UNIVERSITY OF CALIFORNIA SAN DIEGO

Rectified Sparse Bayesian Learning and Effects and Limitations of Nuisance Regression in Functional MRI

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Electrical Engineering (Signal and Image Processing)

by

Alican Nalci

Committee in charge:

Professor Thomas T. Liu, Chair
Professor Bhaskar D. Rao, Co-Chair
Professor Pamela C. Cosman
Professor Truong Q. Nguyen
Professor Eric C. Wong

2019

The dissertation of Alican Nalci is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____
Co-Chair

_____
Chair

University of California San Diego

2019

DEDICATION

To my beloved family.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First of all, I would like to thank my PhD advisors Professor Thomas Liu and Professor Bhaskar Rao for their endless encouragement, patience, and guidance during my PhD studies. I am deeply grateful to them for providing me with the opportunity to freely pursue contemporary problems in functional MRI and sparse signal recovery. I would also like to thank the members of my committee, Professors Pamela Cosman, Truong Nguyen, and Eric Wong for assessing the quality of my research on a regular basis.

I am deeply grateful to my parents Ayla Nalcı and Ahmet Nalcı, and my brother Ozan Nalcı for their unconditional love, support, and encouragement throughout my life. They sacrificed a lot so that I can have a better life and education. I am now appointing my two-year-old niece Can Nalcı as a future PhD candidate. Only a few people in the world are blessed with a soul mate with a beautiful heart. I am extremely fortunate to have met my girlfriend Neşe Demir during my studies at UCSD. Her limitless love and patience have always put a smile on my face.

I would also like to thank my awesome friends and roommates Hasan Al-Rubaye, Doruk Beyter, Ege İşeri, Umut Kodak and Özgür Balkan. My life as a graduate student was significantly more enjoyable thanks to their bad jokes and our late-night burrito parties. I would also like to thank my coworkers at the UCSD Center for Functional MRI and UCSD Digital Signal Processing Lab including Maryam Falahpour, Igor Fedorov, Maher Al-Shoukairi, Eulanca Liu and Gökçe Sarar for their insights, discussions and friendship. I thank Robert Bussell, David Shin, Kun Lu, Aaron Jacobson, Peggy Totzke, and Mary O'Malley for their indispensable help and making everyday at the center incredibly better, and of course I thank my labmates Yacong Ding, Sung-En Chiu, Elina Nayebi, and Bang Nguyen at the DSP lab for their friendship.

Finally, I would like to thank my relatives including Hazal, Emre, Çiğdem, and Şehsuvar Ertürk; Güneş, Ayten, and Korhan Yeletayşi, and my grandparents Sema and Recep; Mehmet and Faika. I also thank my sister-in-law Büke and her parents for their sincere friendship.

Most of the chapters of this dissertation are based on papers which have been published

or submitted for publication in academic journals.

Chapter 2 is, in part, a reprint of the material published in two papers: (1) Alican Nalci, Igor Fedorov, Maher Al-Shoukairi, Thomas T. Liu, and Bhaskar D. Rao. "*Rectified Gaussian Scale Mixtures and the Sparse Non-negative Least Squares Problem.*" IEEE Transactions on Signal Processing 66, No. 12 (2018): 3124-3139, and (2) Alican Nalci, Bhaskar D. Rao, and Thomas T. Liu. "*Sparse Estimation of Quasi-Periodic Spatiotemporal Components in Resting-State FMRI,*" in 24th Annual Scientific Meeting of the ISMRM, 2016, p. 3824. I was the primary author in both papers and B.D. Rao, and T.T. Liu supervised the research.

Chapter 3 is, in full, a reprint of the material as it appears in: Alican Nalci, Bhaskar D. Rao, and Thomas T. Liu. "*Global Signal Regression Acts as a Temporal Downweighting Process in Resting-state fMRI.*" NeuroImage 152 (2017): 602-618. I was the primary author and T.T. Liu supervised the research.

Chapter 4 is, in full, a reprint of the material as it appears in: Alican Nalci, Bhaskar D. Rao, and Thomas T. Liu. "*Nuisance Effects and the Limitations of Nuisance Regression in Dynamic Functional Connectivity fMRI.*" NeuroImage 184 (2019): 1005-1031. I was the primary author and T.T. Liu supervised the research.

Chapter 5 is, in full, a reprint of material in submission as: Alican Nalci, Wenjing Luo, and Thomas T. Liu. "*Nuisance Effects in Inter-scan Functional Connectivity Estimates Before and After Nuisance Regression.*" I was the primary author of this paper.

VITA

2011	Intern, TUBITAK Space Technologies Research Institute, Turkey
2012	Research Intern, Fraunhofer Institute, Germany
2013	B.S. in Electrical Engineering, Bilkent University, Turkey
2014-2019	Research Assistant, University of California, San Diego
2015	M.S. in Electrical Engineering, University of California, San Diego
2019	Ph.D. in Electrical Engineering, University of California, San Diego

PUBLICATIONS

Alican Nalci, Wenjing Luo, and Thomas T. Liu. "Nuisance effects in inter-scan functional connectivity estimates before and after nuisance regression." *In Submission*. 2019.

Alican Nalci, Bhaskar D. Rao, and Thomas T. Liu. "Nuisance effects and the limitations of nuisance regression in dynamic functional connectivity fMRI." *NeuroImage*. 2018; 184:1005-31.

Alican Nalci, Igor Fedorov, Maher Al-Shoukairi, Thomas T. Liu, and Bhaskar D. Rao. "Rectified Gaussian scale mixtures and the sparse non-negative least squares problem." *IEEE Transactions on Signal Processing*. 2018; 66:3124-39.

Igor Fedorov, Alican Nalci, Ritwik Giri, Bhaskar D. Rao, Truong Q. Nguyen, and Harinath Garudadri. "A unified framework for sparse non-negative least squares using multiplicative updates and the non-negative matrix factorization problem." *Signal Processing*. 2018; 146:79-91.

Maryam Falahpour, Alican Nalci, and Thomas T. Liu. "The effects of global signal regression on estimates of resting-state BOLD fMRI and EEG vigilance correlations." *Brain Connectivity*. 2018; 8.10:618-27.

Alican Nalci, Bhaskar D. Rao, and Thomas T. Liu. "Global signal regression acts as a temporal downweighting process in resting-state fMRI." *NeuroImage*. 2017; 152:602-618.

Thomas T. Liu, Alican Nalci and Maryam Falahpour, "The global signal in fMRI: Nuisance or Information?." *NeuroImage*. 2017; 150:213-229.

Alican Nalci, Wenjing Luo, Thomas T. Liu, "Global signal measures are significantly related to functional connectivity variations across scans both before and after global signal regression," in *27th Annual Scientific Meeting of the ISMRM*, Quebec, 2019, p. 5469.

Wenjing Luo, Alican Nalci and Thomas T. Liu, "Resting-state functional connectivity estimates across scans are correlated with nuisance metrics both before and after nuisance regression," in *27th Annual Scientific Meeting of the ISMRM*, Quebec, 2019. p. 5469.

Alican Nalci, Thomas T. Liu, "Regression does not eliminate the effects of nuisance terms in dynamic functional connectivity estimates," in *26th Annual Scientific Meeting of the ISMRM*, Paris, 2018, p. 2394.

Alican Nalci, Thomas T. Liu, "Global signal regression acts as a temporal downweighting process in resting-State fMRI," in *23rd Annual Meeting of the Organization for Human Brain Mapping*, Vancouver, 2017, p. 1888.

Alican Nalci, Thomas T. Liu, "On the relation between global signal normalization, global signal subtraction, and global signal regression in resting-state fMRI," in *25th Annual Scientific Meeting of the ISMRM*, Honolulu, 2017, p. 1655.

Alican Nalci, Maryam Falahpour, Thomas T. Liu, "On dynamic functional connectivity and global signal regression," in *23rd Annual Meeting of the Organization for Human Brain Mapping*, Vancouver, 2017, p. 1919.

Alican Nalci, Thomas T. Liu, "Global signal regression acts as a temporal downweighting process in resting-state fMRI," in *25th Annual Scientific Meeting of the ISMRM*, Honolulu, 2017, p. 1656.

Maryam Falahpour, Alican Nalci, Chi W. Wong, Thomas T. Liu, "Global signal regression alters the correlation between resting-state BOLD fluctuations and EEG vigilance measures," in *25th Annual Scientific Meeting of the ISMRM*, Honolulu, 2017, p. 240.

Alican Nalci, Bhaskar D. Rao, Thomas T. Liu, "Sparse estimation of quasi-periodic spatiotemporal components in resting-state fMRI," in *24th Annual Scientific Meeting of the ISMRM*, Singapore, 2016, p. 3824.

Alican Nalci, Thomas T. Liu, "Deterministic estimation of spatiotemporal motifs in resting-state fMRI," in *24th Annual Scientific Meeting of the ISMRM*, Singapore, 2016, p. 3833.

Alican Nalci, Alireza Khodamoradi, Ozgur Balkan, Fatta Nahab, and Harinath Garudadri. "A computer vision based candidate for functional balance test," in *IEEE Engineering in Medicine and Biology Society*, 2015, p. 3504-3508.

Kanza Khan, Dingli Qin, Daniela Glaser, Alican Nalci, Catherine Printz, Carter McElroy, and Pamela Cosman. "Study of finger force variability for physical therapists and untrained subjects," in *IEEE Signal Processing in Medicine and Biology Symposium*, 2014, p. 1-6.

ABSTRACT OF THE DISSERTATION

Rectified Sparse Bayesian Learning and Effects and Limitations of Nuisance Regression in Functional MRI

by

Alican Nalci

Doctor of Philosophy in Electrical Engineering (Signal and Image Processing)

University of California San Diego, 2019

Professor Thomas T. Liu, Chair
Professor Bhaskar D. Rao, Co-Chair

This dissertation considers the problems of sparse signal recovery (SSR) and nuisance regression in functional MRI (fMRI). The first part of the dissertation introduces a Bayesian framework to recover sparse non-negative solutions in under-determined systems of linear equations. A novel class of probability density functions named Rectified Gaussian Scale Mixtures (R-GSM) is proposed to model the sparse non-negative solution of interest. A Bayesian inference algorithm called Rectified Sparse Bayesian Learning (R-SBL) is developed, which robustly recovers the solution in numerous experimental settings and outperforms the state-of-the-art SSR approaches

by a large margin.

The rest of the dissertation investigates the effects of nuisance regression in fMRI. Chapter 3 proposes a mathematical framework to investigate the effects of global signal regression (GSR). GSR is a widely used nuisance removal approach in resting-state fMRI, however its use has been controversial since it introduces artifactual anti-correlations between pairs of fMRI signals. The proposed framework shows that the main effects of GSR can be well-approximated as a temporal down-weighting or temporal censoring process, in which the data from time points with relatively large GS magnitudes are greatly attenuated (or censored) while data from time points with relatively small GS magnitudes are largely retained. The censoring approximation reveals that the anti-correlated networks are intrinsic to the brain's functional organization and are not simply an artifact of GSR.

In Chapters 4 and 5, the effects of nuisance terms on the relationship between pairs of fMRI signals both before and after nuisance regression are investigated. It is shown that geometric norms of various nuisance regressors can significantly influence the correlation-based functional connectivity (FC) estimates in both static FC and dynamic FC studies. It is demonstrated that nuisance regression is largely ineffective in removing the significant correlations observed between FC estimates and nuisance norms. Consequently, a mathematical bound is derived on the difference between correlation coefficients before and after nuisance regression. This bound restricts the removal of nuisance norm effects from FC estimates.

Chapter 1

Introduction

This chapter provides a concise background that will be helpful in understanding the methods used in this dissertation. Section 1.1 introduces the sparse signal recovery (SSR) problem, probabilistic SSR techniques, and sparse non-negative least squares (S-NNLS) problem. Section 1.2 presents the basics of functional magnetic resonance imaging (fMRI) and shows how linear regression is used as a nuisance removal method in fMRI studies. Finally, Section 1.3 provides an outline of the dissertation and lists the contributions.

1.1 Sparse Signal Recovery

Sparsity arises in a myriad of applications in applied data science, machine learning and signal processing [1,2,3,4,5,6,7]. Specialized algorithms can take advantage of the parsimony in data to provide faster sampling rates in acquisition systems, more efficient digital communications, and better data compression [8,9,10,11,12]. In the context of sparse signal recovery (SSR) the following regression model is considered:

$$\mathbf{y} = \Phi\mathbf{x} + \mathbf{v}, \tag{1.1}$$

where $\Phi \in \mathbb{R}^{N\times M}$ is a fixed dictionary, $\mathbf{y} \in \mathbb{R}^{N\times 1}$ is a measurement vector, $\mathbf{x} \in \mathbb{R}^{M\times 1}$ is the solution of interest, and $\mathbf{v}$ is a noise term often modeled as additive white Gaussian noise (AWGN) with variance σ^2. The goal of SSR is to estimate $\mathbf{x}$ given $\mathbf{y}$ and Φ.

In the simplest case, the linear system in Eq. (1.1) is over-determined and involves more measurements than number of unknowns ($N \geq M$). The solution can then be uniquely recovered by optimizing the least-squares objective:

$$\underset{\mathbf{x}}{\text{minimize}} \quad \|\mathbf{y} - \Phi\mathbf{x}\|_2. \tag{1.2}$$

In cases where the number of measurements is less than the number of unknowns Eq. (1.1) becomes under-determined ($N < M$). This means that for any arbitrary vector $\mathbf{t}$ in the null space $\mathcal{N}(\Phi)$, a projected solution $\mathbf{x}+\mathbf{t}$ may correspond to the same measurement $\mathbf{y} = \Phi\mathbf{x} = \Phi(\mathbf{x}+\mathbf{t})$ [8, 9, 11, 13]. For this reason, a recovered solution is not uniquely distinguishable from an infinite number of other possible solutions. SSR approaches aim to recover a *unique* solution even in under-determined settings.

One useful assumption made by SSR approaches is to constrain the solution subspace by assuming $\mathbf{x}$ is k-sparse, meaning that at most k entries of $\mathbf{x}$ are non-zero where $k \ll N$. The solution now lives in a simpler lower dimensional subspace and the SSR objective can be formulated as:

$$\underset{\mathbf{x}}{\text{minimize}} \quad \|\mathbf{y} - \Phi\mathbf{x}\|_2 + \lambda\|\mathbf{x}\|_0, \tag{1.3}$$

where $\|.\|_0$ is the ℓ_0 pseudo-norm that counts the non-zero elements in $\mathbf{x}$, and $\lambda > 0$ is a regularization term to account for noise. This objective seeks to minimize the number of non-zero elements in $\mathbf{x}$ while still satisfying linear equality conditions.

Theoretically speaking the solution to Eq. (1.3) is unique given that the underlying solution is indeed sparse [13, 14, 15]. In practice Eq. (1.3) is no longer tractable since the ℓ_0 penalty is not a convex objective and the recovery problem is NP-hard [16, 17]. Therefore,

numerous approximations to Eq. (1.3) have been proposed including 'greedy' algorithms such as the Matching Pursuit (MP) and Orthogonal Matching Pursuit (OMP). These algorithms start with an empty support set for the solution and select the non-zero elements of $\mathbf{x}$ iteratively by adding a single element to the support [18, 19, 20, 21, 22, 23]. Greedy algorithms are computationally attractive but may lead to sub-optimal solutions. Therefore, convex relaxations of the ℓ_0 penalty in Eq. (1.3) have been proposed:

$$\underset{\mathbf{x}}{\text{minimize}} \quad \|\mathbf{y} - \Phi\mathbf{x}\|_2 + \lambda g(\mathbf{x}). \tag{1.4}$$

which replaces the ℓ_0 penalty with a convex surrogate function $g(\mathbf{x})$. For example, a choice of ℓ_1 norm for $g(\mathbf{x}) = ||\mathbf{x}||_1 = \sum_{i=1}^{N} |x_i|$ is a convex function and Eq. (1.4) can now be cast as a linear program [24]. Convex relaxations are very attractive due to ease of implementation and their theoretical guarantees for uniqueness of the solution [25, 26].

1.1.1 Bayesian Sparse Signal Recovery

Alternative to the regularization-based approaches the SSR problem can be studied within a Bayesian framework [1, 7]. In this case $\mathbf{x}$ is considered a random variable with entries drawn from a probability density function $p(\mathbf{x})$. This density is commonly referred to as a *prior*. For the SSR problem, a prior $p(\mathbf{x})$ should have a peaky probability mass centered around $\mathbf{x} = 0$. At the same time it should also have heavier tails to model the non-zero coefficients in $\mathbf{x}$. In Figure 1.1a some examples sparse priors are shown including the Gaussian, Laplacian and Student's t distributions. Bayesian SSR approaches are broadly categorized into Type I and Type II.

1.1.2 Type I (MAP Estimation)

Type I approach seeks a maximum-a-posteriori (MAP) estimate of the solution $\mathbf{x}$ given the measurement $\mathbf{y}$ [27]. The distribution $p(\mathbf{x}|\mathbf{y})$ is usually unknown, therefore Bayes' rule can

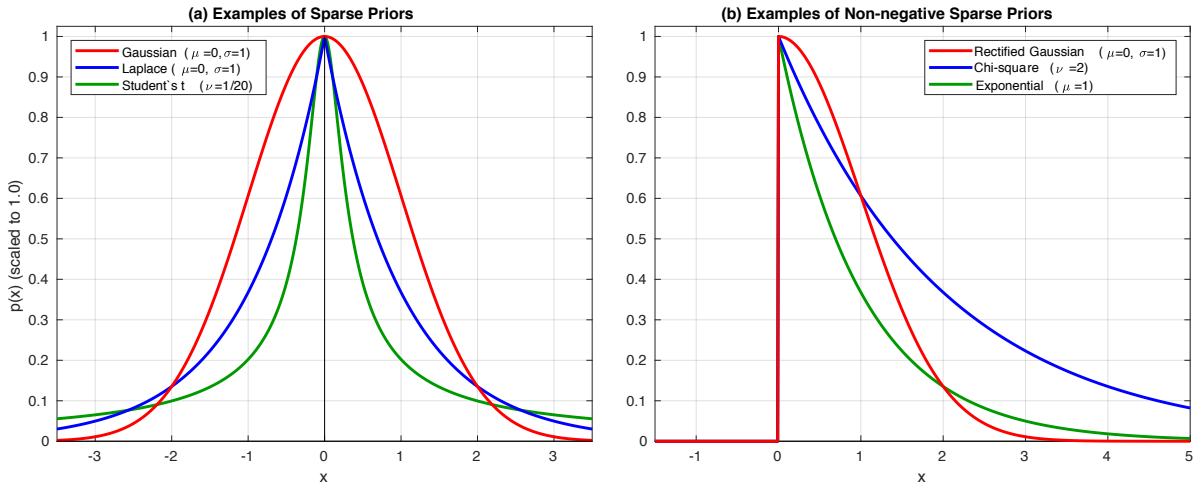


Figure 1.1: Examples of sparsity promoting probability density functions (priors) for the general SSR problem in panel (a) and for the non-negative SSR problem in panel (b). Priors that admit sharp peaks around $x = 0$ and heavier tails are better priors for the SSR problem. All densities are normalized to a maximum value of 1.0 for display.

be invoked to obtain $p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x})p(\mathbf{x})$. In this case we already know $p(\mathbf{x})$ since it is a sparsity promoting prior of our choice. Moreover, $p(\mathbf{y}|\mathbf{x})$ is a Gaussian random variable with mean $\Phi\mathbf{x}$ and variance σ^2 due to the AWGN assumption in Eq. (1.1). Most regularization-based SSR techniques discussed in the previous section can be derived using a Type I approach. For instance, assuming a Laplace prior $p(\mathbf{x}) = (\lambda/2\sigma)e^{-\lambda|\mathbf{x}|/\sigma}$ the MAP estimation is equivalent to the ℓ_1 regularization approach in Eq. (1.4) [28].

Type I method estimates the *mode* of the posterior $p(\mathbf{x}|\mathbf{y})$. If the posterior admits a bi-modal or skewed distribution then mode estimate will not represent the true probability mass of $p(\mathbf{x}|\mathbf{y})$ and will result in a sub-optimal recovery. Therefore, it becomes more important to characterize the posterior distribution of $p(\mathbf{x}|\mathbf{y})$ when it admits a more complicated distribution.

