

# A simple and objective method for reproducible resting state network (RSN) detection in fMRI

Gautam V. Pendse<sup>\*1</sup>, David Borsook<sup>1,2</sup>, and Lino Becerra<sup>1,2</sup>

<sup>1</sup> P.A.I.N Group, Imaging and Analysis Group (IMAG), McLean Hospital, Harvard Medical School

<sup>2</sup> A. A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital

June 15, 2011

---

<sup>\*</sup>To whom correspondence should be addressed. e-mail: gpendse@mclean.harvard.edu

---

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Methods</b>	<b>7</b>
2.1	ICA background . . . . .	7
2.1.1	Estimating the model order $q$ . . . . .	8
2.1.2	Identifiability of the noisy ICA model . . . . .	9
2.1.3	Why is there a run-to-run variability in estimated ICs? . . . . .	14
2.2	ICA algorithms, single subject ICA and group ICA . . . . .	16
2.2.1	Single subject ICA . . . . .	17
2.2.2	Group ICA . . . . .	17
2.3	The original RAICAR algorithm . . . . .	17
2.4	The RAICAR-N enhancement . . . . .	19
2.5	How many subjects should be used per group ICA run in RAICAR-N? . . . . .	23
2.6	How to display the estimated non-Gaussian spatial structure in ICA maps? . . . . .	25
<b>3</b>	<b>Experiments and Results</b>	<b>26</b>
3.1	Human rsfMRI data . . . . .	26
3.2	Preprocessing . . . . .	27
3.3	RAICAR-N analysis with 1 ICA run per subject . . . . .	27
3.4	RAICAR-N on random sets of 5 subjects - 50 group ICA runs . . . . .	28
3.5	RAICAR-N on random sets of 5 subjects - 100 group ICA runs . . . . .	30
<b>4</b>	<b>Group comparison of ICA results</b>	<b>34</b>
4.1	Approaches based on a single group ICA run or no ICA run . . . . .	34
4.1.1	Advantages of single group ICA based approaches . . . . .	35
4.2	Approaches based on multiple single subject or group ICA runs . . . . .	35
4.2.1	Advantages of multiple ICA run approaches . . . . .	36
<b>5</b>	<b>Discussion</b>	<b>36</b>
5.1	Results for publicly available rsfMRI data . . . . .	37
5.2	Single subject ICA vs Group ICA . . . . .	37
5.3	How should subjects be grouped for group ICA? . . . . .	38
5.4	RAICAR-N for group comparisons of <i>reproducible</i> ICs . . . . .	38
5.4.1	Discussion of single group ICA based approaches . . . . .	38
5.4.2	Discussion of multiple ICA run approaches . . . . .	39
<b>6</b>	<b>Conclusions</b>	<b>40</b>
<b>7</b>	<b>Figure Legends</b>	<b>44</b>

---

## Abstract

Spatial Independent Component Analysis (ICA) decomposes the time by space functional MRI (fMRI) matrix into a set of 1-D basis time courses and their associated 3-D spatial maps that are optimized for mutual independence. When applied to resting state fMRI (rsfMRI), ICA produces several spatial independent components (ICs) that seem to have biological relevance - the so-called resting state networks (RSNs). The ICA problem is well posed when the true data generating process follows a linear mixture of ICs model in terms of the identifiability of the mixing matrix. However, the contrast function used for promoting mutual independence in ICA is dependent on the finite amount of observed data and is potentially non-convex with multiple local minima. Hence, each run of ICA could produce potentially different IC estimates even for the same data. One technique to deal with this run-to-run variability of ICA was proposed by Yang et al. [2008] in their algorithm RAICAR which allows for the selection of only those ICs that have a high run-to-run reproducibility. We propose an enhancement to the original RAICAR algorithm that enables us to assign reproducibility  $p$ -values to each IC and allows for an objective assessment of both within subject and across subjects reproducibility. We call the resulting algorithm RAICAR-N (N stands for null hypothesis test), and we have applied it to publicly available human rsfMRI data (<http://www.nitrc.org>). Our reproducibility analyses indicated that many of the published RSNs in rsfMRI literature are highly reproducible. However, we found several other RSNs that are highly reproducible but not frequently listed in the literature.

## Notation

- Scalars variables and functions will be denoted in a non-bold font (e.g.,  $\sigma^2, L, p$  or  $\Psi, f$ ). Vectors will be denoted in a bold font using lower case letters (e.g.,  $\mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\eta}$ ). Matrices will be denoted in bold font using upper case letters (e.g.,  $\mathbf{A}, \boldsymbol{\Sigma}, \mathbf{W}$ ). The transpose of a matrix  $\mathbf{A}$  will be denoted by  $\mathbf{A}^T$  and its inverse will be denoted by  $\mathbf{A}^{-1}$ .  $\mathbf{I}_p$  will denote the  $p \times p$  identity matrix and  $\mathbf{0}$  will denote a vector or matrix of all zeros whose size should be clear from context.  $\binom{N}{L}$  is the number of ways of choosing  $L$  objects from  $N$  objects when order does not matter.
- The  $j$ th component of vector  $\mathbf{t}_i$  will be denoted by  $t_{ij}$  whereas the  $j$ th component of vector  $\mathbf{t}$  will be denoted by  $t_j$ . The element  $(i, j)$  of matrix  $\mathbf{G}$  will be denoted by  $G(i, j)$  or  $G_{ij}$ . Estimates of variables will be denoted by putting a hat on top of the variable symbol. For example, an estimate of  $\mathbf{s}$  will be denoted by  $\hat{\mathbf{s}}$ .
- If  $\mathbf{x}$  is a random vector with a multivariate Normal distribution with mean  $\boldsymbol{\mu}$  and covariance  $\boldsymbol{\Sigma}$  then we will denote this distribution by  $\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ . The joint density of vector  $\mathbf{s}$  will be denoted by  $p_{\mathbf{s}}(\mathbf{s})$  whereas the marginal density of  $s_i$  will be denoted as  $p_{s_i}(s_i)$ .  $\mathbf{E}[f(\mathbf{s}, \boldsymbol{\eta})]$  denotes the expectation of  $f(\mathbf{s}, \boldsymbol{\eta})$  with respect to both random variables  $\mathbf{s}$  and  $\boldsymbol{\eta}$ .

---

# 1 Introduction

Independent component analysis (ICA) [Jutten and Herault, 1991, Comon, 1994, Bell and Sejnowski, 1995, Attias, 1999] models the observed data as a linear combination of a set of statistically independent and unobservable sources. [McKeown et al., 1998] first proposed the application of ICA to the analysis of functional magnetic resonance imaging (fMRI) data. Subsequently, ICA has been applied to fMRI both as an exploratory tool for the purpose of identifying task related components [McKeown et al., 1998] as well as a signal clean up tool for the purpose of removing artifacts from the fMRI data [Tohka et al., 2008]. Recently, it has been shown that ICA applied to resting state fMRI (rsfMRI) in healthy subjects reveals a set of biologically meaningful spatial maps of independent components (ICs) that are consistent across subjects - the so called resting state networks (RSNs) [Beckmann et al., 2005]. Hence, there is a considerable interest in applying ICA to rsfMRI data in order to define the set of RSNs that characterize a particular group of human subjects, a disease, or a pharmacological effect.

Several variants of the linear ICA model have been applied to fMRI data including square ICA (with equal number of sources and sensors) [McKeown et al., 1998], non-square ICA (with more sensors than sources) [Calhoun et al., 2001], and non-square ICA with additive Gaussian noise (noisy ICA) [Beckmann and Smith, 2004]. All of these models are well known in the ICA literature [Jutten and Herault, 1991, Cardoso, 1998, Comon, 1994, Attias, 1999]. Since the other ICA models are specializations of the noisy ICA model, we will assume a noisy ICA model henceforth.

Remarkably, the ICA estimation problem is well posed in terms of the identifiability of the mixing matrix given several non-Gaussian and at most 1 Gaussian source in the overall linear mixture [Rao, 1969, Comon, 1994, Theis, 2004, Davies, 2004]. In the presence of more than 1 Gaussian source, such as in noisy ICA, the mixing matrix corresponding to the non-Gaussian part of the linear mixture is identifiable (upto permutation and scaling). In addition, the source distributions are uniquely identifiable (upto permutation and scaling) given a noisy ICA model with a particular Gaussian co-variance structure, for example, the isotropic diagonal co-variance. For details, see section 2.1.2.

While these uniqueness results are reassuring, a number of practical difficulties prevent the reliable estimation of ICs on real data. These difficulties include (1) true data not describable by an ICA model, (2) ICA contrast function approximations, (3) multiple local minima in the ICA contrast function, (4) confounding Gaussian noise and (5) model order overestimation. See section 2.1.3 for more details. A consequence of these difficulties is that multiple ICA runs on the same data or different subsets of the data produce different estimates of the IC realizations.

One technique to account for this run-to-run variability in ICA was proposed by [Himberg et al., 2004] in their algorithm ICASSO. Using repeated runs of ICA with bootstrapped data using various initial conditions, ICASSO clusters ICs across ICA runs using agglomerative hierarchical clustering and also helps in visualizing the estimated ICs. The logic is that reliable ICs will show up in



---

almost all ICA runs and thus will form a tight cluster well separated from the rest. [Esposito et al., 2005] proposed a technique similar to ICASSO called self-organizing group ICA (sogICA) which allows for clustering of ICs via hierarchical clustering in across subject ICA runs. When applied to multiple ICA runs across subjects, ICASSO does not restrict the IC clusters to contain only 1 IC from each subject per ICA run. In contrast, sogICA allows the user to select the minimum number of subjects for a "group representative" IC cluster containing distinct subjects. By labelling each ICA run as a different "subject" sogICA can also be applied to analyze multiple ICA runs across subjects.

Similar in spirit to ICASSO and sogICA, Yang et al. [2008] proposed an intuitive approach called RAICAR (Ranking and Averaging Independent Component Analysis by Reproducibility) for reproducibility analysis of estimated ICs. The basic idea in RAICAR is to select only those ICs as "interesting" or "stable" which show a high run-to-run "reproducibility". RAICAR uses simple and automated spatial cross-correlation matrix based IC alignment which has been shown to be more accurate compared to ICASSO [Yang et al., 2008]. RAICAR is applicable to both within subject as well as across subjects reproducibility analysis.

A few limitations of ICASSO, sogICA and RAICAR are worth noting:

- ICASSO requires the user to select the number of IC clusters and is inapplicable without modification for across subjects analysis of ICA runs since the IC clusters are not restricted to contain only 1 IC per ICA run.
- sogICA requires the user to select the minimum number of subjects for a "group representative" cluster and also a cutoff on within cluster distances.
- RAICAR uses an arbitrary threshold on the reproducibility indices selected "by eye" or set at an arbitrary value, such as 50% of the maximum reproducibility value.

We propose a simple extension to RAICAR that avoids subjective user decisions and allows for an automatic reproducibility cutoff. The reproducibility indices calculated in RAICAR differ in magnitude significantly depending on whether the input to RAICAR:

- (a) is generated using multiple ICA runs on the same data
- (b) comes from multiple ICA runs on varying data sets (e.g. between and across subject runs)

See Figure 1 for an illustration of this effect. Obviously, the reproducibility indices are much lower in case (b) since we account for both within subject and between subjects variability in estimating ICs. Case (b) is also of great interest from a practical point of view since we are often interested in making statements about a group of subjects. Hence, it is clear that a cutoff on RAICAR reproducibility values for the purposes of selecting the "highly reproducible" components should be data dependent. In this work,

1. We propose a modification of the original RAICAR algorithm by introducing an explicit "null" model of no reproducibility.

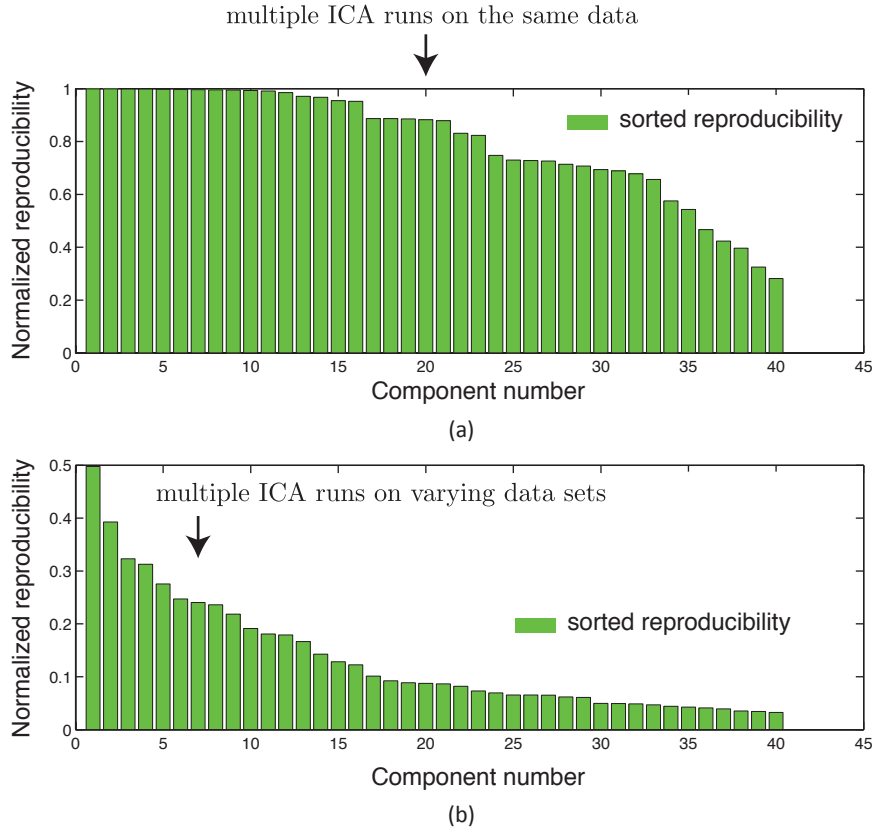


Figure 1: Figure illustrates the variation in normalized reproducibility from RAICAR depending on whether the input to RAICAR is (a) Multiple ICA runs on single subject data or (b) Multiple ICA runs across subjects. Notice that the normalized reproducibility is much lower for across subjects analysis compared to within subject analysis.

- 
2. We use this "null" model to automatically generate  $p$ -values for each IC via simulation. This allows for an objective cutoff specification for extracting reproducible ICs (e.g. reproducible at  $p < 0.05$ ) within and across subjects. We call the resulting algorithm RAICAR-N (N stands for "null" hypothesis test).
  3. We validate RAICAR-N by applying it to publicly available human rsfMRI data.

## 2 Methods

The organization of this article revolves around the following sequence of questions which ultimately lead to the development of RAICAR-N:

1. Why is a reproducibility assessment necessary in ICA analysis? In order to answer this question, we cover the fundamentals of ICA including identifiability issues in sections 2.1 and 2.2.
2. How does the original RAICAR algorithm assess reproducibility? The answer to this question in section 2.3 will set up the stage for RAICAR-N.
3. How does RAICAR-N permit calculation of reproducibility  $p$ -values? In section 2.4, we describe the RAICAR-N "null" model and a simulation based approach for assigning  $p$ -values to ICs.
4. How to promote diversity in group ICA runs given a limited number of subjects when using RAICAR-N and how to display the non-Gaussian spatial structure in estimated ICs? These issues are covered in section 2.5 and 2.6.
5. How can RAICAR-N be extended for between group comparison of ICs and how does it compare to other approaches in the literature? This question is addressed in section 5.4.

### 2.1 ICA background

In this section, we provide a brief introduction to ICA along with a discussion of associated issues related to model order selection, identifiability and run-to-run variability. The noisy ICA model assumes that observed data  $\mathbf{y}$  is generated as a linear combination of unobservable independent sources confounded with Gaussian noise:

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{A} \mathbf{s} + \boldsymbol{\eta} \quad (2.1)$$

---

In this model,

$$\begin{aligned}
\mathbf{y} &= p \times 1 \text{ observed signal vector} \\
\boldsymbol{\mu} &= p \times 1 \text{ mean vector} \\
\mathbf{A} &= p \times q \text{ mixing matrix with } p > q \text{ (more sensors than sources) and rank } q \\
\boldsymbol{\eta} &= p \times 1 \text{ Gaussian noise vector with density } \mathcal{N}(\boldsymbol{\eta} \mid \mathbf{0}, \boldsymbol{\Sigma}) \\
\mathbf{s} &= q \times 1 \text{ vector of independent random variables (the ICs)} \\
&\text{with } \mathbf{E}(\mathbf{s}\mathbf{s}^T) = \mathbf{D} \text{ (diagonal) and } \mathbf{E}(\mathbf{s}) = \mathbf{0} \\
&\text{and with } \mathbf{s} \text{ and } \boldsymbol{\eta} \text{ independent}
\end{aligned} \tag{2.2}$$

If the marginal density of the  $i$ th source  $s_i$  is  $p_{s_i}(s_i)$  then the joint source density  $\mathbf{p}_s(\mathbf{s})$  factorizes as  $\prod_{i=1}^q p_{s_i}(s_i)$  because of the independence assumption but is otherwise assumed to be unknown. Also, since the elements of  $\mathbf{s}$  are independent their co-variance matrix  $\mathbf{D}$  is diagonal. The set of variables  $\mathcal{F} = \{\boldsymbol{\mu}, \mathbf{A}, \mathbf{D}, \boldsymbol{\Sigma}\}$  represents the unknown parameters in the noisy ICA model. Before discussing the identifiability of model 2.1, we briefly discuss the choice of model order or the assumed number of ICs  $q$ .

### 2.1.1 Estimating the model order $q$

Rigorous estimation of the model order  $q$  in noisy ICA is difficult as the IC densities  $p_{s_i}(s_i)$  are unknown. This means that  $\mathbf{p}(\mathbf{y} \mid q, \mathcal{F})$ , the marginal density of the observed data given the model order and the ICA parameters cannot be derived in closed form (by integrating out the ICs) without making additional assumptions on the form of IC densities. Consequently, standard model selection criteria such as Bayes information criterion (BIC) [Kass and Raftery, 1993] cannot be easily applied to the noisy ICA model to estimate  $q$ . One solution is to use a factorial mixture of Gaussians (MOG) joint source density model as in [Attias, 1999], and use the analytical expression for  $\mathbf{p}(\mathbf{y} \mid q, \mathcal{F})$  in conjunction with BIC. This solution is quite general in terms of allowing for an arbitrary Gaussian noise co-variance  $\boldsymbol{\Sigma}$ , but maximizing  $\mathbf{p}(\mathbf{y} \mid q, \mathcal{F})$  with respect to  $\mathcal{F}$  becomes computationally intractable using an expectation maximization (EM) algorithm for  $q > 13$  ICs [Attias, 1999]. Another rigorous non-parametric approach for estimating  $q$  that is applicable to the noisy ICA model with isotropic diagonal Gaussian noise co-variance i.e., with  $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_p$  is the random matrix theory based sequential hypothesis testing approach of Kritchman and Nadler [2009]. To the best of our knowledge, these are the only 2 rigorous approaches for estimating  $q$  in the noisy ICA model.

Approximate approaches for estimating  $q$  commonly used in fMRI literature (e.g., [Beckmann and Smith, 2004]) consist of first relaxing the isotropic diagonal noisy ICA model (with  $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_p$ ) into a probabilistic PCA (PPCA) model of [Tipping, 1999] where the source densities are assumed to be Gaussian i.e., where  $\mathbf{p}_s(\mathbf{s}) = \mathcal{N}(\mathbf{s} \mid \mathbf{0}, \mathbf{I}_q)$ . When using the PPCA model, it becomes possible to integrate out the Gaussian sources to get an expression for  $\mathbf{p}(\mathbf{y} \mid q, \mathcal{F})$  that can be analytically

---

maximized [Tipping, 1999]. Subsequently, methods such as BIC can be applied to estimate  $q$ . Alternative approaches for estimating  $q$  in the PPCA model consist of the Bayesian model selection of Minka [2000], or in data-rich situations such as fMRI, even the standard technique of cross-validation [Hastie et al., 2009].

From a biological point of view, it has been argued [Cole et al., 2010] that the number of extracted ICs simply reflect the various *equally valid* views of the human functional neurobiology - smaller number of ICs represent a coarse view while a larger number of ICs represent a more fine grained view. However, it is worth noting that from a statistical point of view, over-specification of  $q$  will lead to over-fitting of the ICA model which might render the estimated ICs less generalizable across subjects. On the other hand, under-specification of  $q$  will result in incomplete IC separation. Both of these scenarios are undesirable.

