker indicating the level of brain maturation or psychopathology. Through recent innovations in medical imaging and image analysis, the determination of interregional brain connectivity became feasible. Different types of connectivity can be obtained depending on the imaging modality and measure of connectivity, e.g., structural connectivity from diffusion-weighted MRI and fiber tracking (Cammoun et al., 2012; Hagmann et al., 2008), or functional connectivity from functional MRI and statistical dependence on time (Smith et al., 2013; Friston, 2011; van den Heuvel and Hulsoff-Pol, 2010; Achard et al., 2006). Global brain connectivity can be modeled by a network (a weighted graph) called *connectome* (Sporns et al., 2005), where the N nodes stand for brain regions of interest (ROIs), and each edge weight characterizes a measure of connectivity between pairs of ROIs.

Investigating differences in connectivity between distinct populations based on connectivity matrices is attractive, but also comes with a certain number of problems (Fornito et al., 2013; Varoquaux and Craddock, 2013), among them, the high number of multiple comparisons.

Effectively, when the comparison between brain networks are studied at the level of nodes (vertices) ($O(N)$) that represent brain ROIs, or connections (edges) ($O(N^2)$) that link brain ROIs, a huge number of tests has to be performed on the same data, especially, in the case of testing at the level of connections, in which the number of tests basically grows quadratically with the number of nodes. If the multiplicity of tests is ignored, the risk of committing false discoveries increases. As a consequence, erroneous conclusions are frequently drawn (Meskaldji et al., 2013a). On the other hand, considering multiplicity could dramatically decrease the chance of detecting real between-group effects. This is a fact that is commonly reported by researchers especially when the conventional Bonferroni procedure is used for the multiplicity correction and the strong control on the number of false discoveries is exerted. For example, if $N = 100$ nodes, the Bonferroni threshold for significant p-values is of order 10^{-4} when testing at the level of nodes and 10^{-6} when testing at the level of connections.

Depending on the field of application and the nature of the data,

---

*djalel.meskaldji@epfl.ch

many strategies have been adopted in order to face the multiplicity challenge in the presence of positively correlated test statistics. These strategies consist in exploiting the data structure and positive dependence that could be present between tests. This can have an important impact on the power of detecting true alternatives. For example, this concept was adopted in the widely recognized software package for analyzing fMRI data, the Statistical Parametric Mapping (SPM) (Friston et al., 1995; Frackowiak et al., 1997) and in its extensions such as wavelet extensions (Van De Ville et al., 2007, 2004), based on the idea that voxels of a neurological type belonging to a unique anatomical region will usually exhibit positively correlated behavior (Penny and Friston, 2003; Genovese et al., 1999). In this case, the data is supposed to be smooth and follow a multidimensional Gaussian distribution. For this reason, a smoothing has to be applied to the data (Nichols and Hayasaka, 2003). A permutation approach is performed to define active clusters.

The same concept has been followed to derive specific statistical methods in the brain connectivity context. Zalesky et al. (2010) proposed the network based statistic (NBS) as a method to correct for the FWER (the probability of having at least one false positive connection), in the framework of multiple testing applied to the brain network connections. The method relies on a first identification of connected components (in the graph theoretical sense), by thresholding the set of p-values at an arbitrary threshold. An iterative procedure based on permutation testing allows thereafter identifying connected components that carry a between-group effect. These methods have, however, some limitations. First, the inference is obtained at the level of connected components and only exerts a weak FWER control, that is, once a component is declared to be significantly different, nothing could be said at the level of individual nodes or connections belonging to the component. In other words, the type I error metric controlled at the level of nodes/connections is unknown. Second, the results strongly depend on the arbitrary choice of the threshold. The same can be said about the spatial pairwise clustering (SPC), proposed by the same authors, where the definition of components is based on geometrical distance in addition to the connectedness in the graph theoretical sense (Zalesky et al., 2012).

It is commonly admitted that most mental diseases or cognitive trait exhibit changes not in the entire brain uniformly, but rather specific in functional systems or brain regions and this to a different extent. Meskaldji et al. (2011a) proposed an adaptive strategy that exploits the network structure of the brain connectivity by considering brain subnetworks, which results in reducing the number of tests and a considerable improvement in power. The strategy was applied to detect differences in both structural and functional brain connectivity (Owen et al., 2013a,b; Meskaldji et al., 2011a). However, besides the generality of this strategy in terms of summary statistics that could be used, and in terms of the diversity of the brain decomposition methods that could be applied, it suffers from the same drawback as the NBS and the SPC, that is, nothing could be said concerning the statistical evidence of nodes and connections that constitute the significant subnetworks. Nevertheless, the subnetworks could be chosen as small as possible to obtain

statistical evidence at finer scales. We will give throughout this paper some highlights on the differences between these weak control methods.

The question that we will investigate in this paper is to go beyond the cluster/subnetwork level and investigate the differences at the single node/connection level. Inspired by Benjamini and Heller (2007), we propose a screening-filtering strategy that exploits the data structure and positive dependence that could exist between connections/nodes. The advantage of the proposed strategy is that it exerts a strong control of type I error rates under weak assumptions (i.e., weaker than assumptions needed by SPM, NBS and SPC). We study the performance of the screening-filtering approach on simulated networks and on structural brain connectivity matrices. In particular, we examine the influence of the network decomposition and the screening threshold on the statistical inference. We also discuss the conceptual differences between our proposed strategy and the some of existing methods in the literature.

As far as we know, this method is the first adaptive strategy that guarantees the strong control of type I error rate at the level of nodes or connections. For this reason, the performances of the proposed strategy will only be compared to the standard node/connection-wise inference, that is, methods that exerts a strong control, but do not consider neither data structure nor positive dependence between tests.

The paper is organized as follows. We first give the general processing pipeline and the mathematical formulation of the screening-filtering approach. Then, we show by simulations, the benefit of using the proposed strategy. Finally, we present a practical application on real data, from children and adolescents, which consists in comparing structural human brain connectomes between these populations.

Methods

We present in this section the different steps of local procedures that exert a strong false positives control. In particular, we outline two strategies: the standard methods and the screening and filtering methods.

Local network-based measures

Since the imaging measures of connectivity can be used to model the brain as a network, it is worth to locally compare populations not only cell by cell of the connectivity matrices, but also by estimating the network measures that characterize the topological properties of the brain network (Fornito et al., 2013; Meskaldji et al., 2013a; Bassett et al., 2008). The combination of the local and the global inferences gives a better understanding of the network organisation (Meskaldji et al., 2013a). In this paper, we focus on the local measures. Sporns (2011); Rubinov and Sporns (2010) among others are good sources for a comprehensive list of important measures with their interpretations in the brain connectivity context.

For non-homogeneous populations, it is strongly recommended to correct for covariables such as the age or the gender of the subjects, by taking the residuals of a regression as the new
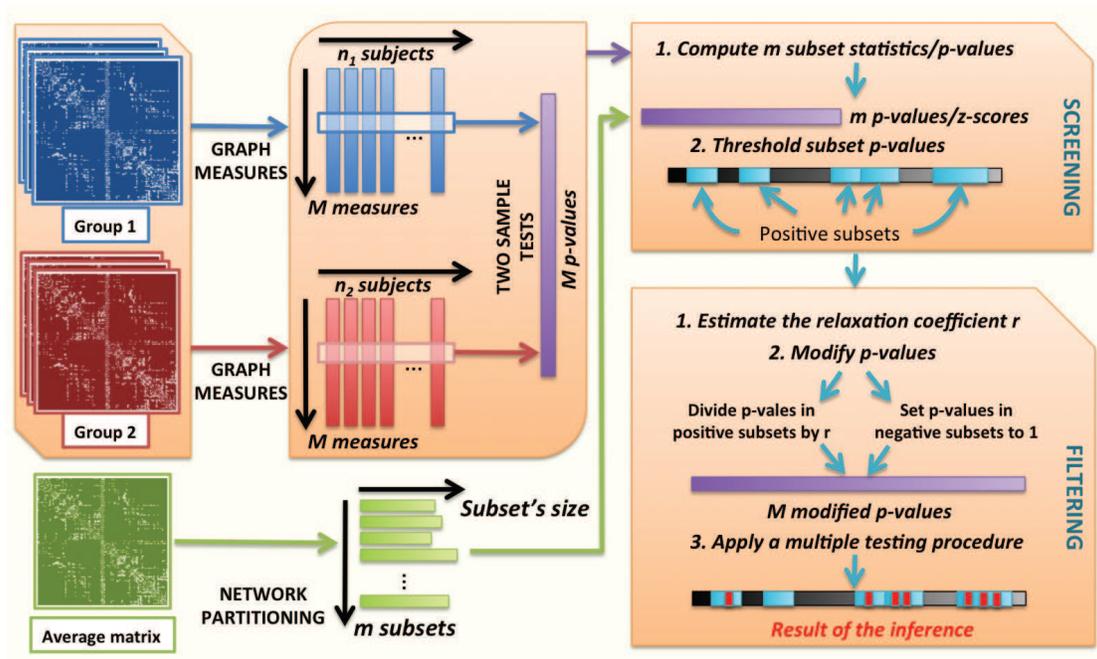
Figure 1: The different steps of a complete local analysis using the screening-filtering algorithm. The algorithm takes as an input two (or more) groups of connectivity matrices to be compared, as well as a prior decomposition of the brain network. If the latter is not available, a decomposition is obtained by applying a data-driven decomposition algorithm. The screening-filtering ends up with a set of modified p-values that could be corrected by a multiple testing procedure to obtain the statistical inference.

observations (Meskaldji et al., 2013a).

This step ends up with a vector of local observations for each node/connection and for each subject.

Testing and p-value computation

Let us assume that the aim of a brain connectivity study is to compare different groups of connectomes. Comparing two populations at the level of nodes or connections or any local unit that we call *atom*, usually consists in performing a (univariate or multivariate) two-sample test for each node/connection (Meskaldji and Van De Ville, 2014). When more than two groups are compared, an analysis of variance (ANOVA) is performed with a predefined contrast. This ends up with M p-values, where M is of the order $O(N)$ when testing at the level of nodes, and M goes like $O(N^2)$ when testing at the level of connections.

Let p_j denotes the p-value of atom $j = 1, \ldots, M$. The standard method (SM) consists in performing a multiple testing procedure to the set of p-values to control a type I error metric. For example, one could apply the Bonferroni procedure to the p-values $\{p_1, \ldots, p_M\}$ by declaring significant p-values that satisfy $p_a \leq \alpha/M$. This guarantees the strong control of the family-wise error rate (FWER).

The SM does not exploit the data structure nor the positive dependence between atoms and typically results into poor sensitivity.

The screening and filtering strategy: a general sketch

The screening-filtering (SF) approach takes as an input the raw p-values and consists in the following. First, we group the family of atoms into m subsets. For each subset, we compute the standardized mean of the z-scores within the subset. Based on these mean scores, we perform a screening at a predefined threshold. In other words, we compare the m p-values corresponding to the mean scores to a predefined threshold. This step results in two classes: positive subsets, that is, the subsets with mean score p-value less than the screening threshold, and negative subsets (the remaining subsets). This step is a particular case of the subnetwork analysis proposed in Meskaldji et al. (2011a), since we are using here a particular case of summary statistics, which is the mean.

Note that if the threshold in the screening step corresponds to a multiplicity correction, then the weak control is guaranteed. For example, using the Bonferroni threshold α/m, in the screening step guarantees the weak control of the FWER, that is, at the level of subsets. We compare the screening step performances to other methods that guarantee the weak control in the application section.