1.1.3 Properties of Prior

The choice of prior plays a very important role in the SSR problem. If a prior is more peaky around 0 with heavier tails (as shown in Figure 1.1a like the Student's t-distribution) then it is a better choice for the SSR problem. In general, priors that are higher-order (super) Gaussian

are suitable priors for the SSR problem. Assuming separable priors of the form $p(\mathbf{x}) = \prod_i p(x_i)$ such super Gaussian priors can be obtained by introducing a hierarchical hyper-parameter γ_i and using a scale mixture representation:

$$p(x_i) = \int_0^\infty p(x_i|\gamma_i)p(\gamma_i)\gamma_i, \tag{1.5}$$

where several choices for $p(x_i|\gamma_i)$ are Laplacian, Gaussian, and Power Exponential (PE) distributions [29, 30, 31]. For instance, when $p(x_i|\gamma_i) = \mathcal{N}(x; 0, \gamma_i)$ the hyper-parameter controls the variance of a zero-mean Gaussian density and when γ_i tends toward 0, sparsity in x_i is achieved.

1.1.4 Type II (Evidence Maximization)

Type II approach approximates the posterior $p(\mathbf{x}|\mathbf{y})$ by first performing a maximum likelihood (ML) estimate of the hyper-parameter γ:

$$\begin{aligned}
\hat{\gamma} &= \underset{\gamma}{\text{maximize}} \;\; p(\gamma|\mathbf{y}) \\
&= \underset{\gamma}{\text{maximize}} \;\; p(\gamma)p(\mathbf{y}|\gamma) \\
&= \underset{\gamma}{\text{maximize}} \;\; p(\gamma) \int p(\mathbf{y}|\mathbf{x})p(\mathbf{x}|\gamma)dx.
\end{aligned} \tag{1.6}$$

where the last line in Eq. (1.6) shows that $\mathbf{x}$ is integrated out (treated as a hidden-variable) to obtain $p(\mathbf{y}|\gamma)$. After an ML estimate for $\hat{\gamma}$ is obtained, the posterior is approximated as $p(\mathbf{x}|\mathbf{y}) \approx p(\mathbf{x}|\mathbf{y}; \hat{\gamma})$ and the mean of $p(\mathbf{x}|\mathbf{y}; \hat{\gamma})$ can used as the minimum-mean-square estimate for the solution.

1.1.5 Sparse Non-Negative Least Squares

In certain applications the solution in Eq. (1.1) is inherently non-negative ($\mathbf{x} \in \mathbb{R}_+^{M\times 1}$) and a non-negative solution can be obtained (assuming an under-complete Φ) by solving the

non-negative least square problem (NNLS):

$$\underset{\mathbf{x}\geq 0}{\text{minimize}} \quad \|\mathbf{y}-\Phi\mathbf{x}\|_2. \tag{1.7}$$

NNLS has a rich history in the context of methods for solving systems of linear equations [32], density estimation [33], and non-negative matrix factorization (NMF) [34]. It is widely used in applications such as text mining [35], image hashing [36], speech enhancement [37], spectral decomposition [38], and impulse response estimation [39].

For over-complete Φ ($N < M$), the regularization-based SSR techniques discussed in previous sections can be constrained to the positive orthant provide non-negative sparse solutions:

$$\underset{\mathbf{x}\geq 0}{\text{minimize}} \quad \|\mathbf{y}-\Phi\mathbf{x}\|_2+\lambda g(\mathbf{x}). \tag{1.8}$$

where $g(\mathbf{x})$ is a sparsity promoting surrogate function as defined previously. This is known as the sparse NNLS (S-NNLS) problem which arises in numerous applications [5, 10, 40, 41, 42]. In terms of the Bayesian framework, some examples of sparsity promoting priors are illustrated in Figure 1.1b including examples of rectified Gaussian, Chi-square and exponential distributions. Chapter 2 of this dissertation provides a Type II framework called Rectified Sparse Bayesian Learning (R-SBL) to solve the S-NNLS problem. R-SBL employs a novel class of priors referred to as the rectified Gaussian scale mixtures (R-GSM).

Assuming separable priors the R-GSM prior has the following form:

$$p(\mathbf{x}) = \prod_{i=1}^{M} \int_0^{\infty} \mathcal{N}^R(x_i; 0, \gamma_i) p(\gamma_i) \gamma_i \tag{1.9}$$

where $\mathcal{N}^R(x_i;\mu_i,\gamma_i)$ is a univariate rectified Gaussian density defined as:

$$\mathcal{N}^R(x_i;\mu_i,\gamma_i) = \sqrt{\frac{2}{\pi\gamma_i}} \frac{e^{-\frac{(x_i-\mu_i)^2}{2\gamma_i}}}{\operatorname{erfc}\left(-\frac{\mu_i}{\sqrt{2\gamma_i}}\right)} u(x_i), \tag{1.10}$$

where μ_i is the location parameter, γ_i is the scale parameter, $u(x_i)$ is the unit step function, and $\operatorname{erfc}(.)$ is the complementary error function.

1.2 Basics of Functional MRI

Functional MRI (fMRI) aims to understand the functional organization of the brain by interpreting neuronal activity. fMRI has been introduced less than 30 years ago when Seiji Ogawa and his colleagues discovered the blood-oxygenation-level-dependent (BOLD) contrast as a way to measure the neuronal activity using MRI scans [43]. fMRI uses the regional concentration of deoxyhemoglobin (dHb) in the brain as a natural contrast agent [43,44]. To briefly summarize the BOLD effect, when neurons in a specific brain region are recruited in accordance with a mental task they consume glucose and oxygen (O_2). Since the neurons do not have internal reserves of glucose and O_2 the body delivers a fresh supply of these nutrients by increasing the cerebral blood flow (CBF) to the involved brain regions.

A relatively small increase in the consumption in O_2 is followed by a much larger increase in local CBF causing an excess of O_2 delivered to the involved brain regions [44,45,46]. The excess O_2 causes a decrease in the regional deoxyhemoglobin (dHb) concentration since free O_2 bounds to the dHb molecules making them oxygeneted [44]. dHb has *paramagnetic* property and disrupts the local magnetic fields induced by MRI scanner due to an effect known as spin dephasing [47,48]. The reduction in local dHb concentration (and reduction of spin dephasing) yields a positive change in the observed MRI (BOLD) signal [43]. An illustration is provided in

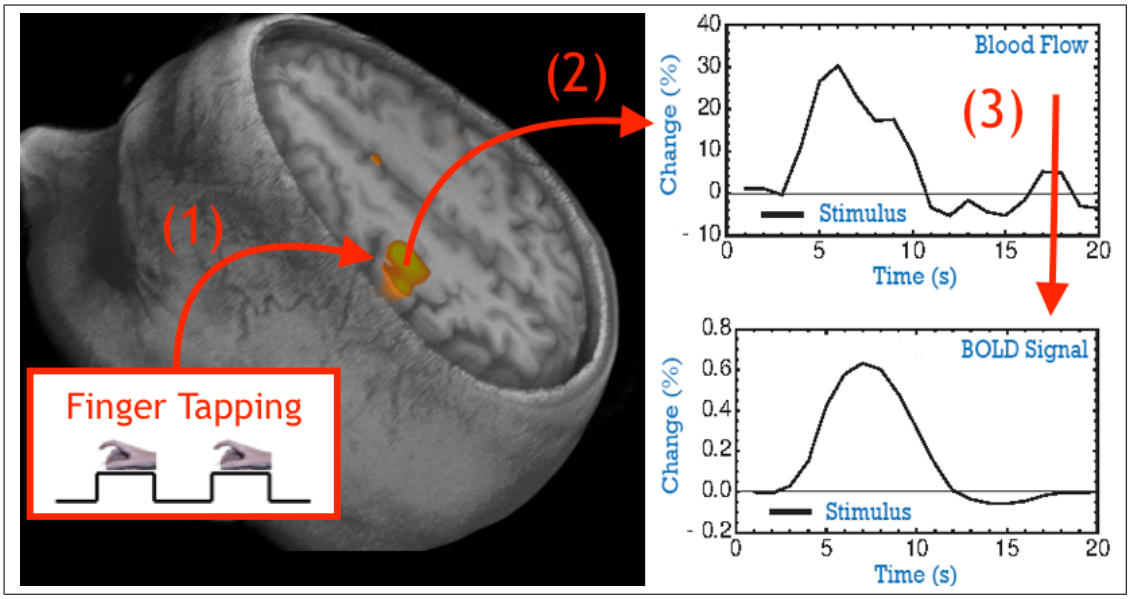


Figure 1.2: A finger tapping experiment induces a change in the local blood flow after a couple of seconds after the neurons are stimulated. This change in blood flow affects the local magnetization around the involved regions, which is then captured by the MRI scanner as the BOLD signal (Image: Courtesy of Dr. Richard Buxton).

Figure 1.2 where we depict how a finger tapping experiment causes an increase in the CBF and MRI BOLD signal in the left motor cortex of human brain.

fMRI studies can be divided into two sub categories. In *task-based* fMRI the goal is to characterize the functional organization of the brain in response to a specific task. The subject typically performs a pre-defined task (e.g. performing arithmetic calculations, finger tapping, watching a video) during the scan. In *resting-state* fMRI the objective is to characterize brain's functional organization in the absence of an explicit task. The brain involves many complicated and dynamic functional relationships even when the subject is at rest.

1.2.1 Functional Connectivity Estimation

Estimating brain's functional organization involves analyzing the relations between BOLD signals across different brain regions. The most widely used form of functional connectivity (FC) estimation is the correlation of BOLD signals across different brain regions [49, 50]. For example,

denoting a pair of zero-mean BOLD signals from two different brain regions in vector notation with x_1 and x_2, a single FC estimate is obtained by computing the Pearson correlation coefficient:

$$r = \frac{x_1^T x_2}{\|x_1\|\|x_2\|} = \cos(\theta), \tag{1.11}$$

where $\|.\|$ denotes the ℓ_2 norm and θ is the inner-angle between vectors x_1 and x_2. The correlation coefficient is bounded between $-1 \leq r \leq +1$. If r is close to $+1$ the BOLD signals from two brain regions have similar fluctuations and are involved in similar functional activity. If r is zero then brain function between these regions is considered unrelated and if r is close to -1, BOLD signals are anti-correlated with each other, meaning that activation in one brain region corresponds to a deactivation in the other region.

To give details about FC estimation approaches we first illustrate two hypothetical brain networks in the first column of Figure 1.3 as *Network A* (blue color) and *Network B* (red color). These networks contain nodes or seed locations (as shown with dots) which show that multiple brain regions can be part of a single functional network. The solid lines (vertices) with same color indicate that the seed pair at the ends of a vertex have similar BOLD signal fluctuations. There are two popular methods for obtaining correlation-based FC estimates.

1.2.2 Static FC Estimation

In *static FC estimation* the correlation coefficient is computed over the whole fMRI scan duration using all time points. This is illustrated in the second column of Figure 1.3, where a pair of BOLD seed signals measured from network A (solid blue line) and B (solid red line) are strongly anti-correlated $r = -0.72$ with each other. This means that networks A and B work in an antagonistic fashion and activation in network A corresponds to a deactivation in network B.

Instead of focusing on a single seed pair, we can correlate a seed signal with every other BOLD signal in the brain, which reveals a spatial FC map that involves varying degrees of

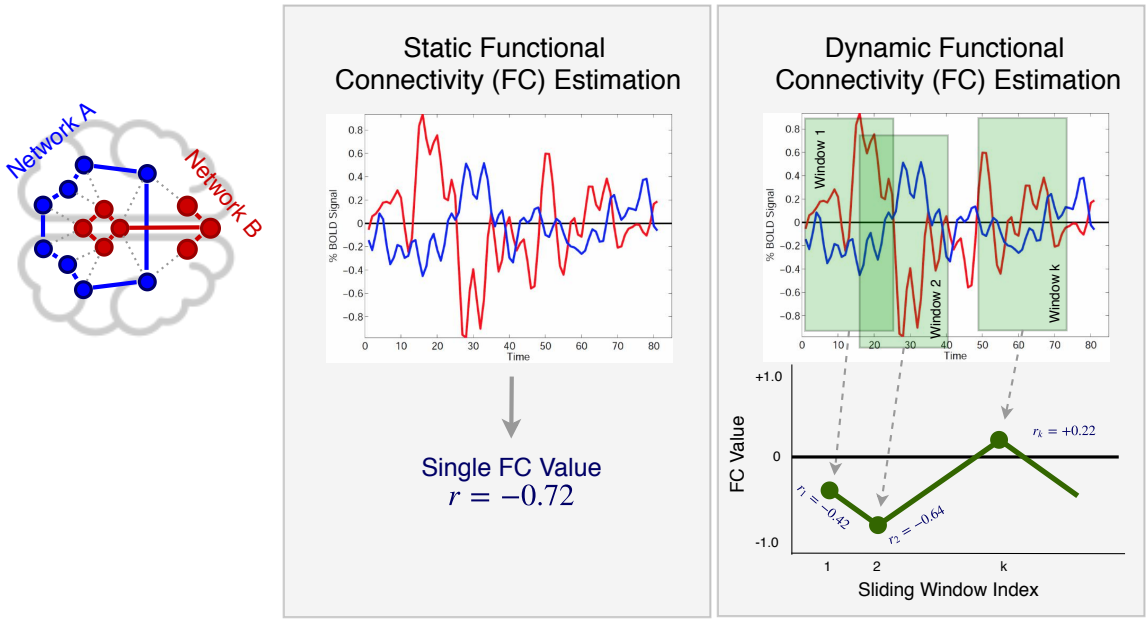


Figure 1.3: Illustration of static and dynamic functional connectivity estimation. Two hypothetical functional brain networks are illustrated on the left-hand side with labels *Network A* (blue color) and *Network B* (red color). These functional networks consist of nodes (seed locations) shown with dots, which means that more than one brain region can be part of a single network. The vertices (solid lines) between nodes illustrate the functional relationship between seed locations. A vertex between two nodes indicates the correlation value computed between the seed signals is similar. In the second column, static FC estimation is illustrated. The correlation coefficient between seed signals from network A (solid blue line) and B (solid red line) is computed across the full scan duration and used as an estimate for FC. The correlation value $r = -0.72$ indicates that two networks are anti-correlated and work in an antagonistic fashion. In the third column, dynamic FC (DFC) estimation is illustrated. In DFC estimation, shorter duration sliding-windows are used to obtain a set of correlation values from networks A and B. The set of correlation values represent estimates of dynamic changes in brain function.

correlation. Examples of FC maps are provided in Figure 1.4a,b for a single fMRI scan using a seed signal from the posterior cingulate cortex (PCC). The FC map in panel (a) is noisy due to various nuisance effects which will be discussed later in Section 1.2.4. The FC map in panel (b) is cleaner as compared to panel (a) since the fMRI data is cleaned prior to computing the correlations between the PCC seed.[1]

In Figure 1.4c,d we computed the average FC maps obtained across 68 fMRI scans. In

[1]Note that we capture BOLD signals over a 3D brain volume. FC maps are usually visualized using a 2D brain-slice format as shown in Figure 1.4. The slices from left to right in each panel correspond to 2D slices sequentially taken from bottom to top parts of the brain.

panel (d) the average FC maps display both positive and negative correlations between the PCC seed signal and the BOLD signal across different brain regions. The correlation-based approach have led to the identification of a number of functional brain networks and provided indispensable insights into brain function.

For instance the red regions in panel (c) correspond to the default-mode network (DMN) of the human brain, which is involved in a set of complex neuronal functions such as mind-wandering [49]. The blue regions in panel (c) braoadly include the task-positive network (TPN) which is activated when the brain is involved in performing a specific task. DMN and TPN are considered as anti-correlated networks [51, 52, 53, 54].

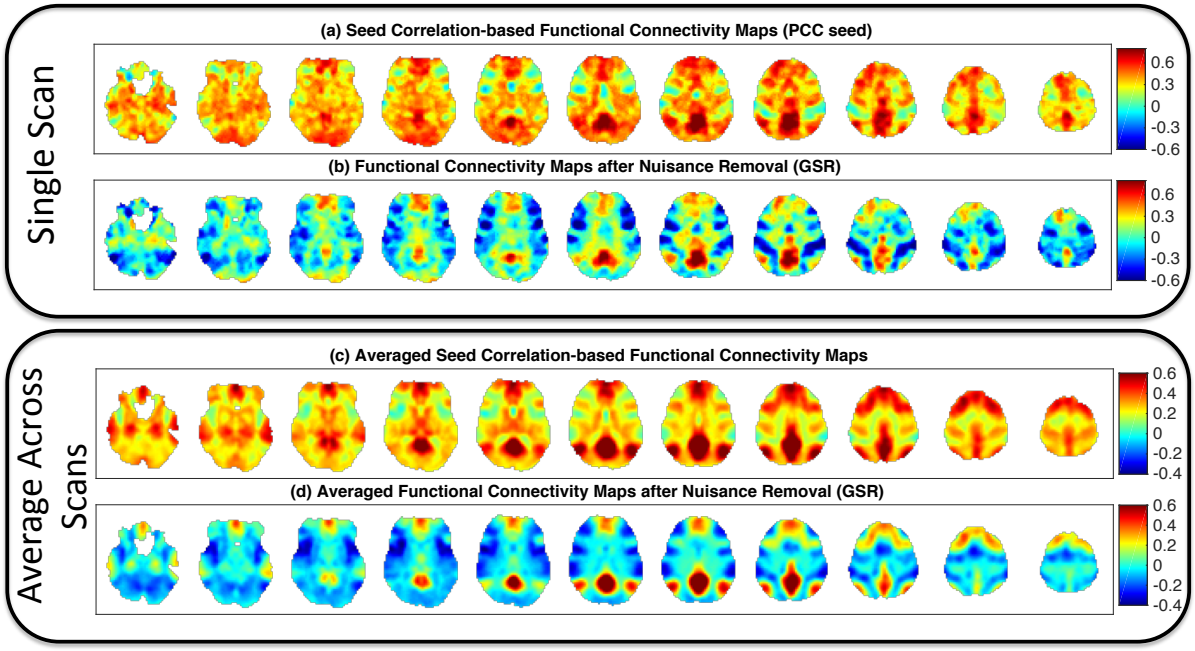


Figure 1.4: Examples of functional connectivity maps obtained for a single scan (panel a) and averaged across 68 scans (panel c). These maps are obtained by correlating a BOLD seed signal from the posterior cingulate cortex (PCC) and every other voxel in the brain. Nuisance effects contaminate the spatial correlations making it hard to observe functional network structures. In panels (b) and (d) we show the maps after performing global signal regression (a nuisance removal approach). The maps are cleaner with increased spatial specificity. The red areas are regions that are positively correlated with the PCC seed. These regions are known as the default mode network of the brain. The blue regions are anti-correlated with the PCC seed and are referred to as the task-positive network.

One drawback of the static FC estimation approach is the assumption of temporal stationarity. To elaborate a single correlation metric computed over the whole scan can only represent a temporally averaged measure of brain connectivity over a typical scan duration of 5-10 minutes. This may lead to failure in capturing important *dynamic* temporal changes in FC measures [55, 56, 57].

1.2.3 Dynamic FC Estimation

An alternative approach uses shorter duration temporal windows (30-60 seconds) to obtain a set of correlation-based FC estimates [56, 57]. This approach is called sliding-window *dynamic FC estimation* and is illustrated in the third column of Figure 1.3. In this approach, the BOLD signals from networks A and B are first windowed across time (green rectangles) and the correlation coefficient is computed for each window separately. The window is then shifted by a certain amount of time samples and a new correlation value is computed for the next window. This process is repeated until all sliding-window correlations are carried out across the scan duration. This approach can capture the dynamic changes in the FC estimates between networks A and B.

1.2.4 Nuisance Regression

The BOLD signal is an indirect measure of neural activity and shows a rather complex relation to CBF and metabolism [58]. The signal-to-noise ratio (SNR) in a typical fMRI experiment is inherently low since the neuronal changes in the BOLD signal happen in small scales and a number of noise factors contaminate the neuronal signal of interest.