### 2.1.2 Identifiability of the noisy ICA model

To what extent is the noisy linear ICA model identifiable? Consider a potentially different decomposition of the noisy ICA model 2.1:

$$\mathbf{y} = \boldsymbol{\mu}_1 + \mathbf{A}_1 \mathbf{s}_1 + \boldsymbol{\eta}_1 \quad (2.3)$$

where

$$\begin{aligned} \mathbf{y} &= p \times 1 \text{ observed signal vector} \\ \boldsymbol{\mu}_1 &= p \times 1 \text{ mean vector} \\ \mathbf{A}_1 &= p \times q \text{ mixing matrix with } p > q \text{ (more sensors than sources) and rank } q \\ \boldsymbol{\eta}_1 &= p \times 1 \text{ Gaussian noise vector with density } \mathcal{N}(\boldsymbol{\eta}_1 \mid \mathbf{0}, \boldsymbol{\Sigma}_1) \\ \mathbf{s}_1 &= q \times 1 \text{ vector of independent random variables (the ICs)} \\ &\text{with } \mathbf{E}(\mathbf{s}\mathbf{s}^T) = \mathbf{D}_1 \text{ (diagonal) and } \mathbf{E}(\mathbf{s}) = \mathbf{0} \\ &\text{and with } \mathbf{s}_1 \text{ and } \boldsymbol{\eta}_1 \text{ independent} \end{aligned} \quad (2.4)$$

What can be said about the equivalence between the parameterizations in 2.1 and 2.3?

**Identifiability of  $\boldsymbol{\mu}$ :** Equating the expectations of the right hand side of 2.3 and 2.1 and noting that  $\mathbf{s}, \boldsymbol{\eta}, \mathbf{s}_1, \boldsymbol{\eta}_1$  have mean  $\mathbf{0}$  we get:

$$\boldsymbol{\mu}_1 = \boldsymbol{\mu} \quad (2.5)$$

Thus the mean vector  $\boldsymbol{\mu}$  is exactly identifiable.

---

**Identifiability of  $\mathbf{A}$ :** A fundamental decomposition result states that the noisy ICA problem is well-posed in terms of the identifiability of the mixing matrix  $\mathbf{A}$  upto permutation and scaling provided that the components of  $\mathbf{s}$  are independent and **non-Gaussian** [Rao, 1969, Comon, 1994, Theis, 2004, Davies, 2004]. If  $\mathbf{\Lambda}$  is a diagonal scaling matrix and  $\mathbf{P}$  is a permutation matrix then the identifiability result can be stated as:

$$\mathbf{A}_1 = \mathbf{A} \mathbf{\Lambda} \mathbf{P} \quad (2.6)$$

where 2.3 is another decomposition of  $\mathbf{y}$  with  $\mathbf{s}_1$  containing independent and non-Gaussian components. In other words, the mixing matrix  $\mathbf{A}$  is identifiable upto permutation and scaling.

**Identifiability of  $\mathbf{D}$  and  $\mathbf{\Sigma}$ :** Equating the second moments of the right hand side of 2.3 and 2.1 and noting the equality of means 2.5 and the independence of  $\mathbf{s}, \boldsymbol{\eta}$  and  $\mathbf{s}_1, \boldsymbol{\eta}_1$  we get:

$$\mathbf{E}[(\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})^T] = \mathbf{A} \mathbf{D} \mathbf{A}^T + \mathbf{\Sigma} = \mathbf{A}_1 \mathbf{D}_1 \mathbf{A}_1^T + \mathbf{\Sigma}_1 \quad (2.7)$$

Let  $\mathbf{W}$  be a  $q \times p$  matrix and  $\tilde{\mathbf{Q}}$  be a  $p \times (p - q)$  orthogonal matrix such that:

$$\begin{aligned} \mathbf{W} &= (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \\ \tilde{\mathbf{Q}}^T \mathbf{A} &= \mathbf{0} \\ \tilde{\mathbf{Q}}^T \tilde{\mathbf{Q}} &= \mathbf{I}_{p-q} \end{aligned} \quad (2.8)$$

From 2.8 and 2.7 we get:

$$\begin{aligned} \mathbf{D} + \mathbf{W} \mathbf{\Sigma} \mathbf{W}^T &= \mathbf{\Lambda} \mathbf{P} \mathbf{D}_1 \mathbf{P}^T \mathbf{\Lambda}^T + \mathbf{W} \mathbf{\Sigma}_1 \mathbf{W}^T \\ \tilde{\mathbf{Q}}^T \mathbf{\Sigma} \tilde{\mathbf{Q}} &= \tilde{\mathbf{Q}}^T \mathbf{\Sigma}_1 \tilde{\mathbf{Q}} \end{aligned} \quad (2.9)$$

**Case 1:**  $\mathbf{\Sigma} = \sigma^2 \mathbf{I}_p$  and  $\mathbf{\Sigma}_1 = \sigma_1^2 \mathbf{I}_p$

The second equation in 2.9 along with the orthogonality of  $\tilde{\mathbf{Q}}$  gives  $\sigma^2 = \sigma_1^2$  and thus  $\mathbf{\Sigma} = \mathbf{\Sigma}_1$ . If we fix the scaling of  $\mathbf{A}_1$  by selecting  $\mathbf{\Lambda}^2 = \mathbf{I}_q$  then from the first equation in 2.9 we get:

$$\begin{aligned} \mathbf{D} &= \mathbf{\Lambda} \mathbf{P} \mathbf{D}_1 \mathbf{P}^T \mathbf{\Lambda}^T \\ &= \mathbf{P} \mathbf{D}_1 \mathbf{P}^T \mathbf{\Lambda}^2 && (\mathbf{P} \mathbf{D}_1 \mathbf{P}^T \text{ is diagonal}) \\ &= \mathbf{P} \mathbf{D}_1 \mathbf{P}^T \end{aligned} \quad (2.10)$$

In other words, the noise co-variance  $\mathbf{\Sigma} = \sigma^2 \mathbf{I}_p$  is uniquely determined and for a fixed scaling  $\mathbf{\Lambda}^2 = \mathbf{I}_q$ , the source variances  $\mathbf{D}$  are also uniquely determined upto permutation.

**Case 2:**  $\mathbf{\Sigma}$  and  $\mathbf{\Sigma}_1$  arbitrary positive definite matrices

Suppose  $\mathbf{X}$  is a square matrix and let  $\text{diag}(\mathbf{X})$  be the diagonal matrix obtained by setting the

non-diagonal elements of  $\mathbf{X}$  to 0 and similarly let  $\text{offdiag}(\mathbf{X})$  be the matrix obtained by setting the diagonal elements of  $\mathbf{X}$  to 0. The noise-covariance is partially identifiable by the following conditions:

$$\begin{aligned} \tilde{\mathbf{Q}}^T \Sigma \tilde{\mathbf{Q}} &= \tilde{\mathbf{Q}}^T \Sigma_1 \tilde{\mathbf{Q}} \\ \text{offdiag}(\mathbf{W} \Sigma \mathbf{W}^T) &= \text{offdiag}(\mathbf{W} \Sigma_1 \mathbf{W}^T) \end{aligned} \quad (2.11)$$

For a fixed scaling  $\Lambda^2 = \mathbf{I}_q$ , the sources variances  $\mathbf{D}, \mathbf{D}_1$  are constrained by:

$$\mathbf{D} + \text{diag}(\mathbf{W} \Sigma \mathbf{W}^T) = \mathbf{P} \mathbf{D}_1 \mathbf{P}^T + \text{diag}(\mathbf{W} \Sigma_1 \mathbf{W}^T) \quad (2.12)$$

In general, the source variances  $\mathbf{D}$  cannot be uniquely determined as noted in [Davies, 2004].

**Identifiability of the distribution of  $\mathbf{s}$ :** Is the distribution of the non-Gaussian components of  $\mathbf{s}$  identifiable? From 2.1 and 2.3:

$$\boldsymbol{\mu} + \mathbf{A} \mathbf{s} + \boldsymbol{\eta} = \boldsymbol{\mu}_1 + \mathbf{A}_1 \mathbf{s}_1 + \boldsymbol{\eta}_1 \quad (2.13)$$

Substituting 2.5 and 2.6 in 2.13 we get:

$$\mathbf{A} \mathbf{s} + \boldsymbol{\eta} = \mathbf{A} \Lambda \mathbf{P} \mathbf{s}_1 + \boldsymbol{\eta}_1 \quad (2.14)$$

Premultiplying both sides by  $\mathbf{W}$  from 2.8 we get:

$$\mathbf{s} + \mathbf{W} \boldsymbol{\eta} = \Lambda \mathbf{P} \mathbf{s}_1 + \mathbf{W} \boldsymbol{\eta}_1 \quad (2.15)$$

Let  $\Psi_{\mathbf{s}}, \Psi_{\boldsymbol{\eta}}, \Psi_{\mathbf{s}_1}, \Psi_{\boldsymbol{\eta}_1}$  be the characteristic functions of  $\mathbf{s}, \boldsymbol{\eta}, \mathbf{s}_1$  and  $\boldsymbol{\eta}_1$  respectively. Then

$$\begin{aligned} \Psi_{\mathbf{s}}(\mathbf{t}) &= \mathbf{E} [\exp(i\mathbf{t}^T \mathbf{s})] & \Psi_{\boldsymbol{\eta}}(\mathbf{t}) &= \mathbf{E} [\exp(i\mathbf{t}^T \boldsymbol{\eta})] \\ \Psi_{\mathbf{s}_1}(\mathbf{t}) &= \mathbf{E} [\exp(i\mathbf{t}^T \mathbf{s}_1)] & \Psi_{\boldsymbol{\eta}_1}(\mathbf{t}) &= \mathbf{E} [\exp(i\mathbf{t}^T \boldsymbol{\eta}_1)] \end{aligned} \quad (2.16)$$

where  $i = \sqrt{-1}$  and  $\mathbf{t}$  is a vector of real numbers of length equal to that of the corresponding random vectors in 2.16. Using 2.15, we can write:

$$\mathbf{E} [\exp(i\mathbf{t}^T \{\mathbf{s} + \mathbf{W} \boldsymbol{\eta}\})] = \mathbf{E} [\exp(i\mathbf{t}^T \{\Lambda \mathbf{P} \mathbf{s}_1 + \mathbf{W} \boldsymbol{\eta}_1\})] \text{ for all } \mathbf{t} \in \mathbf{R}^q \quad (2.17)$$

Noting the independence of  $\mathbf{s}, \boldsymbol{\eta}$  and  $\mathbf{s}_1, \boldsymbol{\eta}_1$ :

$$\begin{aligned} \mathbf{E} [\exp(i\mathbf{t}^T \{\mathbf{s}\})] \mathbf{E} [\exp(i\mathbf{t}^T \{\mathbf{W} \boldsymbol{\eta}\})] &= \mathbf{E} [\exp(i\mathbf{t}^T \{\Lambda \mathbf{P} \mathbf{s}_1\})] \mathbf{E} [\exp(i\mathbf{t}^T \{\mathbf{W} \boldsymbol{\eta}_1\})] \\ \Rightarrow \Psi_{\mathbf{s}}(\mathbf{t}) \Psi_{\boldsymbol{\eta}}(\mathbf{W}^T \mathbf{t}) &= \Psi_{\mathbf{s}_1}(\mathbf{P}^T \Lambda^T \mathbf{t}) \Psi_{\boldsymbol{\eta}_1}(\mathbf{W}^T \mathbf{t}) \text{ for all } \mathbf{t} \in \mathbf{R}^q \end{aligned} \quad (2.18)$$

Now  $\boldsymbol{\eta}$  and  $\boldsymbol{\eta}_1$  are multivariate Gaussian random vectors both with mean  $\mathbf{0}$  and co-variance matrix  $\boldsymbol{\Sigma}$  and  $\boldsymbol{\Sigma}_1$  respectively. Hence, their characteristic functions are given by [Feller, 1966, Wlodzimierz, 1995]:

$$\begin{aligned}\Psi_{\boldsymbol{\eta}}(\mathbf{W}^T \mathbf{t}) &= \exp \left( -\frac{1}{2} \mathbf{t}^T \mathbf{W} \boldsymbol{\Sigma} \mathbf{W}^T \mathbf{t} \right) \text{ for all } \mathbf{t} \in \mathbf{R}^q \\ \Psi_{\boldsymbol{\eta}_1}(\mathbf{W}^T \mathbf{t}) &= \exp \left( -\frac{1}{2} \mathbf{t}^T \mathbf{W} \boldsymbol{\Sigma}_1 \mathbf{W}^T \mathbf{t} \right) \text{ for all } \mathbf{t} \in \mathbf{R}^q\end{aligned}\tag{2.19}$$

**Claim 2.1.** *A sufficient condition for identifiability upto permutation and scaling of the non-Gaussian distributions in  $\mathbf{s}$  given two different parameterizations in 2.1 and 2.3 is:*

$$\text{diag}(\mathbf{W} \boldsymbol{\Sigma} \mathbf{W}^T) = \text{diag}(\mathbf{W} \boldsymbol{\Sigma}_1 \mathbf{W}^T)\tag{2.20}$$

*Proof.* From 2.20 and 2.11, we get:

$$\mathbf{W} \boldsymbol{\Sigma} \mathbf{W}^T = \mathbf{W} \boldsymbol{\Sigma}_1 \mathbf{W}^T\tag{2.21}$$

Thus from 2.19,

$$\Psi_{\boldsymbol{\eta}}(\mathbf{W}^T \mathbf{t}) = \Psi_{\boldsymbol{\eta}_1}(\mathbf{W}^T \mathbf{t}) \text{ for all } \mathbf{t} \in \mathbf{R}^q\tag{2.22}$$

From 2.19,  $\Psi_{\boldsymbol{\eta}}(\mathbf{W}^T \mathbf{t})$  and  $\Psi_{\boldsymbol{\eta}_1}(\mathbf{W}^T \mathbf{t})$  are not equal to 0 for any finite  $\mathbf{t}$ , therefore, from 2.22 and 2.18 we get:

$$\Psi_{\mathbf{s}}(\mathbf{t}) = \Psi_{\mathbf{s}_1}(\mathbf{P}^T \boldsymbol{\Lambda}^T \mathbf{t}) \text{ for all } \mathbf{t} \in \mathbf{R}^q\tag{2.23}$$

Note that  $\boldsymbol{\Lambda}$  is a diagonal scaling matrix with entries  $\lambda_1, \lambda_2, \dots, \lambda_q$  on the diagonal and  $\mathbf{P}$  is a permutation matrix. Thus,

$$\mathbf{P}^T \boldsymbol{\Lambda}^T \mathbf{t} = \begin{pmatrix} \lambda_{i_1} t_{i_1} \\ \lambda_{i_2} t_{i_2} \\ \vdots \\ \lambda_{i_q} t_{i_q} \end{pmatrix}\tag{2.24}$$

where  $i_1, i_2, \dots, i_q$  is some permutation of integers  $1, 2, \dots, q$ . Suppose  $\Psi_{\mathbf{s}(j)}$  is the characteristic function of the  $j$ th component of  $\mathbf{s}$  and  $\Psi_{\mathbf{s}_1(j)}$  is the characteristic function of the  $j$ th component of  $\mathbf{s}_1$ . Since the components of  $\mathbf{s}$  and  $\mathbf{s}_1$  are independent by assumption, the joint characteristic functions  $\Psi(\mathbf{s})$  and  $\Psi(\mathbf{s}_1)$  factorize:

$$\begin{aligned}\Psi_{\mathbf{s}}(\mathbf{t}) &= \Psi_{\mathbf{s}(1)}(t_1) \Psi_{\mathbf{s}(2)}(t_2) \dots \Psi_{\mathbf{s}(j)}(t_j) \dots \Psi_{\mathbf{s}(q)}(t_q) \\ \Psi_{\mathbf{s}_1}(\mathbf{P}^T \boldsymbol{\Lambda}^T \mathbf{t}) &= \Psi_{\mathbf{s}_1(1)}(\lambda_{i_1} t_{i_1}) \Psi_{\mathbf{s}_1(2)}(\lambda_{i_2} t_{i_2}) \dots \Psi_{\mathbf{s}_1(j)}(\lambda_{i_j} t_{i_j}) \dots \Psi_{\mathbf{s}_1(q)}(\lambda_{i_q} t_{i_q})\end{aligned}\tag{2.25}$$



---

From 2.25 and 2.23

$$\Psi_{\mathbf{s}(1)}(t_1) \dots \Psi_{\mathbf{s}(j)}(t_j) \dots \Psi_{\mathbf{s}(q)}(t_q) = \Psi_{\mathbf{s}_1(1)}(\lambda_{i_1} t_{i_1}) \dots \Psi_{\mathbf{s}_1(j)}(\lambda_{i_j} t_{i_j}) \dots \Psi_{\mathbf{s}_1(q)}(\lambda_{i_q} t_{i_q}) \quad (2.26)$$

All characteristic functions satisfy [Feller, 1966, Wlodzimierz, 1995]:

$$\begin{aligned} \Psi_{\mathbf{s}(k)}(0) &= 1 \\ \Psi_{\mathbf{s}_1(k)}(0) &= 1 \text{ for all } k \end{aligned} \quad (2.27)$$

Since  $i_1, i_2, \dots, i_q$  is simply a permutation of integers  $1, 2, \dots, q$ , there exists a  $j$  such that  $i_j = 1$ . Then set  $t_2 = 0, t_3 = 0, \dots, t_q = 0$  in 2.26. Then 2.27 and 2.26 imply:

$$\Psi_{\mathbf{s}(1)}(t_1) = \Psi_{\mathbf{s}_1(j)}(\lambda_{i_j} t_{i_j}) = \Psi_{\mathbf{s}_1(j)}(\lambda_1 t_1) \text{ for all } t_1 \in \mathbf{R} \quad (2.28)$$

Select the scaling matrix as  $\mathbf{\Lambda}^2 = \mathbf{I}_q$  and thus  $\mathbf{\Lambda}$  is a diagonal matrix with elements  $\pm 1$  on the diagonal. Thus  $\lambda_1 = \pm 1$  and 2.28 can be re-written as:

$$\Psi_{\mathbf{s}(1)}(t_1) = \Psi_{\mathbf{s}_1(j)}(\pm t_1) \text{ for all } t_1 \in \mathbf{R} \quad (2.29)$$

Therefore,

$$\Psi_{\mathbf{s}(1)}(t_1) = \Psi_{\mathbf{s}_1(j)}(t_1) \text{ for all } t_1 \in \mathbf{R} \quad (2.30)$$

or

$$\Psi_{\mathbf{s}(1)}(t_1) = \Psi_{\mathbf{s}_1(j)}(-t_1) = \Psi_{-\mathbf{s}_1(j)}(t_1) \text{ for all } t_1 \in \mathbf{R}$$

Hence the characteristic function of the 1st component of  $\mathbf{s}$  is identical to the characteristic function of the (possibly sign-flipped)  $j$ th component of  $\mathbf{s}_1$ . Since characteristic functions uniquely characterize a probability distribution [Feller, 1966], the distribution of  $\mathbf{s}(1)$  and  $\pm \mathbf{s}_1(j)$  is identical. Next, by setting  $t_1 = 0, t_3 = 0, \dots, t_q = 0$ , we can find a distribution from  $\mathbf{s}_1$  that matches the 2nd component  $\mathbf{s}(2)$  of  $\mathbf{s}$ . Proceeding in a similar fashion, it is clear that the distribution of each component of  $\mathbf{s}$  is uniquely identifiable upto sign flips for the choice  $\mathbf{\Lambda}^2 = \mathbf{I}_q$ . For a general  $\mathbf{\Lambda}$ , the source distributions are uniquely identifiable upto permutation and (possibly negative) scaling, as claimed.  $\square$

While the source distributions might not be uniquely identifiable for arbitrary co-variance matrices  $\mathbf{\Sigma}$ , they are indeed uniquely identifiable upto permutation and scaling for the noisy ICA model with isotropic Gaussian noise co-variance. For more general conditions that guarantee uniqueness of source distributions, please see Eriksson and Koivunen [2004, 2006].