Based on the results of the screening step, the practitioner can perform a multiple testing procedure at the level of atoms by computing the p-values conditioned by the statistical results at the first step (see Benjamini and Heller, 2007). However, this solution is time consuming in large data and the control of the FP is not guaranteed for small samples. We propose a procedure where we do not need to estimate the unknown

3

parameters neither to compute the conditional p-values, i.e., we work directly with the unconditioned (original) p-values. We modify the original M p-values such that the modified p-values can be used with any multiple testing procedure to control the desired type I error metric. The modification consists in dividing the p-values belonging to the positive subsets by a coefficient larger then 1, called the *relaxation coefficient*, i.e., the modified p-values are always smaller then the original p-values in the positive subsets. The remaining p-values (that belong to negative subsets), are set to 1. The filtering step follows by applying a multiple testing procedure to the modified p-values, which results in the statistical inference at the level of atoms. The general pipeline of performing group comparison with the screening and filtering strategy is summarized by Figure 1.

The intuition behind this strategy is that the p-value weights (Benjamini and Hochberg, 1997) that were uniform (all equal to 1) before the screening step, are redistributed. The p-value weights in negative subsets are given to positive subsets to enrich p-values therein. The control of the type I error rate is based on conditional probability theory, and hence, the choice of the relaxation coefficient depends on the screening threshold. The relaxation coefficient is chosen such that it guarantees the control of false positives and it becomes larger when the screening coefficient becomes smaller. A mathematical formulation of the algorithm is presented in the following section.

Mathematical formulation

Consider a set of M tests with their corresponding p-values. We group the M tests into m subsets $J_1, \ldots, J_m$ such that $\bigcup_{i=1}^{m} J_i = J$ and $|J_i| = s_i, i = 1, \ldots, m$. The hypotheses h_{ji} are indexed by two indices: the subset index i and the atom index j. The p-values associated with the tests are $p_{ji}, j \in J$ and are related to the corresponding z-scores by $p_{ji} = \overline{\Phi}(Z_{ji})$, where $\overline{\Phi} = 1 - \Phi$ and Φ is the gaussian cumulative distribution function.

Proposition 1
Under the null hypothesis (no real difference), the z-scores follows a normal distribution $\mathcal{N}(0, 1)$.
The proof of this proposition is in the appendix.
Let $\mu_{ji} = \mathbb{E}(Z_{ji})$. Without loss of generality, we consider one-sided tests $h_{ji} = 1$, if $\mu_{ji} > 0$, and $h_{ji} = 0$, otherwise.
We consider the following mixture effect model (Benjamini and Heller, 2007). *Null subsets* are subsets that only contain null hypotheses. Otherwise, the subset is called *affected*. Let $I_0 = \{i : 1, \ldots, m \mid \sum_{j \in J_i} h_{ji} = 0\}$ and $I_1 = \{i : 1, \ldots, m \mid \sum_{j \in J_i} h_{ji} > 0\}$ the indices of null subsets and affected subsets, respectively. The number of null subsets is $|I_0| = m_0$ and the number of affected subsets is $|I_1| = m_1$. The proportion of non-null hypotheses in the subset J_i is $\pi_i = s_i^{-1} \sum_{j \in J_i} h_{ji}$.
The information in each subset is summarized by a summary statistic $T_i = \sqrt{s_i}^{-1} \sum_{j \in J_i} Z_{ji}$. To model the dependence inside each subset J_i, let $\rho_{ji} = corr(T_i; Z_{ji})$ be the correlation between each test $Z_{ji} \in J_i$ with its corresponding subset summary statistic. We assume that $\rho_{ji} > \sqrt{s_i}^{-1}$, which corresponds to

$corr(Z_{ji}, Z_{li}) > 0$ for $(j, l) \in J_i^2$, $i \in I_1$ and $\in j \neq l$. We also assume that $\rho_{ji} = \sqrt{s_i}^{-1}$ for $(j, l) \in J_i^2$, $i \in I_0$ and $j \neq l$.

Proposition 2
According to this model, for a null subset, $T_i \sim \mathcal{N}(0, 1)$.
The proof of this proposition is in the appendix.
Suppose that the p-values corresponding to the summary statistics are $P_i(i = 1, \ldots, m)$, that is, $P_i = \overline{\Phi}(T_i)$. The central limit theorem allows us to make this approximation especially when the size s_i becomes large.
The screening step consists in comparing the subset p-values P_i $(i = 1, \ldots, m)$ to a predefined threshold U. This screening results in two classes of subsets. Let $I^+ = \{i : P_i \leq U\}$ be the positive subsets.
We divide the original p-values inside the positive subsets by the relaxation coefficient r and we apply a filtering using a multiple testing procedure.

Proposition 3
Under the mixture model described above, if the Bonferroni procedure is used in the filtering step at level α, then the expected number of false positives after the filtering step is

$$\mathbb{E}(\mathsf{FP}) = \mathbb{E}\left(\sum_{i \in I^+} s_i \int_c^{\infty} \frac{m_0 \overline{\Phi}(C_0) + m_1(1 - \pi_i) \overline{\Phi}(C_{\mu_i})}{m_0 U + m_1 \overline{\Phi}(\Phi^{-1}(1 - U) - \mu_i)} \varphi(z) dz \right),$$

where φ is the probability density function of the normal distribution, $C_0 = \frac{\Phi^{-1}(1-U) - \rho_{ji} z}{\sqrt{1 - \rho_{ji}^2}}$, $C_{\mu_i} = \frac{\Phi^{-1}(1-U) - \mu_i - \rho_{ji} \sigma_i z}{\sigma_i \sqrt{1 - \rho_{ji}^2}}$, $c = \Phi^{-1}(1 - r\alpha/M)$, $\mu_i = \mathbb{E}(T_i)$ and $\sigma_i = \sqrt{Var(T_i)} \leq 1$.
The proof of this proposition is in the appendix.
The relaxation coefficient r is chosen sufficiently small such that the expected number of false positives $\mathbb{E}(\mathsf{FP}) \leq \alpha$, for all possible values of the unknown parameters, that is,

$$\hat{r} = min\left\{ argmax\{r : \mathbb{E}(\mathsf{FP}) \leq \alpha\}, m_0; \pi_i; \mu_i, \sigma_i, \rho_{ji} \right\}.$$

We still increase the value of the relaxation coefficient r as long as the type I error rate is below the level α. This condition guarantees the strong control of the FWER by Markov's inequality. Consider the particular case as in the upcoming simulations, where $\mu_{ji} = \Delta$ for atoms with effect and $\pi_i = \pi$ for all $i \in I_1$. In this case, $\mu_i = \pi s_i$ for all $i \in I_1$. We give an approximation algorithm to estimate the relaxation coefficient in this case. Let $s = 1/m \sum_{\{i=1,\ldots,m\}} s_i$ be the average subset size. We set the parameter ρ_{ji} to the least favorable value of $1/\sqrt{s}$, which corresponds to the case where the atoms inside the same subset are independent. For the parameter Δ we chose either ∞, which corresponds to the more conservative lower bound, or the mean of the $m_1 s$ largest values of the scores. The latter is the one used in the simulations and the applications presented in this paper. Finally, we approximate the parameter σ_i by 1 for all $i \in I_1$. Algorithm 1 gives an approximated upper bound for the relaxation coefficient r.
Matlab and R scripts that implement this algorithm are available at: http://miplab.epfl.ch/software.
Simulations presented in this paper show that the FDR is controlled if the modified p-values are used in the BH procedure. However, a rigorous proof of the FDR control is presented in Meskaldji (2013) and Meskaldji et al. (2013b).
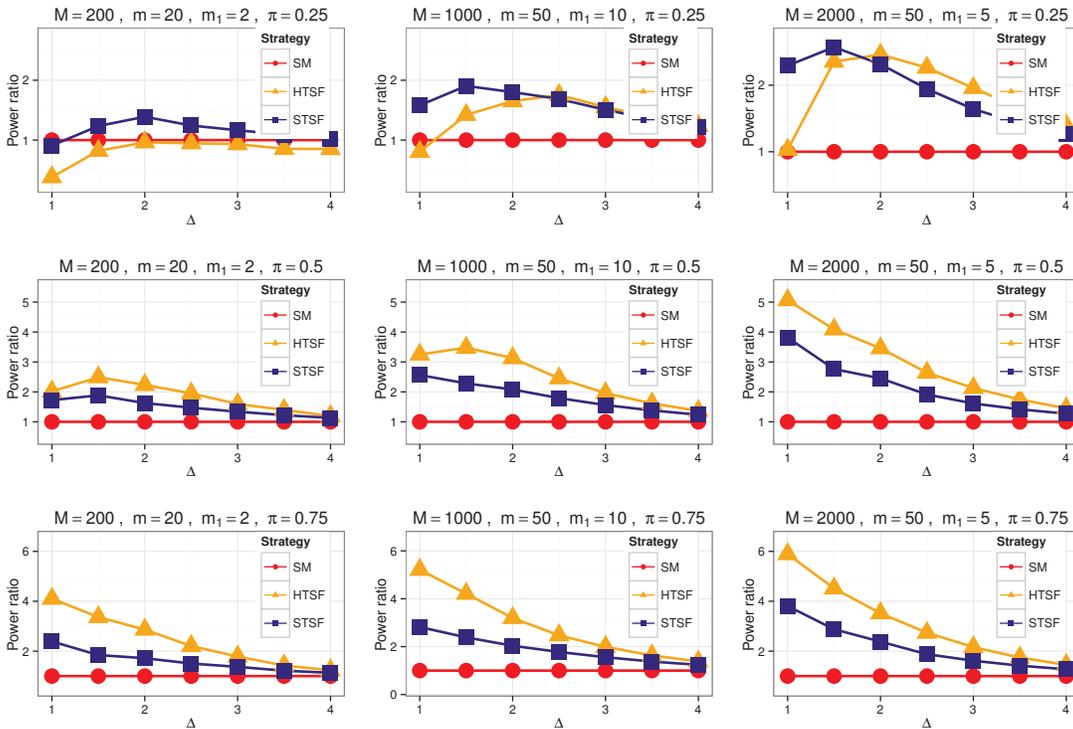
4

Figure 2: Ratio of average power of the SF methods over the average power of the SM against the raw effect, using the Bonferroni procedure. The HTSF (dashed line) and the STSF (dashed-points line). The number of atoms is either $M = 200, 1000$ or 2000. The number of subsets is $m = 20$ when $M = 200$ and 50 when $M = 1000$ and 2000. The number of partially-affected subsets is $m_1 = 2, 10$ and 5 for $M = 200, 1000$ and 2000, respectively. In each situation the proportion π of affected atoms within partially-affected subsets is either $0.25, 0.5$ or 0.75 as indicated. The number of simulations performed to obtain each panel is 1000.

Application to simulated data

We compare the performance of the SF approach with the SM, by considering different simulation settings and one experimental dataset. We also discuss the influence of the different parameters on the statistical inference.

First, we consider simulating brain networks in which we integrate a positive contrast that represents the between-group effect. In this simulation setting, we suppose that we have prior information about the positive dependence between tests. This prior information is available as a predefined decomposition of the global set of tests. We compared the relaxed methods to the SM in terms of average power and average number of false positives.