These noise factors can be broadly classified in two categories as scanner related noise and physiological noise. Scanner related noise factors are typically due to the MRI system or to the interaction between the MRI scanner and subject. This includes thermal noise, radio-frequency

(RF) noise such as RF spikes, scanner drift, and other low-frequency fluctuations [59]. On the other hand, physiological noise includes cardiac noise such as cardiac pulsations that give rise to non-neuronal changes in the BOLD signal, respiratory noise such as perturbations of the main magnetic field due to chest cavity expansions and contractions, and subject motion [59, 60, 61, 62, 63, 64]. If these confounds are not removed from the BOLD signal prior to analysis, they can lead to an increase in the number of false positives and negatives causing erroneous interpretations of the fMRI results [65]. The effect of noise on the correlation-based FC estimates was previously illustrated in Figure 1.4a for a single subject.

In order to remove the nuisance confounds present in the fMRI data, a collection of nuisance terms are used to form a nuisance matrix $\mathbf{N}$, where each column of this matrix consists of a single nuisance term. Typically, columns of $\mathbf{N}$ include Legendre polynomials (e.g. linear and quadratic terms) to model scanner drift and low-frequency artifacts, physiological nuisance regressors which include measures of cardiac and respiratory activity obtained using a pulse oximeter and a respiratory effort transducer, respectively, and head motion measurements obtained during the image registration process.

Other nuisance regressors can be obtained from the raw BOLD data by using the signals derived from non-neuronal brain regions such as the white-matter (WM) and cerebrospinal fluid (CSF) regions. The signals from WM and CSF regions mostly represent an ensemble of various nuisance terms including the effects of scanner related, physiological and other noise.

To remove the contribution of these nuisance terms, a general linear model (GLM) of the BOLD signal $\mathbf{x}$ is considered:

$$\mathbf{x} = \mathbf{N}\mathbf{c} + \hat{\mathbf{x}} \tag{1.12}$$

where $\mathbf{c}$ is a set of regression coefficients to be estimated, and $\hat{\mathbf{x}}$ is the 'cleaner' BOLD signal that cannot be linearly modeled by the nuisance terms. In resting-state fMRI we seek to obtain $\hat{\mathbf{x}}$ by

projecting out the contributions of nuisance regressors as follows:

$$\hat{\mathbf{x}} = \mathbf{x} - \mathbf{N}\mathbf{c} \tag{1.13}$$

$$= \mathbf{x} - \mathbf{N}(\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T\mathbf{x} \tag{1.14}$$

where $\mathbf{c} = (\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T\mathbf{x}$ can be obtained using a least squares fit.

Global Signal Regression: A significant portion of resting-state fMRI studies perform global signal regression (GSR) in which a whole-brain averaged BOLD signal known as the global signal (GS) is treated as a "catch-all" type of nuisance regressor. The underlying assumption of the GS is that nuisance factors in average will dominate over the smaller scale neuronal signals and GS will largely represent nuisance effects present in fMRI data. The GS is computed as $\mathbf{g} = (1/K)\sum_{k=1}^{K}\mathbf{x}_k$ where K is the number of BOLD signals in the brain volume (typically $K \approx 30,000$). Examples of correlation-based FC maps after performing GSR were provided in Figure 1.4c,d. GSR greatly improves the spatial specificity of the FC maps and reveals the anti-correlated networks, which are otherwise obscured by uniform noise-like correlation values [54].

Most of this dissertation (Chapters 3-5) is dedicated to analyze the effects of nuisance terms and the limitation of nuisance regression in static and dynamic FC studies. We will show that geometric ℓ_2 norm of various regressors will be significantly related to the correlation-based FC estimates both in single scans (dynamic FC) and across a collection of scans (static FC). We will derive a theoretical limitation of linear regression which will prevent the removal of nuisance norm effects from the correlation values.

1.3 Outline and Contributions

Chapter 2 studies the sparse non-negative least squares problem from a Bayesian perspective. We introduce a Type II Bayesian framework to recover non-negative sparse solutions in under-determined systems of linear equations. Our main contributions are summarized as follows:

- We propose a novel class of sparsity promoting priors called Rectified Gaussian Scale Mixtures (R-GSM) to model the non-negative solution of interest $p(x_i) = \int_0^\infty p(x_i|\gamma_i)p(\gamma_i)\gamma_i$. The R-GSM prior utilizes a rectified Gaussian density for each mixture component $p(x_i|\gamma_i) = \mathcal{N}^R(0,\gamma_i)$ and encompasses other sparse priors by using a different mixing density $p(\gamma_i)$.

- We detail how the R-GSM prior can be used to solve the sparse non-negative least squares (S-NNLS) problem. We name this technique as rectified sparse Bayesian learning (R-SBL) and provide four expectation maximization (EM) based R-SBL variants that offer a range of options for computational complexity and recovery performance.

- We provide an extensive set of experiments to show the superiority of R-SBL for various distributions of the solution and dictionary coefficients. We demonstrate that R-SBL is extremely robust to the structure of the dictionary and performs very well even when the columns of the dictionary are non-Gaussian or non independent and identically distributed (i.i.d.).

- We provide example applications of R-SBL on real data, including a face recognition application based on sparse representation classifiers, and another application on functional MRI data to estimate sparsely repeating spatio-temporal patterns of brain activations.

Chapter 3 investigates the temporal effects of global signal regression (GSR) on the fMRI data. GSR has been a controversial nuisance removal approach in resting-state fMRI since it introduces artifactual negative correlations between BOLD signals across different brain regions.

This has led to a questioning of the existence of anti-correlations observed between functional brain networks [51, 52]. Chapter 3 provides simpler mathematical approximations to GSR to characterize its main effects on the fMRI data. Our main contributions are summarized as follows:

- We introduce a new mathematical framework to approximate the effects of GSR on the fMRI data. We show that the GSR can be well approximated by a temporal downweighting of the fMRI data, where the weighting factor varies with time but is uniform across space. The weighting factor decreases with the magnitude of the global signal, so that fMRI data from time points with large global signal magnitudes are greatly attenuated whereas those time points with small global signal magnitudes remain mostly unaffected.

- We introduce a second GS censoring (GSC) approximation to GSR in which the fMRI data is set to zero for time points where the GS magnitude exceeds a specified threshold. GSC approach partitions the fMRI data into two orthogonal temporal subsets. A temporal subset that corresponds to small GS magnitudes and a complementary subset that corresponds to large GS magnitudes.

- We show that anti-correlated networks are inherent to brain's functional organization and are not simply an artifact of GSR. These anti-correlated networks reside in a temporal subset of the fMRI data where the GS magnitudes are relatively small.

Chapter 4 investigates the effects of nuisance terms on the sliding-window correlations (DFC estimates) obtained between pairs of fMRI signals. We show that nuisance regression is largely ineffective in removing the relation between the nuisance terms and DFC estimates due to a theoretical bound of linear regression on the correlation coefficients obtained before and after regression. Our specific contributions are as follows:

- We reveal that sliding-window correlations (DFC estimates) obtained between pairs of fMRI signals (x_1 and x_2) are significantly correlated with sliding-window fluctuations in the ℓ_2 norm of various nuisance regressors.

- We demonstrate that significant correlations between the DFC estimates and nuisance norms can persist even after performing nuisance regression. We derive a theoretical bound of linear regression that limits the difference in the correlation values obtained before and after nuisance regression:

$$|\tilde{r} - r| \leq 2\left(\frac{1-\sqrt{|n_O|^2/|n|^2}}{1+\sqrt{|n_O|^2/|n|^2}}\right) \tag{1.15}$$

  where $r = x_1^T x_2/(|x_1||x_2|)$ and $\tilde{r} = \tilde{x}_1^T \tilde{x}_2/(|\tilde{x}_1||\tilde{x}_2|)$ are the correlation values obtained before and after nuisance regression, respectively. The metric $|n_O|^2/|n|^2$ is the relative amount of energy of a nuisance regressor n which is orthogonal to both x_1 and x_2.

- This work highlights the fact that while linear regression can model the dependencies between a regressor time course and multiple measurements, it does not consider the effects of a regressor on the relationship between multiple measurements. We provide a potential nuisance norm regression (NNR) approach to remove nuisance norm fluctuations from the correlation values.

Chapter 5 provides an analysis of nuisance effects on the variability of static FC estimates across different scans both before and after nuisance regression. A better understanding of nuisance effects in comparing different scans is critical considering the increasing use of fMRI to examine the differences in FC measures between healthy controls and important disease populations such as Alzheimer's disease [66, 67, 68], Parkinson's disease [69], depression [70], and schizoprenia [71]. Our specific contributions are as follows:

- We show that inter-scan variations in FC estimates are significantly correlated with the norms of various nuisance terms including head motion measurements, signals derived from the white-matter and cerebrospinal regions and the whole-brain global signal (GS) both before and after nuisance regression. We introduce nuisance contamination maps,

which illustrate the spatial distribution and the extent of nuisance contamination present in the FC estimates across scans.

- We demonstrate that performing nuisance regression using non-GS regressors to be largely ineffective in reducing the significant correlations observed between FC estimates and nuisance norms.

- We generalize the theory developed in Chapter 4 to static FC measures and confirm the validity of the theoretical limitation of nuisance regression for correlation-based static FC estimates. Finally, we provide a detailed analysis of the limited efficacy of GSR and show how GSR can introduce GS norm-related fluctuations in the FC estimates across scans.

Chapter 2

Rectified Gaussian Scale Mixtures and the Sparse Non-Negative Least Squares Problem

In this chapter, we develop a Bayesian evidence maximization framework to solve the sparse non-negative least squares (S-NNLS) problem. We introduce a family of probability densities referred to as the Rectified Gaussian Scale Mixture (R-GSM) to model the sparsity enforcing prior distribution for the solution. The R-GSM prior encompasses a variety of heavy-tailed densities such as the rectified Laplacian and rectified Student-t distributions with a proper choice of the mixing density. We utilize the hierarchical representation induced by the R-GSM prior and develop an evidence maximization framework based on the Expectation-Maximization (EM) algorithm. Using the EM based method, we estimate the hyper-parameters and obtain a point estimate for the solution. We refer to the proposed method as rectified sparse Bayesian learning (R-SBL). We provide four R-SBL variants that offer a range of options for computational complexity and the quality of the E-step computation. These methods include the Markov chain Monte Carlo EM, linear minimum mean-square-error estimation, approximate message passing and a

diagonal approximation. Using numerical experiments, we show that the proposed R-SBL method outperforms existing S-NNLS solvers in terms of both signal and support recovery performance, and is very robust against the structure of design matrix. At the of this chapter, we apply R-SBL to functional MRI data and present an approach for estimating repeating spatiotemporal brain activation patterns.

2.1 Introduction

This work considers the following signal model

$$\mathbf{y} = \Phi\mathbf{x} + \mathbf{v}, \tag{2.1}$$

where the solution vector $\mathbf{x} \in \mathbb{R}_+^M$ is assumed to be non-negative, the matrix $\Phi \in \mathbb{R}^{N\times M}$ is fixed and obtained from the physics of the underlying problem, $\mathbf{y} \in \mathbb{R}^N$ is the measurement, and $\mathbf{v}$ is the additive noise modeled as a zero mean Gaussian with uncorrelated entries $v_i \sim \mathcal{N}(0,\sigma^2)$.

Recovering $\mathbf{x}$ using the signal model in Eq. (2.1) is known as solving the non-negative least squares (NNLS) problem. NNLS has a rich history in the context of methods for solving systems of linear equations [32], density estimation [33], and non-negative matrix factorization (NMF) [5, 34, 41, 42]. NNLS is also widely used in text mining [35], image hashing [36], speech enhancement [37], spectral decomposition [38], magnetic resonance chemical shift imaging [72], and impulse response estimation [39].

The maximum-likelihood solution for the signal model in Eq. (2.1) is given by

$$\underset{\mathbf{x}\geq\mathbf{0}}{\text{minimize}} \quad \|\mathbf{y} - \Phi\mathbf{x}\|_2. \tag{2.2}$$

In many applications, $N < M$ and Eq. (2.1) is under-determined. This means that a unique solution for $\mathbf{x}$ may not exist. Recovering a unique solution is possible if more information is

known *a-priori* about the solution vector. For example, a useful assumption is that the solution vector is *sparse* and contains only a few non-zero elements [73, 74, 75]. In this case, the sparsest solution (assuming a noiseless case) can be recovered by modifying Eq. (2.2) to

$$\underset{\mathbf{x}\geq\mathbf{0},\ \mathbf{y}=\Phi\mathbf{x}}{\text{minimize}} \quad \|\mathbf{x}\|_0, \tag{2.3}$$

where $\|.\|_0$ is the ℓ_0 pseudo-norm, which counts the non-zero elements in $\mathbf{x}$. The count of non-zero elements is also referred to as the cardinality of the solution. Then, the recovery objective in Eq. (2.3) is to minimize the cardinality of $\mathbf{x}$ while satisfying the optimization constraints. This approach is commonly referred to as solving the sparse NNLS (S-NNLS) problem.

The S-NNLS problem is becoming increasingly popular in certain applications where the non-negative solution needs to be recovered from a limited number of measurements. For example, in [40] an S-NNLS method was applied to magnetic resonance imaging (MRI) data to reconstruct narrow fiber-crossings from a limited number of acquisitions. In [76], another method was used to uncover regulatory networks from micro-array mRNA expression profiles from breast cancer data. In [77, 78], an S-NNLS method was applied to functional MRI data to estimate sparsely repeating spatio-temporal activation patterns in the human brain. S-NNLS solvers are also used in applied mathematics for designing dictionaries for sparse representations, such as sparse NMF and non-negative K-SVD [5, 10].

The objective function in Eq. (2.3) is not tractable since the ℓ_0 penalty is not convex and the problem is NP-hard [16, 17]. Therefore, 'greedy' algorithms have been proposed to approximate the solution [19, 20, 21, 22, 23]. An example is the class of algorithms known as Orthogonal Matching Pursuit (OMP) [18, 19], which greedily selects the non-zero elements of $\mathbf{x}$. In order to adapt OMP to the S-NNLS problem, the criterion by which a new non-zero element of $\mathbf{x}$ is selected is modified to select the one having the largest *positive* value [23]. Another approach in this class of algorithms finds an $\mathbf{x}$ such that $\|\mathbf{y}-\Phi\mathbf{x}\|_2 \leq \varepsilon$ and $\mathbf{x} \geq 0$ using the active-set

Lawson-Hanson algorithm [32] and then prunes $\mathbf{x}$ until $\|\mathbf{x}\|_0 \leq K$, where K is a pre-specified cardinality [5].

Greedy algorithms are computationally attractive but may lead to sub-optimal solutions. Therefore, convex relaxations of the ℓ_0 penalty have been proposed [17,79,80,81,82]. One simple alternative replaces the ℓ_0 norm with the ℓ_1 norm and reformulates the problem in Eq. (2.3) as

$$\underset{\mathbf{x}\geq\mathbf{0}}{\text{minimize}} \quad \|\mathbf{y}-\Phi\mathbf{x}\|_2+\lambda\|\mathbf{x}\|_1, \tag{2.4}$$

where $\lambda > 0$ is a regularization parameter to account for the measurement noise. The advantage of the formulation in Eq. (2.4) is that it is a convex optimization problem and can be solved by a number of methods [82, 83, 84, 85]. One approach is to estimate $\mathbf{x}$ with projected gradient descent [86]. In fact, the ℓ_1 penalty in Eq. (2.4) can be replaced by any arbitrary sparsity inducing surrogate function $g(\mathbf{x})$, thus leading to alternative methods based on solving

$$\underset{\mathbf{x}\geq\mathbf{0}}{\text{minimize}} \quad \|\mathbf{y}-\Phi\mathbf{x}\|_2+\lambda g(\mathbf{x}). \tag{2.5}$$

For example, a surrogate $g(\mathbf{x}) = \sum_{i=1}^{M} \log\left(x_i^2+\beta\right)$ leads to an iterative reweighted optimization approach [87, 88].

A promising view on the S-NNLS problem is to cast the entire problem in a Bayesian framework and consider the maximum a-posteriori (MAP) estimate of $\mathbf{x}$ given $\mathbf{y}$

$$\mathbf{x}_{MAP} = \arg\max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}). \tag{2.6}$$

There is a strong connection between the MAP framework and the previous deterministic formulations. Recently, it has been shown that formulations of the form in Eq. (2.5) can be represented by using the formulation in Eq. (2.6) with a proper choice of $p(\mathbf{x})$ [27]. For example, considering

a separable $p(\mathbf{x})$ of the form

$$p(\mathbf{x}) = \prod_{i=1}^{M} p(x_i), \tag{2.7}$$

the ℓ_1 regularization approach in Eq. (2.4) (i.e. a choice of $g(\mathbf{x}) = \|\mathbf{x}\|_1$ in Eq. (2.5)) is equivalent to the Bayesian formulation in Eq. (2.6) with an exponential prior for x_i. In this work our emphasis will be on Bayesian approaches for solving Eq. (2.1).

2.1.1 Contributions of the paper

- We introduce a family of non-negative probability densities referred to as the rectified Gaussian scale mixture (R-GSM) to model non-negative and sparse solutions.
- We discuss how the R-GSM prior encompasses other sparsity inducing non-negative priors, such as the rectified Laplacian and rectified Student-t distributions through a proper choice of the mixing density.
- We detail how the R-GSM prior can be utilized to solve the S-NNLS problem using an evidence maximization based estimation procedure that utilizes the expectation-maximization (EM) framework. We refer to this technique as rectified sparse Bayesian learning (R-SBL).
- We provide four alternative R-SBL methods that offer a range of options for computational complexity and the quality of the E-step computation. These methods include the Markov Chain Monte Carlo EM, linear minimum mean-square-error estimation, approximate message passing and a diagonal approximation.
- We use extensive empirical results to show the robustness and superiority of the R-GSM priors and R-SBL algorithm for the S-NNLS problem. Especially, under various i.i.d. and non-i.i.d. settings for the design matrix Φ.

2.1.2 Organization of the paper

In Section 2.2, we discuss the advantages of using scale mixture priors for $p(\mathbf{x})$ and introduce the R-GSM prior. In Section 2.3, we define the Type I and Type II Bayesian approaches to solve the S-NNLS problem and introduce the R-SBL framework. We provide the details of an evidence maximization based estimation procedure in Section 2.3.2. We present empirical results comparing the proposed R-SBL algorithm to the baseline S-NNLS solvers in Section 2.5.

2.2 Rectified Gaussian Scale Mixtures

We assume separable priors of the form in Eq. (2.7) and focus on the choice of $p(x_i)$. The choice of prior plays a central role in the Bayesian inference [1, 89, 90]. For the S-NNLS problem, the prior must induce sparsity and satisfy the non-negativity constraints. Consequently, we consider the hierarchical scale mixture prior

$$p(x_i) = \int_0^\infty p(x_i|\gamma_i)p(\gamma_i)d\gamma_i. \tag{2.8}$$

The scale mixture prior was first considered in the form of Gaussian Scale Mixtures (GSM) with $p(x_i|\gamma_i) = \mathcal{N}(x_i; 0, \gamma_i)$ [91]. Super-gaussian densities are suitable priors for promoting sparsity [1, 7] and can be represented in the form shown in Eq. (2.8) with a proper choice of mixing density $p(\gamma_i)$ [29, 92, 93, 94, 95]. This has made scale mixture priors valuable for the standard sparse signal recovery problem. Another advantage of the scale mixture prior is that, it establishes a Markovian structure of the form

$$\boldsymbol{\gamma} \rightarrow \mathbf{x} \rightarrow \mathbf{y}, \tag{2.9}$$

where inference can be performed in the $\mathbf{x}$ domain (referred to as Type I) and in the $\boldsymbol{\gamma}$ domain (Type II). Experimental results for the standard sparse signal recovery problem show that performing

inference in the γ domain consistently achieves superior performance [1, 27, 96, 97].

The Type II procedure involves finding a maximum-likelihood (ML) estimate of γ using evidence maximization and approximating the posterior $p(\mathbf{x}\,|\,\mathbf{y})$ by $p(\mathbf{x}\,|\,\mathbf{y},\gamma_{\text{ML}})$. The performance gains can be understood by noting that γ is deeper than $\mathbf{x}$ in Eq. (2.9), so the influence of errors in performing inference in the γ domain may be diminished [27, 96]. Also, γ is close enough to $\mathbf{y}$ such that meaningful inference about γ can still be performed, mitigating the problem of local minima that is more prevalent when seeking a Type I estimate of $\mathbf{x}$ [96].