**Corollary 2.2.** *If  $\mathbf{\Sigma} = \sigma^2 \mathbf{I}_p$  and  $\mathbf{\Sigma}_1 = \sigma_1^2 \mathbf{I}_p$ , then the source distributions are uniquely identifiable upto sign flips for  $\mathbf{\Lambda}^2 = \mathbf{I}_q$ .*

---

*Proof.* Suppose  $\Sigma = \sigma^2 \mathbf{I}_p$  and  $\Sigma_1 = \sigma_1^2 \mathbf{I}_p$ . Then from 2.9  $\Sigma = \Sigma_1$  and thus  $\text{diag}(\mathbf{W}\Sigma\mathbf{W}^T) = \text{diag}(\mathbf{W}\Sigma_1\mathbf{W}^T)$ . The corollary then follows from Claim 2.1.  $\square$

**Corollary 2.3.** *If  $\mathbf{D} = \mathbf{D}_1 = \mathbf{I}_q$ , then the source distributions are uniquely identifiable up to sign flips for  $\Lambda^2 = \mathbf{I}_q$ .*

*Proof.* If  $\mathbf{D} = \mathbf{D}_1 = \mathbf{I}_q$ , then noting that  $\mathbf{P}\mathbf{P}^T = \mathbf{I}_q$ , we get  $\mathbf{D} = \mathbf{P}\mathbf{D}_1\mathbf{P}^T$ . Hence from 2.12, we get  $\text{diag}(\mathbf{W}\Sigma\mathbf{W}^T) = \text{diag}(\mathbf{W}\Sigma_1\mathbf{W}^T)$ . The corollary then follows from Claim 2.1.  $\square$

### 2.1.3 Why is there a run-to-run variability in estimated ICs?

From the discussion in section 2.1.2, it is clear that for a noisy ICA model with isotropic diagonal additive Gaussian noise co-variance:

1. The noisy ICA parameters  $\mathcal{F} = \{\boldsymbol{\mu}, \mathbf{A}, \mathbf{D}, \Sigma\}$  are uniquely identifiable up to permutation and scaling.
2. The source distributions in  $\mathbf{s}$  are uniquely identifiable upto permutation and scaling.

While the above theoretical properties of ICA are reassuring, there are a number of practical difficulties that prevent the reliable estimation of ICs on real data:

#### 1. Validity of the ICA model:

The assumption that the observed real data is generated by an ICA model is only that - an "assumption". If this assumption is not valid, then the uniqueness results do not hold anymore.

#### 2. Mutual information approximations:

From an information theoretic point of view, the ICA problem is solved by minimizing a contrast function which is an *approximation* to the mutual information [Hyvarinen, 1998] between the ICs that depends on the finite amount of observed data. Such an approximation is necessary, since we do not have access to the marginal source densities  $p_{s_i}$ . Different approximations to mutual information will lead to different objective functions and hence different solutions. This is one of the reasons why different ICA algorithms often produce different IC estimates even for the same data.

#### 3. Non-convexity of ICA objective functions:

The ICA contrast function is potentially non-convex and hence has multiple local minima. Since global minimization is a challenging problem by itself, most ICA algorithms will only converge to local minima of the ICA contrast function. The run-to-run variability of IC estimates will also depend on the number of local minima in a particular ICA contrast function.

#### 4. IC estimate corruption by Gaussian noise:

For noisy ICA, the IC realizations cannot be recovered exactly even if the true mixing matrix  $\mathbf{A}$  and mean vector  $\boldsymbol{\mu}$  are known in 2.1. Commonly used estimators for recovering realization of ICs include the least squares [Beckmann and Smith, 2004] as well as the minimum mean square error (MMSE) [Davies, 2004]. Consider the least squares estimate  $\hat{\mathbf{s}}$  of a realization of  $\mathbf{s}$  based on  $\mathbf{y}$ :

$$\hat{\mathbf{s}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T (\mathbf{y} - \boldsymbol{\mu}) = \mathbf{s} + (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \boldsymbol{\eta} \quad (2.31)$$

This means that even for known parameters, IC realization estimates  $\hat{\mathbf{s}}$  will be corrupted by correlated Gaussian noise. Hence using different subsets of the data under the true model will also lead to variability in estimated ICs.

#### 5. Over-fitting of the ICA model:

Over specification of the model order leads to the problem of over-fitting in ICA. As we describe below, this can lead to (1) the phenomenon of IC "splitting" and (2) an increase in the variance of the IC estimates.

##### 1. IC "splitting"

Suppose that the true model order or the number of non-Gaussian sources in an ICA decomposition of  $\mathbf{y}$  such as 2.1 is  $q$ . Then a fundamental result in [Rao, 1969, Theorem 1] states that for any other ICA decomposition of  $\mathbf{y}$ , the number of non-Gaussian sources remains the same while the number of Gaussian sources can change. In other words,  $\mathbf{y}$  cannot have two different ICA decompositions containing different number of non-Gaussian sources.

In view of this fact, how can a model order  $q$  ICA decomposition containing  $q$  non-Gaussian sources be "split" into a  $(q + 1)$  ICA decomposition containing  $(q + 1)$  non-Gaussian sources when performing ICA estimation using an assumed model order of  $(q + 1)$ ? As we describe below, the order  $(q + 1)$  ICA decomposition is only an *approximation* to the order  $q$  ICA decomposition.

Let  $\mathbf{a}_i$  be the  $i$ th column of  $\mathbf{A}$  in 2.1. In the presence of noise, it might be possible to *approximate*:

$$\mathbf{a}_i s_i \approx \mathbf{a}_i^1 s_i^1 + \mathbf{a}_i^2 s_i^2 \quad (2.32)$$

Here:

- $\mathbf{a}_i s_i$  is the contribution of the  $i$ th non-Gaussian source  $s_i$  to the ICA model 2.1.
- $s_i^1$  and  $s_i^2$  are independent non-Gaussian random variables that are also independent with respect to all non-Gaussian sources  $s_j, j \neq i$  in 2.1.
- $\mathbf{a}_i^1$  and  $\mathbf{a}_i^2$  are the basis time courses corresponding to  $s_i^1$  and  $s_i^2$  respectively.

- The time courses  $\mathbf{a}_i^1$  and  $\mathbf{a}_i^2$  look similar to each other.

Note that if  $\mathbf{a}_i^1 = \mathbf{a}_i^2$ , then 2.32 can be made into an equality by choosing  $s_i = s_i^1 + s_i^2$ . By replacing  $\mathbf{a}_i s_i$  in 2.1 using 2.32, we arrive at an *approximate* model order  $(q+1)$  decomposition of  $\mathbf{y}$ . In this decomposition, the component  $s_i$  from a model order  $q$  decomposition appears to be "split" into two sub-components:  $s_i^1$  and  $s_i^2$ .

## 2. Inflated variance of IC estimates

Overestimation of model order will lead to over-fitting of the mixing matrix  $\mathbf{A}$ . In other words,  $\mathbf{A}$  could have several columns that are highly correlated with each other. This could happen as a result of IC "splitting" as discussed above. Now, for a given realization  $\mathbf{s}$ , the variance of  $\hat{\mathbf{s}}$  is given by  $\text{Var}(\hat{\mathbf{s}}) = \sigma^2(\mathbf{A}^T \mathbf{A})^{-1}$  (for isotropic Gaussian co-variance). An increase in number of columns of  $\mathbf{A}$  and the fact that many of them are highly correlated implies that the variability of IC estimates  $\text{Var}(\hat{\mathbf{s}})$  is inflated.

In other words, running ICA multiple times on the same data or variations thereof with random initialization could produce different ICs.

## 2.2 ICA algorithms, single subject ICA and group ICA

In this section, we give a brief summary of how the ICA parameters are estimated in practice and also summarize the two most common modes of ICA application to fMRI data - single subject ICA (section 2.2.1) and temporal concatenation based group ICA (section 2.2.2).

Given several independent observations  $\mathbf{y}$  as per the noisy ICA model 2.1, most ICA algorithms estimate the ICA parameters  $\mathcal{F} = \{\boldsymbol{\mu}, \mathbf{A}, \mathbf{D}, \boldsymbol{\Sigma}\}$  and the realizations of  $\mathbf{s}$  in 2 steps. We only consider the case with  $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_p$ , since as shown in section 2.1.2, the mixing matrix  $\mathbf{A}$  and source distributions of  $\mathbf{s}$  are identifiable upto permutation and scaling for this case.

1. First, the diagonal source co-variance is arbitrarily set as  $\mathbf{D} = \mathbf{I}_q$ . The mean vector  $\boldsymbol{\mu}$  is estimated as  $\mathbf{E}(\mathbf{y})$ . Then, using PCA or PPCA [Tipping, 1999], the mixing matrix  $\mathbf{A}$  is estimated, upto an orthogonal rotation matrix  $\mathbf{O}$ , to be in a signal subspace which is spanned by the principal eigenvectors corresponding to the largest eigenvalues of the data co-variance matrix  $\mathbf{E}[(\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})^T]$ . The noise variance  $\sigma^2$  is estimated in this step as well.
2. Next, an estimator  $\hat{\mathbf{s}}$  for the source realizations is defined using techniques such as least squares or MMSE. The only unknown involved in these estimates is the orthogonal rotation matrix  $\mathbf{O}$ .
3. Finally, the non-Gaussianity of the empirical density of components of  $\hat{\mathbf{s}}$  is optimized with respect to  $\mathbf{O}$  using algorithms such as fixed point ICA Hyvarinen [1998, 1999].

For more details on noisy ICA estimation, please see [Beckmann and Smith, 2004] and for more details on ICA algorithms, please see [Hyvarinen et al., 2001].

### 2.2.1 Single subject ICA

How is ICA applied to single subject fMRI data? Suppose we are given a single subject fMRI scan which we rearrange as a  $p \times n$  2D matrix  $\mathbf{Y}$  in which column  $i$  is the  $p \times 1$  observed time-course  $\mathbf{y}_i$  in the brain at voxel  $i$ . Observed time-courses  $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$  are considered to be  $n$  independent realizations of  $\mathbf{y}$  as per the linear ICA model 2.1. Suppose  $\hat{\mathbf{S}} = [\hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2, \dots, \hat{\mathbf{s}}_n]$  is the  $q \times n$  matrix containing the estimated source realizations at the  $n$  voxels. The  $j$ th row of  $\hat{\mathbf{S}}$  is the  $j$ th IC. In other words, we decompose the time by space fMRI 2D matrix into a set of basis time-courses and a set of  $q$  3D IC maps using ICA.

### 2.2.2 Group ICA

How is ICA applied to data from a group of subjects in fMRI? Suppose we collect fMRI images from  $m$  subjects. First, we register all subjects to a common space using a registration algorithm (e.g., affine registration). Next, we rearrange each of the fMRI scans into  $m$  2D matrices  $\mathbf{Y}_1 \dots \mathbf{Y}_m$ , each of size  $p \times n$ . Column  $j$  in  $\mathbf{Y}_i$  is the demeaned time-course observed at voxel location  $j$  for subject  $i$ . The matrices  $\mathbf{Y}_1 \dots \mathbf{Y}_m$  are temporally concatenated to get a  $pm \times n$  matrix  $\mathbf{Z}$  as follows:

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_i \\ \vdots \\ \mathbf{Y}_m \end{pmatrix} \quad (2.33)$$

Column  $i$  of  $\mathbf{Z}$  is the  $pm \times 1$  vector  $\mathbf{z}_i$  which is assumed to follow a linear ICA model 2.1.  $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n$  are considered to be independent realizations of the model 2.1. Suppose  $\hat{\mathbf{S}}_G = [\hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2, \dots, \hat{\mathbf{s}}_n]$  is a  $q \times n$  matrix containing the estimated source realizations at the  $n$  voxels. The  $j$ th row of  $\hat{\mathbf{S}}_G$  is the  $j$ th group IC. In group ICA, the joined time-series across subjects is modeled using noisy linear ICA. In practice,  $\mathbf{Y}_i$  is the PCA reduced data set for subject  $i$ . The PCA reduction is either done separately for each subject using subject specific data co-variance [Calhoun et al., 2001] or an average data co-variance across subjects [Beckmann et al., 2005]. The average co-variance approach requires each subject to have the same number of time points in fMRI scans.

## 2.3 The original RAICAR algorithm

In this section, we give a brief introduction to the RAICAR algorithm of [Yang et al., 2008]. Suppose we are given a data set which we decompose into  $n_C$  ICs using ICA (e.g., single subject or group ICA). Our goal is to assess which ICs consistently show up in multiple ICA runs i.e.,

the reproducibility of each of these  $n_C$  ICs. To that extent, we run the ICA algorithm  $K$  times. Suppose  $\mathbf{x}_j^{(m)}$  is the  $n \times 1$  vector (e.g. spatial ICA map re-arranged into a vector) of the  $j$ th IC from  $m$ th ICA run. Suppose  $\mathbf{G}_{lm}$  is a  $n_C \times n_C$  absolute spatial cross-correlation coefficient matrix between the ICs from runs  $l$  and  $m$ :

$$\mathbf{G}_{lm}(i, j) = |\text{corrcoef}(\mathbf{x}_i^{(l)}, \mathbf{x}_j^{(m)})| \quad (2.34)$$

where  $|\cdot|$  denotes absolute value.  $\mathbf{G}_{lm}(i, j)$  is the absolute spatial cross-correlation coefficient between IC  $i$  from run  $l$  and IC  $j$  from run  $m$ . The matrices  $\mathbf{G}_{lm}$  are then arranged as elements of a  $K \times K$  *block-matrix*  $\mathbf{G}$  such that the  $l$ th row and  $m$ th column of  $\mathbf{G}$  is  $\mathbf{G}(l, m) = \mathbf{G}_{lm}$  (see Figure 2). This block matrix  $\mathbf{G}$  is the starting point for a RAICAR across-run component matching process.

Since ICs within a particular run cannot be matched to each other, the  $n_C \times n_C$  matrices  $\mathbf{G}(l, l)$ ,  $l = 1 \dots K$  along the block-diagonal of  $\mathbf{G}$  are set to  $\mathbf{0}$  as shown in Figure 2 with a gray color. The following steps are involved in a RAICAR analysis:

1. Find the maximal element of  $\mathbf{G}$ . Suppose this maximum occurs in matrix  $\mathbf{G}_{lm}$  at position  $(i, j)$ . Hence component  $i$  from run  $l$  matches component  $j$  from run  $m$ . Let us label this matched component by  $MC_1$  (the first matched component).
2. Next, we attempt to find from each run  $s$  ( $s \neq l$  and  $s \neq m$ ) a component that matches with component  $MC_1$ . Suppose element  $(a_s, j)$  is the maximal element in the  $j$ th column of  $\mathbf{G}_{sm}$ . Then component  $a_s$  is the best matching component from run  $s$  with the  $j$ th component from run  $m$ .

Similarly, suppose element  $(i, b_s)$  is the maximal element in the  $i$ th row of  $\mathbf{G}_{ls}$ . Then component  $b_s$  is the best matching component from run  $s$  with component  $i$  from run  $l$ . As noted in [Yang et al., 2008], in most cases  $a_s = b_s$ . However, it is possible that  $a_s \neq b_s$ . Hence the component number  $e_s$  matching  $MC_1$  from run  $s$  is defined as follows:

$$e_s = \begin{cases} a_s & \text{if } \mathbf{G}_{sm}(a_s, j) \geq \mathbf{G}_{ls}(i, b_s), \\ b_s & \text{if } \mathbf{G}_{sm}(a_s, j) < \mathbf{G}_{ls}(i, b_s). \end{cases} \quad (2.35)$$

We would also like to remove component  $e_s$  of run  $s$  from further consideration during the matching process. To that extent, we zero out the  $e_s$ th row from  $\mathbf{G}_{sr}$ ,  $r = 1 \dots K$  and the  $e_s$ th column from  $\mathbf{G}_{rs}$ ,  $r = 1 \dots K$ .

3. Once a matching component  $e_s$  has been found for all runs  $s \neq l, m$ , we also zero out the  $i$ th row from  $\mathbf{G}_{lr}$ ,  $r = 1 \dots K$  and the  $i$ th column from  $\mathbf{G}_{rl}$ ,  $r = 1 \dots K$ . Similarly, we zero out the  $j$ th column from  $\mathbf{G}_{rm}$ ,  $r = 1 \dots K$  and the  $j$ th row from  $\mathbf{G}_{mr}$ ,  $r = 1 \dots K$ . This eliminates component  $i$  from run  $l$  and component  $j$  from run  $m$  from further consideration during the matching process.

- 
4. Steps 1-3 complete the matching process for one IC component across runs. These steps are repeated until  $n_C$  components are matched across the  $K$  runs. We label the matched component  $s$  as  $MC_s$  which contains a set of  $K$  matching ICs one from each of the  $K$  ICA runs.

Suppose matched component  $s$ ,  $MC_s$  consists of the matched ICs  $\mathbf{x}_{i_1}^{(1)}, \mathbf{x}_{i_2}^{(2)}, \dots, \mathbf{x}_{i_K}^{(K)}$ . Form the  $K \times K$  cross-correlation matrix  $H_{MC_s}$  between the matched components in  $MC_s$ . The  $(a, b)$ th element of this matrix is simply:

$$H_{MC_s}(a, b) = |\text{corrcoef}(\mathbf{x}_{i_a}^{(a)}, \mathbf{x}_{i_b}^{(b)})| \quad (2.36)$$

The normalized reproducibility of  $MC_s$  is then defined as:

$$\text{Reproducibility}(MC_s) = \left( \frac{2}{(K-1)K} \right) \sum_{a=1}^K \sum_{b=a+1}^K H_{MC_s}(a, b) \quad (2.37)$$

The double sum in 2.37 is simply the sum of the upper triangular part of  $H_{MC_s}$  excluding the diagonal. The normalizing factor  $\frac{(K-1)K}{2}$  is simply the maximum possible value of this sum. Hence the normalized reproducibility satisfies:  $\text{Reproducibility}(MC_s) \leq 1$ .

Note that our definition of normalized reproducibility is slightly different from that in Yang et al. [2008]. Whereas Yang et al. [2008] averages the *thresholded* absolute correlation coefficients, we simply average the *un-thresholded* absolute correlation coefficients to compute reproducibility thereby avoiding the selection of a threshold on the absolute correlation coefficients.

## 2.4 The RAICAR-N enhancement

In this section, we describe how to compute reproducibility  $p$ -values for each matched component in RAICAR. Note that the RAICAR "component matching" process can be used to assess the reproducibility of *any* spatial component maps - not necessarily ICA maps. For instance, RAICAR can be used to assess the reproducibility of a set of PCA maps across subjects.

In order to generate reproducibility  $p$ -values for the matched component maps:

1. We need to determine the distribution of normalized reproducibility that we get from the RAICAR "component matching" process when the input to RAICAR represents a set of "non-reproducible component maps" across the  $K$  runs.
2. In addition, we would also like to preserve the *overall structure* seen in the *observed* sets of spatial component maps across the  $K$  runs when generating sets of "non-reproducible component maps" across the  $K$  runs.