We simulated a set of M hypotheses (corresponding to M atoms), which we divided into m subsets with different sizes $s_i \geq 1, i = 1, \ldots, m$, randomly chosen. Among the m subsets, we randomly choose m_1 subsets to contain the effect. We call these subsets partially-affected subsets. In each of these subsets we randomly selected a set of hypotheses for which we simulated a test score Z as a random realization of the shifted standard normal distribution with mean Δ, that is, $Z \sim \mathcal{N}(\Delta, 1)$. These atoms contain the contrast, the between-group effect. The average proportion of atoms with effect in the m_1 partially-affected subsets is π. For all the remaining hypotheses, either in

the m_1 subsets (containing the effect) or in the remaining $m - m_1$ subsets (without effect), the test scores are random normal $\mathcal{N}(0, 1)$ realizations. Positive dependence was modelled by the proportion of affected atoms in each partially-affected subset.

We simulated global sets with $M = 200, 1000$ and 2000 atoms, with $m = 20$ or 50 subsets, and $m_1 = 2, 5$ or 10 partially-affected subsets. The average proportion of atoms with effect within partially-affected subsets was either $\pi = 0.25, 0.5$ or 0.75.

The value of $M = 200$ represents a moderate number of atoms in brain connectivity studies at the nodal level. However, for connection-wise studies, the number of atoms to be considered is of order $O(N(N-1)/2)$ for an undirected network with N nodes. For example, a network with 90 nodes and density of 0.25 or 0.5 has approximately 1000 or 2000 connections, respectively. This situation is highlighted by the values $M = 1000$ and 2000. For $M = 200$, the number of subsets is 20, which corresponds to an average subset size of 10. Among the 20 subsets, only two of them contain the contrast. However, when $m = 1000$ or 2000 the number of subsets is the same $m = 50$, which gives two different average subset sizes of 20 and 40, respectively. Although the number of subsets is the same in the two latter cases, the number of partially-affected subsets is 10 and 5, respectively. To measure the influence of the quality of the decomposition on the performance of the
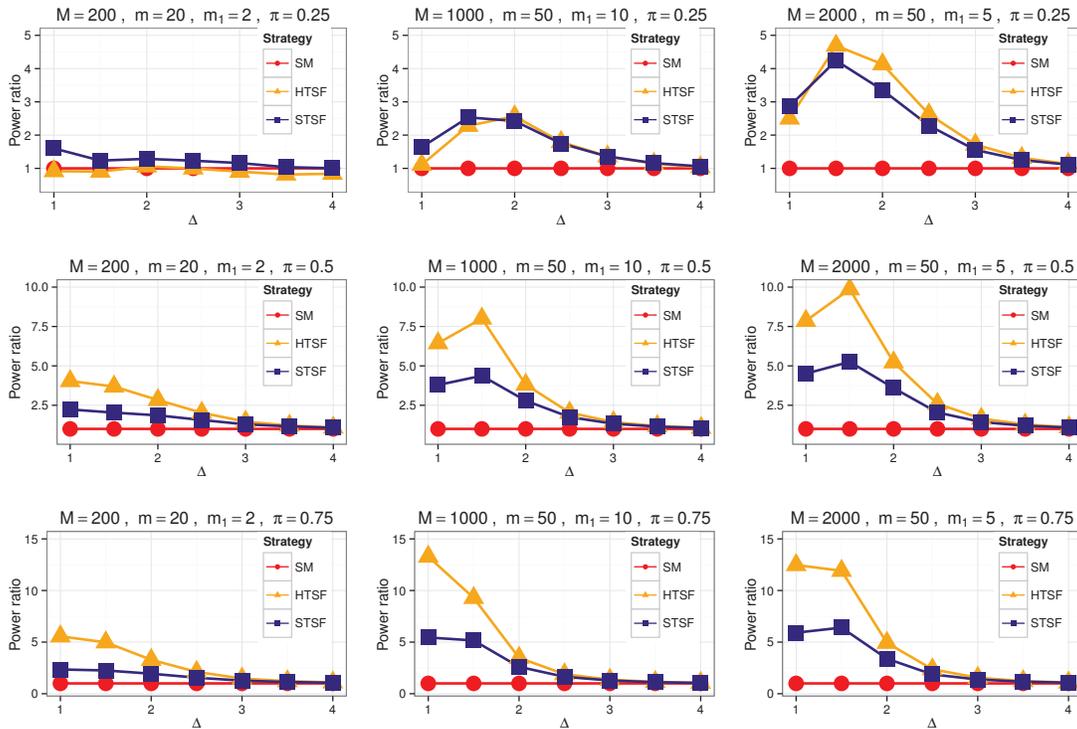
5

Figure 3: Ratio of average power of the SF methods over the average power of the SM against the raw effect, using the BH procedure. The setting is the same as in Figure 2.

proposed strategy, we considered three different values of the average number of affected atoms within partially-affected subsets (in the m_1 subsets) $\pi = 0.25, 0.5$ or 0.75. It is not impossible to have almost a perfect decomposition in practice, however, we did not consider the case where $\pi \approx 1$.

We applied the SM and the SF algorithm to detect the positive contrast. We carefully evaluated the influence of the choice of the screening threshold. In this simulation and in the next application, we used two different screening thresholds. The first choice is based on a multiplicity correction. For example, if the Bonferroni procedure is used for the screening, the threshold is α/m. The second choice is simply to perform a screening at threshold α. Based on this choice, we obtain two specific SF algorithms that we term *hard thresholding SF* (HTSF) and *soft thresholding* SF (STSF) algorithms, respectively, see Table 1. Of course, the estimated relaxation coefficient with the HTSF algorithm is always larger than the one estimated with the STSF algorithm for the same data and the same decomposition. We used two different type-I error metrics; the expected number of false positives $\mathbb{E}(\mathsf{FP})$ (which closely relates to the FWER via Markov's inequality) and the false discovery rate FDR $= \mathbb{E}(\mathsf{FP}/R)$, where R is the number of rejections (atoms declared to be significant). These two type-I error metrics are particular cases of the general family $\mathbb{E}(\mathsf{FP}/s(R))$, proposed by Meskaldji et al. (2011b), where s is any non-decreasing function. To control the expected number of false

positives, we used the Bonferroni procedure that performs each single test at level α/M, that is, it compares each single p-value to the threshold α/M. The false discovery rate is controlled by applying the Benjamini and Hochberg (1995) (BH) procedure, which consists in ordering the p-values from smallest to largest and chose as a significance threshold the largest p-value situated under the increasing line with slope α/M. Since the thresholds form a line, this procedure is called the linear step-up procedure. Note that to control the $\mathbb{E}(\mathsf{FP}/s(R))$, the ordered p-values $p_i, i = 1, \ldots, M$, are compared to the sequence $s(i)\alpha/M$ Meskaldji et al. (2011b).

Figures 2 and 3 show the ratio of the average power of the relaxed methods over the average power of the SM, in different situations, when using the Bonferroni or the BH procedures, respectively, both used with $\alpha = 0.05$. On the other hand, Table 2 gives the estimated expected number of false positive $\mathbb{E}(\mathsf{FP})$ in the Bonferroni case, and the estimated FDR in the BH case.

The simulations illustrate the power gain obtained by using the relaxed methods. The SF methods almost always perform better than the usual SM. The gain is realized even though less false positives are committed. The only case in which the relaxed methods does not seem to perform better than SM, especially the HTSF, is when the number of tests is moderate and the proportion of affected atoms within the partially-affected subsets is relatively small. This issue rapidly disappears when

Data: A set of M p-values, a decomposition of the global set of p-values (or the average subset size and the number of subsets), a global error rate α and a screening threshold U.

Result: An approximation of the relaxation coefficient r.

Initialization old$r = 0$; $\delta = 100$.

while $\delta > \epsilon$ **do**
  $r = $ old$r$
  temp$=0$;
  while $temp \leq \alpha$ **do**
    $r = r + \delta$;
    EFP = []; (an empty vector)
    $c = \Phi^{-1}(1 - r\alpha/M)$
    $\Gamma_0 = \int_c^\infty \overline{\Phi}\left(\frac{\Phi^{-1}(1-U)-(1/\sqrt{s})z}{\sqrt{1-1/s}}\right)\varphi(z)dz$
    for $m_1 \in \{1, ..., m\}$ **do**
      for $\pi \in \{1/s, ..., \lfloor s \rfloor/s\}$ **do**
        $\Delta = \infty$ or the mean of the $m_1 \pi s$ largest scores.
        $\Gamma_1 = \int_c^\infty \overline{\Phi}\left(\frac{\Phi^{-1}(1-U)-\pi\Delta-(1/\sqrt{s})z}{\sqrt{1-1/s}}\right)\varphi(z)dz$
        EFP$=$[EFP; $s(m_0\Gamma_0 + m_1(1-\pi)\Gamma_1)$]
      end
    end
    temp$=$max(EFP)
  end
  old$r = r - \delta$
  $\delta = \delta/10$
end
return $r - \delta$

Algorithm 1: Proposed algorithm to approximate the relaxation coefficient.

Name	Screening	Filtering
SM	–	Bonferroni/BH
STSF	α (No correction)	Bonferroni/BH
HTSF	Bonferroni/BH	Bonferroni/BH

Table 1: The different steps of the local methods that are compared in this paper. Th SM has only one step.

	SM	HTSF	STSF
$\mathbb{E}(\text{FP})$	0.046	0.035	0.044
FDR	0.057	0.053	0.055

Table 2: The estimated expected number of false positives $\mathbb{E}(\text{FP})$ in the Bonferroni case and the estimated FDR in the BH case, for the different methods SM, HTSF and STSF. The number of simulations is 1000 for each case.

the proportion becomes larger and the gain obtained by the SF methods reaches more than 5 times the power of the SM when the raw effect Δ is small. This corresponds to situations with small between-group effect or small sample size. When Δ becomes large, all methods, including the SM are equivalent. We can also observe when comparing the panels corresponding to $M = 1000$ and 2000, that the gain increases as the size of the subsets increases. Other simulations (not shown here) indicate that the advantage of the SF methods over the SM increases with the number of partially-affected subsets, m_1. This is because the screening will detect more positive subsets on aver-

age. Finally, the relaxed methods behave almost in a similar way when using either the Bonferroni procedure or the BH procedure.

To conclude this simulation study, we can say that the STSF is more stable in terms of gain because small proportions are easily detected. The HTSF seems to perform well when the proportion π becomes larger, which is directly related to the appropriate choice of the decomposition. The HTSF should be chosen when we have more confidence on the network decomposition. Otherwise, STSF is preferable, as it has a less strict screening in the first step. The choice of the decomposition and the screening threshold are discussed in more details in the application section.

Application to brain connectomes

The screening and filtering strategy could be applied to any modality of brain connectivity or any complex networks in general.

Here we present a comparison study of whole-brain structural connectivity matrices derived from diffusion MRI tractography.

Description of the data

We consider two groups of brain connectivity matrices based on the dataset used in Hagmann et al. (2010b), which consists of 30 connectivity matrices. We define two groups based on the age of the subjects: 16 pre-school children and 14 adolescent children.

The connectivity matrices are derived from diffusion-weighted MRI acquisitions and well-established algorithmic procedures as described in Hagmann et al. (2008); Cammoun et al. (2012). These matrices have been normalized by eliminating the linear bias towards longer fibers introduced by the tractography algorithm, and by dividing the fiber density between two ROIs by the average of their areas to account for brain volume normalization (Hagmann et al., 2010b).