Although priors of the form shown in Eq. (2.8) have been used in the compressed sensing literature (where the signal model is identical to Eq. (2.1) without the non-negativity constraint) [27, 98, 99], such priors have not been extended to solve the S-NNLS problem. Considering the findings that the scale mixture prior has been useful for the development of sparse signal recovery algorithms [27, 96, 100], we propose a R-GSM prior for the S-NNLS problem, where $p(x_i|\gamma_i)$ in Eq. (2.8) is a rectified Gaussian (RG) distribution. We refer to the proposed Type II inference framework as R-SBL.

The univariate RG distribution is defined as

$$\mathcal{N}^R(x;\mu,\gamma) = \sqrt{\frac{2}{\pi\gamma}}\frac{e^{-\frac{(x-\mu)^2}{2\gamma}}u(x)}{\text{erfc}\left(-\frac{\mu}{\sqrt{2\gamma}}\right)}, \tag{2.10}$$

where μ is the location parameter (and not the mean), γ is the scale parameter, $u(x)$ is the unit step function, and $\text{erfc}(x)$ is the complementary error function[1].

As noted in previous works [101, 102], closed form inference computations using a multivariate RG distribution are tractable only if the location parameter is zero (by effectively getting rid of the $\text{erfc}(.)$ term). Although a non-zero μ could provide a richer class of priors, possibly to model approximately sparse or non-sparse solutions, considering the tractability issues

[1] $\text{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}dt$

and the potential overfitting problems (twice as many parameters), we focus on the R-GSM priors with $\mu = 0$ to promote *sparse* non-negative solutions. It is a pragmatic choice and adequate for the problem at hand. When $\mu = 0$, the RG density simplifies to

$$\mathcal{N}^R(x;0,\gamma) = \sqrt{\frac{2}{\pi\gamma}} e^{-\frac{x^2}{2\gamma}} u(x). \tag{2.11}$$

Thus, the R-GSM prior introduced in this work have the form

$$p(x) = \int_0^\infty \mathcal{N}^R(x;0,\gamma) p(\gamma) d\gamma. \tag{2.12}$$

Different choices of $p(\gamma)$ lead to different options for $p(x)$ and some examples are presented below.

2.2.1 R-GSM representation of sparse priors

We can utilize the proposed R-GSM framework to obtain a variety of non-negative sparse priors. For instance, consider the rectified Laplace prior $p(x) = \lambda e^{-\lambda x} u(x)$. By using an exponential prior for $p(\gamma) = \frac{\lambda^2}{2} e^{-\frac{\lambda^2 \gamma}{2}} u(\gamma)$, we can express $p(x)$ in the R-GSM framework as [103]

$$p(x) = 2u(x) \int_0^\infty \mathcal{N}(x|0,\gamma) \frac{\lambda^2}{2} e^{-\frac{\lambda^2\gamma}{2}} u(\gamma) d\gamma = \lambda e^{-\lambda x} u(x). \tag{2.13}$$

Similarly, by considering a Gamma(a,b) distribution for $p(\gamma)$, we obtain a rectified Student-t distribution for $p(x)$ and Eq. (2.8) simplifies to [1]

$$p(x) = 2u(x) \int_0^\infty \mathcal{N}(x|0,\gamma) \frac{\gamma^{a-1} e^{\frac{-\gamma}{b}}}{a^b \Gamma(\mathsf{a})} d\gamma \tag{2.14}$$

$$= \frac{2b^a \Gamma(a+\frac{1}{2})}{(2\pi)^{\frac{1}{2}} \Gamma(a)} \left(b + \frac{x^2}{2}\right)^{-(a+\frac{1}{2})} u(x), \tag{2.15}$$

where Γ is defined as $\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}dt$. More generally, all of the distributions represented by the GSM family have a corresponding rectified version represented by the R-GSM family (e.g. contaminated Normal and slash densities, symmetric stable and logistic, hyperbolic, etc.) [29, 91, 92, 93, 94, 95].

2.2.2 Relation to other Bayesian works

In [101], a modified Gaussian prior was considered for the NNLS problem. The authors used a Gaussian prior of arbitrary mean and variance and performed non-negative rectification using a 'cut' function. Their goal was to better represent *non-sparse* signals by avoiding the selection of $\mu = 0$, as we consider in our work. Our R-GSM prior substantially differs from this work as we consider a *mixture* of zero-location RG distributions for the prior, as opposed to a single Gaussian density with the 'cut' rectification. Our design objective is to induce sparsity by using a hierarchical hyper-parameter γ.

In [104], a non-negative generalized approximate message passing (GAMP) approximation was proposed, using a Bernoulli non-negative Gaussian mixture prior of arbitrary location and scale parameters. This extends the prior given in [101] but uses a fixed number of mixture components e.g. $L = 3$. The sparsity is enforced by using a Dirac delta function and an additional sparsity rate λ that would 'favor' the Dirac function and attenuate other mixture components simultaneously. The authors infer a bulk of parameters including the scale, location, and mixture weights as well as the sparsity rate simultaneously. Our R-SBL approach differs from [104] as we only consider a single sparsity inducing hyper-parameter vector γ, and our mixture components are strictly located at zero. Our approach simplifies the overall inference procedure and the problem formulation. We also consider an infinite number of mixture components as opposed to considering a fixed number of components.

Finally, we consider a more general class of priors than the existing methods since the R-GSM prior is based on an arbitrary mixing density $p(\gamma)$. As indicated in Section 2.2.1, different

selections of $p(\gamma)$ lead to more flexible and generalized priors for the sparse solution.

2.3 Bayesian Inference with Scale Mixture Prior

We detail the Type I and Type II methods for solving the S-NNLS problem with the R-GSM prior. Though this paper is dedicated to Type II estimation because of its superior performance in sparse signal recovery problems [27, 96], we briefly introduce Type I in the following section for the sake of completeness.

2.3.1 Type I estimation

Using Type I to solve the S-NNLS problem translates into calculating the MAP estimate of $\mathbf{x}$ given $\mathbf{y}$

$$\arg\min_{\mathbf{x}} \| \mathbf{y} - \Phi \mathbf{x} \|_2^2 - \lambda \sum_{i=1}^{M} \ln p(x_i). \tag{2.16}$$

Some of the ℓ_0 relaxation methods described in Section 2.1 can be derived from a Type I perspective. For instance, by choosing an exponential prior for $p(x_i)$, Eq. (2.16) reduces to the ℓ_1 regularization approach in Eq. (2.4) with the interpretation of λ as being determined by the parameters of the prior and the noise variance. Similarly, by choosing a Gamma prior for $p(x_i)$, Eq. (2.16) reduces to

$$\arg\min_{\mathbf{x}} \| \mathbf{y} - \Phi \mathbf{x} \|_2^2 + \lambda \sum_{i=1}^{M} \ln \left(b + \frac{x_i^2}{2} \right), \tag{2.17}$$

which leads to the reweighted ℓ_2 approach to the S-NNLS problem described in [87,88]. A unified Type I approach for the R-GSM prior can be readily derived using the approaches discussed in [27, 29].

2.3.2 Type II estimation

The Type II framework involves finding a ML estimate of $\boldsymbol{\gamma}$ using evidence maximization and approximating the posterior $p(\mathbf{x}\,|\,\mathbf{y})$ by $p(\mathbf{x}\,|\,\mathbf{y},\boldsymbol{\gamma}_{\mathrm{ML}})$. Then, appropriate point estimates and the solution $\mathbf{x}$ can be obtained. We refer to this approach as the *rectified sparse Bayesian learning* (R-SBL).

Several strategies exist for estimating $\boldsymbol{\gamma}$. The first strategy considers the problem of forming a ML estimate of $\boldsymbol{\gamma}$ given $\mathbf{y}$ [1, 27, 105, 106]. In our case, $p(\boldsymbol{\gamma}\,|\,\mathbf{y})$ does not admit a closed form expression making this strategy difficult. The second strategy investigated here, aims to estimate $\boldsymbol{\gamma}$ by using the EM algorithm [27, 98, 106]. In the EM approach, we treat $(\mathbf{x},\mathbf{y},\boldsymbol{\gamma})$ as the complete data and $\mathbf{x}$ as the hidden variable. Utilizing the current estimate $\boldsymbol{\gamma}'$, where t refers to the iteration index, the expectation step (E-step) involves finding the expectation of the log-likelihood, $Q(\boldsymbol{\gamma},\boldsymbol{\gamma}')$ given by

$$Q(\boldsymbol{\gamma},\boldsymbol{\gamma}') = E_{\mathbf{x}|\,\mathbf{y};\boldsymbol{\gamma}'}\left[\ln p(\mathbf{y}|\mathbf{x}) + \ln p(\mathbf{x}|\boldsymbol{\gamma}) + \ln p(\boldsymbol{\gamma})\right] \tag{2.18}$$

$$\doteq \sum_{i=1}^{M} E_{\mathbf{x}|\,\mathbf{y};\boldsymbol{\gamma}'}\left[-\frac{1}{2}\ln\gamma_i - \frac{x_i^2}{2\gamma_i} + \ln p(\gamma_i)\right], \tag{2.19}$$

where $\doteq$ indicates that constant terms, and terms that do not depend on $\boldsymbol{\gamma}$ have been dropped since they do not affect the consequent M-step. For simplicity, we assume a non-informative prior on $\boldsymbol{\gamma}$ [1]. In the M-step, we maximize $Q(\boldsymbol{\gamma},\boldsymbol{\gamma}')$ with respect to $\boldsymbol{\gamma}$ by taking the derivative and setting it equal to zero, which yields the update rule

$$\gamma_i'^{+1} = E_{\mathbf{x}|\mathbf{y},\boldsymbol{\gamma}',\sigma^2}[x_i^2] := \langle x_i^2\rangle. \tag{2.20}$$

To compute $\langle x_i^2\rangle$, we consider the multivariate posterior density $p(\mathbf{x}|\mathbf{y},\boldsymbol{\gamma},\sigma^2)$ which has the form

(see Appendix 2.8.2)

$$p(\mathbf{x}|\mathbf{y},\gamma) = c(\mathbf{y}) e^{-\frac{(\mathbf{x}-\mu)^T \Sigma^{-1} (\mathbf{x}-\mu)}{2}} u(\mathbf{x}), \tag{2.21}$$

where μ and Σ are given by [1, 98, 107]

$$\mu = \Gamma \Phi^T (\sigma^2 \mathbf{I} + \Phi \Gamma \Phi^T)^{-1} \mathbf{y} \tag{2.22}$$

$$\Sigma = \Gamma - \Gamma \Phi^T (\sigma^2 \mathbf{I} + \Phi \Gamma \Phi^T)^{-1} \Phi \Gamma, \tag{2.23}$$

and $\Gamma = \text{diag}(\gamma)$. The posterior in Eq. (2.21) is known as a multivariate RG (or a multivariate truncated normal [108]). The normalizing constant $c(\mathbf{y})$ does not admit a closed form expression. However, the M-step in Eq. (2.20) only requires the marginal density. Unfortunately, the marginals of a multivariate RG are not univariate-RG's and do not admit closed form expressions [108], which also means no immediate expressions for the marginal moments.

However, we can approximate the first and the second moments $\langle x_i \rangle$ and $\langle x_i^2 \rangle$ of the multivariate RG posterior. In the following, we propose four different approaches for this purpose that offer a trade-off between computational complexity and theoretical accuracy.

Markov Chain Monte Carlo EM (MCMC-EM)

Advances in numerical methods made it possible to sample from complex multivariate distributions [109, 110, 111]. Numerical methods are particularly useful when the first and second order statistics of a posterior density do not have a closed form expressions. In this case, the E-step can be performed by drawing samples using numerical Markov Chain Monte Carlo (MCMC) and then calculating the sample statistics. This approach is usually referred to as MCMC-EM [112, 113].

First, we consider the Gibbs sampling approach in [114, 115]. We use hat notation to refer

to the empirical estimates of various parameters (e.g. $\hat{\Sigma}$, $\hat{\mu}$). We use the multivariate truncated normal (TN) definition in [115] and write

$$\mathsf{TN}(\mathbf{x};\hat{\mu},\hat{\Sigma},\mathbf{R},\alpha_L,\alpha_U) = \left(c_{tn} e^{-\frac{(\mathbf{x}-\hat{\mu})^T \hat{\Sigma}^{-1}(\mathbf{x}-\hat{\mu})}{2}} \right) \mathbf{1}_{\alpha_L \leq \mathbf{R}\mathbf{w} \leq \alpha_U}, \tag{2.24}$$

where $\mathbf{1}_{(\cdot)}$ is the indicator function and c_{tn} is the normalizing constant for the density. In the case of a multivariate rectified Gaussian, the truncation bounds are $\alpha_L = \mathbf{0}$ and $\alpha_U = \infty$, and $\mathbf{R} = \mathbf{I}$. By introducing the transformation, $\mathbf{w} = \hat{\mathbf{L}}^{-1}(\mathbf{x} - \hat{\mu})$ where $\hat{\mathbf{L}}$ is the lower triangular Cholesky decomposition of $\hat{\Sigma}$, it can be shown that $\mathbf{w}$ is $\mathsf{TN}(\mathbf{w};\mathbf{0},\mathbf{I},\hat{\mathbf{L}},\alpha_L^*,\alpha_U^*)$ with new truncation bounds $\alpha_L^* = \alpha_L - \hat{\mu} = -\hat{\mu}$ and $\alpha_U^* = \alpha_U - \hat{\mu} = \infty$.

The Gibbs sampler then proceeds by iteratively drawing samples from the conditional distribution $p(w_i|\mathbf{y},\hat{\gamma},\sigma^2,\mathbf{w}_{-\mathbf{i}})$, where $\mathbf{w}_{-\mathbf{i}}$ refers to the vector containing all but the ith element of $\mathbf{w}$. Given a set of samples drawn from $\mathbf{w}$, we can obtain the samples for the original distribution of interest by inverting the transformation: $\{\mathbf{x}^n\}_{n=1}^N = \{\hat{\mathbf{L}}\mathbf{w}^n + \hat{\mu}\}_{n=1}^N$. Then, the first and second empirical moments can be calculated from the drawn samples using

$$\langle x_i \rangle \approx \frac{1}{N} \sum_{n=1}^{N} (x_i^n), \tag{2.25}$$

$$\langle x_i^2 \rangle \approx \frac{1}{N} \sum_{n=1}^{N} (x_i^n)^2, \tag{2.26}$$

and the EM can be iterated by updating $\hat{\gamma}_i^{t+1} = \langle x_i^2 \rangle$.

After convergence, a point estimate for $\mathbf{x}$ is needed. The optimal estimator of $\mathbf{x}$ in the minimum mean-square-error (MMSE) sense is simply $\hat{\mathbf{x}}_{mean} = \langle x_i \rangle$. An alternative point estimate

is to use $\hat{\mathbf{x}}_{mode}$ given by

$$\hat{\mathbf{x}}_{mode} = \arg\max_{\mathbf{x}} p(\mathbf{x} \,|\, \mathbf{y}, \hat{\gamma}, \sigma^2) \tag{2.27}$$

$$= \arg\min_{\mathbf{x} \geq 0} \| \mathbf{y} - \Phi \mathbf{x} \|_2^2 + \lambda \sum_{i=1}^{M} \frac{x_i^2}{\hat{\gamma}_i}, \tag{2.28}$$

where Eq. (2.28) can be solved by any NNLS solver. The estimate $\hat{\mathbf{x}}_{mode}$ could be a favorable point estimate because it chooses the peak of $p(\mathbf{x}|\,\mathbf{y}, \hat{\gamma}, \sigma^2)$, which may not be well-characterized by its mean.

For the sparse recovery problem at hand, we experienced very slow convergence with Gibbs sampling. Convergence was particularly slow for higher problem dimensions and at larger cardinalities. The latter was expected as a sparse solution is harder to recover in those cases. Thus, we resorted to Hamiltonian Monte Carlo (HMC) which is designed specifically for target spaces constrained by linear or quadratic constraints [111]. HMC improves the MCMC mixing performance by using the gradient information of the target distribution [112].

Despite use of the state of the art MCMC techniques, MCMC-EM might still converge to poor local minima solutions and result in sub-optimal performance [116, 117, 118]. Particularly, performance may be poorer for under-determined problems. Though MCMC-EM is not thoroughly investigated for the sparse recovery problem, here we list four major issues for consideration:

I. Convergence: MCMC-EM based algorithms can get stuck in a local minima depending on the problem dimensions and complexity of the search space. This is true even for well-posed problems [113, 119]. In under-determined problems, the solution set for Eq. (2.2) may contain many local minima and thus, a good MCMC-EM implementation should try to avoid local minima.

II. Computational Limits: Current MCMC sampling techniques are not optimal for drawing

large sample sizes from high dimensional multivariate posterior densities. Therefore, the number of available samples is often limited by computational constraints [109, 110, 111].

III. Quality of Parameter Estimates: Since the MCMC samples are determined by random sampling at each iteration, the estimates of $\hat{\gamma}$, $\hat{\mu}$, and $\hat{\Sigma}$ depend highly on the quality of the MCMC estimates $\hat{\mathbf{x}}$, which in turn affects the quality of next cycle of MCMC samples. This may lead the EM algorithm to converge to a sub-optimal solution.

IV. Structure of the Empirical $\hat{\Sigma}$: When M is large and the dimensions of the empirical scale matrix are also large, $\hat{\Sigma}$ may no longer be a good numerical estimate [117, 120, 121]. This issue could be exacerbated when the problem is inherently under-determined with $N < M$, and reveals itself as $\hat{\Sigma}$ being close to singular. Therefore, regularization methods for $\hat{\Sigma}$ are often used to alleviate this problem [117, 118].

The scale matrix $\hat{\Sigma}$ has direct control over the search space for MCMC and spurious off-diagonal values tend to increase the number of local-minima. Therefore, to address the issues listed above, we incorporated ideas from prior work to regularize the estimates of $\hat{\Sigma}$:

- As in [117, 118], we assume that $\hat{\Sigma}$ is sparse and we prune its off-diagonal entries when they drop below a certain threshold T_p. This prevents the spurious off-diagonal values in $\hat{\Sigma}$ from affecting the next cycle of MCMC samples and improves future estimates of $\hat{\gamma}$.

- We incorporate the shrinkage estimation idea presented in [118, 120] and regularize $\hat{\Sigma}$ as a convex sum of the empirical $\hat{\Sigma}$ and a target matrix $\mathbf{T}$ such that, $\hat{\Sigma} = \lambda\hat{\Sigma} + (1-\lambda)\mathbf{T}$. A simple selection for $\mathbf{T}$ is the matrix $\hat{\Sigma}_\beta$, which is equal to the original $\hat{\Sigma}$ with diagonal elements scaled by a factor β. Though this approach does not guarantee convergence to a global minimum and the solution could still be a local minima or a saddle point solution, we empirically observed better recovery performance.

Linear minimum mean-square-error (LMMSE)

The LMMSE estimation approach is motivated by the complexity of the MCMC-EM approach. Examining the parameters being computed, one can interpret them as finding the MMSE estimate of $\mathbf{x}$ and the associated MSE. This motivates replacing the MMSE estimate by the simple LMMSE estimate of $\mathbf{x}$. The affine LMMSE estimate for $\mathbf{x}$ is

$$\hat{\mathbf{x}} = \mu_{\mathbf{x}} + \mathbf{R}_{\mathbf{x}}\Phi^T(\Phi\mathbf{R}_{\mathbf{x}}\Phi^T + \sigma^2\mathbf{I})^{-1}(\mathbf{y} - \Phi\mu_{\mathbf{x}}), \tag{2.29}$$

where $\mathbf{R}_{\mathbf{x}}$ is the covariance matrix of $\mathbf{x}$ (a diagonal matrix). The estimation error covariance matrix is given by [122]

$$\mathbf{R}_{\mathbf{e}} = \mathbf{R}_{\mathbf{x}} - \mathbf{R}_{\mathbf{x}}\Phi^T(\Phi\mathbf{R}_{\mathbf{x}}\Phi^T + \sigma^2\mathbf{I})^{-1}\Phi\mathbf{R}_{\mathbf{x}}. \tag{2.30}$$

To elaborate, in the E-step where γ is fixed at γ', the entries of $\mathbf{x}$ are independent, and the prior mean and the prior covariance will be equal to the mean and variance of the independent univariate RG distributions with $p(x_i|\gamma_i) = \mathcal{N}^R(0,\gamma_i)$. The mean of a univariate rectified Gaussian density with zero location parameter is given by [123]

$$\mu_{x,i} = \sqrt{\frac{2\gamma_i}{\pi}}, \tag{2.31}$$

and the variances which are the diagonal entries of the diagonal matrix $\mathbf{R}_{\mathbf{x}}$ are given by

$$R_{x,ii} = \gamma_i\,(1 - 2/\pi)\,. \tag{2.32}$$

Using the values of $\mu_{\mathbf{x}}$ and $\mathbf{R}_{\mathbf{x}}$ from Eq. (2.31) and Eq. (2.32) in Eq. (2.29) we obtain the LMMSE point estimate for the solution vector. Similarly, the update for γ (M-step) is given by

$$\gamma_i = \hat{x}_i^2 + R_{e,ii}. \tag{2.33}$$

This is sufficient to implement the EM algorithm. Upon convergence, the mean point estimate is simply $\hat{\mathbf{x}}_{mean} = \hat{\mathbf{x}}$, and the mode point estimate can be obtained by utilizing the converged values γ_i in Eq. (2.28).