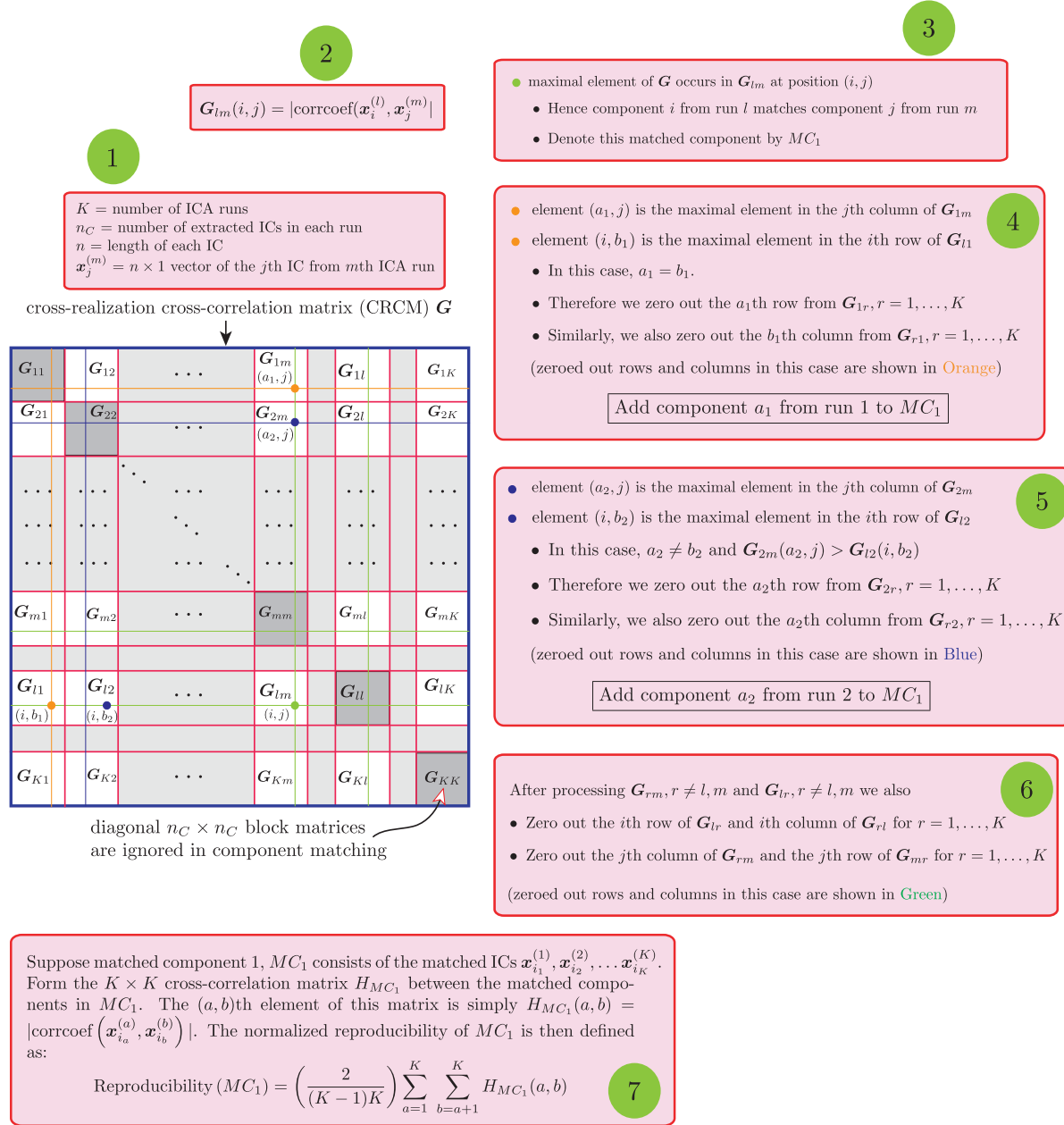


Figure 2: Pictorial depiction of the original RAICAR algorithm [Yang et al., 2008]. The ICA algorithm is run  $K$  times with each run producing  $n_C$  ICs.  $\mathbf{G}$  is a  $K \times K$  block matrix with elements  $\mathbf{G}(l, m) = \mathbf{G}_{lm}$  where  $\mathbf{G}_{lm}$  is the  $n_C \times n_C$  absolute spatial cross-correlation matrix between ICs from runs  $l$  and  $m$ . The numbered green circles indicate the sequence of steps in applying RAICAR to a given data set. Our definition of normalized reproducibility in box 7 averages un-thresholded correlation coefficients thereby avoiding the selection of a correlation coefficient threshold prior to averaging.



Hence for IC reproducibility assessment, we propose to use the original set of ICs across the  $K$  runs to generate the "non-reproducible component maps" across the  $K$  runs.

Suppose  $K$  ICA runs are submitted to RAICAR which gives us a  $n_C \times 1$  vector of *observed* normalized reproducibility values  $\text{Reproducibility}(MC_i), i = 1 \dots n_C$  - one for each IC. We propose to attach  $p$ -values for measuring the reproducibility of each IC in a data-driven fashion as follows:

1. First, we label the  $Kn_C$  ICs across the  $K$  runs using unique integers. In run 1, the ICs are labelled using integers  $1, \dots, n_C$ . In run 2, the ICs are labelled using integers  $(n_C+1), \dots, 2n_C$  and so on. In run  $K$ , the ICs are labelled using integers  $(K-1)n_C + 1, \dots, Kn_C$ .
2. Our "null" hypothesis is:

$$\mathbf{H}_0 : \text{None of the ICs are reproducible} \quad (2.38)$$

Hence, we can randomly label component  $i$  from run  $l$  as component  $d$  from run  $s$

To do this, we randomly permute the integers  $1, 2, \dots, Kn_C$  to get the permuted integers  $p(1), p(2), \dots, p(Kn_C)$ . Obviously  $p(i) \neq p(j)$  if  $i \neq j$ .

3. The  $K$  sets "non-reproducible component runs under  $\mathbf{H}_0$ " are constructed by assigning components with labels:
  - $p(1), \dots, p(n_C)$  to run 1 under  $\mathbf{H}_0$ .
  - $p(n_C + 1), \dots, p(2n_C)$  to run 2 under  $\mathbf{H}_0$
  - $p((K-1)n_C + 1), \dots, p(Kn_C)$  to run  $K$  under  $\mathbf{H}_0$
4. After  $K$  runs have been generated under  $\mathbf{H}_0$ , we subject these to a RAICAR analysis. This gives us  $n_C$  values of normalized reproducibility, one for each matched component under  $\mathbf{H}_0$ .
5. Steps 1-4 are repeated  $R$  times to build up a *pooled*  $Rn_C \times 1$  vector of normalized reproducibility  $\mathbf{Reproducibility}_{Null}$  under  $\mathbf{H}_0$ .
6. Finally, we assign a  $p$ -value for reproducibility to each matched IC across the  $K$  runs. The observed reproducibility for  $i$ th matched IC is  $\text{Reproducibility}(MC_i)$  and its  $p$ -value is:

$$\text{Reproducibility}_{pval}(MC_i) = \frac{\{\text{no. of } \mathbf{Reproducibility}_{Null} \geq \text{Reproducibility}(MC_i)\} + 1}{Rn_C + 1} \quad (2.39)$$

7. Only those components with  $\text{Reproducibility}_{pval}(MC_i) < p_{crit}$  are considered to be significantly reproducible. We can use a fixed and objective value for  $p_{crit}$  such as 0.05. Note that this fixed cutoff is independent of the amount of variability in the input to RAICAR-N. Please see Figure 3 for a pictorial depiction of this process.

8

Once we calculate Reproducibility ( $MC_1$ ), Reproducibility ( $MC_2$ ), ..., Reproducibility ( $MC_{n_C}$ ), how can we calculate a cutoff to separate the high reproducibility ICs from the low reproducibility ICs?

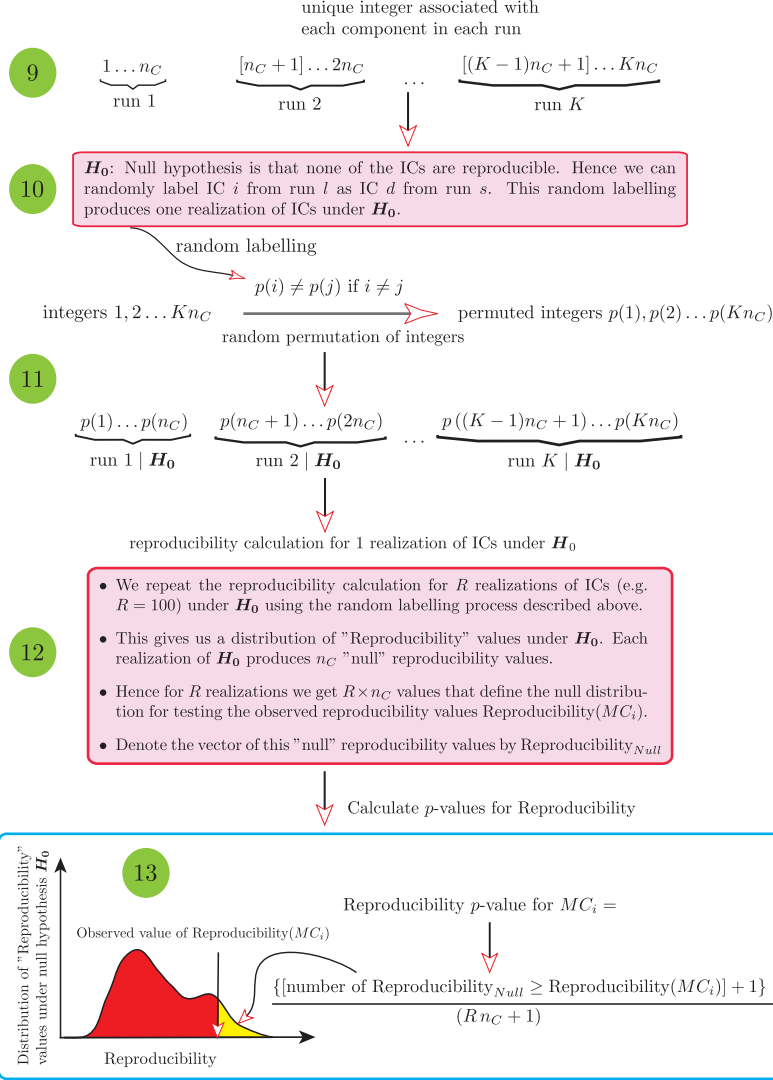


Figure 3: Pictorial depiction of the process for generating a "null" distribution in RAICAR-N. Our "null" hypothesis is: " $H_0$ : None of the ICs are reproducible. Hence, we can randomly label IC  $i$  from run  $l$  as IC  $d$  from run  $s$ ". Therefore we randomly split the  $Kn_C$  ICs across  $K$  runs into  $K$  parts and run the RAICAR algorithm on each set of randomly split ICs. This gives us a set of "null" reproducibility values which can be used to compute  $p$ -values for the observed reproducibility of ICs in the original RAICAR run. The green circles indicate the sequence of steps for generating the "null" distribution after the steps in Figure 2.

---

## 2.5 How many subjects should be used per group ICA run in RAICAR-N?

The input to RAICAR-N can either be single subject ICA runs or group ICA runs across a set of subjects. Note that the individual subject ICA runs are spatially unconstrained whereas a group ICA spatially constrains the group ICs across a set of subjects. Hence the number of ICs that can be declared as significantly reproducible at the group level are usually more than those that can be declared significantly reproducible at the single subject level. Hence the following question is relevant:

Suppose we have a group of  $N$  subjects. We randomly select  $L$  subjects and form a single group of subjects. We repeat this process  $K$  times to get  $K$  groups of  $L$  subjects each of which is subjected to a group ICA analysis. Given the number of subjects  $N$ , how should we choose  $L$  and  $K$ ?

First, we discuss the choice of  $L$ . If  $L = N$  then each of the  $K$  groups will contain the same  $N$  subjects and hence there will be no diversity in the  $K$  groups. We would like to control the amount of diversity in the  $K$  groups of  $L$  subjects. Consider any 2 subjects  $X$  and  $Y$ . The probability  $P_{XY}(L)$  that both  $X$  and  $Y$  appear in a set of  $L$  randomly chosen subjects from  $N$  subjects is given by:

$$P_{XY}(L) = \frac{\binom{N-2}{L-2}}{\binom{N}{L}} \quad (2.40)$$

The expected number of times that  $X$  and  $Y$  appear together in sets of  $L$  subjects out of  $K$  independently drawn sets is:

$$E_{XY}(L) = K P_{XY}(L) \quad (2.41)$$

Ideally, we would like  $E_{XY}(L)$  to be only a small fraction of  $K$ . Hence we impose the restriction:

$$E_{XY}(L) = K P_{XY}(L) \leq \alpha_{max} K \quad (2.42)$$

where  $\alpha_{max}$  is a user defined constant such as  $\alpha_{max} = 0.05$ . This implies that the chosen value of  $L$  must satisfy:

$$P_{XY}(L) \leq \alpha_{max} \quad (2.43)$$

In practice, we choose the largest value of  $L$  that satisfies this inequality. As shown in Figure 5, if  $N = 23$  and  $\alpha_{max} = 0.05$  then the largest value of  $L$  that satisfies 2.43 is  $L = 5$ .

The number of group ICA runs  $K$  should be as large as possible. From our experiments on real fMRI data we can roughly say that values of  $K > 50$  give equivalent results.

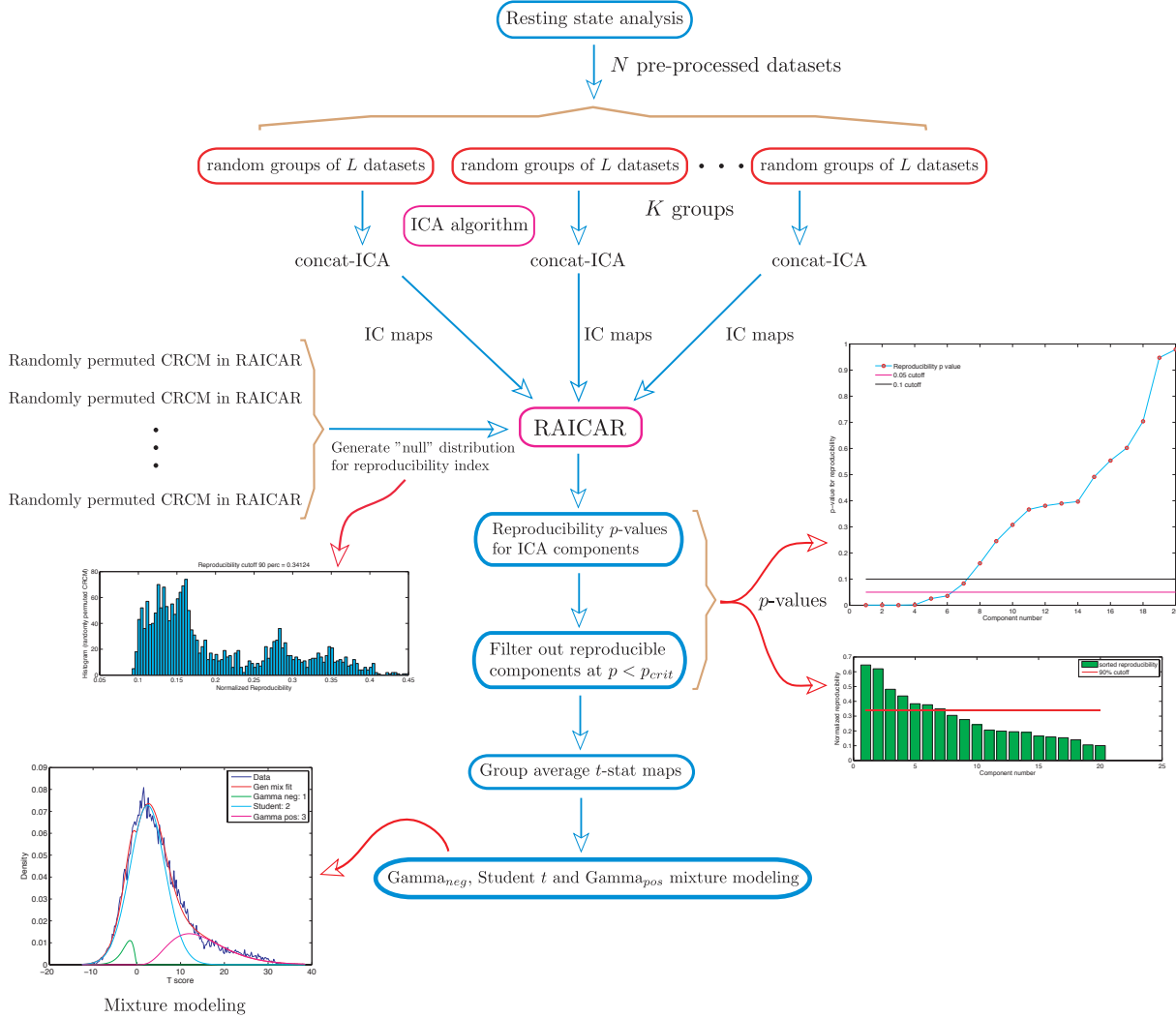


Figure 4: Flowchart for a group ICA based RAICAR-N analysis. The  $N$  single subject data sets are first pre-processed and subsequently bootstrapped to create  $K$  groups, each group containing  $L$  distinct subjects. Each group of  $L$  subjects is submitted to a temporal concatenation group ICA analysis. The resulting IC maps (either raw ICs or ICs scaled by noise standard deviation) are subjected to a RAICAR analysis. The cross-realization cross correlation matrix (CRCM) is randomly permuted multiple times:  $\mathbf{G} \rightarrow \mathbf{G}(\mathbf{g}, \mathbf{g})$  where  $\mathbf{g}$  is a random permutation of integers from  $1, \dots, Kn_C$ . The permuted CRCMs are subjected to a RAICAR analysis to generate a realization of reproducibility values under the "null" hypothesis. The computed "null" distribution of reproducibility values is used to assign  $p$  values to the observed reproducibility of the original RAICAR run. Finally, reproducible ICs are averaged using a random effects analysis and the resulting  $t$ -statistic images are subjected to Gamma<sub>neg</sub>, Student  $t$  and Gamma<sub>pos</sub> mixture modeling.

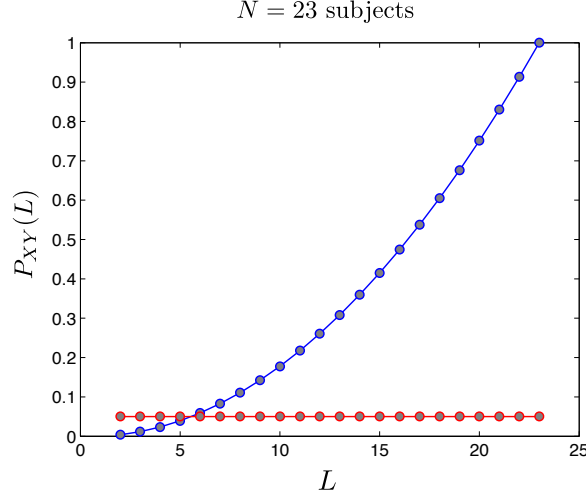


Figure 5: Figure shows a plot of  $P_{XY}(L)$  vs  $L$  for  $N = 23$  in blue. The red line shows the  $\alpha_{max} = 0.05$  cutoff. The largest value of  $L$  for which  $P_{XY}(L) \leq 0.05$  is  $L = 5$ .

## 2.6 How to display the estimated non-Gaussian spatial structure in ICA maps?

The ICs have been optimized for non-Gaussianity. However, there can be many types of non-Gaussian distributions. It has been empirically found that the non-Gaussian distributions of ICs found in fMRI data have the following structure:

1. A central Gaussian looking part and
2. A tail that extends out on either end of the Gaussian

It has been suggested in [Beckmann and Smith, 2004] that a Gaussian/Gamma mixture model can be fitted to this distribution and the Gamma components can be thought of as representatives of the non-Gaussian structure. We follow a similar approach:

1. The output of a RAICAR-N analysis is a set of spatial ICA maps (either  $z$ -transformed maps or raw maps) concatenated into a 4-D volume.
2. We do a voxelwise transformation to Normality using the voxelwise empirical cumulative distribution function as described in [van Albada and Robinson, 2007].
3. Next, we submit the resulting 4-D volume to a voxelwise group analysis using ordinary least squares. The design matrix for group analysis depends on the question being considered. In our case, the design matrix was simply a single group average design.
4. The resulting  $t$ -statistic maps are subjected to Student  $t$ ,  $\text{Gamma}_{pos}$  and  $\text{Gamma}_{neg}$  mixture modeling. The logic is that if the original ICA maps are pure Gaussian (i.e., have no interest-

ing non-Gaussian structure) then the result of a group average analysis will be a pure Student  $t$  map which will be captured by a single Student  $t$  (i.e., the  $\text{Gamma}_{pos}$  and  $\text{Gamma}_{neg}$  will be driven to 0 class fractions). Hence the "null" hypothesis will be correctly accounted for.

5. If the Gamma distributions have  $> 0.5$  posterior probability at some voxels then those voxels are displayed in color to indicate the presence of significant non-Gaussian structure over and above the background Student  $t$  distribution.

Examples of Student  $t$ ,  $\text{Gamma}_{pos}$  and  $\text{Gamma}_{neg}$  mixture model fits are shown in Figure 6.

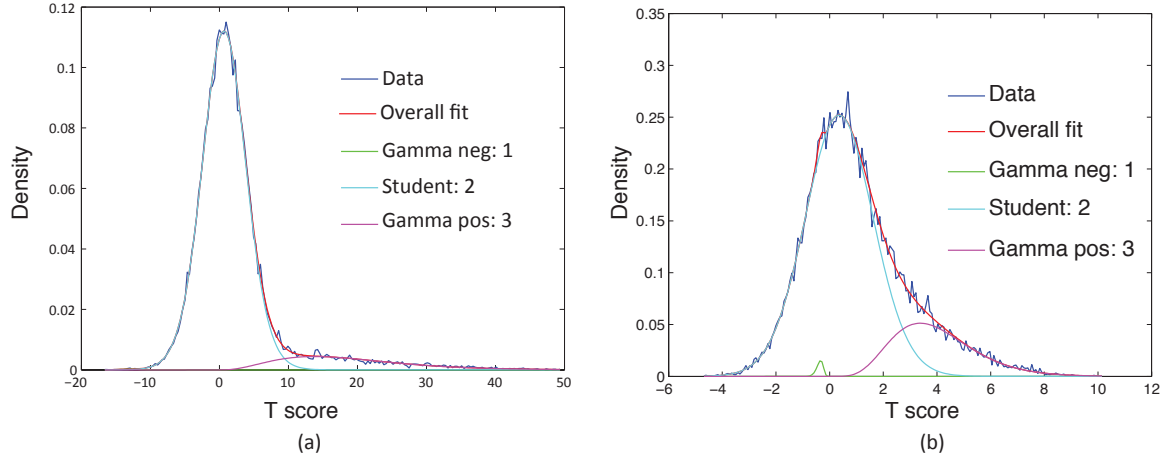


Figure 6: Examples of displaying non-Gaussian spatial structure using a Student  $t$ ,  $\text{Gamma}_{pos}$  and  $\text{Gamma}_{neg}$  mixture model. Notice how the  $\text{Gamma}_{neg}$  density is driven to near 0 class fraction in the absence of significant negative non-Gaussian structure.