Local measures

Many possible nodal and connection measures could be used to compare brain connectomes. In this study, we use one connection measure: connection weight (CW); and two nodal measures: nodal strength (NS) and nodal efficiency (NE). The CW measures the strength of connectivity between two nodes which is related to the capacity of transfer of information between two ROIs. The NS is defined as the sum of weights of edges connected to a specific node. It is computed by the sum of row/column values corresponding to that node in the connectivity matrix of subject k, that is,

$$\text{NS}_i^k = \sum_{j=1}^N w_{ij}^k,$$

where w_{ij}^k is the weight of the connection between nodes i and j in matrix (subject) k. The NS can be interpreted as the capacity of transfer between not only two nodes, but the capacity of transfer from a specific node to its neighbors. Finally, the NE

7

of a node is the inverse of the average minimum weighted path length from that specific node to all other nodes, that is,

$$\text{NE}_i^k = \left\{ \frac{1}{N-1} \sum_{j=1, j \neq i}^{N} d_{ij}^k \right\}^{-1} ,$$

where d_{ij}^k is the weighted distance; i.e., the length of the weighted shortest path between node i and node j for subject k. Note that in this case, the value of each cell is inverted because the distance between two connected nodes is supposed to be the inverse of the connection weight (Rubinov and Sporns, 2010) . The nodal efficiency represents the speed and the capacity of transfer of information from a node not only to its neighbors but to all other nodes in the network. The nodal efficiency is related to the nodal strength because the first step of shortest paths from a specific node starts by its neighbors.

Decomposition of the brain network

There are many ways to decompose a network into subnetworks. For example, one could group connections that share the same node as a subnetwork. In doing so, the number of subnetworks will be the number of nodes in the brain network. The choice is very large and studying all possibilities exceeds the scope of this paper. Here we choose decomposing a network based on the node communities, that is, subnetworks are defined on the basis of groups of nodes. Once groups of nodes (communities) are specified, subnetworks are defined either by connections between nodes of the same community (the intra-community connections) or the connections between two communities (inter-community connections) (Meskaldji et al., 2011a).

Here we present the different network community decomposition methods that we used in our application study. Our choice is large because we would like to study the robustness of the method against the network decomposition misspecification.

We used two different prior decompositions of the global set of nodes. The first one corresponds to the decomposition of the brain into lobes (LO). The second is based on a recent study (Chen et al., 2012), which uses a different approach to cortical localisation, we call it Chen decomposition (CH).

To highlight the large range of applicability of the screening and filtering strategy, we also decompose the set of nodes into communities using different data-driven algorithms, that is, algorithms based only on the data without any prior grouping information. A large number of decomposition methods is based on the optimization of the quantity known as modularity. Modularity is a quantity that measures the goodness of decomposition of a network into communities. A good community decomposition will have many edges within communities and a few between them (Newman and Girvan, 2004).

The four decomposition methods we used in this study are the following. The first one is called walktrap (WT), which is based on the fact that random walks on a graph tend to get trapped into densely connected subgraphs corresponding to communities (Pons and Latapy, 2005). The second one is the fast greedy algorithm (FG), which is based on the greedy optimization of the modularity (Clauset et al., 2004). The third algorithm is the edge betweenness algorithm (EB). It is based on iterative removal of edges from the network and edge betweenness recalculation (Newman and Girvan, 2004). Finally, the leading eigenvector (LE) in which the modularity function is rewritten in matrix terms which leads to express the optimisation task as a spectral problem in linear algebra (Newman, 2006).

Figure 1, in supplementary information (SI) illustrates the result of the six decomposition methods on the average group 1 (the pre-school children group) connection matrix. The decompositions obtained and the number of communities discovered by data-driven algorithms are not the same (Table 3). Nevertheless, the communities discovered are almost always situated within the same hemisphere. The variety of decompositions will afford a careful study of the influence of the decomposition on the final statistical inference.

Note that, depending on the application, the inference could be performed only on the intra-connectivity within communities which results in reducing the number of hypotheses tested. However, in our study, we keep all the connections.

Statistical inference

To compare the two groups of connectomes, we applied the SM with the Bonferroni and the BH procedures, the NBS and the permutation testing FDR (NBS and permutation FDR are implemented in NBS Matlab Software downloaded from: http://www.nitrc.org/projects/nbs/), and the screening and filtering methods, i.e., HTSF and STSF, both with Bonferroni and BH procedures. We distinguish between two types of control: the weak control of the FWER and the strong control of the FWER. The weak control is guaranteed by the SM with BH procedure (SM-BH), the permutation testing FDR (PFDR), the screening step of the HTSF with the Bonferroni procedure (HTS-Bonf), the filtering step of both HTSF and STSF with the BH procedure (HTSF-BH and STSF-BH, respectively). The strong control, however is guaranteed by the SM with the Bonferroni procedure (SM-Bonf) and both HTSF and STSF with the Bonferroni procedure (HTSF-Bonf and STSF-Bonf, respectively).

Results and Discussion

Weak control of false positives

Table 3 gives the number of so-defined subnetworks obtained on the basis of the different decomposition methods. It gives as well the number of positive subsets after the screening step for each decomposition method. In table 4, we report the relaxation coefficients estimated by Algorithm 1 for different SF methods and different decompositions. The number of significantly different connections under weak control are reported in Table 5 for SM-BH, PFDR, HTS-Bonf, HTSF-BH and STSF-BH. The number of connections within the component detected by the NBS method for different threshold values is reported in Table 6.

According to these results, we comment some important points.

	LO	CH	WT	FG	EB	LE
Number of subnetworks	91	91	45	28	10	28
HTS-Bonf	30	36	23	17	9	19
HTS-BH	62	72	38	25	9	23
STS	65	72	38	25	9	24

Table 3: The first line represents number of the so-defined subnetworks based on the node community decomposition methods. The remaining lines give the number of positive subnetworks obtained at the screening step for each of the SF methods for the different decompositions.

	LO	CH	WT	FG	EB	LE
HTSF-Bonf	7.22	7.22	7.21	7.18	7.01	7.18
STSF-Bonf	3.66	3.66	4.25	4.55	5.00	4.55
HTSF-BH	4.20	3.99	4.50	4.72	5.16	4.85
STSF-BH	3.66	3.66	4.25	4.55	5.00	4.55

Table 4: The relaxation coefficient estimated using Algorithm 1 in different situations for different SF methods. The relaxation coefficients corresponding to the HTSF are always lager than the STSF case.

First, we observe in Table 5 the improvement gained by the screening and filtering methods against the SM, almost twice the number of significant connections. Second, although the different number of positive subnetworks obtained in the screening step corresponding to different decompositions and two different screening thresholds (HTSF-BH and STSF-BH, Table 3) and different relaxation coefficients, the improvement seems to be stable. This does not hold in the case of the HTS-Bonf for the different decomposition methods. The same could be said concerning the NBS results, which dramatically depend on the arbitrary threshold (Table 6). It is important to mention a fundamental difference between the NBS and HTS-Bonf, and PFDR, SM-BH, HTSF-BH and STSF-BH. All these methods guarantee the weak control of the FWER. However, PFDR, SM-BH, HTSF-BH and STSF-BH guarantee as well the strong control of the FDR, whereas the strong control exerted by the NBS and the HTS-Bonf is completely unknown. For all the reasons mentioned above, we suggest for the weak control of false positives, either the HTSF-BH or the STSF-BH.

Note that the NBS method ends up with one single p-value for the whole detected component whereas the screening step of the SF methods ends up with different p-values corresponding to the subnetworks.

Strong control of false positives

In the case of strong control of the false positives, we used the Bonferroni procedure and we compared the SF methods (HTSF-Bonf and STSF-Bonf) with the SM-Bonf in terms of the number of connections/nodes declared to be significantly different between the two compared groups, which we reported in Table 7. For each of the relaxed methods, STSF and HTSF, we give two values that correspond to the prior decompositions: Lobes (LO) and Chen (CH) as well as the number of common results obtained when using these two prior decompositions ($\cap_p$); four values that correspond to the different data-driven decompositions: Walk trap (WT), Fast greedy (FG), Edge betweenness (EB) and Leading eigenvector (LE). We also report the number of common rejections between the results obtained with the four data-driven algorithms ($\cap_{dd}$), and, finally, the

common rejections by all the decomposition methods ($\cap$). The last column corresponds to the common results between the common ones obtained by the HTSF and the common ones obtained by the STSF ($\cap\cap$). The common values $\cap_p$, $\cap_{dd}$ and $\cap$ highlight the dependence of the SF methods on the choice of the decomposition and the common values $\cap\cap$ indicate the influence of the screening threshold.

The results show the potential gain due to the SF algorithm and its relevance in brain connectivity analysis; i.e., a considerable improvement under the same control exerted, the strong control of the FWER. The results are in accordance with the performances obtained by simulations, but in this case, one cannot determine the number of false positives since the ground truth effect is unknown. For example, in the nodal case (NS and NE), the number of tests is $M = 83$ (a moderate number) and the SM already detects many nodes as being significantly different between the two groups, especially, when using the NE as nodal topological measure. This means that the between-group effect is strong, which corresponds to a high raw effect Δ. The between-group effect seems to be weaker in the NS case than in the NE case and the advantage of the SF methods is more remarkable. Now, if we observe the results of the connection-wise analysis in which the number of tests is much larger, we see that the advantage of the SF methods is more pronounced. The number of significant connections is almost twice in all cases. The STSF behaves almost in the same way as the HTSF with a slightly small advantage for the HTSF, which indicates that the decomposition methods are quite appropriate. Moreover, the common detections $\cap_p$, $\cap_{dd}$ and $\cap$ indicate that both the HTSF and the STSF have a non negligable dependence on the choice of the decomposition. In addition, the SF methods seem to be influenced by the screening threshold in the Bonferroni case more than in the BH case. This is because the number of positive subsets detected by the BH procedure in the screening step is almost the same as the one obtained with no multiplicity correction corresponding to the screening step of the STSF.

Figure 4 shows the common significant connections obtained

9

	LO	CH	WT	FG	EB	LE
SM-BH	432	432	432	432	432	432
PFDR	348	348	348	348	348	348
HTS-Bonf	1329	1486	1911	2328	3458	2330
HTSF-BH	747	780	847	854	875	858
STSF-BH	737	761	828	852	867	843

Table 5: The number of connections declared as significantly different between the two groups using different weak control strategies and different decomposition methods (when it's applicable). Concerning the PFDR case, the number of permutation was set to 5000.