Generalized approximate message passing (GAMP)

In this section, we present an EM implementation using the generalized approximate message passing (GAMP) algorithm [97, 124]. A different GAMP based approach was used in [104], which uses an i.i.d. Bernoulli non-negative Gaussian mixture prior with a fixed mixture order that is independent of M. To overcome the convergence issues with the type of GAMP algorithm in [104] e.g. when a non-i.i.d. design matrix Φ is used [125, 126, 127], we incorporate the damping technique in [97, 127] into the proposed R-SBL GAMP algorithm.

GAMP is a low complexity iterative inference algorithm. The low complexity is achieved by applying quadratic and Taylor series approximations to loopy belief propagation. GAMP can approximate the MMSE estimate when used in the sum-product version, or can approximate the MAP estimate when used in the max-sum version. The sum-product version computes the mean and variance of the approximate marginal posteriors on x_i which are given by

$$p(x_i|r_i;\tau_{r_i}) \propto p(x_i)\mathcal{N}(x_i;r_i,\tau_{r_i}), \tag{2.34}$$

where r_i approximates an AWGN corrupted version of the true x_i as

$$r_i \approx x_i + \bar{r}_i \tag{2.35}$$

$$\bar{r}_i \sim \mathcal{N}(0, \tau_{r_i}). \tag{2.36}$$

In the large system limit and when the design matrix Φ is i.i.d sub-Gaussian, the approximation in Eq. (2.35) was shown to be exact [124, 128]. Therefore, in the sum-product version of GAMP, the estimate $\hat{x}_i$ in Eq. (2.37) corresponds to the MMSE estimate of x_i given r_i, and similarly the conditional variance of x_i given r_i is defined in Eq. (2.38).

$$\hat{x}_i = \mathbb{E}\{x_i | r_i; \tau_{r_i}\} \tag{2.37}$$

$$\tau_{x_i} = \text{var}\{x_i | r_i; \tau_{r_i}\}. \tag{2.38}$$

In the max-sum version of GAMP, the MAP estimate $\hat{x}_i$ given r_i is obtained in Eq. (2.39) using the proximal operator defined in Eq. (2.41), while τ_{x_i} given in Eq. (2.40) corresponds to the sensitivity of the proximal thresholding.

$$\hat{x}_i = \text{prox}_{-\ln p(x_i)}(r_i; \tau_{r_i}) \tag{2.39}$$

$$\tau_{x_i} = \tau_r \text{prox}'_{-\ln p(x_i)}(r_i; \tau_{r_i}) \tag{2.40}$$

$$\text{prox}_f(\hat{a}, \tau^a) \triangleq \underset{x \in R}{\arg\min}\, f(x) + \frac{1}{2\tau^a}|x - \hat{a}|^2. \tag{2.41}$$

When implementing the EM algorithm, the approximate posterior computed by the sum-product GAMP can be used to efficiently approximate the E-step [129]. Moreover, in the case of max-sum GAMP, in the large system limit and under i.i.d sub-Gaussian Φ an extra step can be added as in [104] to compute the marginal distributions using Eq. (2.34). These marginals then can be used to approximate the E-step. For the rectified Gaussian scale mixture prior $p(x|\gamma)$ the details of finding $\hat{x}_i$ and τ_{x_i} estimates in both the sum-product and max-sum cases are shown in

Appendix 2.8.1.

Upon convergence of the GAMP algorithm, the approximate E-step of the EM algorithm is complete and we can evaluate the M-step in Eq. (2.20) as

$$\langle x_i^2 \rangle = \int_{x_i} x_i^2 p(x|r_i;\tau_{r_i}) = \hat{x}_i^2 + \tau_{x_i}. \tag{2.42}$$

The EM-based R-SBL GAMP algorithm is summarized in Table 2.1. Here, the steps used by the GAMP algorithm to evaluate s and τ_s are the same for both sum-product and max-sum versions (for AWGN case) [124]. In Table 2.1, all mathematical operations are element wise. $K_{\max}$ is the maximum allowed number of GAMP iterations, $\varepsilon_{\text{gamp}}$ is the GAMP tolerance parameter, $I_{\max}$ is the maximum number of EM iterations, and ε_{em} is the EM tolerance parameter. Also, $\theta_s \in (0,1]$ is the damping factor which can be selected according to the empirical criteria in [97], and η, ν, $h(.)$, and $g(.)$ are defined in Appendix 2.8.1.

Diagonal approximation (DA)

We know *a-priori* that the posterior in Eq. (2.21) does not admit a closed form expression. However, to implement the EM algorithm we only need the marginal moments of the posterior. We first note that, if the scale matrix Σ is diagonal then we could evaluate the normalizing constant $c(\mathbf{y})$ in closed form since the multivariate RG posterior can be written as a product of univariate marginals (see Appendix 2.8.2).

In the diagonal approximation (DA) approach, we resort to approximating the posterior in Eq. (2.21) with a suitable posterior density $p(\mathbf{x}|\mathbf{y},\gamma) \approx \tilde{p}(\mathbf{x}|\mathbf{y},\gamma)$, which could be written as a product of independent marginal densities i.e. $\tilde{p}(x_i|\mathbf{y},\gamma)$. This approximate posterior density is

Table 2.1: R-SBL GAMP Algorithm

Initialization
$S \leftarrow |\Phi|^2$ (component wise magnitude squared)
Initialize $\dot{\tau}_x^0, \gamma^0 > 0$
$\dot{s}^0, \dot{x}^0 \leftarrow 0$

for $i = 1, 2, ..., I_{\max}$
  Initialize $\tau_x^1 \leftarrow \dot{\tau}_x^{i-1}, \hat{x}^1 \leftarrow \dot{x}^{i-1}, s^1 \leftarrow \dot{s}^{i-1}$
  // E-Step Approximation
  for $k = 1, 2, ..., K_{\max}$
    $1/\tau_p^k \leftarrow S\tau_x^k$
    $p^k \leftarrow s^{k-1} + \tau_p^k \Phi \hat{x}^k$
    $\tau_s^k \leftarrow \frac{\sigma^{-2}\tau_p^k}{\sigma^{-2}+\tau_p^k}$
    $s^k \leftarrow (1-\theta_s)s^{k-1} + \theta_s(p^k/\tau_p^k - y)/(\sigma^2 + 1/\tau_p^k)$
    $1/\tau_r^k \leftarrow S^\top \tau_s^k$
    $r^k \leftarrow \hat{x}^k - \tau_r^k \Phi^\top s^k$
    if MaxSum then
      $\tau_x^{k+1} \leftarrow \nu^k$
      $\hat{x}^{k+1} \leftarrow \eta^k u(r^k)$
    else
      $\tau_x^{k+1} \leftarrow \nu^k g(\frac{\eta^k}{\nu^k})$
      $\hat{x}^{k+1} \leftarrow \eta^k + \sqrt{\nu^k} h(\frac{\eta^k}{\nu^k})$
    end if
    if $\|\hat{x}^{k+1} - \hat{x}^k\|^2 / \|\hat{x}^{k+1}\|^2 < \varepsilon_{\text{gamp}}$, break
  end for %end of k loop
  $\dot{s}^i \leftarrow s^k$
  if MaxSum
    $\dot{x}^i \leftarrow \eta^{k+1} + \sqrt{\nu^{k+1}} h(\frac{\eta^{k+1}}{\nu^{k+1}})$, $\dot{\tau}_x^i \leftarrow \nu^{k+1} g(\frac{\eta^{k+1}}{\nu^{k+1}})$
  else
    $\dot{x}^i \leftarrow \hat{x}^{k+1}$, $\dot{\tau}_x^i \leftarrow \tau_x^{k+1}$
  end if
  // M-Step
  $\gamma^{i+1} \leftarrow |\dot{x}^i|^2 + \dot{\tau}_x^i$
  if $\|\dot{x}^i - \dot{x}^{i-1}\|^2 / \|\dot{x}^i\|^2 < \varepsilon_{\text{em}}$, break
end for %end of i loop

derived in Appendix 2.8.2 as

$$\tilde{p}(\mathbf{x}\,|\,\mathbf{y},\boldsymbol{\gamma}) = \prod_{i=1}^{M} \tilde{p}(x_i|\mathbf{y},\boldsymbol{\gamma}) \tag{2.43}$$

$$= \prod_{i=1}^{M} \sqrt{\frac{2}{\pi\Sigma_{ii}}} \frac{e^{-\frac{(x_i-\mu_i)^2}{2\Sigma_{ii}}} u(x_i)}{\operatorname{erfc}\left(-\frac{\mu_i}{\sqrt{2\Sigma_{ii}}}\right)}, \tag{2.44}$$

where μ_i is the ith element of μ and Σ_{ii} is the ith diagonal element of Σ obtained using Eqs. (2.22) and (2.23). The marginal $\tilde{p}(x_i|\mathbf{y},\boldsymbol{\gamma})$ in Eq. (2.43) is the univariate RG density defined in Eq. (2.10), where $\tilde{p}(x_i|\mathbf{y},\boldsymbol{\gamma}) = \mathcal{N}^R(x_i;\mu_i,\Sigma_{ii})$. Then, the univariate RG marginals are well-characterized by their first and second moments given in [123], with the first moment given as

$$\langle x_i \rangle = \mu_i + \sqrt{\frac{2\Sigma_{ii}}{\pi}} \frac{e^{-\frac{\mu_i^2}{2\Sigma_{ii}}}}{\operatorname{erfc}\left(-\frac{\mu_i}{\sqrt{2\Sigma_{ii}}}\right)}, \tag{2.45}$$

and the second moment given as

$$\langle x_i^2 \rangle = \mu_i^2 + \Sigma_{ii} + \mu_i \sqrt{\frac{\Sigma_{ii}}{\pi}} \frac{e^{-\frac{\mu_i^2}{2\Sigma_{ii}}}}{\operatorname{erfc}\left(-\frac{\mu_i}{\sqrt{2\Sigma_{ii}}}\right)}. \tag{2.46}$$

Note that the moments of $\tilde{p}(x_i|\mathbf{y},\boldsymbol{\gamma})$ are approximations to the moments of the true marginals which do not admit closed form. However, we can perform EM using the approximate moments to approximate the true solution. EM can be carried out by setting $\gamma_i^{t+1} = \langle x_i^2 \rangle$ and iterating over t. After convergence of γ_is, the mean point estimate is obtained as $\hat{\mathbf{x}}_{mean} = \langle x_i \rangle$. The mode point estimate $\hat{\mathbf{x}}_{mode}$ can be calculated by using converged values of γ_is in Eq. (2.28).

If the diagonal elements of Σ are large valued or become large over EM iterations as compared to the off-diagonals, then DA is expected to work well. Note that assuming a diagonal Σ was also motivated by prior work [117, 120, 121, 130, 131] for various applications. In this

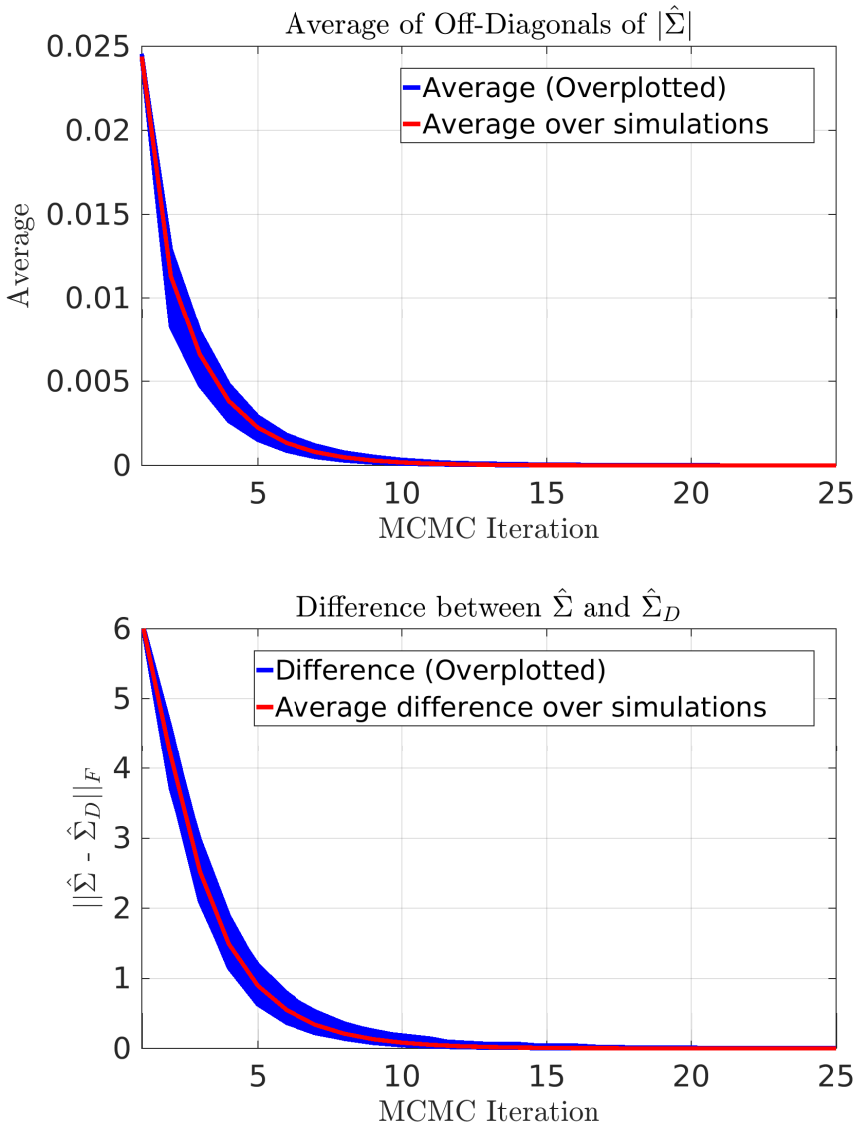


Figure 2.1: *Top*: Empirical observations for the structure of Σ. We performed S-NNLS recovery using MCMC-EM (without regularizing the estimates of Σ) and monitored the average value of off-diagonals for $|\Sigma|$. We simulated for 1,000 runs and overplotted the results (blue lines). The average of average off-diagonals for $|\Sigma|$ over 1,000 results is shown with the red line. The exponentially decreasing behavior suggests that the off-diagonal magnitudes of Σ decrease over MCMC iterations, indicating that true Σ is approaching to a diagonal form. *Bottom*: The distance between the true Σ and a diagonal matrix formed by its diagonal entries Σ_D. This suggests that the true Σ approaches to a diagonal form over MCMC iterations.

work, we empirically report that DA has very good sparse recovery performance and has low complexity.

To further support the DA approximation, we present empirical findings regarding the structure of Σ. We performed sparse recovery simulations using Eq. (2.1) with the MCMC-EM approach as the ground truth (without regularizing the MCMC estimates of $\hat{\Sigma}$). We assumed that $\mathbf{x}$ was of size 200 with 10 non-zero elements drawn from $\mathcal{N}^R(0,1)$. The dictionary $\Phi \in \mathbb{R}^{50\times 200}$ columns were normally distributed $\Phi \sim \mathcal{N}(0,\mathbf{I})$. We solved this problem for 1,000 simulations and overlay plots of the average absolute value of the off-diagonals of $\hat{\Sigma}$ as a function of MCMC-EM iteration in the first row of Figure 2.1 (blue lines).

We see that the average off-diagonal elements of $|\hat{\Sigma}|$ exponentially approach 0 as a function of MCMC-EM iteration. The average of this behavior over 1,000 simulations (red line) has a final value of 10^{-4} after 10 iterations. This indicates that the off-diagonals of $\hat{\Sigma}$ of the true posterior (with MCMC sampling) approach zero. Moreover, in the second row of Figure 2.1 we overlay plots of the Frobenius norm of the difference between $\hat{\Sigma}$ and $\hat{\Sigma}_D$, where $\hat{\Sigma}_D$ is the diagonal matrix consisting of diagonal elements from $\hat{\Sigma}$. This shows that as MCMC-EM converges $\hat{\Sigma}$ approaches a diagonal form.

These results suggest that, if there is flexibility in choosing the dictionary Φ as in compressed sensing, then proper choice of Φ can lead to the DA approach producing high quality approximate marginals $\tilde{p}(x_i|\mathbf{y},\gamma)$ that are close to the true marginals.

2.3.3 Computational complexity of proposed methods

For computational comparisons, we assume that $N \leq M$. Under this assumption, the time complexity of the DA algorithm is $O(N^2M)$ per EM iteration. This complexity is similar to the original SBL algorithm in [7, 98] and is due to the computationally intensive matrix inversion step $(\sigma^2\mathbf{I}+\Phi\Gamma\Phi^T)^{-1}$ given in Eq. (2.22). Time complexity of the LMMSE algorithm is also $O(N^2M)$ per EM iteration. This complexity is determined from a similar matrix inversion step

$(\Phi \mathbf{R_x} \Phi^T + \sigma^2 \mathbf{I})^{-1}$ in Eq. (2.30) (note that $\mathbf{R_x}$ is diagonal). The GAMP algorithm bypasses the computationally intensive matrix inversion and the resulting complexity is $O(NM)$ time [97]. This is linear in both problem dimensions and significantly faster than the both the DA and LMMSE methods. For the MCMC-EM algorithm, the actual computational cost is determined by the random Hamiltonian MCMC sampling, which is explained in more detail in [111].

2.4 Experiment Design

In this section we provide the layout of our numerical experiments. We provide extensive comparisons between the proposed R-SBL variants LMMSE, GAMP, MCMC and DA and the baseline S-NNLS solvers, including NNGM-AMP [104], SLEP-ℓ_1 [132], and NN-OMP [133]. In all of the experiments below, we generate sparse vectors $\mathbf{x^{gen}} \in \mathbb{R}_+^{400}$, such that $||\mathbf{x^{gen}}||_0 = K$, and random dictionaries $\Phi \in \mathbb{R}^{100 \times 400}$. We normalize the columns of Φ by $1/\sqrt{N}$ [134]. For a fixed Φ and $\mathbf{x}^{gen}$, we compute the measurements $\mathbf{y} = \Phi \mathbf{x}^{gen}$ and use the baseline algorithms and the proposed R-SBL variants to approximate $\mathbf{x}^{gen}$.

In the first set of experiments, we simulate a 'noiseless' recovery scenario, where the noise variance is set as $\sigma^2 = 10^{-6}$, the non-zero entries of the solution vector are drawn from a rectified Gaussian density $\mathcal{N}^R(0,1)$ and the dictionary columns are i.i.d. Normal distributed $\Phi \sim \mathcal{N}(0, \mathbf{I})$. We experiment with cardinalities $K = \{10, 20, 30, 35, 40, 45, 50\}$.