### 3 Experiments and Results

#### 3.1 Human rsfMRI data

rsfMRI data titled: Baltimore (Pekar, J.J./Mostofsky, S.H.;  $n = 23$  [8M/15F]; ages: 20-40; TR = 2.5; # slices = 47; # timepoints = 123), a part of the 1000 functional connectomes project, was downloaded from the Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC): [http://www.nitrc.org/projects/fcon\\_1000/](http://www.nitrc.org/projects/fcon_1000/).

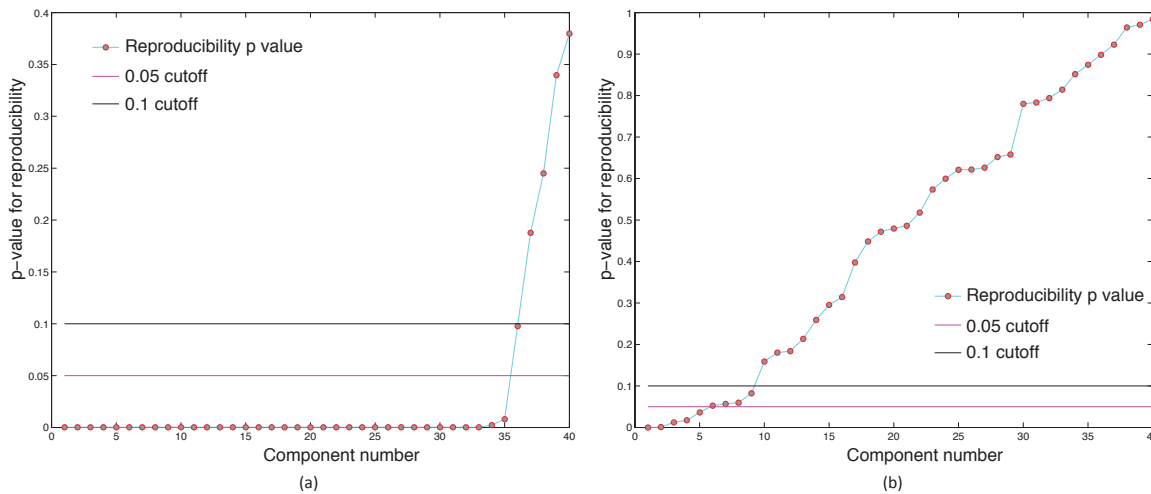


Figure 7:  $p$ -value cutoffs for within and across single subject analysis using RAICAR-N. This figure illustrates the intuitive fact that within subject ICA runs are much more reproducible compared to across subject ICA runs.

## 3.2 Preprocessing

Data was analyzed using tools from the FMRIB software library (FSL: <http://www.fmrib.ox.ac.uk/fsl/>). Preprocessing steps included motion correction, brain extraction, spatial smoothing with an isotropic Gaussian kernel of 5mm FWHM and 100s high-pass temporal filtering. Spatial ICA was performed using a noisy ICA model as implemented in FSL MELODIC [Beckmann and Smith, 2004] in either single subject or multi-subject temporal concatenation mode also called group ICA. Please see section 2.2 for a brief summary of single subject ICA and group ICA. In each case, we fixed the model order of ICA at  $q = 40$  to be consistent with the model order range typically extracted in rsfMRI and fMRI [Smith et al., 2009, Esposito et al., 2005]. For temporal concatenation based group ICA, single subject data was first affinely registered to the MNI 152 brain and subsequently resampled to 4x4x4 resolution (MNI 4x4x4) to decrease computational load.

## 3.3 RAICAR-N analysis with 1 ICA run per subject

Spatial ICA was run once for each of the  $N = 23$  subjects in their native space. The resulting set of ICA components across subjects were transformed to MNI 4x4x4 space and were submitted to a RAICAR-N analysis.<sup>1</sup> ICA components were sorted according to their reproducibility and  $p$ -values

<sup>1</sup>In all RAICAR-N analyses reported in this article, we used the  $z$ -transformed IC maps - which are basically the raw IC maps divided by a voxelwise estimate of noise standard deviation (named as `melodic_IC.nii.gz` in

were computed for each ICA component. Please see Figure 8.

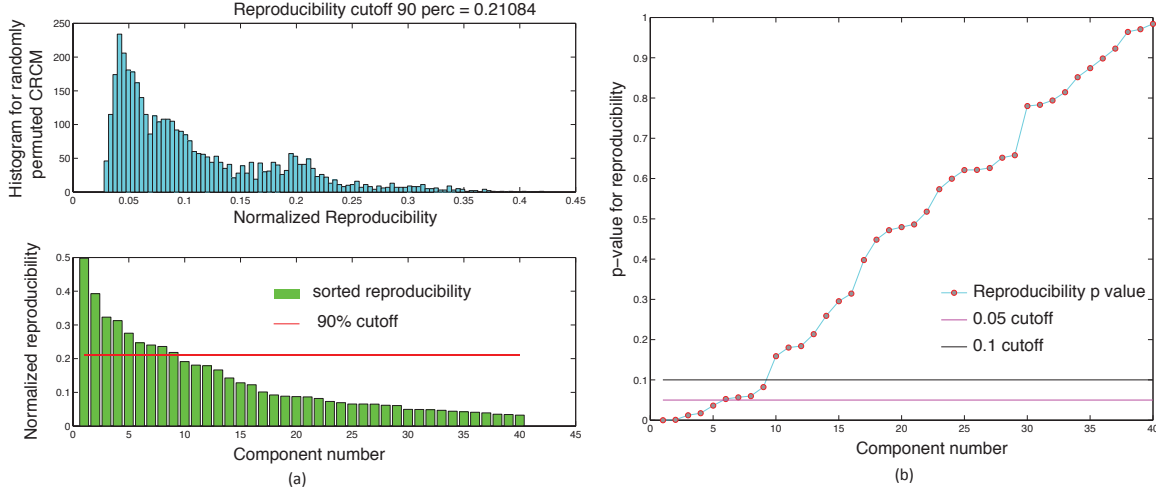


Figure 8: Single subject rsfMRI ICA runs across 23 subjects were combined using a RAICAR-N analysis. Figure (a) shows the observed values of normalized reproducibility (bottom) as well as the "null" distribution of normalized reproducibility across  $R = 100$  simulations (top). Figure (b) shows the  $p$ -values for each IC along with the 0.05 and 0.1 cutoff lines.

We compared the reproducible RSNs from the single subject RAICAR-N analysis to the group RSN maps reported in literature [Beckmann et al., 2005]. Please see Figure 9.

To summarize, when single subject ICA runs are combined across subjects:

- We are able to declare 4 "standard" RSNs as significantly reproducible at a  $p$ -value  $< 0.05$ .
- There are 2 other "standard" RSNs that achieve a reproducibility  $p$ -value between 0.05 and 0.06.
- There are 2 other "non-standard" RSNs that are of interest: one achieves a  $p$ -value of 0.0125 and the other achieves a  $p$ -value of 0.05699.

### 3.4 RAICAR-N on random sets of 5 subjects - 50 group ICA runs

To promote diversity across the group ICA runs, as discussed in section 2.5,  $L = 5$  subjects were drawn at random from the group of  $N = 23$  subjects and submitted to a temporal concatenation based group ICA. This process was repeated  $K = 50$  times and the resulting set of 50 group ICA maps were submitted to a RAICAR-N analysis. ICA components were sorted according to their reproducibility and  $p$ -values were computed for each ICA component. Please see Figure 10.

MELODIC). It is also possible to use the raw IC maps as inputs to RAICAR-N.



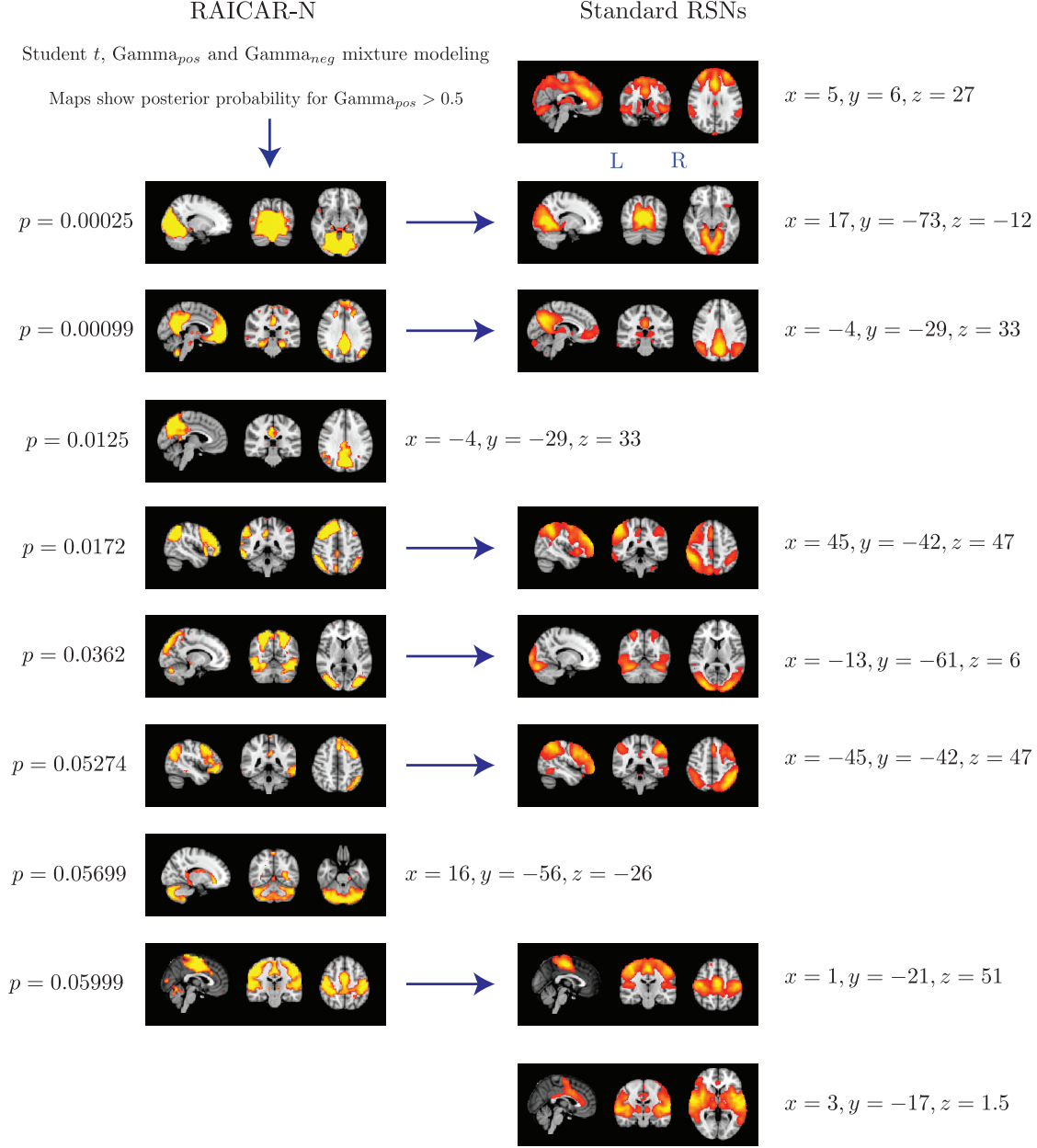


Figure 9: The top 8 "reproducible" ICs from a RAICAR-N analysis on single subject ICA runs compared with standard RSN maps reported in literature [Beckmann et al., 2005]. We are able to declare 4 "standard" RSNs as significantly reproducible at a  $p$ -value  $< 0.05$ . There are 2 other "standard" RSNs that achieve a reproducibility  $p$ -value between 0.05 and 0.06 as well as 2 "non-standard" RSNs that achieve  $p$ -values of 0.0125 and 0.05699 respectively. We also could not find 2 of the published RSNs in [Beckmann et al., 2005] as reproducible in single subject ICA runs.

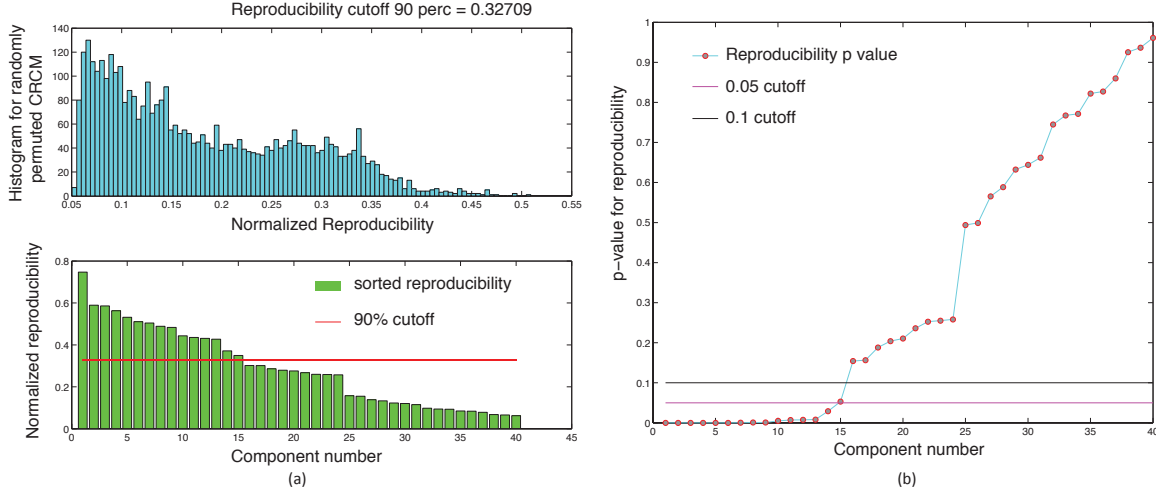


Figure 10:  $L = 5$  subjects were randomly drawn from the set of  $N = 23$  subjects and submitted to a temporal concatenation based group ICA. This process was repeated  $K = 50$  times and the resulting ICA maps were submitted to a RAICAR-N analysis. Figure (a) shows the observed values of normalized reproducibility (bottom) as well as the "null" distribution of normalized reproducibility across  $R = 100$  simulations (top). Figure (b) shows the  $p$ -values for each IC along with the 0.05 and 0.1 cutoff lines.

We compared the reproducible RSNs from the single subject RAICAR-N analysis to the RSN maps reported in literature [Beckmann et al., 2005]. Please see Figure 11.

In summary, when 50 random 5 subject group ICA runs (from a population of 23 subjects) are combined using RAICAR-N:

- We are able to declare 8 "standard" RSNs as significantly reproducible at a  $p$ -value  $< 0.05$ .
- There are 6 other "non-standard" RSNs that can be declared as significantly reproducible at a  $p$ -value  $< 0.05$ .
- There is 1 other "non-standard" RSN that achieves a  $p$ -value of 0.05299.

### 3.5 RAICAR-N on random sets of 5 subjects - 100 group ICA runs

To promote diversity across the group ICA runs, as discussed in section 2.5,  $L = 5$  subjects were drawn at random from the group of  $N = 23$  subjects and submitted to a temporal concatenation based group ICA. This process was repeated  $K = 100$  times and the resulting set of 100 group ICA maps were submitted to a RAICAR-N analysis. ICA components were sorted according to their reproducibility and  $p$ -values were computed for each ICA component. Please see Figure 12.

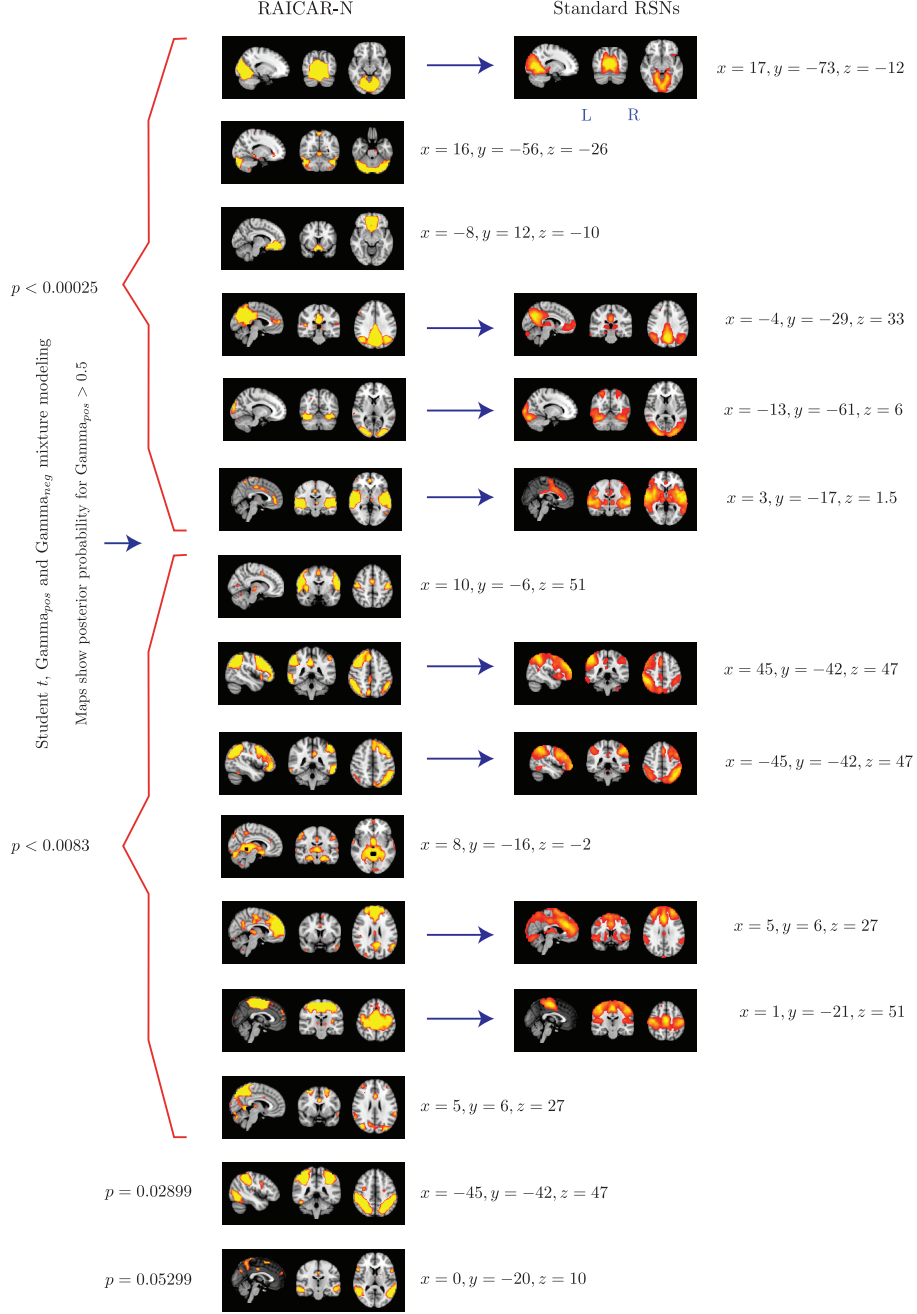


Figure 11: The top 15 "reproducible" ICs from  $K = 50$  runs of  $L = 5$  subject group ICA RAICAR-N analysis compared with standard RSN maps reported in literature [Beckmann et al., 2005]. We are able to declare 8 "standard" RSNs as significantly reproducible at a  $p$ -value of  $< 0.05$ . There are 6 other "non-standard" RSNs that can be declared as significantly reproducible at a  $p$ -value of  $< 0.05$  and 1 other "non-standard" RSN that achieves a  $p$ -value of 0.05299.