Threshold	0.05	0.25	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7	7.5
Nb. connections	1051	1024	966	841	677	508	368	260	171	101	66	31	5	1	1	2	0

Table 6: The number of connections within the component detected by the NBS method for different values of the first screening threshold. The number of permutations was set to 5000.
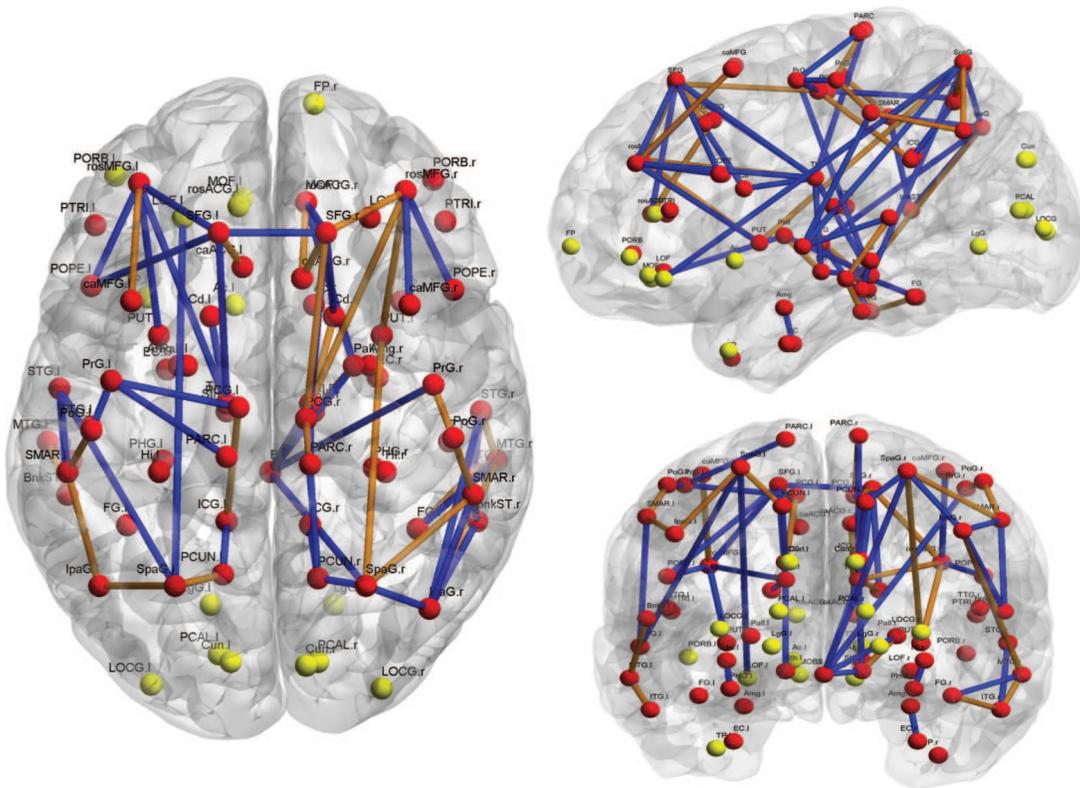

Figure 4: The common significant nodes/connections between all significant results obtained by the HTSF and the STSF, and all the decomposition methods (the last column of Table PracticalExampleResults). The red nodes are significantly different in terms of NS and NE. However, the yellow nodes significantly different only in terms of NE. The blue connections correspond to significantly different connections with Cohen's d between 1 and two, whereas, orange connections have a Cohen's d larger than 2. We used the Brain Net viewer tool (Xia et al., 2013) for this visualisation.

by both the STSF and the HTSF, corresponding to the last column of Table 7. A detailed list of the significant nodes and connections is available as supplementary information.

A permutation based evaluation

We present here an experimental study that could be seen as an intermediate evaluation between simulations and the previous practical example. Our comparison consists of the following. Among the 16 pre-school children and the 14 adolescent children, we randomly select 5 or 10 subjects from each group, apply the different methods (SM, HTSF and

	SM	HT									ST									∩∩
		LO	CH	∩$_p$	WT	FG	EB	LE	∩$_{dd}$	∩	LO	CH	∩$_p$	WT	FG	EB	LE	∩$_{dd}$	∩	
NS	61	67	64	64	69	68	69	69	66	64	66	66	66	66	66	68	66	66	66	64
NE	80	83	83	83	83	83	83	83	82	82	82	82	82	82	82	82	82	82	82	82
CW	52	91	86	82	97	99	102	101	95	81	81	81	81	88	91	94	91	88	81	69

Table 7: The number of nodes/connections declared to be significantly different using SM, HTSF and STSF. For the SF methods, we give different values that correspond to the different decompositions: LO, CH, WT, FG, EB and LE, the common rejections obtained by the prior decompositions ($\cap_p$), the data-driven algorithms ($\cap_{dd}$) and the common rejections obtained by all decompositions ($\cap$). $\cap\cap$ represents the global intersection of all cases. Three different network measures are used: nodal strength (NS), nodal efficiency (NE) and connection weight (CW). In both cases, the Bonferroni procedure is used to control the FWER.
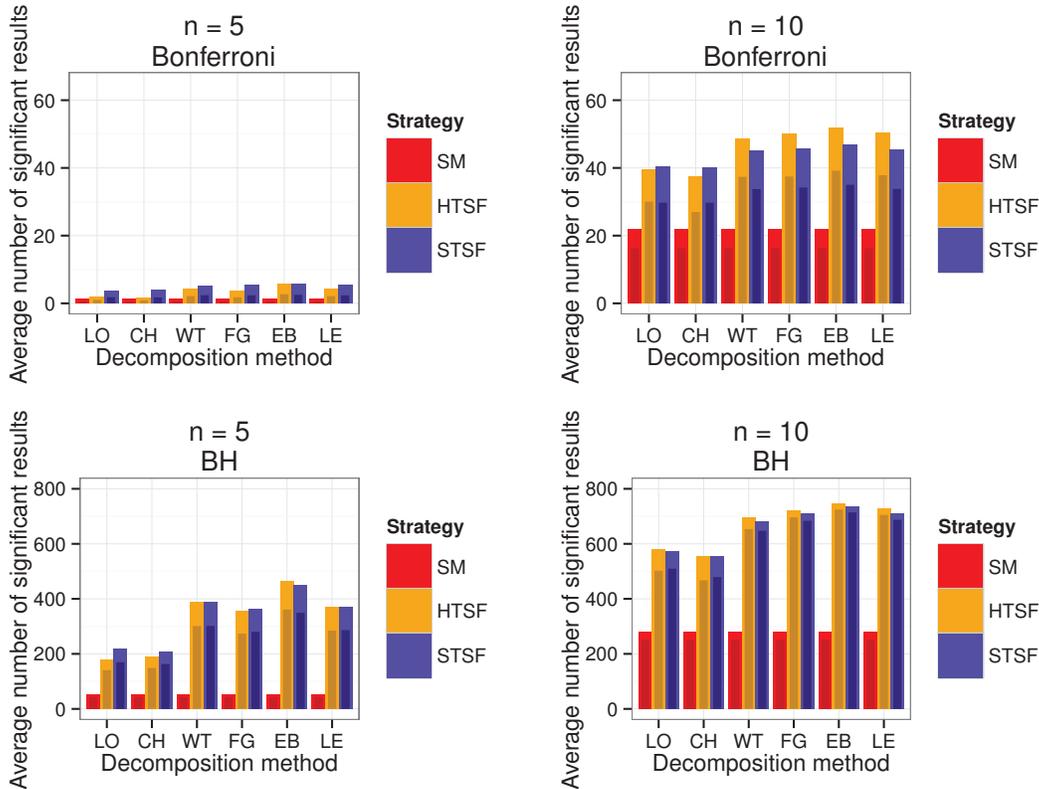

Figure 5: Average number of connections detected as significantly different between two groups randomly chosen among the initial groups, using different strategies and different decomposition methods, over 1000 random simulations (group selections). The Bonferroni procedure (FWER) is used in the first row and the BH procedure (FDR) in the second row. The number of subjects in each group is either $n = 5$ or $n = 10$ as indicated. The dark bars show the common rejections between the rejections obtained by each method (SM, HTSF or STSF) and the rejections obtained by the SM when using the complete sample.

STSF) and then estimate the number of significant connections obtained with the different decompositions. The operation is repeated 1000 times.

In Figure 5, we show the average number of significant connections for each method, using two different multiple testing procedures: the Bonferroni procedure and the BH procedure. For the relaxed methods, we give different values that correspond to the different decompositions. We also reported the average number of the common rejections between the rejections obtained by each method (SM, HTSF and STSF) and the rejections obtained by the SM when using the complete sample (the dark bars in the figure).

Figure 5 clearly shows the advantage of the SF methods, especially, when the between-group effect is weak. It also

indicates that the STSF and the HTSF are quite equivalent. According to the simulation results, this is due to a moderate quality of the decomposition of the global brain network which is based on node communities, and a decomposition based on the global set of connections may improve the results of the HTSF in practice. We plan to investigate this specific question in a future work. In addition, the data-driven decompositions lead to better performances and may outperform the prior decompositions. This increases the degree of applicability and validation of the proposed strategy.

Besides these observations concerning the performance of the SF methods, we clearly see the difference between the Bonferroni procedure and the BH procedure. The latter, by the definition of the FDR, detects much more connections but at

the price of a considerable number of false positives, whereas the expected number of false positives in the Bonferroni case is always majored by α. Meskaldji et al. (2011b, 2013c) studied this specific issue in an optimality framework and proposed to use either a truncated version of the FDR or a concave threshold curve instead of the linear thresholds of the BH procedure. These choices limit the number of false positives without losing much power.

Neurological substrate of adolescence and childhood

Childhood, adolescence and maturation of brain connections

The life-history stages of postnatal human development are divided into infancy (first three years), childhood (2 to 4 years), juvenile (3 to 4 years), adolescence (3 to 5 years) and youth stage (3-4 years), respectively. Each one of these stages marks a unique set of human characteristics, biological as well as behavioral, required for gradual transition into adulthood.
Profound differences in cognitive skills, as well as behavior, between preschool children (Piaget, 1964), adolescents (Spear, 2000) and adults have been studied extensively. Moreover, modern neuroimaging techniques allowed identification of specific biomarkers related to each life-history stage. To give an example, while the brain reaches its final size during the juvenile stage (White, 1996) the gray matter and white matter still continue to mature. From infancy trough childhood, juvenile, adolescence and the youth stage volume of the cerebral white matter exponentially increases while gray volume decreases linearly (Lebel and Beaulieu, 2011). This increase in cerebral white matter volume parallels the increase in white matter density (Paus et al., 1999) and is most likely caused by an increase of axonal diameter and myelin sheet thickness. Our results are in agreements with these findings (SI. Table 1-3). When comparing adolescents with preschool children we have identified microstructural differences in fibre pathways. All of the connections (69 connections reported in Figure 4) found to be significantly different in adolescent subjects and had significantly stronger structural connectivity strength (SI. Table 1-3). Certain cerebral regions (ROIs) were associated with increased number of significantly stronger connections (SI. Table 1-3, red) allowing us to anatomically and qualitatively analyze these results.