In the second set, we construct various dictionary types to analyze the robustness of our R-SBL method and the baseline solvers for the S-NNLS problem. The dictionary types considered here are not necessarily i.i.d. Gaussian and are similar to the ones used in [97, 135]. These dictionaries can be low-rank, coherent, ill-posed, and non-negative as detailed below:

A. *Coherent dictionaries:* We introduce coherence among the columns of an original dictionary $\Phi = \mathcal{N}(0, \mathbf{I})$ and report recovery performances for a fixed $K = 50$. This was done by multiplying Φ with a coherence matrix $\mathbf{C}$ to obtain a new dictionary Φ_c with coherent columns. Here,

$\mathbf{C}$ is the Cholesky factor of the Toeplitz(ρ) matrix with a coherence parameter ρ. We experiment with different coherence values by selecting $\rho = \{0.1, 0.2, ..., 0.80, 0.85, 0.90, 0.95\}$.

B. *Low-rank dictionaries:* We construct rank-deficient dictionaries such that $\Phi = \mathbf{AB}$, where $\mathbf{A} \in \mathbb{R}^{N \times R}$, $\mathbf{B} \in \mathbb{R}^{R \times M}$ and $R < N$. The entries of $\mathbf{A}$ and $\mathbf{B}$ are i.i.d. Normal. The rank ratio R/N is considered as a measure of rank deficiency, where smaller values indicate more deviation from an i.i.d. dictionary. We experiment with $R/N = \{1, 0.95, ..., 0.4\}$ and report recovery performances for a fixed $K = 50$.

C. *Ill-conditioned dictionaries:* We experiment with ill-conditioned dictionaries with a condition number $\kappa > 1$. For a fixed κ, the dictionary is constructed as $\Phi = \mathbf{USV}^T$. Here, $\mathbf{U}$ and $\mathbf{V}$ contain the left and right singular vectors of an i.i.d. Gaussian matrix, and $\mathbf{S}$ is a diagonal matrix containing the eigenvalues. We decay the elements of $\mathbf{S}$ with $\mathbf{S}_{i+1,i+1} = \kappa^{-1/(N-1)}\mathbf{S}_{i,i}$ for $i = 1, 2, ..., N-1$. The value of κ measures the deviation from an i.i.d. Gaussian dictionary, with larger κ values indicate more deviation. We experiment using the condition numbers $\kappa = \{8, 10, ..., 28\}$.

D. *Non-negative dictionaries:* Non-negative dictionaries are used in sparse recovery applications such as sparse NMF [5] and NN K-SVD [10], where a positive mapping is required on the solution vector. We construct non-negative dictionaries Φ with columns that are drawn according to $\Phi \sim RG(0, \mathbf{I})$. We experiment with cardinalities $K = \{10, 20, 30, 35, 40, 45, 50\}$.

In the third set of experiments, we set the noise variance σ^2 for $\mathbf{v}$ such that the signal-to-noise ratio (SNR) is 20 dB and repeat the first set of experiments. This experiment was meant to assess the robustness of R-SBL variants under noisy conditions.

In the fourth set of experiments, we investigate recovery performances for a variety of distributions for Φ, and for the non-zero elements of $\mathbf{x}$. We randomly draw the nonzero elements of $\mathbf{x}^{\mathbf{gen}}$ according to the following distributions:

I. NN-Cauchy (Location: 0, Scale: 1)

II. NN-Laplace (Location: 0, Scale: 1)

III. Gamma (Location: 1, Scale: 2)

IV. Chi-square with $\nu = 2$

V. Bernoulli with $p(0.25) = 1/2$ and $p(1.25) = 1/2$

where the prefix 'NN' stands for non-negative. These distributions are obtained by taking the absolute value of the respective probability densities. We also generate random dictionaries Φ according to the following densities:

I. Normal (Location: 0, Scale: 1)

II. ± 1 with $p(1) = 1/2$ and $p(-1) = 1/2$

III. $\{0,1\}$ with $p(0) = 1/2$ and $p(1) = 1/2$

In all of the experiments detailed here, the results were averaged over 1,000 simulations. Moreover, the R-SBL MCMC approach was only used in the first set of experiments to demonstrate the high quality of the parameter estimates obtained with the lower complexity approaches such as DA, LMMSE and GAMP. We omit the MCMC in other experiments due to computational constraints.

2.4.1 Performance metrics

To evaluate the performance of various S-NNLS algorithms, we used the normalized mean square error (NMSE) and the probability of error in the recovered support set (PE) [17]. We computed the NMSE between the recovered signal $\hat{\mathbf{x}}$ and the ground truth $\mathbf{x}^{gen}$ using

$$\text{NMSE} = \|\hat{\mathbf{x}} - \mathbf{x}^{\mathbf{gen}}\|^2 / \|\mathbf{x}^{\mathbf{gen}}\|^2. \tag{2.47}$$

The PE metric was computed using

$$\text{PE} = \frac{max\{|S|, |\hat{S}|\} - |S \cap \hat{S}|}{max\{|S|, |\hat{S}|\}}, \tag{2.48}$$

where the support of the true solution was S and the recovered support of $\hat{\mathbf{x}}$ was $\hat{S}$. A value of $\text{PE} = 0$ indicates that the ground truth and recovered supports are the same, whereas $\text{PE} = 1$ indicates no overlap between supports. Averaging the PE over multiple trials gives the empirical probability of making errors in the recovered support. The averaged values of NMSE and PE over 1,000 simulations and for each experiment are reported in the Experiment Results section.

2.4.2 MCMC implementation

We used the MCMC implementation presented in [111]. The MCMC parameters explained in Section 2.3.2 were selected as follows. The off-diagonal pruning of the empirical scale parameter $\hat{\Sigma}$ was performed with a threshold of $T_p = 5 \times 10^{-2}$. Diagonal scaling was performed with a factor of $\beta = 1.7$, and a shrinkage parameter of $\lambda = 0.5$. These values were empirically determined to minimize the NMSE for the first set of experiments.

2.5 Experiment Results

Here, we show that in all of the sparse recovery experiments detailed above, the proposed R-SBL variants outperform the baseline solvers in terms of NMSE and PE. The R-SBL variants outperform the baseline solvers when the dictionary is non-i.i.d., coherent, low-rank, ill-posed or even non-negative, showing the robustness of R-SBL to different characteristics of the dictionary Φ.

In Figure 2.2(a) we show the sparse recovery performance of the R-SBL variants and the baseline solvers as a function of the cardinality for the first set of experiments. As the cardinality

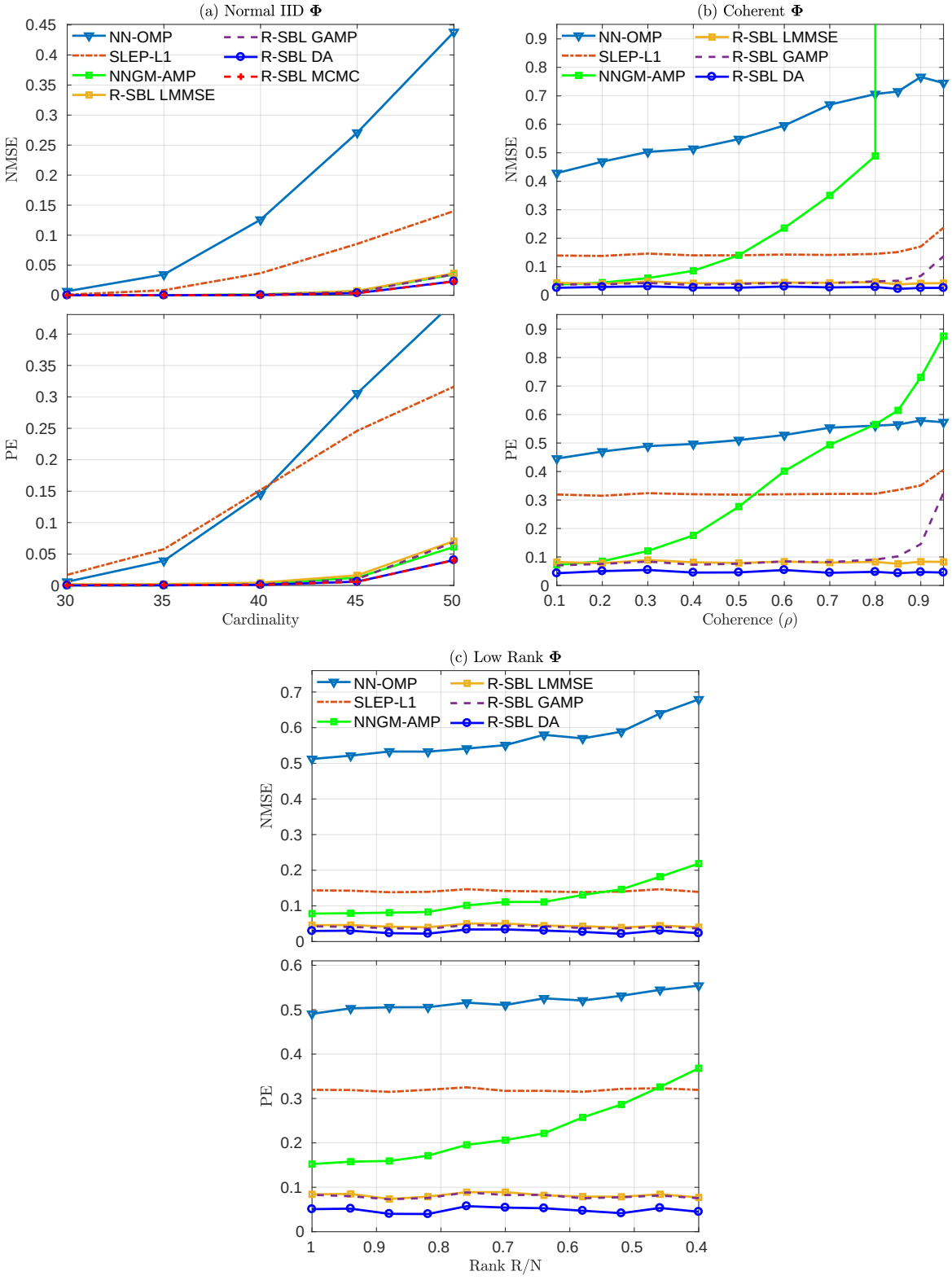
of the ground truth solution increases (after $K = 30$) the performances of NN-OMP and SLEP-ℓ_1 deteriorate both in terms of NMSE and PE. On the other hand, R-SBL variants and NNGM-AMP are quite robust with very small recovery error. For the largest cardinality of $K = 50$, we see that R-SBL DA and MCMC outperform other methods. The DA variant is nearly identical to MCMC in terms of NMSE and PE. This is expected since MCMC prunes off-diagonal elements of the scale matrix Σ iteratively, when they drop below a certain threshold.

2.5.1 Coherent Dictionaries

In Figure 2.2(b) we show the recovery performances when the dictionary is coherent. The degree of dictionary coherence is shown on the horizontal axis with ρ which ranges from 0.1 to 0.95. The proposed R-SBL variants are extremely robust to increasing coherence and outperform the baseline solvers in terms of both NMSE and PE. SLEP-ℓ_1 is robust to increasing coherence but performs worse when compared to the R-SBL variants. NNGM-AMP breaks down after $\rho = 0.3$ and performs worse than SLEP-ℓ_1 after $\rho = 0.5$, and worse than NN-OMP after $\rho = 0.8$. The LMMSE and DA variants are not affected by the coherence level and achieve better recovery even for $\rho = 0.95$. The performance of R-SBL GAMP slightly deteriorates after an extreme coherence of $\rho = 0.90$, but is still better than the baseline solvers.

These results demonstrate that the proposed R-SBL variants are robust to dictionary coherence and are superior to the baseline solvers. The robustness of our R-SBL framework seems to be inherited from the robustness of the original SBL algorithm to the structure of Φ [97, 136], which uses a GSM prior on $\mathbf{x}$. Our R-GSM prior on $\mathbf{x}$ seems to provide a similar robustness to the R-SBL algorithm.

Figure 2.2: (Figure on the following page.) Sparse recovery performances (NMSE and PE) of the R-SBL variants and the baseline S-NNLS solvers for various Φ. In (a) the dictionary elements were i.i.d Normal and the sparse recovery results are shown for cardinalities $K = 30$ to $K = 50$. R-SBL DA achieves the best recovery performance. R-SBL LMMSE and GAMP are similar to NNGM-AMP and are much better than SLEP-ℓ_1 and NN-OMP. In (b) the dictionary columns are coherent with the coherence degree ρ indicated in the x-axis. R-SBL variants are extremely robust to increasing coherence and result in a very small NMSE and PE across all ρ values. NNGM-AMP breaks down after $\rho = 0.2$ with deteriorating performance with increasing ρ and SLEP-ℓ_1 is better than NNGM-AMP after $\rho = 0.5$. In (c) the dictionary is rank-deficient with rank-ratio R/N indicated in the x-axis. R-SBL variants are superior to baseline methods across all R/N values.



(Figure 2.2 caption on the previous page.)

2.5.2 Low-rank Dictionaries

In Figure 2.2(c) we show the recovery performances for rank-deficient dictionaries. The degree of rank deficiency is shown on the horizontal axis with the rank ratio R/N. The R-SBL variants outperform the baseline solvers in terms of both NMSE and PE for all values of R/N. The recovery performances of the R-SBL variants are extremely robust against the changes in R/N. Among the R-SBL variants, DA performs slightly better than LMMSE and GAMP, and GAMP performs similar to LMMSE. The recovery performance of NNGM-AMP is better than NN-OMP and SLEP-ℓ_1, however its performance degrades as R/N gets smaller.

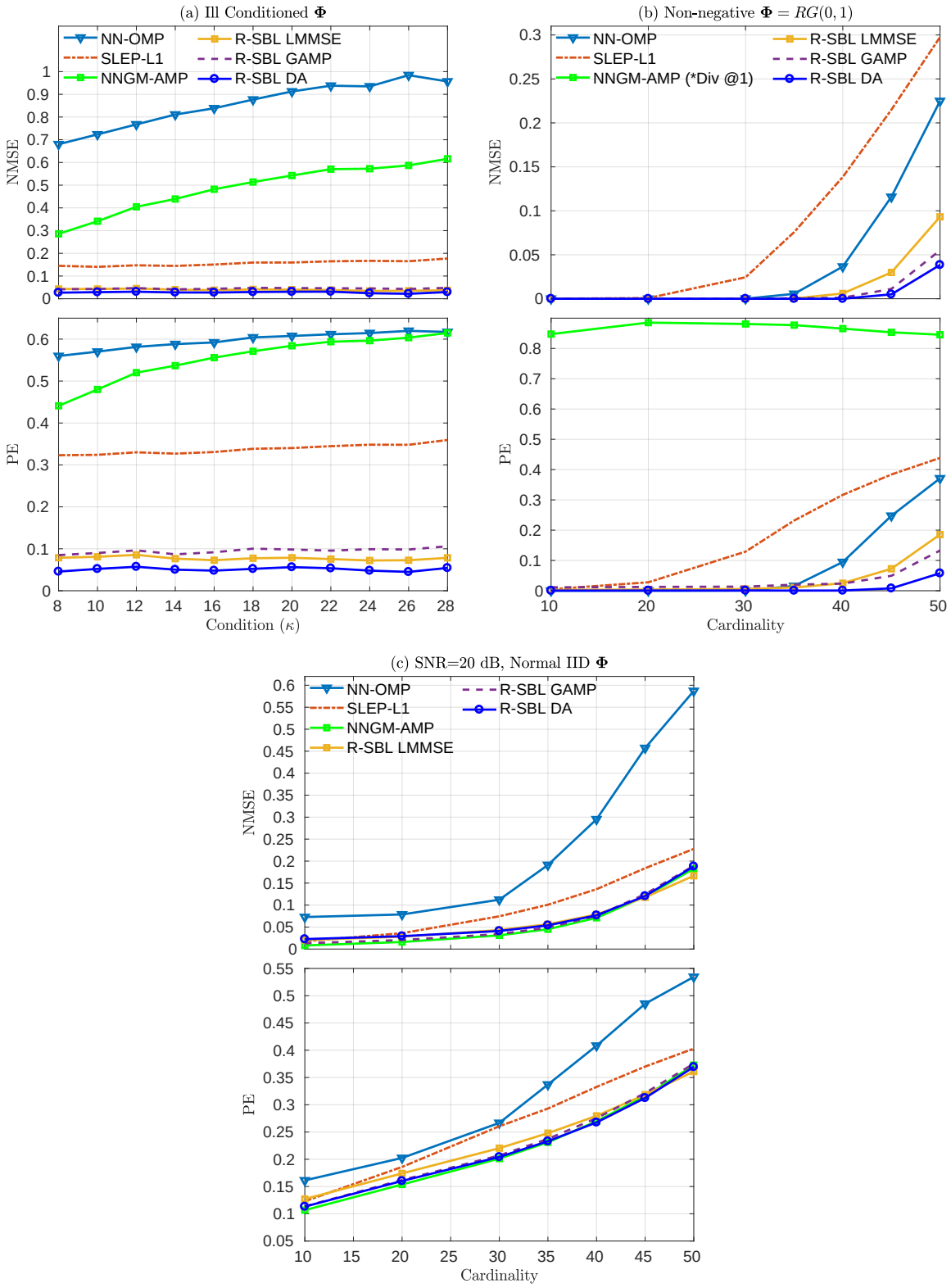
2.5.3 Ill-conditioned Dictionaries

In Figure 2.3(a) we demonstrate the recovery performances for ill-conditioned dictionaries. The condition number on the horizontal axis varies from $\kappa = 8$ to $\kappa = 28$. The proposed R-SBL variants perform significantly better than the baseline solvers across different κ values in terms of NMSE and PE. The recovery performances of the R-SBL variants are also extremely robust to different selections of κ. SLEP-ℓ_1 is better than NN-OMP and NNGM-AMP and is also robust to the selection of κ. The performances of NN-OMP and NNGM-AMP methods rapidly deteriorate with increasing κ values.

2.5.4 Non-negative Dictionaries

In Figure 2.3(b) we show the recovery performances when the dictionary is non-negative with elements drawn from i.i.d. $RG(0,1)$. The cardinality K on the horizontal axis of Figure 2.3(b) varies from $K = 10$ to $K = 50$. The NNGM-AMP approach was not able to recover feasible solutions for non-negative dictionaries and the point estimates for $\mathbf{x}$ diverged for different K. Therefore, the NMSE values for NNGM-AMP were not shown in Figure 2.3(b). Unlike in Figure 2.2(a), where the dictionary can be both positive and negative, NN-OMP performs better than

Figure 2.3: (Figure on next page.) Sparse recovery performances of the S-NNLS solvers for various Φ. In (a) the dictionary is ill-conditioned with condition number κ given in the x-axis. R-SBL variants outperform the baseline solvers for various κ and are very robust to the selection of κ. R-SBL DA achieves the lowest NMSE and PE. SLEP-ℓ_1 is superior to NNGM-AMP. In (b) the dictionary is non-negative with elements drawn from i.i.d. $\mathrm{RG}(0,1)$. The recovery performances are given for various cardinality K in the x-axis. R-SBL variants achieve superior recovery across all values of K. NNGM-AMP diverges regardless of the value of K and is unable to recover a feasible solution. In (c) the dictionary is i.i.d. Normal and SNR is 20 dB. The R-SBL variants perform similar to NNGM-AMP under noisy conditions, but are superior to SLEP-ℓ_1 and NN-OMP at larger cardinalities.



(Figure 2.3 caption on the previous page.)

SLEP-ℓ_1. The proposed R-SBL variants outperform the baseline approaches. Among the R-SBL variants, DA performs slightly better than GAMP, and GAMP is slightly better than LMMSE.

2.5.5 Noisy Conditions

We compared the recovery performances in a noisy setting, where the dictionary is i.i.d. Normal distributed. In this case, the observations were contaminated with additive white Gaussian noise to have a signal-to-noise ratio (SNR) of 20 dB. Figure 2.3(c) shows the NMSE and PE versus the cardinality. Compared with the noiseless case in Figure 2.2(a), the performances of all of the methods noticeably reduced. However, the proposed R-SBL variants performed better as compared to the NN-OMP and SLEP-ℓ_1 solvers, and performed similar to the NNGM-AMP approach.

2.5.6 Other types of $\mathbf{x}^{gen}$ and Φ

Here, the dictionary Φ was drawn according to i.i.d. Normal, ± 1 Bernoulli, and $\{0, 1\}$ Bernoulli distributions. We experimented with different distributions for the non-zero entries of $\mathbf{x}^{gen}$, as detailed in Tables 2.2, 2.3 and 2.4.