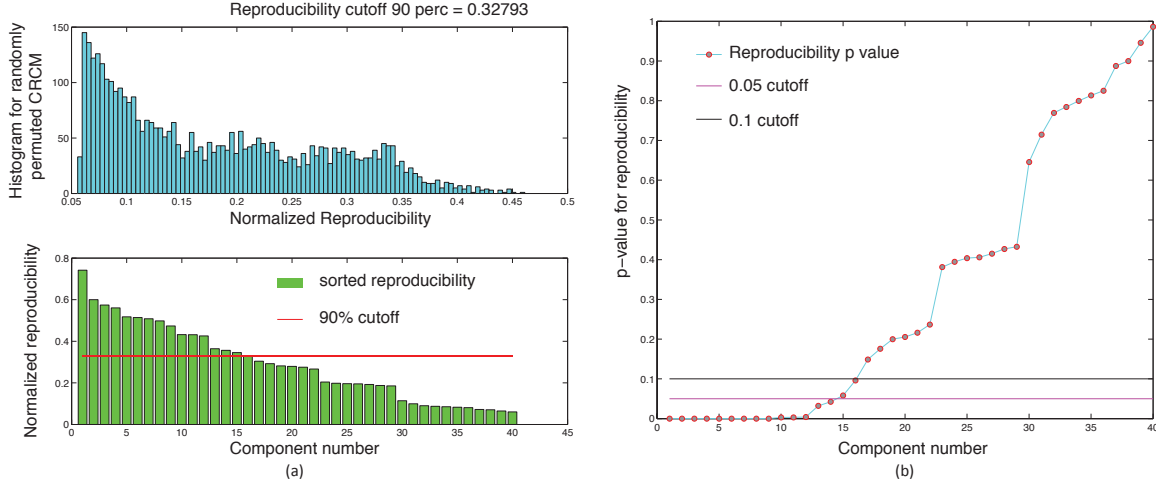


Figure 12:  $L = 5$  subjects were randomly drawn from the set of  $N = 23$  subjects and submitted to a temporal concatenation based group ICA. This process was repeated  $K = 100$  times and the resulting ICA maps were submitted to a RAICAR-N analysis. Figure (a) shows the observed values of normalized reproducibility (bottom) as well as the "null" distribution of normalized reproducibility across  $R = 100$  simulations (top). Figure (b) shows the  $p$ -values for each IC along with the 0.05 and 0.1 cutoff lines.

We compared the reproducible RSNs from the single subject RAICAR-N analysis to the RSN maps reported in literature [Beckmann et al., 2005]. Please see Figure 13.

In summary, when 100 random 5 subject group ICA runs (from a population of 23 subjects) are combined using RAICAR-N:

- We are able to declare 8 "standard" RSNs as significantly reproducible at a  $p$ -value  $< 0.05$ .
- There are 6 other "non-standard" RSNs that can be declared as significantly reproducible at a  $p$ -value  $< 0.05$ .
- There is 1 other "non-standard" RSN that achieves a  $p$ -value of 0.05824.



---

## 4 Group comparison of ICA results

In this section, we summarize the main approaches for group analysis of ICA results which can be broadly classified into two categories: (1) Approaches based on a single ICA run or no ICA run and (2) Approaches based on multiple ICA runs. To make things concrete, suppose we have two groups of subjects  $A$  and  $B$ .

### 4.1 Approaches based on a single group ICA run or no ICA run

The main idea in these approaches is to use the results of a group ICA using all subjects to derive subject specific spatial maps for group comparison. A typical sequence of steps is as follows:

1. The first step involves extraction of a set of template IC maps or a set of template mixing matrix time courses. This can be accomplished using two techniques:
  - (a) **Group ICA based template IC maps or time courses:**  
A temporal concatenation based group ICA is run using data from all subjects in group  $A$  and  $B$ . This usually involves two PCA data reductions. The first reduction is based on a subject wise PCA decomposition [Calhoun et al., 2001] or an average PCA decomposition [Beckmann et al., 2005] as discussed in section 2.2.2. The next reduction is based on PCA reduced temporally concatenated data. Subsequently, the group ICs and the dual PCA reduced mixing matrix time courses are estimated using an ICA algorithm.
  - (b) **User supplied set of template IC maps:**  
The user supplies a set of spatial maps, perhaps corresponding to an ICA decomposition on an independent data set.
2. The next step either uses template IC maps or time courses.
  - (a) **Template time course based approach:**  
First, the mixing matrix is PCA back projected and partitioned into subject specific sub matrices. Next, subject specific spatial maps corresponding to the group ICs are estimated via least-squares and a second PCA back projection is used to estimate the corresponding subject specific time courses. This is the approach proposed in [Calhoun et al., 2001], which we will refer to as the group ICA back projection approach.
  - (b) **Template IC based approach:**  
First, spatial multiple regression using the template ICs as regressors is used against the original data of each subject to derive subject specific time courses corresponding to each template IC. Next, a second multiple regression using the subject specific time courses is used against the original data of each subject to derive subject specific spatial maps corresponding to each template IC. This approach called "dual-regression" has been proposed by [Beckmann et al., 2009]. A similar approach called fixed average spatial

---

ICA (FAS ICA) had also been proposed earlier in [Calhoun et al., 2004]. Both dual-regression and FAS ICA involve the first spatial regression stage, but dual-regression also includes a second temporal regression stage.

3. Once subject specific spatial maps and time courses corresponding to group ICs have been determined, they are entered into a random effects analysis for group comparison.

#### 4.1.1 Advantages of single group ICA based approaches

1. Much reduced computational load compared to multiple ICA based approaches.
2. Ability to take advantage of constrained spatial IC estimation across all subjects via group ICA.

Please see section 5 for discussion.

## 4.2 Approaches based on multiple single subject or group ICA runs

In these approaches results of multiple ICA runs in groups  $A$  and  $B$  are used for a between group analysis. A typical sequence of steps is as follows:

1. The first step involves:
  - running a separate single subject ICA for all subjects from groups  $A$  and  $B$  (possibly with multiple runs per subject) or
  - running a set of group ICA runs across various sets of subjects separately, with each set containing subjects either from group  $A$  or group  $B$
2. The next step is to establish a correspondence between the ICs within and across groups. There are two main techniques of establishing this correspondence:

(a) **Template based methods:**

In these approaches, the user defines a template or a spatial map containing the network of interest. Examples of templates include a spatial map of the default mode network (DMN) derived from a separate ICA analysis, a spatial map from a separate PCA analysis, or even a binary mask defining the regions of interest. The template is then used to select from each run of ICA (single subject or group ICA) in each group ( $A$  and  $B$ ), an IC that best matches the template using a predefined metric such as spatial correlation coefficient or goodness of fit (GOF) [Greicius et al., 2004].

(b) **Template free methods:**

These approaches do not need a pre-defined template from the user, but instead attempt to match or cluster all ICs simultaneously within and across groups. Examples

---

of such approaches include self organizing group ICA (sogICA, [Esposito et al., 2005]) and RAICAR [Yang et al., 2008]. Each matched component or IC cluster includes one IC from each ICA run (single subject or group ICA) in each group ( $A$  and  $B$ ).

3. Finally, the selected ICs in template based methods or ICs from a selected IC cluster/matched component in template free methods are then entered into a random effects group analysis (with repeated measures for multiple single subject ICA runs) for between group comparison.

#### 4.2.1 Advantages of multiple ICA run approaches

1. They account for both algorithmic and data set variability of ICA.
2. Group comparisons happen on true ICs i.e., optimal solutions for the ICA problem.

Please see section 5 for discussion.

## 5 Discussion

As discussed in section 2.1.2, in the noisy linear ICA model with isotropic diagonal Gaussian noise co-variance, for a given true model order, the mixing matrix and the source distributions are identifiable upto permutation and scaling. However, as pointed out in section 2.1.3, various factors prevent the convergence of ICA algorithms to unique IC estimates. These factors include ICA model not being the true data generating model, approximations to mutual information used in ICA algorithms, multiple local minima in ICA contrast functions, confounding Gaussian noise as well as variability due to model order over-estimation. A practical implication of these factors is that ICA algorithms converge to different IC estimates depending on how they are initialized and on the specific data used as input to ICA. Hence, there is a need for a rigorous assessment of reproducibility or generalizability of IC estimates. A set of reproducible ICs can then be used as ICA based characteristics of a particular group of subjects.

We proposed an extension to the original RAICAR algorithm for reproducibility assessment of ICs within or across subjects. The modified algorithm called RAICAR-N builds up a "null" distribution of normalized reproducibility values under a random assignment of observed ICs across the  $K$  runs. This "null" distribution is used to compute reproducibility  $p$ -values for each observed matched component from RAICAR. An objective cutoff such as  $p < 0.05$  can be used to detect "significantly reproducible" components. This avoids subjective user decisions such as selection of the number of clusters in ICASSO or the reproducibility cutoff in RAICAR or a cutoff on intra cluster distance in sogICA.



---

## 5.1 Results for publicly available rsfMRI data

We applied RAICAR-N to publicly available  $N = 23$  subject rsfMRI data from <http://www.nitrc.org/>. We analyzed the data in 2 different ways:

1.  $n_C = 40$  ICs were extracted for each of the  $N = 23$  subjects. The  $K = 23$  single subject ICA runs were subjected to a RAICAR-N analysis (after registration to standard space).

In single subject ICA based RAICAR-N analysis (see Figures 8 - 9), we are able to declare 6 out of the 8 ICs reported in [Beckmann et al., 2005] (which used group ICA) as "reproducible" (4 ICs have  $p$ -values  $< 0.05$  and 2 ICs have  $p$ -values  $< 0.06$ ). This is consistent with the 5 reproducible RSNs reported in [DeLuca et al., 2005] using single subject ICA analysis.

2.  $L = 5$  subjects were randomly drawn from  $N = 23$  subjects to create one group of subjects which was subjected to a group ICA analysis in which  $n_C = 40$  components were extracted. This process was repeated  $K = 50$  or 100 times and the resulting group ICA runs were subjected to a RAICAR-N analysis.

In group ICA based RAICAR-N analysis (see Figures 10 - 13), we are able to declare all 8 components reported in [Beckmann et al., 2005] as "reproducible" (at  $p < 0.05$ ). Some of the ICs detected as "reproducible" in the group ICA based RAICAR-N on human rsfMRI data are not shown in [Beckmann et al., 2005] but do appear in the more recent paper [Smith et al., 2009]. RAICAR-N results for  $K = 50$  are almost identical to those for  $K = 100$  suggesting that  $K = 50$  runs of group ICA are sufficient for a RAICAR-N reproducibility analysis.

## 5.2 Single subject ICA vs Group ICA

Based on our results, it appears that single subject ICA maps are less reproducible compared to group ICA maps as illustrated in Figures 8 and 10. A single subject ICA based analysis is more resistant to subject specific artifacts. On the other hand, a group ICA based analysis makes the strong assumption that ICs are spatially identical across subjects. If this assumption is true, group ICA takes advantage of temporal concatenation to constrain the ICs spatially across subjects thereby reducing their variance. Hence, when there are no gross artifacts in individual rsfMRI data sets, group ICA is expected to be more sensitive for reproducible IC detection. As seen in Figures 9 and 11, our results agree with this proposition. All ICs declared as "reproducible" in the single subject based RAICAR-N analysis continue to remain "reproducible" in the group ICA based RAICAR-N analysis.

---

### 5.3 How should subjects be grouped for group ICA?

This raises the question of how the subjects should be grouped together for individual group ICA runs in preparation for RAICAR-N. If all  $N$  subjects are used in all group ICA analyses then there is no diversity in the individual group ICA runs. In this case, a RAICAR-N analysis will capture algorithmic variability due to non-convexity of ICA objective function but not dataset variability. Hence, our conclusions might not be generalizable to a different set of  $N$  subjects.

Another option is to randomly select  $L$  subjects out of  $N$  for each group ICA run and submit the resulting  $K$  group ICA runs to RAICAR-N. In this case, we will account for both algorithmic and data set variability via a RAICAR-N analysis. In other words, we will be able to determine those ICs that are "reproducible" across different sets of  $L$  subjects and across multiple ICA runs. A key question is: How should we choose  $L$  and  $K$ ? In section 2.5, we proposed a simple method to determine the number of subjects  $L$  to be used in a single group ICA run out of the  $N$  subjects - the key idea is to form groups with enough "diversity". Multiple such group ICA runs can then be submitted to a RAICAR-N analysis for reproducibility assessment. Clearly, the larger the value of  $N$ , the larger the value of  $L$ . Hence, increasing the number of subjects  $N$  in a study will allow us to make conclusions that are generalizable to a larger set of  $L$  subjects. Also, conclusions generalizable to  $L_1$  subjects are expected to hold for  $L_2 > L_1$  subjects but not vice versa.

### 5.4 RAICAR-N for group comparisons of *reproducible* ICs

In the present work, our focus was on enabling the selection of reproducible ICs for a given single group of subjects. However, RAICAR-N can be extended for between group analysis of reproducible components as well. Before we describe how to do so, it is useful to discuss other approaches for group analysis of RSNs described in section 5.4. Suppose we have two groups of subjects  $A$  and  $B$ .

#### 5.4.1 Discussion of single group ICA based approaches

1. Subject specific maps corresponding to group ICA maps derived using ICA back projection or dual regression are not true ICs, i.e., they are not solutions to an ICA problem.
2. These approaches do not account for either the algorithmic or the data set variability of an ICA decomposition. The single group ICA decomposition will contain both reproducible and non-reproducible ICs, but there is no systematic way to differentiate between the two.
3. Both dual regression and ICA back projection using data derived IC templates are circular analyses. First, group ICA using **all** data is used to derive template IC maps or template time courses. Next least-squares based ICA back projection or dual regression using a subset of the same data is used to derive subject specific maps and time courses corresponding to

---

each IC. Thus model 1 (group ICA) on data  $\mathcal{D}$  is used to learn an assumption  $\mathcal{A}$  (template IC maps or template time courses) that is then used to fit model 2 (dual regression or ICA back projection) on a subset of the **same** data  $\mathcal{D}$ . This is circular analysis [Kriegeskorte et al., 2009, Vul and Kanwisher, 2010].

It is easy to avoid circular analysis in a dual regression approach via cross-validation. For example, one can split the groups  $A$  and  $B$  into two random parts, a "training" set and a "test" set. First, the "training" set can be used to derive template IC maps using group ICA. Next, the "training" set based template IC maps can be used as spatial regressors for dual regression on the "test" set. Alternatively, the template ICs for dual regression can also come from a separate ICA decomposition on a independent data set unrelated to groups  $A$  and  $B$  such as human rsfMRI data. This train/test approach cleanly avoids the circular analysis problem. It is not clear how to use cross-validation for an ICA back projection approach since template time courses cannot be assumed to remain the same across ICA decompositions.

4. Subject specific structured noise is quite variable in terms of its spatial structure. Hence, a group ICA analysis cannot easily model or account for subject specific structured noise via group level ICs. Consequently, subject specific spatial maps in ICA back projection or dual regression will have a noise component that is purely driven by the amount of structured noise in individual subjects. On the other hand, a single subject ICA based analysis can accurately model subject specific structured noise via single subject ICs.

#### 5.4.2 Discussion of multiple ICA run approaches

1. [Zuo et al., 2010] report that using different sets of template ICs in template based methods using spatial correlation such as [Harrison et al., 2008] can result in the selection of different ICs in individual ICA runs. This is not surprising since IC correspondence derived from template based methods **does** depend on the particular template used. This is similar to a seed based correlation analysis being dependent on the particular seed ROI used. It is worth noting that template free approaches such as sogICA and RAICAR do not rely on any template.
2. [Cole et al., 2010] state that individual runs across subjects (or groups of subjects) can be quite variable in terms of the spatial structure of the estimated ICs. For example, [Cole et al., 2010] point out that an IC might be apparently split into two sub-components in some subjects but not others. The real problem is that the same model order could lead to over-fitting in some subjects (or groups of subjects) but not in others. Hence, the observed differences in a group comparison might be biased by the unknown difference in the amount of over-fitting across groups  $A$  and  $B$ .

As described in 2.1.3, over-fitting can lead to the phenomenon of component "splitting" in ICA. This is not limited to single subject ICA but can also occur in group ICA. For

---

instance, [Zuo et al., 2010] report the "default mode" network as split into three sub networks using group ICA and note that component "splitting" can also reflect functional segregation or hierarchy within a particular IC and is not necessarily a consequence of model order overestimation in every case.

Over-fitting can be correctly accounted for by a reproducibility analysis. This is because we expect the real and stable non-Gaussian sources to be reproducible across multiple ICA runs (algorithmic variability) and across different subjects or groups of subjects (data set variability).

If we want the results of a between group ICA analysis to be generalizable to an independent group of subjects then we must account for both the algorithmic and data variability of ICA. We propose to modify RAICAR-N for enabling between group comparisons of "reproducible" ICs as follows:

1. Enter multiple within and across subject (or within and across sets of subjects) ICA runs for groups  $A$  and  $B$  into a RAICAR analysis. Perform the RAICAR component matching process across groups  $A$  and  $B$ .
2. Use RAICAR-N to compute reproducibility  $p$ -values separately for group  $A$  and  $B$  for each matched component across groups  $A$  and  $B$ .
3. Only ICs that are *separately* reproducible in both groups  $A$  and  $B$  and that are *maximally* similar to each other are used for between group comparisons.

To summarize, a RAICAR-N analysis:

- ✓ can be applied for "reproducible" component detection either within or across subjects in any component based analysis - not necessarily ICA. For instance, a set of PCA maps across subjects can be submitted to a RAICAR-N analysis.
- ✓ is simple to implement and accounts for both algorithmic and data set variability of an ICA decomposition.
- ✓ avoids any user decisions except the final  $p$ -value cutoff which can be objectively pre-set at standard values such as 0.05.
- ✓ can be extended to enable comparisons of reproducible ICs between groups  $A$  and  $B$ .

## 6 Conclusions

Multiple group ICA runs using groups of subjects with enough "diversity" can be used to account for the run-to-run variability in ICA algorithms both due to the non-convex ICA objective function as well as across subjects data variability. These group ICA runs can be subjected to a RAICAR-N

---

”reproducibility” analysis. RAICAR-N enables the objective detection of ”reproducible components” in any component based analysis of fMRI data such as ICA and can also be used for a between group comparison of ”reproducible” ICs.

## Acknowledgements

We gratefully acknowledge financial support from the Pain and Analgesia Imaging and Neuroscience (P.A.I.N) group, McLean Hospital, Harvard Medical School, Belmont MA, USA under the grants K24NS064050 (DB) and R01NS065051 (DB). We would also like to thank Dr. Christian Beckmann for making the IC image files from his 2005 paper [Beckmann et al., 2005] available to us.

## References

- H. Attias. Independent factor analysis. *Neural Computation*, 11:803–851, 1999.
- C. Beckmann, C. Mackay, N. Filippini, and S. Smith. Group comparison of restingstate group comparison of resting state fmri data using multi-subject ica and dual regression. Number poster 441 SU-AM. 15 Annual Meeting of Organization for Human Brain Mapping, 2009.
- C. F. Beckmann and S. M. Smith. Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Transactions on Medical Imaging*, 23:137152, 2004.
- C. F. Beckmann, M. DeLuca, J. T. Devlin, and S. M. Smith. Investigations into resting-state connectivity using independent component analysis. *Philos. Trans. R. Soc. Lond. B. (Biol Sci)*, 360(1457):1001–1013, 2005.
- A. J. Bell and T. J. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7:1129–1159, 1995.
- V. D. Calhoun, T. Adali, G. D. Pearlson, and J. J. Pekar. A method for making group inferences from functional MRI data using independent component analysis. *Human Brain Mapping*, 14(3):140151, 2001.
- V. D. Calhoun, J. J. Pekar, and G. D. Pearlson. Alcohol intoxication effects on simulated driving: exploring alcohol-dose effects on brain activation using functional MRI. *Neuropsychopharmacology*, 29(11):2097–2107, 2004.
- J. Cardoso. Blind signal separation: statistical principles. *Proceedings of the IEEE*, 9(10):2009–2025, 1998.
- D. M. Cole, S. M. Smith, and C. F. Beckmann. Advances and Pitfalls in the Analysis and Interpretation of Resting-State FMRI Data. *Frontiers in Systems Neuroscience*, 4(8), 2010.