Prefrontal Cortex

Among these regions were thalamus and superior frontal gyrus of the right hemisphere and dorsolateral prefrontal cortex (rostral part of middle frontal gyrus) of both hemispheres (SI. Table 1-2). Although our results corroborate the results reported by Lebel and Beaulieu (2011) and are in agreement with reports on prolonged maturation of frontal lobe (Gogtay et al., 2004; Giedd et al., 1999) the whole brain connectomic approach allowed us to define fine circuit differences in more detailed fashion.
Connections between superior frontal gyrus and gyrus cinguli (anterior and posterior segments) were found to be stronger in

adolescents (SI. Table 1-2). Moreover, dorsolateral prefrontal cortex (rostral portion of middle frontal gyrus) and superior frontal gyrus had significantly stronger connections with thalamus in both hemispheres (SI. Table 1-2), suggesting protracted maturation of anterior thalamic radiation (Bava et al., 2010). The significantly stronger connections between prefrontal cortex and limbic structures (gyrus cinguli and limbic nuclei of thalamus in both hemispheres), found in adolescents, suggest fine-tuning of fronto-limbic circuitry throughout childhood and juvenile stage (SI. Table 1-2). During the adolescence prefrontal cortex undergoes substantial reorganization; the volume and thickness of prefrontal cortex rapidly decrease (Gogtay et al., 2004; Giedd et al., 1999) which parallels the process of synaptic pruning (Petanjek et al., 2011). Our results show stronger connectivity between dorsolateral prefrontal cortex (rostral middle frontal gyrus) and neighboring cortical areas in both hemispheres (such as superior frontal gyrus and rostral portion of middle frontal gyrus) (SI. Table 1-2). Therefore, compared to children, prefrontal cortex in adolescents favors stronger connectivity with its neighboring areas, structural integration. Similarly it also favors stronger connectivity with limbic areas (cingulate gyrus).
Lateral prefrontal cortex is a key neural substrate of executive function. In order to carry out executive functions the connections between lateral prefrontal cortex and other cerebral areas allow influx/outflow of information. These information (e.g., cognitive representation of perception and of action) are temporally integrated in lateral prefrontal cortex which is crucial for goal-directed behavior (Fuster, 2001). That is, integrating executive memory (sequences of behavior, action schemas) or working memory (Goldman-Rakic, 1995, 1991) with preparatory set finally leads to central function of lateral prefrontal cortex: temporal organization of behavior linking the organism with its environment (Fuster, 2001).
The anterior cingulate cortex, on the other hand, serves as an *anterior attentional system* (Posner and Petersen, 1990). That is, it has a key role in concentrating attention on behavioral or cognitive tasks while at the same time serves as a key player for integrating the motivational value (Fuster, 2001). Furthermore, it plays an important role in control of basic drives and emotional behavior (Fuster, 2001).
There is a shift in predominance of dopaminergic activity in dorsolateral prefrontal cortex over anterior cingulate cortex during adolescence (Spear, 2000). This might relate to the reward deficiency leading to reckless behavior, sensation seeking and risk taking seen in adolescence (Spear, 2000). While there is no direct evidence that changes in connectivity strength between mesocortical and prefrontal regions underlie adolescent specific behavior, our results suggest that this circuit might be one of its neurological substrates. Furthermore, during adolescence major transformation of cognitive processing occurs (transformation of cognitive thought leading to abstract reasoning (Graber and Petersen, 1991)).
Left prefrontal cortex has a role in encoding the new memory (Gabrieli et al., 1998) thought semantic organization of encoded material (Fletcher et al., 1998a). Contrarily, right prefrontal cortex is more involved during retrieval of stored mem-

ory and internal monitoring of retrieved material (Fletcher et al., 1998b). Our results suggest that from childhood to adolescence, strengthening of associational fibers of the prefrontal cortex occurs in hemisphere specific fashion (SI. Table 1-2). Reorganization of frontal connectivity, thus, might underlie the changes seen in cognitive skills (processing of information) during childhood and adolescence (SI. Table 1-2).

Parietal Cortex

Choudhury et al. (2006) suggested that efficiency, and possibly strategy, of perspective taking develop in parallel with brain maturation during adolescence. Moreover, socio-cognitive development during adolescence is related to maturation of prefrontal, inferior parietal and superior temporal cortex.

In line with the current concepts of development, parietal and frontal cortexes reach the peak volume in pre-adolescent stage (approx. 10-12 years) which if followed by decline during adolescence (Gogtay et al., 2004; Giedd et al., 1999). It has been suggested that synaptic reorganization of frontal and parietal cortices during adolescence has an impact on cognitive processes such as mentalising and perspective taking in the motor, conceptual and emotional domains (Ruby and Decety, 2004, 2003) as well as in the process of abstract reasoning (Luna, 2004; Qin et al., 2004). Our results show that the connections of superior and inferior parietal cortex, especially with superior temporal cortex, strengthen considerably in adolescence (Tables 1, 2). Furthermore, parietal cortex is involved in distinguishing between self and others, in terms of imagining how someone would think or feel (Ruby and Decety, 2004, 2003, 2001) while superior temporal cortex is associated with the prediction of actions based on past ones. Reorganization of these circuits might explain differences in socio-cognitive functioning between children and adolescents (Choudhury et al., 2006) while developmental alteration of these circuits could lead to poorer socio-cognitive outcome (Fischi-Gómez et al., 2014).

Brain Stem Connections

Interestingly, our results show that, compared to children, adolescents have stronger connections of the brain stem and right hemisphere (precentral gyrus, superior parietal, basal ganglia and thalamus).

Paus et al. (1999) showed that cortico spinal tract matures until late adolescence, being a neurological substrate for elaboration of fine finger movements. Indeed, our results also reveal prolonged maturation of pyramidal tract of the right hemisphere (SI. Table 1-2). Having in mind that we did not take into account handedness of subjects, we cannot rule out that hemispheric differences in maturation of this tract could be contributed to the handedness of our subjects.

Finally, careful analysis of our results (Figure 4) revealed that majority of the connections found to be stronger in adolescence where associational connections of the right hemisphere between the areas of frontal cortex (rostral portion of the middle frontal gyrus, superior frontal gyrus), parietal cortex (superior and inferior parietal lobule) and their neighboring areas (SI. Table 1). Within left hemisphere majority of connections found to be stronger were connections between prefrontal cortical areas

(rostral portion of the middle frontal gyrus) and its neighboring areas (SI. Table 2).

In summary, results of our structural connectivity analysis show that adolescents, compared to children, favor structural integration of prefrontal cortices with the neighboring areas and have more elaborated pyramidal tract of the right hemisphere. In addition, the structural integration of the parietal cortex of the right hemisphere and superior temporal cortex of the left hemisphere with their neighboring structures confirms hemispheric differences in structural integration from childhood to adolescence.

Graph analysis reveals fine-tuning of brain connectivity during childhood and puberty

Qualitative analysis of our structural connectivity results identified some of the networks associated with maturation of axonal pathways (SI. Table 1-3). Nevertheless, new imaging analysis tools (like graph analysis) nowadays allow in depth quantitative analysis of brain development on global level (Park and Friston, 2013; Bullmore and Sporns, 2009). When assessing nodal efficiency, our results showed that in adolescents, all of the cerebral nodes (except left frontal pole) have significantly stronger nodal efficiency compared to children (SI. Table 4, 5; 1). Moreover, in adolescence, majority of nodes have also significantly stronger nodal strength (SI. Table 4, 5; 1). NS suggests that adolescents favor increased integration and decreased segregation of structural connectivity as shown previously (Hagmann et al., 2010a). This global trend of increased efficiency might reflect maturation of axonal pathways and consequential structural network refinement (Hagmann et al., 2010a). It is known that processes of axonal myelinisation and axonal diameter increase are protracted through adolescence into early adulthood (Paus, 2010). These biological changes influence the speed of neural transmission (Aboitiz et al., 1992) and are crucial for integration of information across spatially segregated brain regions (Paus et al., 1999). Spatio-temporal differences in maturation of different cerebral areas have been linked to certain aspects of behavior during development. As an example, Galvan et al. (2006) reported that earlier development of nucleus accumbens relative to orbito-frontal cortex underlies risk-taking behaviors seen in adolescents. Within this context, our results suggest that significantly higher node strength (the sum of weights attached to ties belonging to a node) seen in adolescent subjects (SI. Table 4, 5; 1) reflects the developmental changes (myelinisation or increase of axonal diameter) that could underlie adolescent behavioral. Similarly, as major changes occur during late childhood and adolescence (SI. Table 1-5), even a slight mismatch in their maturation pattern (Paus et al., 2008), could potentially lead to emergence of psychiatric disorders (e.g., schizophrenia) or obsessive compulsive (Anticevic et al., 2013a,b; Paus et al., 2008).

Interestingly, in adolescents, relative to children, the occipital lobe (calcarine, cuneus and lateral occipital areas) and lingual gyrus of both hemispheres did not show significant differences in nodal strength. While calcarine cortex contains the primary visual area (Brodmann area 17), cuneus and lateral occipital cortices contain extrastriate components of visual cortex. From

functional point of view, Brodmann areas 18 and 19 (mostly occupying lateral occipital cortex) are classified as unimodal visual association cortex (Mesulam, 1998). They are monosynaptically connected with V1 and constitute upstream visual association areas; in plain words, they are only one synapse away from the associated primary sensory area. These upstream unimodal sectors encode basic features of sensation. Downstream sectors (e.g., inferior temporal cortex), on the other hand, are at a distance of two or more synaptic units from associated primary sensory area and they encode complex contents of sensory experience (Mesulam, 1998). Our results reveal that almost all of the nodes are significantly stronger in adolescents relative to children (SI. Table 4, 5) except the nodes belonging to upstream unimodal sectors of visual information processing. Therefore our results suggest that neural substrate of basic features of visual sensation might be unchanged during development, at least from structural point of view.

Conclusion

We presented a screening-filtering strategy that exploits the structure of the data and positive dependence that could exist between tests corresponding to nodes or connections, without relying on strong assumptions. The strategy is adapted to strongly control either the FWER, the FDR or any error rate based on the modified p-values. We showed via simulations and practical examples that the proposed strategy almost always performs better than the usual node/connection wise analysis. The power obtained by the proposed strategy reaches several times the power of the standard methods that do not exploit the positive dependence and the data structure, and this holds under a strong control of the false positives. We showed as well that the screening and filtering strategy improvement is quite robust against network decomposition misspecification and screening thresholds. Although, the SF methods presented in this paper do not exploit the information of positive dependence in an optimal way, the gain obtained in the simulations should be satisfactory for many applications, especially, in neuroimaging. We also compared the screening and filtering strategy to the NBS, which is the first and most used method to detect differences in brain connectivity. We showed how the SF methods in the FDR case compete with the NBS in terms of performance, robustness and strong control.

Using the novel strategy for comparison of connectomes we have demonstrated stronger connectivity of associational connections of prefrontal cortex in adolescents relative to preschool children. Connections between prefrontal cortex and limbic structures were also identified as stronger in adolescence. Additionally, adolescents have significantly stronger NS across the brain with the exception of the nodes belonging to upstream unimodal sectors of visual information processing. Therefore, our results corroborate hypothesis that prefrontal associational connections and fronto-limbic circuitry have protracted development that continues into adolescence. Similarly, fine-tuning of connectivity, seen as significantly stronger node degree across almost all areas of the brain, support the hypothesis that significant modification of information processing occurs in almost all areas of the brain from childhood to adolescence. In conclusion, we suggest that these findings might represent biological blueprint related to cognition and behavior seen in adolescents.

Appendix

Proof of Proposition 1

We have to show that, under the null hypothesis, the c.d.f. of $Z_{ji} = \Phi^{-1}(1 - p_{ji})$ is the normal c.d.f., that is, $\mathbb{P}(Z_{ji} < z) = \Phi(z)$. Under the null hypothesis, $p_{ji} \sim \mathcal{U}(0, 1)$. We have $\mathbb{P}(Z_{ji} < z) = \mathbb{P}(\Phi^{-1}(1 - p_{ji}))$. Since Φ is an increasing function, then $\mathbb{P}(Z_{ji} < z) = \mathbb{P}(1 - p_{ji} < \Phi(z)) = \mathbb{P}(p_{ji} > 1 - \Phi(z)) = 1 - (1 - \Phi(z)) = \Phi(z)$.