For i.i.d. Normal Φ in Table 2.2, the R-SBL DA generally outperforms the baseline solvers and other R-SBL variants when $\mathbf{x}^{gen}$ is RG, NN-Cauchy, NN-Laplace, Gamma and Chi-square distributed. The LMMSE variant achieves slightly better performance in terms of PE for the NN-Cauchy distribution. The NNGM-AMP is better than LMMSE and GAMP variants, when $\mathbf{x}^{gen}$ is RG, however it fails in terms of PE when $\mathbf{x}^{gen}$ is NN-Cauchy. The NNGM-AMP approach shows better performance when $\mathbf{x}^{gen}$ is Bernoulli. This is expected since the prior density for NNGM-AMP is a Bernoulli non-negative Gaussian mixture. The R-GSM prior, on the other hand, is not well matched to the Bernoulli distribution, as it is a mixture of continuous distributions. Overall, we see that R-SBL DA approach results in the best recovery performance.

Table 2.2: NMSE and PE results for various distributions for $\mathbf{x}^{gen}$. The dictionary is i.i.d Normal distributed.

		Φ is i.i.d Normal					
	$\mathbf{x}^{gen}$	NN-OMP	SLEP-ℓ_1	NNGM AMP	R-SBL (LMMSE)	R-SBL (GAMP)	R-SBL (DA)
NMSE	RG	0.4460	0.1439	0.0389	0.0488	0.0428	**0.0313**
	NN-Cauchy	0.0097	0.0086	0.0020	0.0004	0.0003	**0.0002**
	NN-Laplace	0.1566	0.0693	0.0091	0.0066	0.0059	**0.0034**
	Gamma	0.1476	0.0661	0.0074	0.0065	0.0045	**0.0024**
	Chi-square	0.1583	0.0673	0.0091	0.0077	0.0066	**0.0035**
	Bernoulli	0.5845	0.1265	**0.0052**	0.0524	0.0416	0.0339
PE	RG	0.4601	0.3208	0.0711	0.0873	0.0823	**0.0549**
	NN-Cauchy	0.2307	0.3509	0.2142	**0.0187**	0.0200	0.0408
	NN-Laplace	0.3202	0.3137	0.0407	0.0292	0.0229	**0.0118**
	Gamma	0.3091	0.3093	0.0416	0.0260	0.0207	**0.0080**
	Chi-square	0.3200	0.3086	0.0473	0.0307	0.0280	**0.0133**
	Bernoulli	0.4852	0.3283	**0.0101**	0.1714	0.1514	0.1264

In Table 2.3, we present the results for when Φ is ± 1 Bernoulli. The recovery performances observed in Table 2.3 are very similar to Table 2.2 and overall, the R-SBL DA approach enjoys better recovery performance.

In Table 2.4, we show recovery results for $\{0,1\}$ Bernoulli distributed Φ. The R-SBL DA and LMMSE variants achieve superior recovery when compared to the baseline solvers. The NNGM-AMP approach diverges for different $\mathbf{x}^{gen}$. This is consistent with our previous observation that NNGM-AMP failed when the dictionary elements were positive e.g. drawn from i.i.d. $RG(0,1)$ in Figure 2.3(b).

2.5.7 Recovery time analysis

In Section 2.3.3, we presented the worst case computational complexity of the DA, LMMSE and GAMP variants per EM iteration. As the execution time also depends on how fast an EM approach converges to the final solution, we provide an analysis of the average execution times for different cardinality values. First, we provide a simple way to speed up the proposed R-SBL algorithms. We prune the problem size when the elements of $\boldsymbol{\gamma}$ become smaller than a

Table 2.3: NMSE and PE results for various distributions for $\mathbf{x}^{gen}$. The dictionary is i.i.d ± 1 Bernoulli distributed.

		Φ is ± 1 Bernoulli					
	$\mathbf{x}^{gen}$	NN-OMP	SLEP-ℓ_1	NNGM AMP	R-SBL (LMMSE)	R-SBL (GAMP)	R-SBL (DA)
NMSE	RG	0.3996	0.1387	0.0409	0.0504	0.0415	**0.0332**
	NN-Cauchy	0.0083	0.0077	0.0023	0.0005	0.0004	**0.0003**
	NN-Laplace	0.1368	0.0712	0.0101	0.0096	0.0090	**0.0050**
	Gamma	0.1294	0.0665	0.0079	0.0061	0.0051	**0.0023**
	Chi-square	0.1267	0.0667	0.0109	0.0083	0.0094	**0.0055**
	Bernoulli	0.5610	0.1180	**0.0113**	0.0466	0.0412	0.0363
PE	RG	0.4272	0.3182	0.0794	0.0950	0.0824	**0.0568**
	NN-Cauchy	0.1810	0.3475	0.2307	0.0187	**0.0175**	0.0321
	NN-Laplace	0.2909	0.3131	0.0508	0.0369	0.0333	**0.0163**
	Gamma	0.2682	0.3072	0.0472	0.0274	0.0248	**0.0093**
	Chi-square	0.2769	0.3104	0.0532	0.0357	0.0369	**0.0195**
	Bernoulli	0.4734	0.3290	**0.0154**	0.1727	0.1571	0.1345

Table 2.4: NMSE and PE results for various distributions for $\mathbf{x}^{gen}$. The dictionary is i.i.d $\{0,1\}$ Bernoulli distributed.

		Φ is $\{0,1\}$ Bernoulli					
	$\mathbf{x}^{gen}$	NN-OMP	SLEP-ℓ_1	NNGM AMP	R-SBL (LMMSE)	R-SBL (GAMP)	R-SBL (DA)
NMSE	RG	0.2063	0.2497	Diverged	0.0873	0.0520	**0.0386**
	NN-Cauchy	0.0085	0.0188	Diverged	0.0031	0.0286	**0.0002**
	NN-Laplace	0.0960	0.1406	Diverged	0.0296	0.0070	**0.0043**
	Gamma	0.0901	0.1335	Diverged	0.0283	0.0047	**0.0022**
	Chi-square	0.0894	0.1360	Diverged	0.0327	0.0077	**0.0054**
	Bernoulli	0.2203	0.2586	Diverged	0.0747	0.0682	**0.0558**
PE	RG	0.3558	0.4070	0.8404	0.1782	0.0950	**0.0581**
	NN-Cauchy	0.2651	0.4434	0.8314	0.1131	0.4480	**0.0354**
	NN-Laplace	0.3140	0.4071	0.8398	0.1193	0.0371	**0.0134**
	Gamma	0.3102	0.4016	0.8354	0.1240	0.0275	**0.0087**
	Chi-square	0.3126	0.4072	0.8377	0.1341	0.0363	**0.0171**
	Bernoulli	0.3803	0.4120	0.8399	0.2689	0.1705	**0.1455**

given threshold. For example, when an index of the vector $\boldsymbol{\gamma}$ becomes smaller than i.e. $\gamma_i \leq \varepsilon_\gamma$, we ignore the computations regarding that index in the next iterations. This effectively reduces the problem dimensions and improves execution time.

In Figure 2.4, we included the average execution times of the proposed algorithms in units of seconds. The pruning threshold was selected as $\varepsilon_\gamma = 10^{-5}$ for all methods. For the EM based

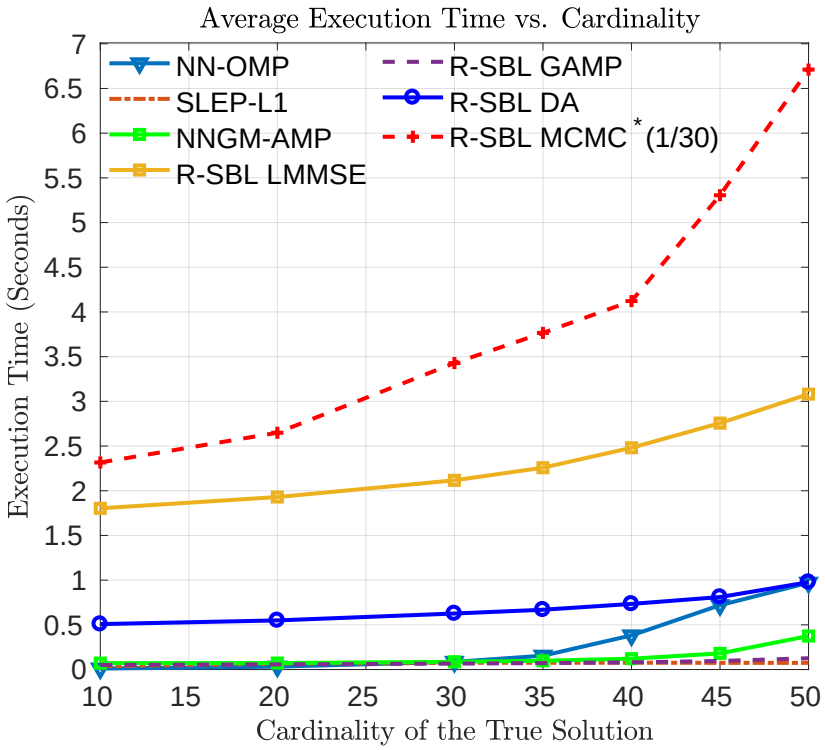


Figure 2.4: Execution times of the S-NNLS solvers as a function of cardinality for the noiseless scenario.

methods, we monitored the convergence of the γ's in EM iterations. We stopped the EM updates when $\|\gamma^t - \gamma^{t-1}\|_2 \leq 10^{-3}$, where t is the current EM iteration index. For other approaches, we monitored the linear equality constraints and stopped the algorithms when $\|\mathbf{y} - \mathbf{\Phi}\hat{\mathbf{x}}^t\|_2 \leq 10^{-3}$, where $\hat{\mathbf{x}}^t$ is the solution estimate at iteration t. As expected due to computationally intensive random sampling, R-SBL MCMC is the slowest method. For display purposes, we scaled down the average MCMC execution time values by 30. The LMMSE approach takes about 3 seconds for $K = 50$ to recover the optimal solution and is the second slowest method. Even though the complexity of DA and LMMSE is similar, DA achieves much faster convergence and takes about 0.5 to 1 seconds as K increases.

For this particular experiment, GAMP is the fastest R-SBL variant regardless of the cardinality and is similar to SLEP-ℓ_1. However, since the complexity of GAMP is $O(NM)$, for very large problem sizes (e.g. large N and M) GAMP may become slower despite superior recovery

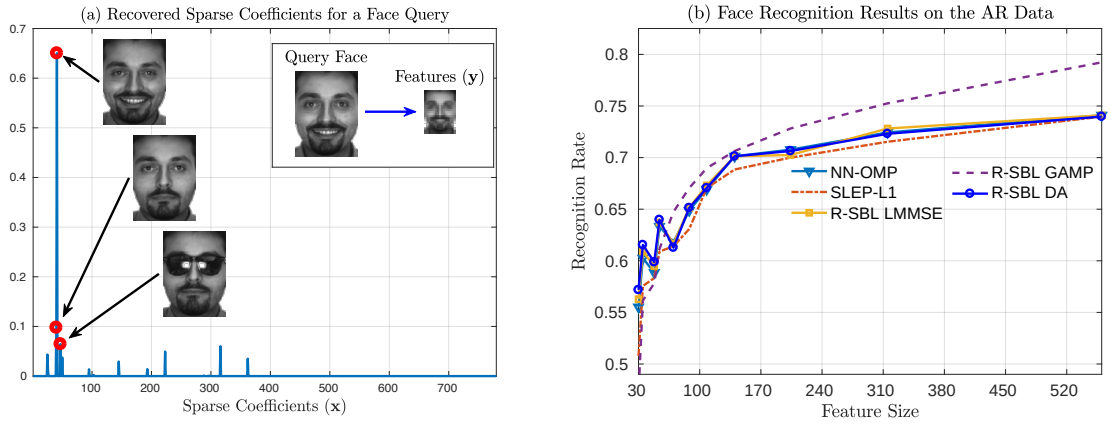


Figure 2.5: (a) Illustration of the sparse FR process. A query face is down-sampled to obtain an observation $\mathbf{y}$. Using the training dictionary Φ, a sparse solution is obtained using the R-SBL variants and baseline solvers to satisfy $\mathbf{y} = \Phi\mathbf{x}$. The index that corresponds to the maximum positive value in $\mathbf{x}$ is used to select a corresponding column in Φ. The image in this column corresponds to the correct individual. (b) FR accuracy for different feature sizes using all test samples. The R-SBL GAMP enjoys better FR performance for different feature sizes.

performance. In this case, a convex solver may be preferable depending on the desired recovery performance. R-SBL GAMP is faster than NNGM-AMP at larger cardinalities. Finally, NN-OMP is similar to SLEP-ℓ_1 but its execution time increases for larger cardinalities. Considering the fast recovery speed and good recovery performance of R-SBL GAMP under various Φ types, the R-SBL GAMP variant is a very good candidate for time sensitive sparse recovery applications.

2.5.8 Application: Face Recognition

Here, we present a face recognition (FR) application based on the non-negative sparse representations considered in [137, 138, 139]. Our goal is to show that the R-SBL approach works well in real-world applications involving real-data. A sparse representation classifier (SRC) for FR was initially proposed in [140] using the ℓ_1 penalty without the non-negativity constraints. The SRC approach was found to be robust against occlusion, disguise, pixel corruptions, and achieved superior results as compared to well-known FR algorithms [138, 140, 141, 142].

In the SRC framework, the dictionary Φ represents the training samples and each column

of Φ contains training features from a single face image. A single person may have more than one training image, and hence multiple columns of Φ might correspond to the same person. For a given test face $\mathbf{y}$ in vectorized form, a vector $\mathbf{x}$ is obtained by solving Eq. (2.1) using ℓ_1 sparsity, with the assumption that only a few non-zero entries will exist in the solution $\mathbf{x}$. Ideally, the index of the maximal non-negative entry in $\mathbf{x}$ is used to select the corresponding column in Φ. This column should correspond to one of the training samples for the correct person. In [138], the SRC performance was further improved by adding the non-negativity constraint on $\mathbf{x}$ in addition to the ℓ_1 sparsity. The authors have shown their algorithm to be more robust against noise and to be computationally more efficient as compared to the original SRC approach.

In our experiment, we consider the R-SBL framework for the FR problem and compare it with the baseline solvers. Note that SLEP-ℓ_1 was considered as the non-negative ℓ_1 minimization counterpart of R-SBL in place of [138]. We used the public AR dataset [143] and selected the first 30 males and 30 females for the FR problem. Each person in the dataset has 26 face images with different facial expression, illumination, and disguise (e.g. sunglasses and scarves). The first 13 images of each person ($M = 13 \times 60 = 780$) were selected as the training set, and the remaining 780 face images were used for testing. For feature selection, we used the down-sampling method used in [137, 138, 140], where the pixel dimensions of each face image were down-sampled to have a total of N pixels. In separate experiments, each 165×120 pixel image was down-sampled by a factor of $\{1/28, 1/26, ..., 1/6\}$, yielding feature dimensions of minimum of $N = 30$ to a maximum of $N = 650$.

The overall process is shown in Figure 2.5(a), where a query face is shown in the top right-hand side panel. This image was then down-sampled and the original feature dimension was reduced from $19,800$ to 512. After sparse recovery with R-SBL, the original faces belonging to several largest non-zero elements of $\mathbf{x}$ are shown. As desired, the maximal positive index of $\mathbf{x}$ belongs to the same person in the query face.

In Figure 2.5(b) we performed FR using all 780 samples in the test set and measured the

recognition rate for different feature sizes. The recognition rate was computed by counting the number of test samples for which R-SBL recovered the correct individual from $\mathbf{x}$. This count was normalized by 780. Overall, the R-SBL variants with the exception of R-SBL GAMP performed similar to the baseline solvers for large feature sizes. This is expected since the recovery problem was highly sparse, and the cardinality was very small $K = 13$ as compared to the length of $\mathbf{x}$ (i.e. largest length of $\mathbf{x}$ is 780). R-SBL GAMP was superior to all algorithms for large feature sizes and performed significantly better in identifying the correct individual. NNGM-AMP diverged for this application and did not yield reportable results.

2.6 Conclusion

In this work we introduced a hierarchical Bayesian method to solve the S-NNLS problem. We proposed the rectified Gaussian scale mixture model as a general and versatile prior to promote sparsity in the solution of interest. Since the marginals of the posterior were not tractable, we constructed our R-SBL algorithm using the EM framework with four different approaches. We demonstrated that our R-SBL approaches outperformed the available S-NNLS solvers, in most cases by a large margin. The proposed R-SBL framework is very robust to the structure of Φ and performed well regardless of Φ being i.i.d. and non-i.i.d. distributed. The performance gains achieved by the R-SBL variants are consistent across different non-negative data distributions for $\mathbf{x}$, and different structures for the design matrix Φ in coherent, low-rank, ill-posed and non-negative settings. The DA variant was found to be an easy to implement S-NNLS solver with simple closed-form moment expressions.

2.7 Application on fMRI data: Sparse Estimation of Quasi-periodic Spatiotemporal Components in functional MRI

The complex spatiotemporal nature of resting-state functional connectivity has received increasing attention over the recent years [144, 145]. Quasi-periodic spatiotemporal patterns have been observed in both animal and human fMRI studies [146, 147, 148]. These patterns are space-time brain activity patterns with a duration of tens of seconds that repeat in a somewhat periodic manner across the course of an fMRI scan.

2.7.1 Problem Definition

We define a spatiotemporal template across space (s) and time (t) as $P(s,t)$. This template repeats itself across the fMRI data at different time points such that sum of scaled and temporally delayed versions of the template form a global spatiotemporal component: $ST(s,t) = \sum_{i=1}^{k} x_i P(s,t-t_i)$, where t_i is the amount of delay for the ith recurring copy of the template and k is the number of counts the template appears within an fMRI scan. Based on this assumption we decompose the whole fMRI data $Y(s,t)$ as follows:

$$Y(s,t) = ST(s,t) + R(s,t) \tag{2.49}$$

$$= \sum_{i=1}^{k} x_i P(s,t-t_i) + R(s,t), \tag{2.50}$$

where the residual component $R(s,t)$ in Eq. (2.50) represents the part of the fMRI data that could not be explained by the spatiotemporal template $P(s,t)$. The estimation objective is to recover a set of non-negative weights $\mathbf{x}$ and estimate $ST(s,t)$ given $Y(s,t)$ and $P(s,t)$.