- 
- P. Comon. ‘Independent Component Analysis, a new concept? *Signal Processing*, 36(3):287–314, 1994.
- M. Davies. Identifiability Issues in Noisy ICA. *IEEE Signal Processing Letters*, 11(5):470–473, 2004.
- M. DeLuca, C. F. Beckmann, N. DeStefano, P. M. Matthews, and S. M. Smith. fMRI resting state networks define distinct modes of long-distance interactions in the human brain. *Neuroimage*, 29(4):1359–1367, 2005.
- J. Eriksson and V. Koivunen. Identifiability, separability, and uniqueness of linear ICA models. *Signal Processing Letters, IEEE*, 11(7):601–604, 2004.
- J. Eriksson and V. Koivunen. Complex random vectors and ICA models: identifiability, uniqueness, and separability. *IEEE Transactions on Information Theory*, 52(3):1017 – 1029, 2006.
- F. Esposito, T. Scarabino, A. Hyvarinen, J. Himberg, E. Formisano, S. Comani, G. Tedeschi, R. Goebel, E. Seifritz, and F. Di Salleb. Independent component analysis of fMRI group studies by self-organizing clustering. *NeuroImage*, 25(1):193–205, 2005.
- W. Feller. *An Introduction to Probability Theory, Vol. II*. New York:Wiley, 1966.
- M. D. Greicius, G. Srivastava, A. L. Reiss, and V. Menon. Default-mode network activity distinguishes Alzheimer’s disease from healthy aging: Evidence from functional MRI. *Proc. Natl. Acad. Sci. U. S. A.*, 101(13):4637–4642, 2004.
- B. J. Harrison, J. Pujol, M. Lopez-Sola, R. Hernandez-Ribas, J. Deus, H. Ortiz, C. Soriano-Mas, M. Yucel, C. Pantelis, and N. Cardoner. Consistency and functional specialization in the default mode brain network. *Proc. Natl. Acad. Sci. U. S. A.*, 105(28):9781–9786, 2008.
- T. Hastie, R. Tibshirani, and J. H. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York:Springer, 2009.
- A. Himberg, A. Hyvarinen, and F. Esposito. Validating the independent components of neuroimaging time-series via clustering and visualization. *NeuroImage*, 22(3):1214–1222, 2004.
- A. Hyvarinen. New approximations of differential entropy for independent component analysis and projection pursuit. *Advances in Neural Information Processing Systems*, 10:273–279, 1998.
- A. Hyvarinen. Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. *IEEE Transactions on Neural Networks*, 10(3):626–634, 1999.
- A. Hyvarinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. John Wiley and Sons, 2001.
- C. Jutten and J. Herault. Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, 24:1–10, 1991.

- 
- R. E. Kass and A. E. Raftery. Bayes factors and model uncertainty. Technical Report Report 254, University of Washington, USA, 1993.
- N. Kriegeskorte, W. K. Simmons, P. S. F. Bellgowan, and C. I. Baker. Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience*, 12:535–540, 2009.
- S. Kritchman and B. Nadler. Non-Parametric Detection of the Number of Signals: Hypothesis Testing and Random Matrix Theory. *IEEE Transactions on Signal Processing*, 57(10):3930–3941, 2009.
- M. J. McKeown, S. Makeig, G. G. Brown, T. P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski. Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping*, 6:160–188, 1998.
- T. P. Minka. Automatic choice of dimensionality for PCA. Technical Report 514, MIT, Media Laboratory Vision and Modeling Group, 2000.
- C. R. Rao. A Decomposition Theorem for Vector Variables with a Linear Structure. *The Annals of Mathematical Statistics*, 40(5):1845–1849, 1969.
- S. M. Smith, P. T. Fox, K. L. Miller, D. C. Glahn, P. M. Fox, C. E. Mackay, N. Filippini, K. E. Watkins, R. Toro, A. R. Laird, and C. F. Beckmann. Correspondence of the brains functional architecture during activation and rest. *Proc. Natl. Acad. Sci. U.S.A.*, 106(31):13040–13045, 2009.
- F. J. Theis. A New Concept for Separability Problems in Blind Source Separation. *Neural Computation*, 16:1827–1850, 2004.
- M. E. Tipping. Probabilistic Principal Component Analysis. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 61(3):611–622, 1999.
- J. Tohka, K. Foerde, A. R. Aron, S. M. Tom, A. W. Toga, and R. A. Poldrack. Automatic independent component labeling for artifact removal in fMRI. *NeuroImage*, 39:1227–1245, 2008.
- S. J. van Albada and P. A. Robinson. Transformation of arbitrary distributions to the normal distribution with application to EEG test-retest reliability. *Journal of Neuroscience Methods*, 161:205–211, 2007.
- E. Vul and N. Kanwisher. In *Hanson, S. and Bunzl, M., Foundational issues in human brain mapping*, chapter 6, pages 71–92. MIT Press, 2010.
- B. Włodzimierz. *The normal distribution: Characterizations with applications*. New York:Springer-Verlag, 1995.
- Z. Yang, S. LaConte, X. Weng, and X. Hu. Ranking and Averaging Independent Component Analysis by Reproducibility (RAICAR). *Human Brain Mapping*, 29:711–725, 2008.

---

X. Zuo, C. Kelly, J. S. Adelstein, D. F. Klein, F. X. Castellanos, and M. P. Milham. Reliable Intrinsic Connectivity Networks: Test-Retest Evaluation Using ICA and Dual Regression Approach. *NeuroImage*, 49(3):21632177, 2010.

## 7 Figure Legends

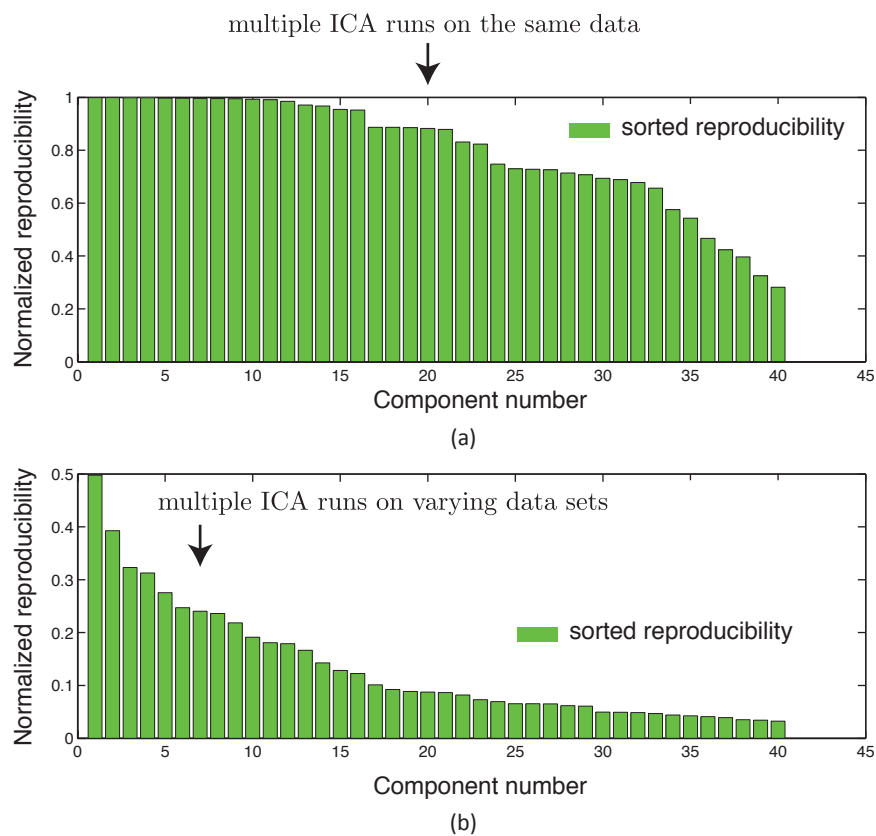
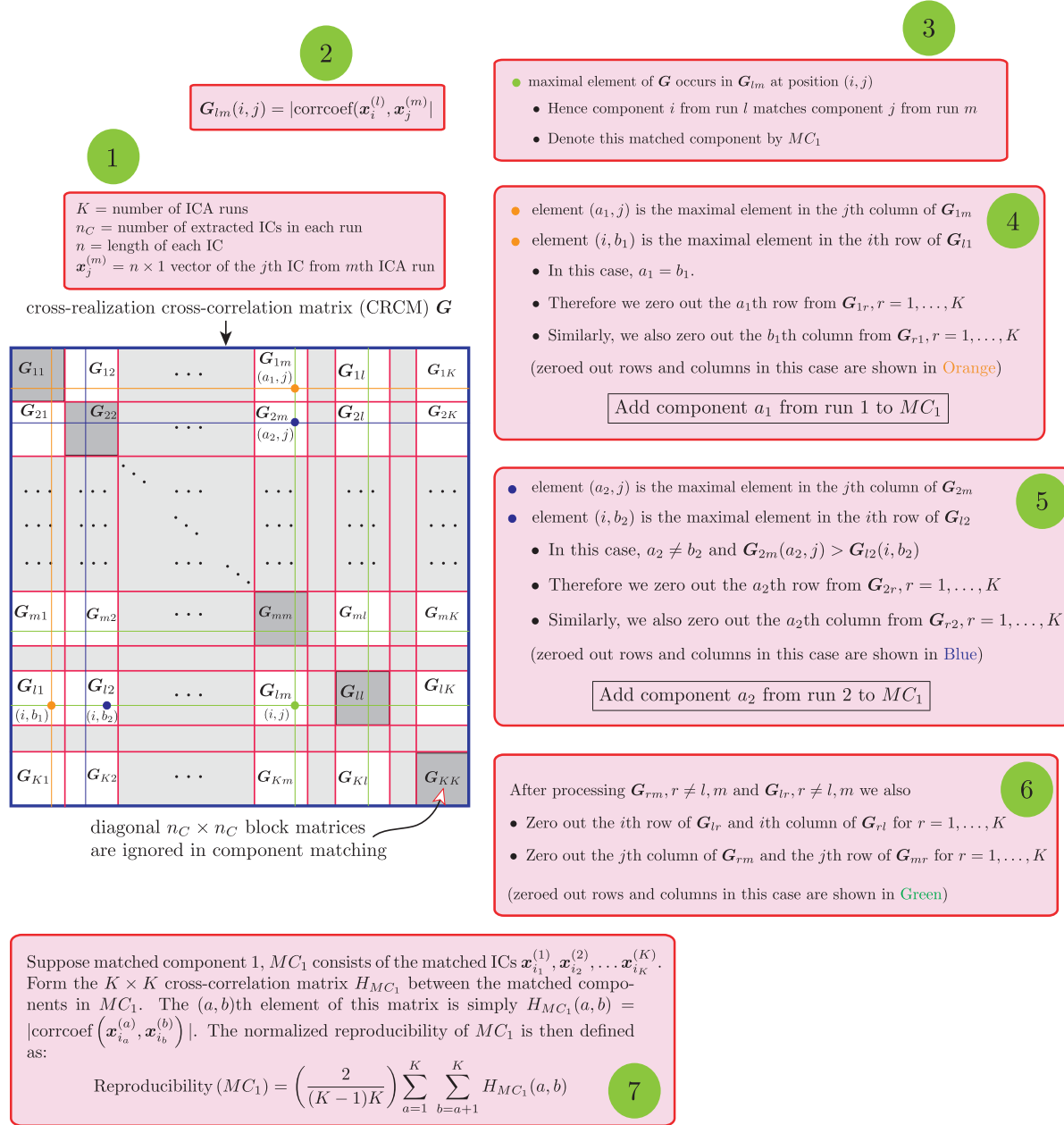


Figure 1: Figure illustrates the variation in normalized reproducibility from RAICAR depending on whether the input to RAICAR is (a) Multiple ICA runs on single subject data or (b) Multiple ICA runs across subjects. Notice that the normalized reproducibility is much lower for across subjects analysis compared to within subject analysis.





8

Once we calculate  $\text{Reproducibility}(MC_1), \text{Reproducibility}(MC_2), \dots, \text{Reproducibility}(MC_{n_C})$ , how can we calculate a cutoff to separate the high reproducibility ICs from the low reproducibility ICs?

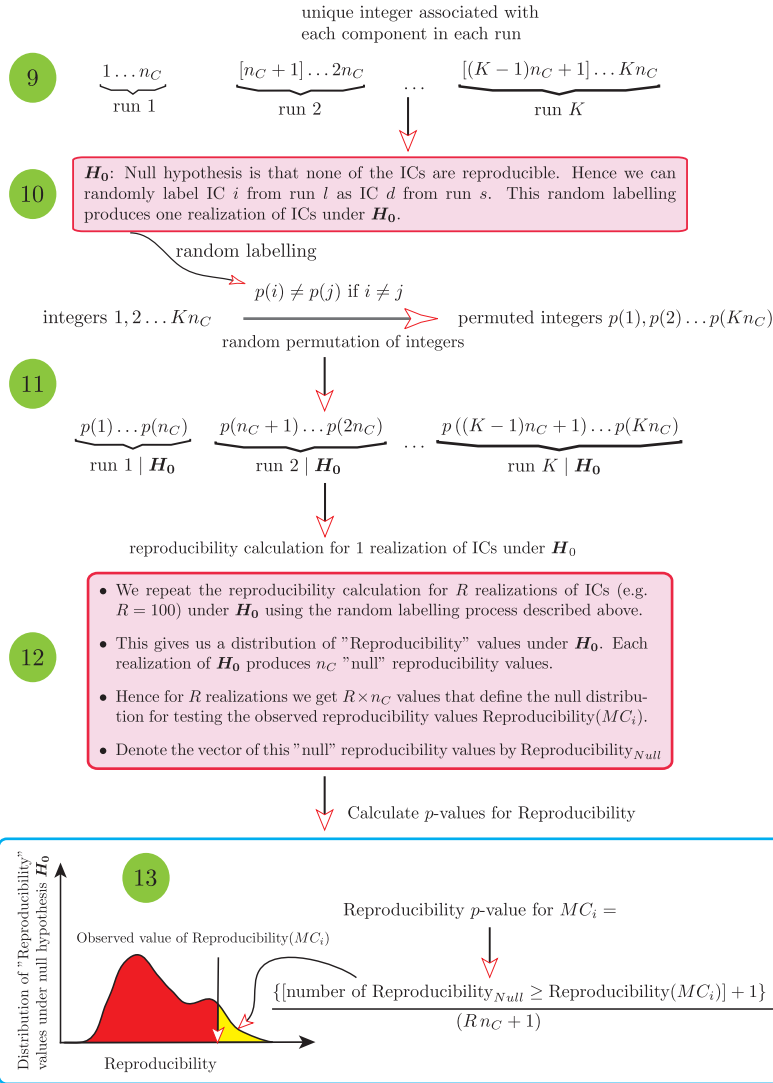


Figure 3: Pictorial depiction of the process for generating a "null" distribution in RAICAR-N. Our "null" hypothesis is: " $H_0$ : None of the ICs are reproducible. Hence, we can randomly label IC  $i$  from run  $l$  as IC  $d$  from run  $s$ ". Therefore we randomly split the  $Kn_C$  ICs across  $K$  runs into  $K$  parts and run the RAICAR algorithm on each set of randomly split ICs. This gives us a set of "null" reproducibility values which can be used to compute  $p$ -values for the observed reproducibility of ICs in the original RAICAR run. The green circles indicate the sequence of steps for generating the "null" distribution after the steps in Figure 2.

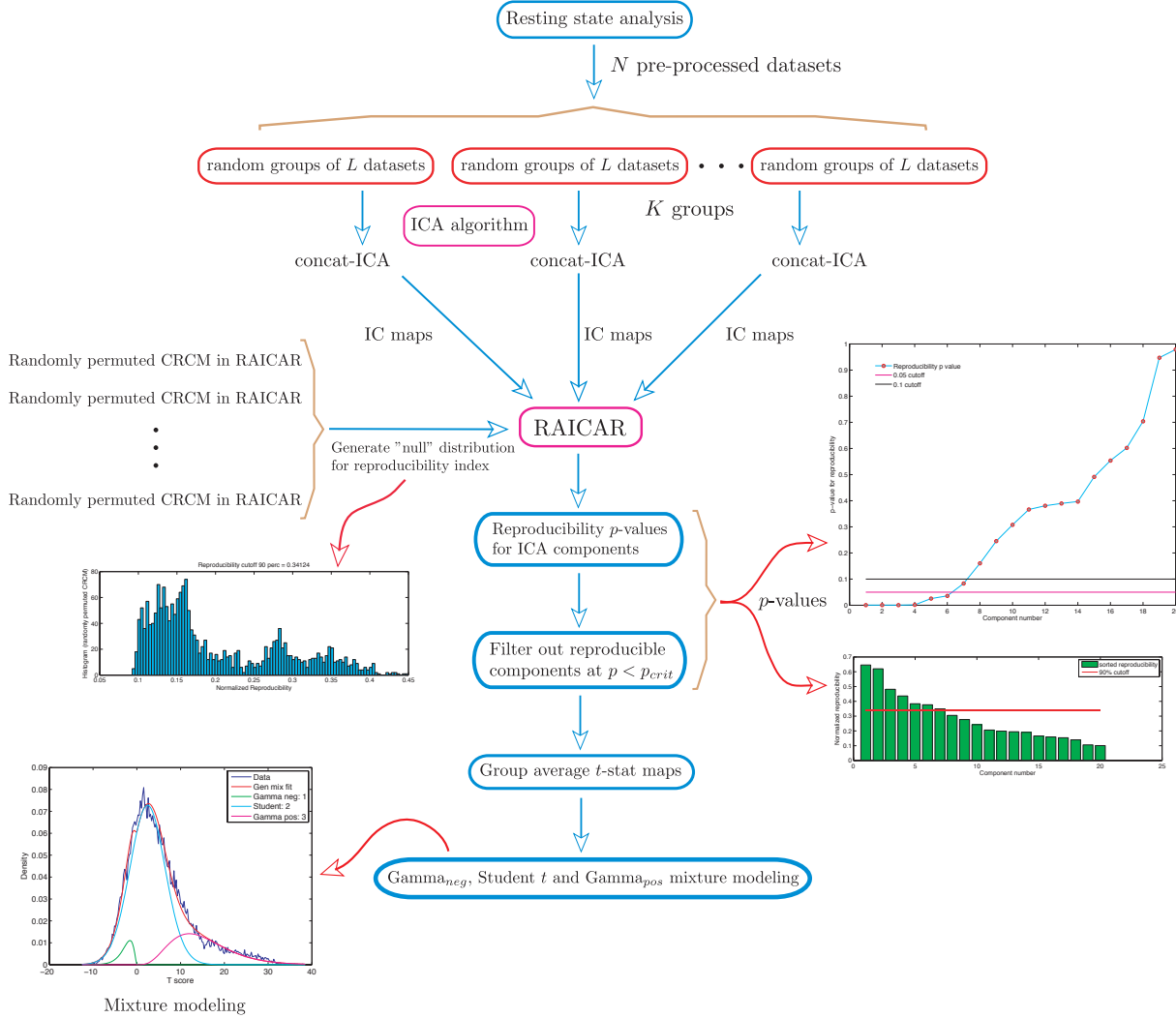


Figure 4: Flowchart for a group ICA based RAICAR-N analysis. The  $N$  single subject data sets are first pre-processed and subsequently bootstrapped to create  $K$  groups, each group containing  $L$  distinct subjects. Each group of  $L$  subjects is submitted to a temporal concatenation group ICA analysis. The resulting IC maps (either raw ICs or ICs scaled by noise standard deviation) are subjected to a RAICAR analysis. The cross-realization cross correlation matrix (CRCM) is randomly permuted multiple times:  $\mathbf{G} \rightarrow \mathbf{G}(\mathbf{g}, \mathbf{g})$  where  $\mathbf{g}$  is a random permutation of integers from  $1, \dots, Kn_C$ . The permuted CRCMs are subjected to a RAICAR analysis to generate a realization of reproducibility values under the "null" hypothesis. The computed "null" distribution of reproducibility values is used to assign  $p$  values to the observed reproducibility of the original RAICAR run. Finally, reproducible ICs are averaged using a random effects analysis and the resulting  $t$ -statistic images are subjected to Gamma<sub>neg</sub>, Student  $t$  and Gamma<sub>pos</sub> mixture modeling.