Proof of Proposition 2

$T_i = \sqrt{s_i}^{-1} \sum_{j \in J_i} Z_{ji}$. Since T_i is a linear combination of gaussian random variables, then it is also a gaussian random variable. Let us compute its mean and variance.
$\mathbb{E}(T_i) = \mathbb{E}\left(\sqrt{s_i}^{-1} \sum_{j \in J_i} Z_{ji}\right) = \sqrt{s_i}^{-1} \sum_{j \in J_i} \mathbb{E}(Z_{ji}) = \sqrt{s_i}^{-1} \sum_{j \in J_i} 0 = 0$. Here we used the fact that the expectation is a linear operator and that $Z_{ji} \sim \mathcal{N}(0, 1)$.
$Var(T_i) = Var\left(\sqrt{s_i}^{-1} \sum_{j \in J_i} Z_{ji}\right) = \left(\sqrt{s_i}^{-1}\right) \sum_{j \in J_i} Var(Z_{ji}) = \left(\sqrt{s_i}^{-1}\right) \sum_{j \in J_i} 1 = \frac{1}{s_i} s_i = 1$. Here we used the fact that the variance is a bi-linear operator and that $Z_{ji} \sim \mathcal{N}(0, 1)$.

Proof of Proposition 3

According to the mixture model assumed for the data, we have the following distributions:

$$\left(T_i | Z_{ji} = z\right) \sim \mathcal{N}\left(\rho_{ji} z, \left(1 - \rho_{ji}^2\right)\right) = \mathcal{N}\left(\frac{1}{\sqrt{s_i}} z, \left(1 - \frac{1}{s_i}\right)\right),$$

for $h_{ji} = 0$ and $i \in I_0$, and

$$\left(T_i | Z_{ji} = z\right) \sim \mathcal{N}\left(\mu_i + \rho_{ji} \sigma_i z, \sigma_i^2 \left(1 - \rho_{ji}^2\right)\right),$$

for $h_{ji} = 0$ and $i \in I_1$, where $\mu_i = \mathbb{E}(T_i)$ and $\sigma_i = \sqrt{Var(T_i)} \leq 1$. If the Bonferroni procedure is used in the filtering step, the modified p-values $\{p_{ji}/r \mid i \in I^+\}$ are compared to α/M. This means that the original p-values (belonging to the positive subsets) are compared to $r\alpha/M$, or equivalently, the scores Z_{ji} are compared to $c = \Phi^{-1}(1 - r\alpha/M)$.
Let $\mathbb{E}(\mathsf{FP})$ be the expected number of false positives after the filtering step. Note that the expected number of FP in the negative subsets is zero because all p-values are set to 1.
Let J^N be the set of null hypotheses, that is, $J^N = \{j : h_{ji} = 0\}$. Given $\rho_{ji}\{j = 1, \ldots, M\}, m_0, m, s_i\{i = 1, \ldots, m\}, \mu_i\{i = 1, \ldots, m\}, \sigma_i\{i = 1, \ldots, m\}, \pi_i\{i = 1, \ldots, m\}, U$ and c, and that

$\mathbb{E}(I^+) = m\mathbb{P}(P_i \leq U)$, we have

$$\mathbb{E}(\text{FP}) = \mathbb{E}\left(\mathbb{E}\left(\text{FP}|I^+\right)\right)$$

$$= \mathbb{E}\left(\sum_{i \in I^+} \sum_{j \in J_i \cap J^N} \mathbb{P}_{h_{ji}=0}(Z_{ji} > c|P_i \leq U)\right)$$

$$= \mathbb{E}\left(\sum_{i \in I^+} s_i \sum_{j \in J_i \cap J^N} \frac{\mathbb{P}\left(P_i \leq U|Z_{ji}=z\right)\mathbb{P}_{h_0}\left(Z_{ji} \geq z\right)}{\mathbb{P}(P_i \leq U)}\right)$$

$$= \mathbb{E}\left(\sum_{i \in I^+} s_i \int_c^\infty \frac{m_0\overline{\Phi}(C_0) + m_1(1-\pi_i)\overline{\Phi}\left(C_{\mu_i}\right)}{m_0 U + m_1\overline{\Phi}(\Phi^{-1}(1-U)-\mu_i)}\varphi(z)dz\right),$$

where φ is the probability density function of the normal distribution, $C_0 = \frac{\Phi^{-1}(1-U)-\rho_{ji}z}{\sqrt{1-\rho_{ji}^2}}$ and $C_{\mu_i} = \frac{\Phi^{-1}(1-U)-\mu_i-\rho_{ji}\sigma_i z}{\sigma_i\sqrt{1-\rho_{ji}^2}}$.

Acknowledgments

References

Aboitiz, F., Scheibel, A. B., Fisher, R. S., Zaidel, E., 1992. Individual differences in brain asymmetries and fiber composition in the human corpus callosum. Brain research 598 (1), 154–161.

Achard, S., Salvador, R., Whitcher, B., Suckling, J., Bullmore, E., 2006. A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. The Journal of Neuroscience 26 (1), 63–72.

Anticevic, A., Cole, M. W., Repovs, G., Savic, A., Driesen, N. R., Yang, G., Cho, Y. T., Murray, J. D., Glahn, D. C., Wang, X.-J., et al., 2013a. Connectivity, pharmacology, and computation: toward a mechanistic understanding of neural system dysfunction in schizophrenia. Frontiers in psychiatry 4.

Anticevic, A., Hu, S., Zhang, S., Savic, A., Billingslea, E., Wasylink, S., Repovs, G., Cole, M. W., Bednarski, S., Krystal, J. H., et al., 2013b. Global resting-state functional magnetic resonance imaging analysis identifies frontal cortex, striatal, and cerebellar dysconnectivity in obsessive-compulsive disorder. Biological psychiatry.

Barrat, A., Barthelemy, M., Pastor-Satorras, R., Vespignani, A., 2004. The architecture of complex weighted networks. Proceedings of the National Academy of Sciences of the United States of America 101 (11), 3747–3752.

Bassett, D. S., Bullmore, E., Verchinski, B. A., Mattay, V. S., Weinberger, D. R., Meyer-Lindenberg, A., 2008. Hierarchical organization of human cortical networks in health and schizophrenia. The Journal of Neuroscience 28 (37), 9239–9248.

Bava, S., Thayer, R., Jacobus, J., Ward, M., Jernigan, T. L., Tapert, S. F., 2010. Longitudinal characterization of white matter maturation during adolescence. Brain research 1327, 38–46.

Benjamini, Y., Heller, R., 2007. False Discovery Rates for Spatial Signals. Journal of the American Statistical Association 102 (480), 1272–1281.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B. Methodological 57 (1), 289–300.

Benjamini, Y., Hochberg, Y., 1997. Multiple hypotheses testing with weights. Scandinavian Journal of Statistics 24 (3), 407–418.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nature Reviews Neuroscience 10 (3), 186–198.

Cammoun, L., Gigandet, X., Meskaldji, D.-E., Thiran, J. P., Sporns, O., Do, K. Q., Maeder, P., Meuli, R., Hagmann, P., 2012. Mapping the human connectome at multiple scales with diffusion spectrum MRI. Journal of Neuroscience Methods 203 (2), 386–397.

Chen, C.-H., et al., 2012. Hierarchical Genetic Organization of Human Cortical Surface Area. Science 335 (6076), 1634–1636.

Choudhury, S., Blakemore, S.-J., Charman, T., 2006. Social cognitive development during adolescence. Social cognitive and affective neuroscience 1 (3), 165–174.

Clauset, A., Newman, M. E. J., Moore, C., Dec 2004. Finding community structure in very large networks. Phys. Rev. E 70, 066111.

Fischi-Gómez, E., Vasung, L., Meskaldji, D.-E., Lazeyras, F., Borradori-Tolsa, C., Hagmann, P., Barisnikov, K., Thiran, J.-P., Hüppi, P. S., 2014. Structural brain connectivity in school-age preterm infants provides evidence for impaired networks relevant for higher order cognitive skills and social cognition. Cerebral Cortex, bhu073.

Fletcher, P., Shallice, T., Dolan, R., 1998a. The functional roles of prefrontal cortex in episodic memory. i. encoding. Brain 121 (7), 1239–1248.

Fletcher, P., Shallice, T., Frith, C., Frackowiak, R., Dolan, R., 1998b. The functional roles of prefrontal cortex in episodic memory. ii. retrieval. Brain 121 (7), 1249–1256.

Fornito, A., Zalesky, A., Breakspear, M., 2013. Graph analysis of the human connectome: Promise, progress, and pitfalls. NeuroImage 80 (0), 426 – 444, mapping the Connectome.

Frackowiak, R., Friston, K., Frith, C., Dolan, R., Mazziotta, J., 1997. Human Brain Function. Academic Press.

Friston, K. J., 2011. Functional and effective connectivity in neuroimaging: a review. Brain connectivity 1 (1), 13–36.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., Frackowiak, R. S. J., 1995. Statistical parametric maps in functional imaging: A general linear approach. Human Brain Mapping 2 (4), 189–210.

Fuster, J. M., 2001. The prefrontal cortex-an update-time is of the essence. Neuron 30 (2), 319–333.

Gabrieli, J. D., Poldrack, R. A., Desmond, J. E., 1998. The role of left prefrontal cortex in language and memory. Proceedings of the national Academy of Sciences 95 (3), 906–913.

Galvan, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., Casey, B., 2006. Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. The Journal of Neuroscience 26 (25), 6885–6892.

Genovese, C. R., Dawid, A. P., Bernardo, J. M., Berger, J. O., Smith, A. F. M., 1999. Functional magnetic resonance imaging and spatio-temporal inference. Bayesian Statistics 6, 255–274.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., Paus, T., Evans, A. C., Rapoport, J. L., 1999. Brain development during childhood and adolescence: a longitudinal mri study. Nature neuroscience 2 (10), 861–863.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., Nugent, T. F., Herman, D. H., Clasen, L. S., Toga, A. W., et al., 2004. Dynamic mapping of human cortical development during childhood through early adulthood. Proceedings of the National Academy of Sciences of the United States of America 101 (21), 8174–8179.

Goldman-Rakic, P., 1995. Cellular basis of working memory. Neuron 14 (3), 477–485.

Goldman-Rakic, P. S., 1991. Chapter 16 cellular and circuit basis of working memory in prefrontal cortex of nonhuman primates. In: The Prefrontal Its Structure, Function and Cortex Pathology. Vol. 85 of Progress in Brain Research. Elsevier, pp. 325 – 336.

Graber, J. A., Petersen, A. C., 1991. Cognitive changes at adolescence: Biological perspectives. Brain maturation and cognitive development: Comparative and cross-cultural perspectives, 253–279.

Hagmann, P., Cammoun, L., Gigandet, X., Gerhard, S., Grant, P. E., Wedeen, V., Meuli, R., Thiran, J.-P., Honey, C. J., Sporns, O., 2010a. MR con-

nectomics: Principles and challenges. Journal of Neuroscience Methods 194 (1), 34–45.

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., Sporns, O., 07 2008. Mapping the Structural Core of Human Cerebral Cortex. PLoS Biol 6 (7), e159.

Hagmann, P., Sporns, O., Madan, N., Cammoun, L., Pienaar, R., Wedeen, V. J., Meuli, R., Thiran, J.-P., Grant, P. E., 2010b. White matter maturation reshapes structural connectivity in the late developing human brain. Proceedings of the National Academy of Sciences 107 (44), 19067–19072.

Lebel, C., Beaulieu, C., 2011. Longitudinal development of human brain wiring continues from childhood into adulthood. The Journal of Neuroscience 31 (30), 10937–10947.

Luna, B., 2004. Algebra and the adolescent brain. Trends in Cognitive Sciences 8 (10), 437–439.