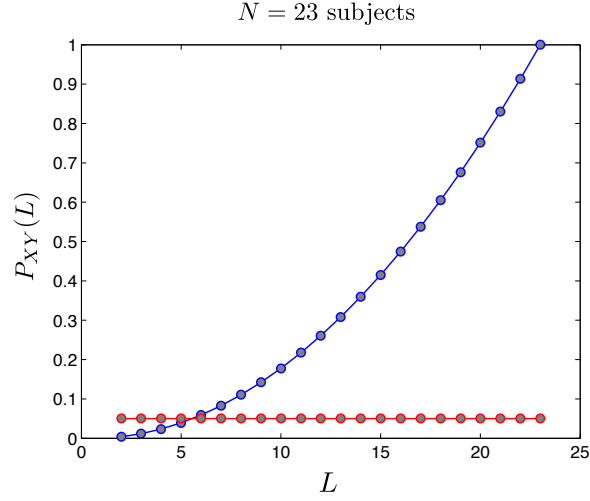


Figure 5: Figure shows a plot of  $P_{XY}(L)$  vs  $L$  for  $N = 23$  in blue. The red line shows the  $\alpha_{max} = 0.05$  cutoff. The largest value of  $L$  for which  $P_{XY}(L) \leq 0.05$  is  $L = 5$ .

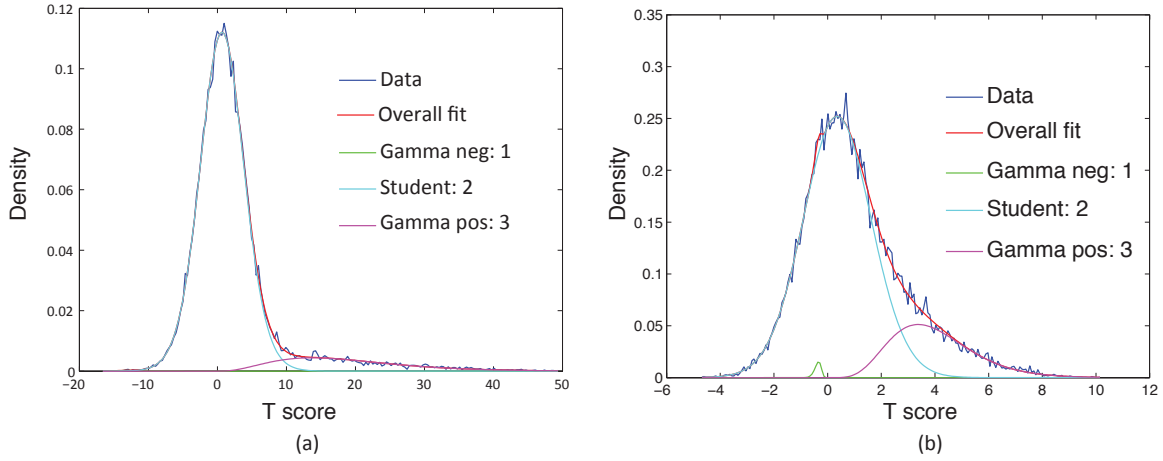


Figure 6: Examples of displaying non-Gaussian spatial structure using a Student  $t$ ,  $\text{Gamma}_{pos}$  and  $\text{Gamma}_{neg}$  mixture model. Notice how the  $\text{Gamma}_{neg}$  density is driven to near 0 class fraction in the absence of significant negative non-Gaussian structure.

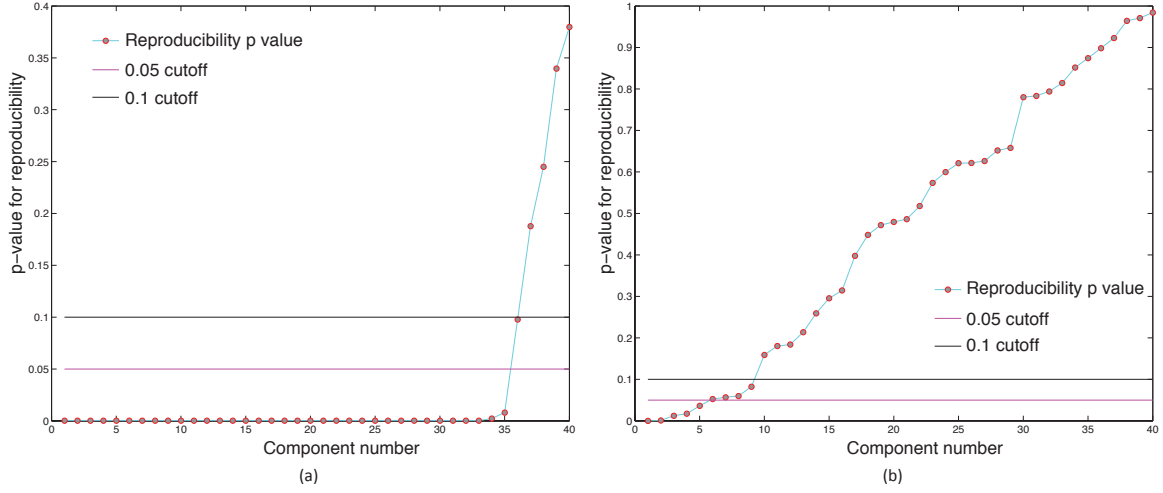


Figure 7:  $p$ -value cutoffs for within and across single subject analysis using RAICAR-N. This figure illustrates the intuitive fact that within subject ICA runs are much more reproducible compared to across subject ICA runs.

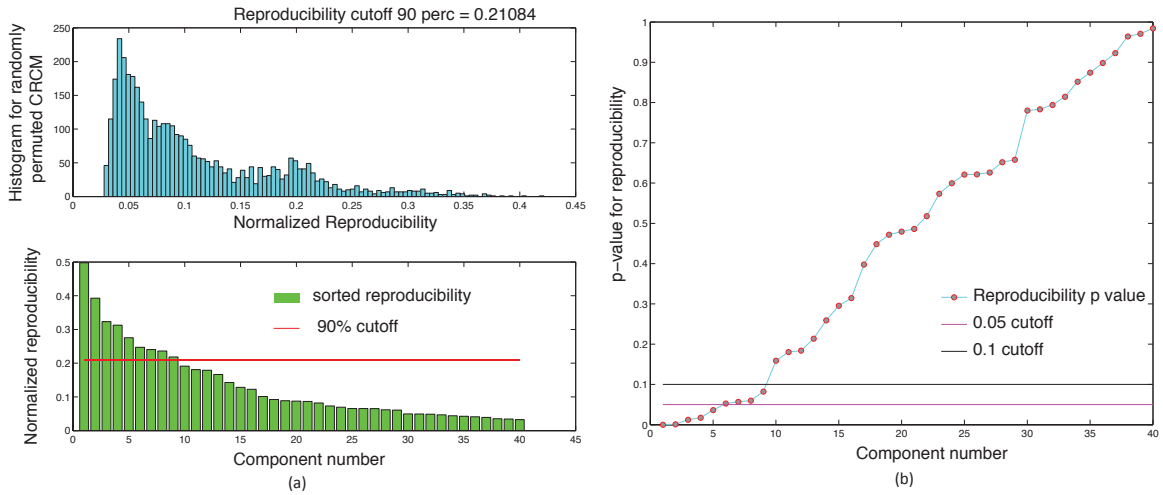


Figure 8: Single subject rsfMRI ICA runs across 23 subjects were combined using a RAICAR-N analysis. Figure (a) shows the observed values of normalized reproducibility (bottom) as well as the "null" distribution of normalized reproducibility across  $R = 100$  simulations (top). Figure (b) shows the  $p$ -values for each IC along with the 0.05 and 0.1 cutoff lines.

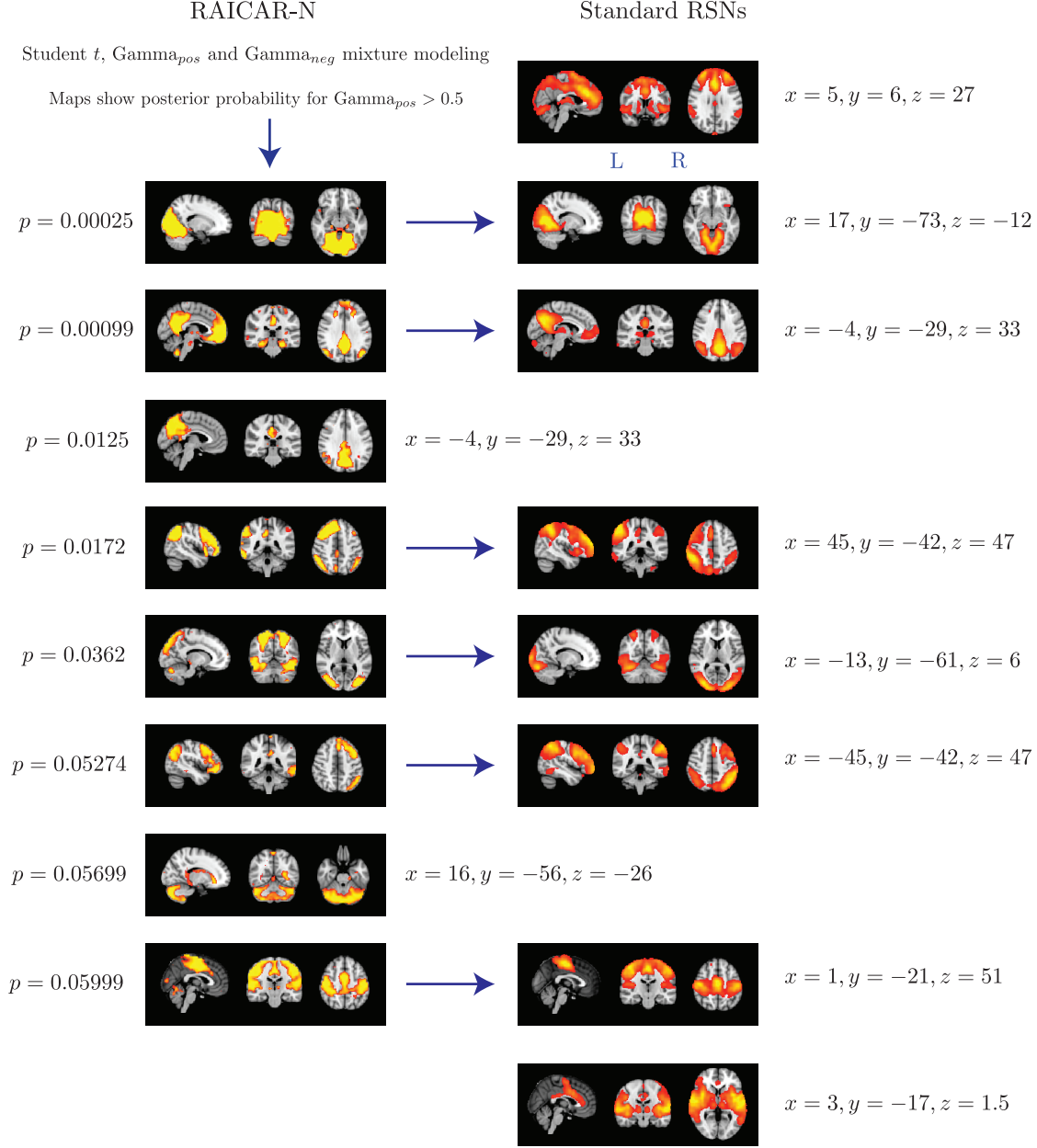


Figure 9: The top 8 "reproducible" ICs from a RAICAR-N analysis on single subject ICA runs compared with standard RSN maps reported in literature Beckmann et al. [2005]. We are able to declare 4 "standard" RSNs as significantly reproducible at a  $p$ -value  $< 0.05$ . There are 2 other "standard" RSNs that achieve a reproducibility  $p$ -value between 0.05 and 0.06 as well as 2 "non-standard" RSNs that achieve  $p$ -values of 0.0125 and 0.05699 respectively. We also could not find 2 of the published RSNs in Beckmann et al. [2005] as reproducible in single subject ICA runs.

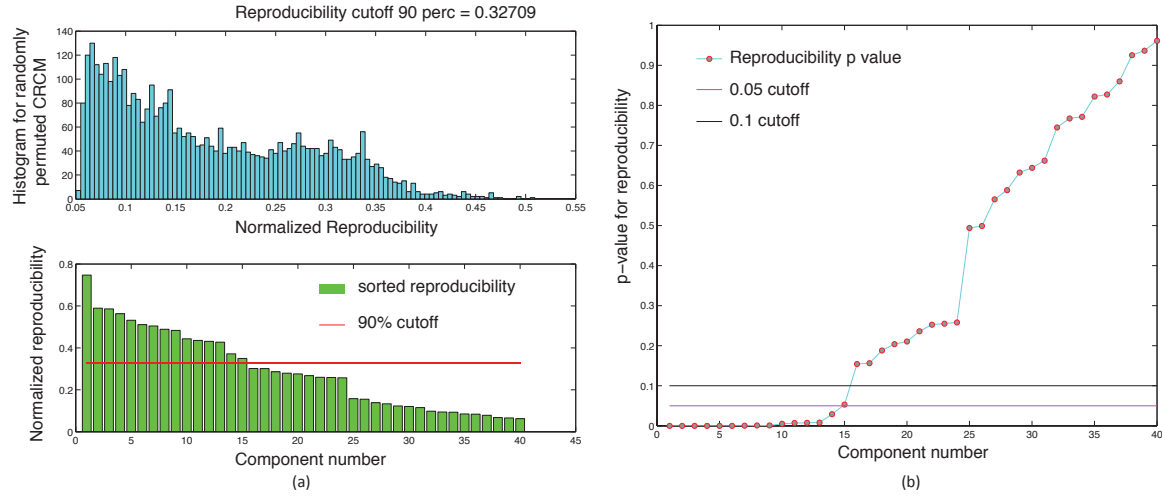


Figure 10:  $L = 5$  subjects were randomly drawn from the set of  $N = 23$  subjects and submitted to a temporal concatenation based group ICA. This process was repeated  $K = 50$  times and the resulting ICA maps were submitted to a RAICAR-N analysis. Figure (a) shows the observed values of normalized reproducibility (bottom) as well as the "null" distribution of normalized reproducibility across  $R = 100$  simulations (top). Figure (b) shows the  $p$ -values for each IC along with the 0.05 and 0.1 cutoff lines.

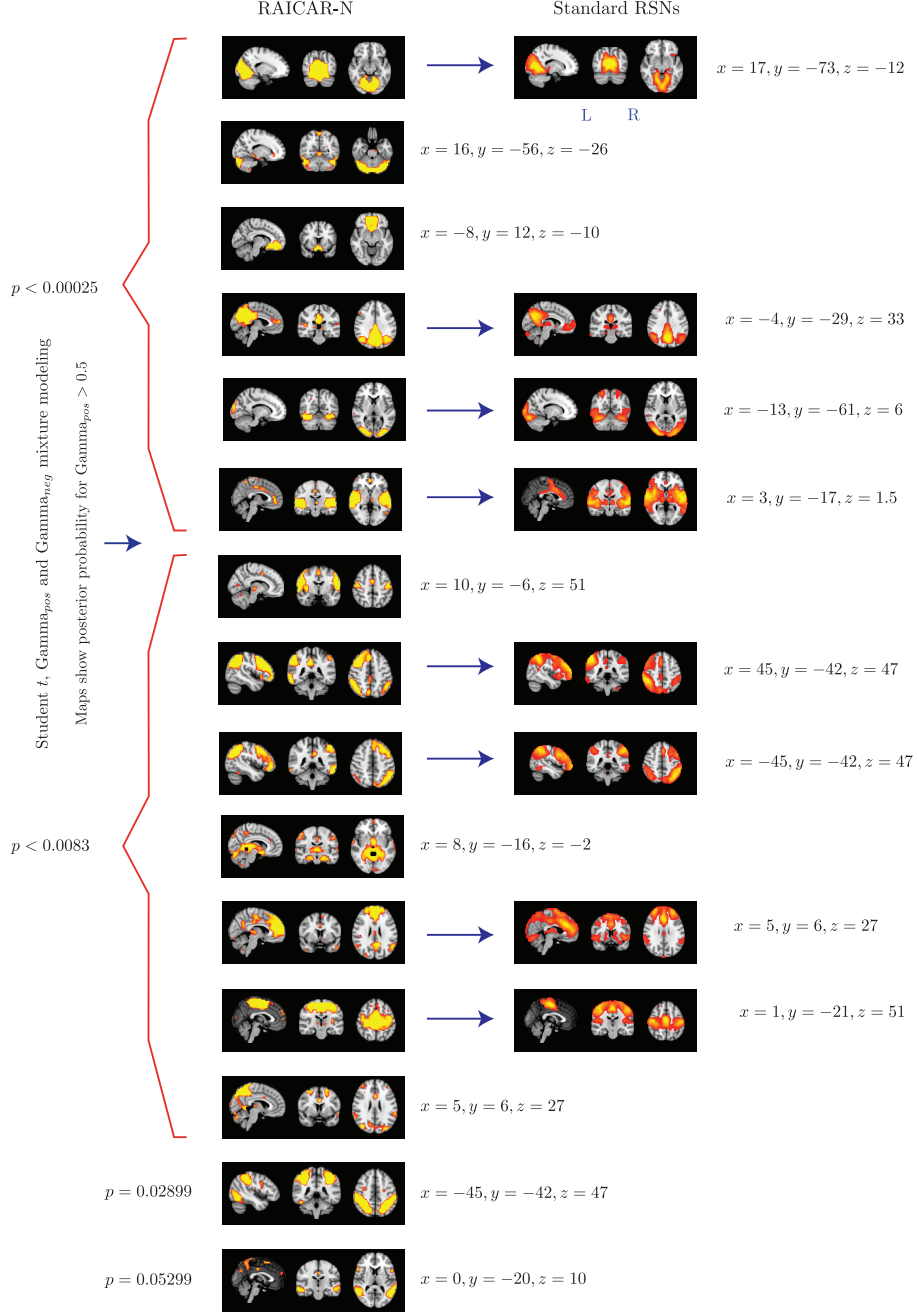


Figure 11: The top 15 "reproducible" ICs from  $K = 50$  runs of  $L = 5$  subject group ICA RAICAR-N analysis compared with standard RSN maps reported in literature [Beckmann et al. \[2005\]](#). We are able to declare 8 "standard" RSNs as significantly reproducible at a  $p$ -value of  $< 0.05$ . There are 6 other "non-standard" RSNs that can be declared as significantly reproducible at a  $p$ -value of  $< 0.05$  and 1 other "non-standard" RSN that achieves a  $p$ -value of 0.05299.



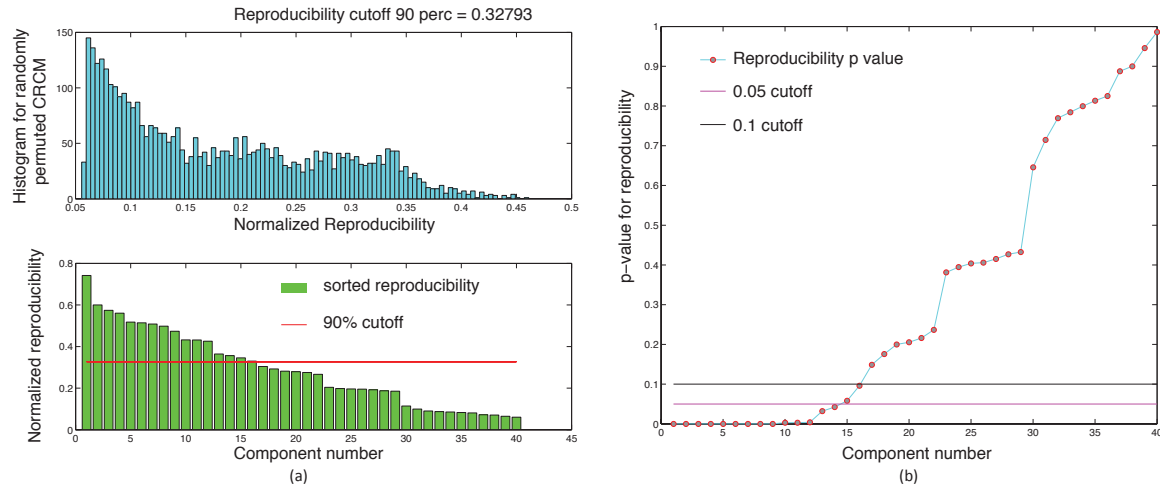


Figure 12:  $L = 5$  subjects were randomly drawn from the set of  $N = 23$  subjects and submitted to a temporal concatenation based group ICA. This process was repeated  $K = 100$  times and the resulting ICA maps were submitted to a RAICAR-N analysis. Figure (a) shows the observed values of normalized reproducibility (bottom) as well as the "null" distribution of normalized reproducibility across  $R = 100$  simulations (top). Figure (b) shows the  $p$ -values for each IC along with the 0.05 and 0.1 cutoff lines.