Meskaldji, D.-E., 2013. Multiple comparison procedures for large correlated data with application to brain connectivity analysis. Ph.D. thesis, Ecole Polytechnique Fédérale de Lausanne.

Meskaldji, D.-E., Fischi-Gomez, E., Griffa, A., Hagmann, P., Morgenthaler, S., Thiran, J.-P., 2013a. Comparing connectomes across subjects and populations at different scales. NeuroImage 80 (0), 416 – 425, mapping the Connectome.

Meskaldji, D.-E., Hagmann, P., Thiran, J.-P., Morgenthaler, S., 2013b. Two step multiple comparison procedures for positively dependent data. ArXiv e-prints 1307.3286.

Meskaldji, D.-E., Ottet, M.-C., Cammoun, L., Hagmann, P., Meuli, R., Eliez, S., Thiran, J.-P., Morgenthaler, S., 08 2011a. Adaptive Strategy for the Statistical Analysis of Connectomes. PLoS ONE 6 (8), e23009.

Meskaldji, D.-E., Thiran, J.-P., Morgenthaler, S., dec 2011b. A comprehensive error rate for multiple testing. ArXiv e-prints 1112.4519.

Meskaldji, D.-E., Thiran, J.-P., Morgenthaler, S., 2013c. Optimality in multiple comparison procedures. ArXiv e-prints 1307.2614.

Meskaldji, D.-E., Van De Ville, D., 2014. Multimodal graph theoretical analysis of functional brain connectivity using adaptive two-step strategy. Proceedings of the Eleventh IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 919–922.

Mesulam, M.-M., 1998. From sensation to cognition. Brain 121 (6), 1013–1052.

Newman, M. E., 2006. Finding community structure in networks using the eigenvectors of matrices. Physical review E 74 (3), 036104.

Newman, M. E. J., Girvan, M., Feb 2004. Finding and evaluating community structure in networks. Phys. Rev. E 69, 026113.

Nichols, T., Hayasaka, S., October 2003. Controlling the familywise error rate in functional neuroimaging: a comparative review. Stat Methods Med Res 12 (5), 419–446.

Owen, J. P., Li, Y.-O., Yang, F.-P. G., Bukshpun, P., Vora, S., Wakahiro, M., Hinkley, L. B., Nagarajan, S., Sherr, E. H., Mukherjee, P., 2013a. Resting state networks and the functional connectome of the human brain in agenesis of the corpus callosum. Brain connectivity.

Owen, J. P., Li, Y.-O., Ziv, E., Strominger, Z., Gold, J., Bukhpun, P., Wakahiro, M., Friedman, E. J., Sherr, E. H., Mukherjee, P., 2013b. The structural connectome of the human brain in agenesis of the corpus callosum. NeuroImage 70 (0), 340 – 355.

Park, H.-J., Friston, K., 2013. Structural and functional brain networks: from connections to cognition. Science 342 (6158), 1238411.

Paus, T., 2010. Growth of white matter in the adolescent brain: myelin or axon? Brain and cognition 72 (1), 26–35.

Paus, T., Keshavan, M., Giedd, J. N., 2008. Why do many psychiatric disorders emerge during adolescence? Nature Reviews Neuroscience 9 (12), 947–957.

Paus, T., Zijdenbos, A., Worsley, K., Collins, D. L., Blumenthal, J., Giedd, J. N., Rapoport, J. L., Evans, A. C., 1999. Structural maturation of neural pathways in children and adolescents: in vivo study. Science 283 (5409), 1908–1911.

Penny, W., Friston, K., 2003. Mixtures of general linear models for functional neuroimaging. IEEE Trans Med Imaging 22 (4), 504–14.

Petanjek, Z., Judaš, M., Šimić, G., Rašin, M. R., Uylings, H. B., Rakic, P., Kostović, I., 2011. Extraordinary neoteny of synaptic spines in the human prefrontal cortex. Proceedings of the National Academy of Sciences 108 (32), 13281–13286.

Piaget, J., 1964. Part i: Cognitive development in children: Piaget development and learning. Journal of research in science teaching 2 (3), 176–186.

Pons, P., Latapy, M., 2005. Computing communities in large networks using random walks. In: Computer and Information Sciences-ISCIS 2005. Springer, pp. 284–293.

Posner, M. I., Petersen, S. E., 1990. The attention system of the human brain. Annual Review of Neuroscience 13 (1), 25–42, pMID: 2183676.

Qin, Y., Carter, C. S., Silk, E. M., Stenger, V. A., Fissell, K., Goode, A., Anderson, J. R., 2004. The change of the brain activation patterns as children learn algebra equation solving. Proceedings of the National Academy of Sciences of the United States of America 101 (15), 5686–5691.

Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. Neuroimage 52 (3), 1059–1069.

Ruby, P., Decety, J., 2001. Effect of subjective perspective taking during simulation of action: a pet investigation of agency. Nature neuroscience 4 (5), 546–550.

Ruby, P., Decety, J., 2003. What you believe versus what you think they believe: a neuroimaging study of conceptual perspective-taking. European Journal of Neuroscience 17 (11), 2475–2480.

Ruby, P., Decety, J., 2004. How would you feel versus how do you think she would feel? a neuroimaging study of perspective-taking with social emotions. Cognitive Neuroscience, Journal of 16 (6), 988–999.

Smith, S. M., Vidaurre, D., Beckmann, C. F., Glasser, M. F., Jenkinson, M., Miller, K. L., Nichols, T. E., Robinson, E. C., Salimi-Khorshidi, G., Woolrich, M. W., Barch, D. M., Uğurbil, K., Essen, D. C. V., 2013. Functional connectomics from resting-state fmri. Trends in Cognitive Sciences 17 (12), 666–682.

Spear, L. P., 2000. The adolescent brain and age-related behavioral manifestations. Neuroscience & Biobehavioral Reviews 24 (4), 417–463.

Sporns, O., 2011. Networks of the Brain. MIT Press, Cambridge.

Sporns, O., Giulio, T., Rolf, K., 09 2005. The Human Connectome: A Structural Description of the Human Brain. PLoS Comput Biol 1 (4), e42.

Van De Ville, D., Blu, T., Unser, M., 2004. Integrated wavelet processing and spatial statistical testing of fmri data. NeuroImage 23 (4), 1472 – 1485.

Van De Ville, D., Seghier, M. L., Lazeyras, F., Blu, T., Unser, M., 2007. Wspm: Wavelet-based statistical parametric mapping. NeuroImage 37 (4), 1205 – 1217.

van den Heuvel, M. P., Hulsoff-Pol, H. E., 2010. Exploring the brain network: a review on resting-state fMRI functional connectivity. European neuropsychopharma. 20, 519–534.

Varoquaux, G., Craddock, R. C., 2013. Learning and comparing functional connectomes across subjects. NeuroImage 80 (0), 405 – 415, mapping the Connectome.

White, S. H., 1996. The child's entry into the age of reason. The five to seven year shift: The age of reason and responsibility, 17–30.

Xia, M., Wang, J., He, Y., 07 2013. Brainnet viewer: A network visualization tool for human brain connectomics. PLoS ONE 8 (7), e68910.

Zalesky, A., Cocchi, L., Fornito, A., Murray, M. M., Bullmore, E., 2012. Connectivity differences in brain networks. NeuroImage 60 (2), 1055–1062.

Zalesky, A., Fornito, A., Bullmore, E. T., 2010. Network-based statistic: Identifying differences in brain networks. NeuroImage 53 (4), 1197–1207.
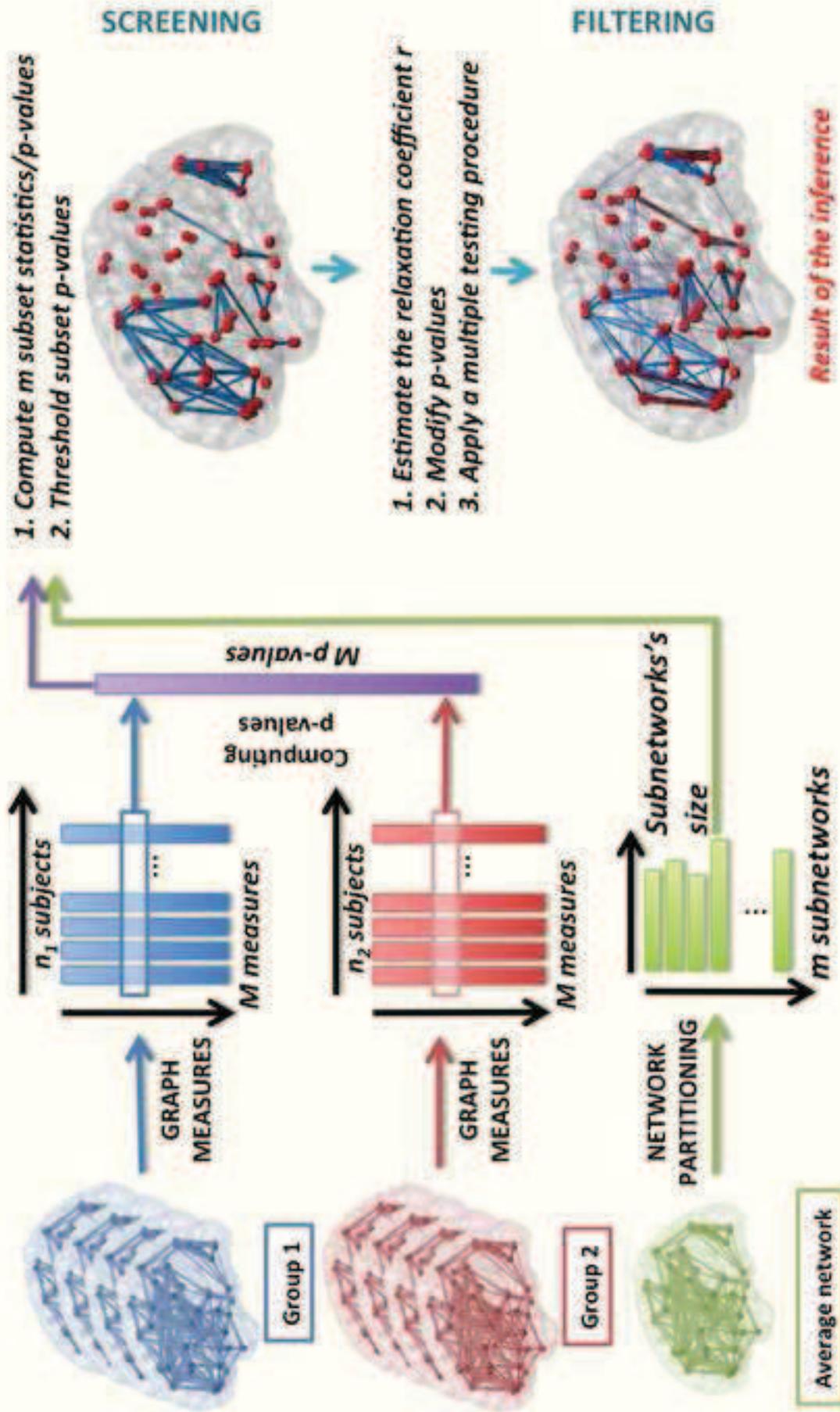
Highlights:

We propose an algorithm to locally compare connectomes

The algorithm properties are illustrated with analytical proofs and simulations

We compare our method against state-of-the-art methods

We compare structural connectomes between preschool children and adolescents