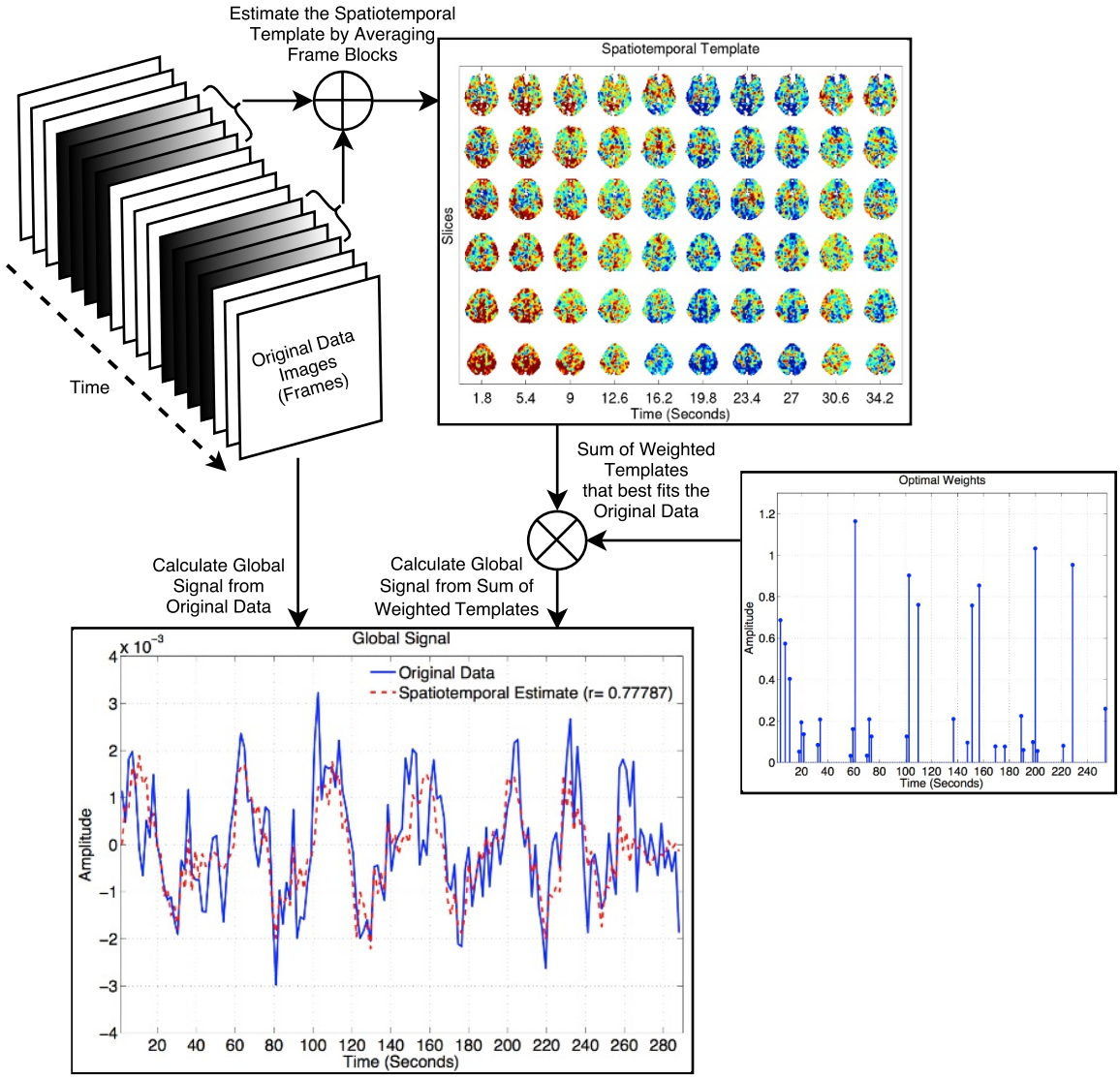


Figure 2.6: First, a spatiotemporal template (upper right) is estimated using the approach of [146]. Then, R-SBL (DA) is used to estimate the optimal weighted sum of templates that best fits the original data, with the recovered non-negative weights shown in lower right of the figure. The fMRI global signal of the original data (blue) is highly correlated ($r = 0.78$) with the global signal of the weighted sum.

2.7.2 Spatiotemporal Component Estimation

In Figure 2.6 we illustrate the spatiotemporal estimation approach. First, using the pattern matching algorithm as described in [146], we identified a basis spatiotemporal template $P(s,t)$

with a window duration of 36 seconds.[2] This template is a 4D matrix where the first 3 dimensions correspond to spatial coordinates for each voxel and the 4th dimension corresponds to time. We then vectorized the basis template to form a column vector $\mathbf{p}$. The fMRI data $\mathbf{Y}$ (also 4D matrix but longer in duration) were also vectorized to form a column vector $\mathbf{y}$. We constructed a large dictionary $\mathbf{D}$ (not shown in the figure) where each column contains a shifted and zero-padded version of $\mathbf{p}$. The number of rows in $\mathbf{D}$ was equal to the length of $\mathbf{y}$ and the shift between adjacent columns was equal to the number of voxels (number of spatial elements in the first 3 dimensions of $\mathbf{Y}$). We assumed that the spatiotemporal template repeated itself sparsely across the whole fMRI scan duration. Thus, we sought a representation of the data of the form: $\mathbf{y} \approx \mathbf{D}\mathbf{x}$ where $\mathbf{x}$ is an unknown vector of non-negative coefficients from Eq. (2.50). A solution to this problem can be obtained by solving the sparse signal recovery problem. We used the R-SBL framework (R-SBL DA variant) to estimate the sparse non-negative solution $\mathbf{x}$.

2.7.3 Data

Resting-state functional MRI data (5 minutes) were acquired from a healthy subject using a 3 Tesla GE MR750 scanner. We used echo-planar imaging (EPI) with 166 volumes, 30 slices, $3.4 \times 3.4 \times 5$ mm^3 voxel size, 64×64 matrix size, with TR $= 1.8$s and TE $= 30$ms. We discarded the initial 6 volumes to minimize longitudinal relaxation effects, leaving 160 frames for further analysis. All images were then slice-time corrected and co-registered. The resultant images were converted to coordinates of Talairach and Tournoux (TT), resampled to 3 mm cubic voxels, and spatially smoothed using a 6 mm full-width-at-half-maximum isotropic Gaussian kernel. The 1^{st} and 2^{nd} order Legendre polynomials (a constant term to model the temporal mean and a linear trend) were projected out from each voxel's time course. Each voxel time series was then converted into a percent change BOLD time series through demeaning and division by its mean

[2]Note that we describe an alternative way of obtaining a basic spatiotemporal template in [149] which uses singular value decomposition of the spatio-temporal data.

value.

2.7.4 Results

The recovered non-negative solution $\hat{\mathbf{x}}$ by the R-SBL framework is shown in Figure 2.6 on the right-hand side. This solution is 26% sparse and large non-zero coefficients appear in quasi-periodic intervals with approximately about 40 to 50 second gaps between the large peaks. This means that the spatiotemporal template repeats itself quasi-periodically across an fMRI scan of 5 minutes.

To assess whether the estimated spatiotemporal component $\mathbf{D}\hat{\mathbf{x}}$ captured the true spatiotemporal structure in $\mathbf{y}$ we computed empirical measures of the spatiotemporal correlation of the original fMRI data, the estimated spatiotemporal component $\mathbf{D}\hat{\mathbf{x}}$ and the residual data $\mathbf{y} - \mathbf{D}\hat{\mathbf{x}}$. This was achieved by computing the spatiotemporal correlation between all possible pairs of vectorized space-time blocks of the respective data, where the duration of each spatiotemporal block was fixed as 36 seconds. The results are displayed as a spatiotemporal correlation matrix (STCM) where the $(i,j)^{th}$ entry corresponds to the correlation between the space-time blocks at times i and j.

Figure 2.7 shows the STCM for the original data, the estimated spatiotemporal component, and the residual term. The STCM for the estimated spatiotemporal component largely captures the quasi-periodic structure seen in the STCM of the original data. This structure is greatly attenuated in the STCM of the residual data.

Finally, we computed the whole-brain averaged global signal (GS) of the original data $\mathbf{y}$ and compared it to the whole-brain GS of the estimated spatiotemporal component $\mathbf{D}\hat{\mathbf{x}}$. These time courses are shown at the bottom of Figure 2.7 with blue and dashed red lines for the original and spatiotemporal data, respectively. The GS of the original data was significantly correlated ($r = 0.78$, $p < 10^{-3}$) with the GS of the weighted sum of templates. This indicates that in average the spatio-temporal part of data $ST(s,t)$ as estimated by R-SBL accounts for a large variance 60%

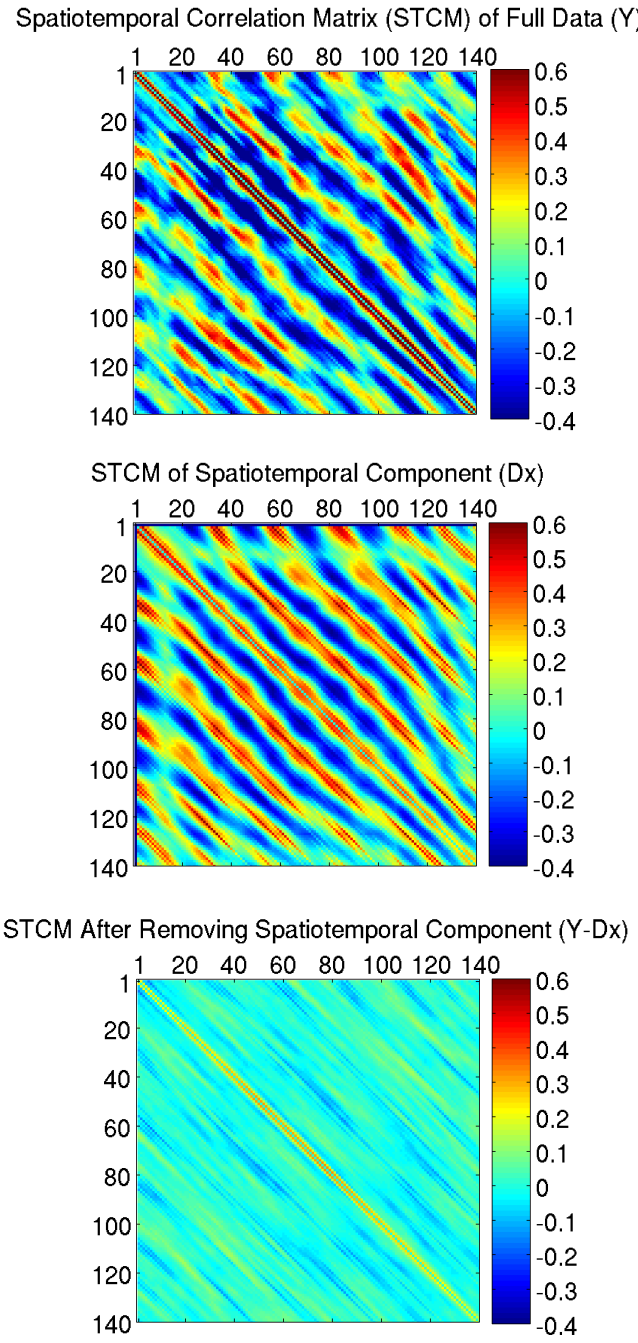


Figure 2.7: Spatiotemporal correlation matrix (STCM) for the original data (first row), the estimated spatiotemporal component (second row), and the residual term (last row). Color maps on the right-hand side show the correlation values obtained between the space time blocks taken from the respective data at time indexes i and j. The blue and red diagonal stripes indicate that a space-time data block at time i correlates quasi-periodically well with other blocks with an approximate periodicity of 20 to 40 time points where the duration between consecutive time points is 1.8 seconds. The estimated spatiotemporal component $\mathbf{D}\hat{\mathbf{x}}$ in the second row has a clear quasi-periodic correlation structure that largely captures the correlation structure from the original data. Removal of this data results in a "flat" or cleaned up STCM in the residual data.

in the entire fMRI data $Y(s,t)$.

2.7.5 Discussion

We have presented an approach for estimating repeating spatiotemporal patterns in resting-state fMRI data and demonstrated the approach using a sample resting-state fMRI data. This preliminary finding suggests that the activity represented by sparsely repeating spatiotemporal pattern may account for a significant fraction of the fMRI global signal and a large variance in the fMRI data. However, further work is needed to understand both the origins of the quasi-periodic patterns and their potential link with the global signal.

2.8 Appendix

2.8.1 Full derivation of GAMP

We use the R-GSM prior $p(x|\gamma)$ and evaluate Eq. (2.37) and Eq. (2.38) to find the first two moments of the approximate marginal posterior under the sum-product GAMP mode

$$\hat{x}_i = \mathbb{E}\{x_i|r_i;\tau_{r_i}\} = \int_{x_i} x_i p(x_i|r_i;\tau_{r_i}) \tag{2.51}$$

$$= \int_{+} x_i \mathcal{N}^R(x_i|0,\gamma_i)\mathcal{N}(x_i,r_i,\tau_{r_i}), \tag{2.52}$$

then using the Gaussian multiplication rule[3], we obtain

$$\hat{x}_i = \int_{+} x_i \Upsilon \mathcal{N}^R(x_i|\eta_i,\nu_i), \tag{2.53}$$

where η_i and ν_i are given in Eq. (2.55) and Eq. (2.56), respectively.

[3] $\mathcal{N}(x;\mu_a,\tau_a)\mathcal{N}(x;\mu_b,\tau_b) = \Upsilon\mathcal{N}(x;\frac{\frac{\mu_a}{\tau_a}+\frac{\mu_b}{\tau_b}}{\frac{1}{\tau_a}+\frac{1}{\tau_b}},\frac{1}{\frac{1}{\tau_a}+\frac{1}{\tau_b}})$, where Υ is a scaling factor.

We then find the mean of the resulting rectified Gaussian

$$\hat{x}_i = \eta_i + \sqrt{\nu_i} h(\frac{\eta_i}{\nu_i}) \tag{2.54}$$

$$\eta_i = \frac{r_i \gamma_i}{\tau_{r_i} + \gamma_i} \tag{2.55}$$

$$\nu_i = \frac{\tau_{r_i} \gamma_i}{\tau_{r_i} + \gamma_i} \tag{2.56}$$

$$h(a) = \frac{\varphi(a)}{\Phi_c(a)}, \tag{2.57}$$

where φ refers to the pdf and Φ_c refers to the complementary cdf of a zero-mean and unit-variance Gaussian distribution. The conditional variance of x_i given r_i is simply

$$\tau_{x_i} = var\{x_i | r_i; \tau_{r_i}\} = \int_{x_i} x_i^2 p(x_i | r_i; \tau_{r_i}) - \hat{x}_i^2 \tag{2.58}$$

$$= \int_+ x_i^2 \mathcal{N}^R(x_i | 0, \gamma_i) \mathcal{N}(x_i, r_i, \tau_{r_i}) - \hat{x}_i^2, \tag{2.59}$$

using the Gaussian multiplication rule

$$\tau_{x_i} = \int_+ x_i^2 \Upsilon \mathcal{N}^R(x_i | \eta_i, \nu_i), \tag{2.60}$$

we find the variance of the resulting rectified Gaussian as

$$\tau_{x_i} = \nu_i g(\frac{\eta_i}{\nu_i}) \tag{2.61}$$

$$g(a) = 1 - h(a)\left(h(a) - a\right). \tag{2.62}$$

In the case of max-sum GAMP implementation, we evaluate Eq. (2.39) and Eq. (2.40)

$$\hat{x}_i = \arg\min_{\hat{x}_i \geq 0} \frac{x_i^2}{2\gamma_i} + \frac{1}{2\tau_{r_i}} |\hat{x}_i - r_i|^2 \tag{2.63}$$

$$\hat{x}_i = \begin{cases} \frac{r_i \gamma_i}{\tau_{r_i} + \gamma_i} = \eta_i & \text{if } \hat{x}_i \geq 0 \\ 0 & \text{if } \hat{x}_i < 0 \end{cases} \tag{2.64}$$

Using Eq. (2.40)

$$\tau_{x_i} = \begin{cases} \frac{\tau_{r_i} \gamma_i}{\tau_{r_i} + \gamma_i} = \nu_i & \text{if } \hat{x}_i \geq 0 \\ 0 & \text{if } \hat{x}_i < 0^4 \end{cases} \tag{2.65}$$

Upon convergence of the max-sum, the approximate marginals are obtained using Eq. (2.54) and Eq. (2.61).

2.8.2 Approximate marginals and moments using DA

We derive the approximate moments used in the R-SBL DA approximation. We start with the posterior $p(\mathbf{x}\,|\,\mathbf{y},\boldsymbol{\gamma})$ and use chain rule to write

$$p(\mathbf{x}\,|\,\mathbf{y},\boldsymbol{\gamma}) = \frac{p(\mathbf{y}\,|\,\mathbf{x},\boldsymbol{\gamma})p(\mathbf{x}\,|\,\boldsymbol{\gamma})}{\int_{\mathbf{x}} p(\mathbf{y}\,|\,\mathbf{x},\boldsymbol{\gamma})p(\mathbf{x}\,|\boldsymbol{\gamma})d\,\mathbf{x}}. \tag{2.66}$$

Here $p(\mathbf{y}\,|\,\mathbf{x},\boldsymbol{\gamma})$ is a Gaussian density due to the Gaussian noise assumption. Since $p(\mathbf{x}\,|\,\boldsymbol{\gamma})$ is a rectified Gaussian density the numerator of Eq. (2.66) is a Gaussian multiplied by a rectified

[4] Practically it was found that setting $\tau_{x_i} = 0$ when $\hat{x}_i < 0$ increases the chances of the algorithm getting stuck at a local minimum. Instead, we set $\tau_{x_i} = \frac{\tau_{r_i}\gamma_i}{\tau_{r_i}+\gamma_i} = \nu_i$.

Gaussian, which results in a rectified Gaussian density. Then, we can simply write

$$p(\mathbf{x}\,|\,\mathbf{y},\gamma) = c(\mathbf{y}) e^{-\frac{(\mathbf{x}-\mu)^T \Sigma^{-1} (\mathbf{x}-\mu)}{2}} u(\mathbf{x}), \tag{2.67}$$

where $c(\mathbf{y})$ is the normalizing constant for the posterior density and μ and Σ are given by Eqs. (2.22) and (2.23), respectively. Let $\Sigma = \mathbf{L}\mathbf{L}^T$ and $\mathbf{r} = \mathbf{x} - \mu$, so that $d\,\mathbf{x} = d\mathbf{r}$ and $\Sigma^{-1} = \mathbf{L}^{-T}\mathbf{L}^{-1}$. Therefore, we have

$$1 = c(\mathbf{y}) \int_{-\mu}^{\infty} e^{\frac{-\mathbf{r}^T \mathbf{L}^{-T} \mathbf{L}^{-1} \mathbf{r}}{2}} d\mathbf{r}. \tag{2.68}$$

Now, let $\mathbf{z} = \mathbf{L}^{-1}\mathbf{r}$, which implies that $d\mathbf{r} = |\mathbf{L}| d\mathbf{z}$ and

$$c(\mathbf{y}) = \frac{1}{|\mathbf{L}| \int_{-\beta}^{\infty} e^{-\mathbf{z}^T \mathbf{z}/2} d\mathbf{z}}, \tag{2.69}$$

where $\beta = \mathbf{L}^{-1}\mu$ is the lower limit of the new integral in vector form. The lower limit β depends on a linear combination of elements of μ since $\mathbf{L}$ is not diagonal. Thus, the integral in the denominator of Eq. (2.69) is **not tractable** as the integration limits are not separable and the multidimensional integral over $\mathbf{z}$ in Eq. (2.69) is not separable as a product of one dimensional integrals.

Assume that, we are interested in an approximate density $\tilde{p}(\mathbf{x}\,|\,\mathbf{y},\gamma)$, instead of the exact posterior. We calculate an approximate $\tilde{c}(\mathbf{y})$ by approximating Σ with its diagonal i.e. $\Sigma_d = diag(\Sigma) \approx \Sigma$. In this case, the new $\mathbf{L}$ is diagonal with entries $\sqrt{\Sigma_{ii}}$. Thus, the integral in Eq. (2.69) is separable and the approximate normalizing constant $\tilde{c}(\mathbf{y})$ has closed form

$$\tilde{c}(\mathbf{y}) = \frac{1}{|\Sigma_d|^{1/2} \prod_{i=1}^{M} \sqrt{\frac{\pi}{2}}\,\mathrm{erfc}\left(-\frac{\mu_i}{\sqrt{2\Sigma_{ii}}}\right)}. \tag{2.70}$$

Approximating the actual normalizing constant with $\tilde{c}(\mathbf{y})$, we write the approximate posterior as

$$\tilde{p}(\mathbf{x}\,|\,\mathbf{y},\boldsymbol{\gamma}) = \frac{e^{-\frac{(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}_d^{-1}(\mathbf{x}-\boldsymbol{\mu})}{2}}u(\mathbf{x})}{\prod_{i=1}^{M}\sqrt{\frac{\pi\Sigma_{ii}}{2}}\,\mathrm{erfc}\left(-\frac{\mu_i}{\sqrt{2\Sigma_{ii}}}\right)} \tag{2.71}$$

$$= \prod_{i=1}^{M}\sqrt{\frac{2}{\pi\Sigma_{ii}}}\frac{e^{-\frac{(x_i-\mu_i)^2}{2\Sigma_{ii}}}u(x_i)}{\mathrm{erfc}\left(-\frac{\mu_i}{\sqrt{2\Sigma_{ii}}}\right)} \tag{2.72}$$

$$= \prod_{i=1}^{M}\tilde{p}(x_i|\mathbf{y},\boldsymbol{\gamma}) \tag{2.73}$$

Eq. (2.73) shows that multivariate $\tilde{p}(\mathbf{x}\,|\,\mathbf{y},\boldsymbol{\gamma})$ is separable into product of univariate densities. The univariate density $\tilde{p}(x_i|\mathbf{y},\boldsymbol{\gamma})$ is the univariate RG density defined in Eq. (2.10) e.g. $\tilde{p}(x_i|\mathbf{y},\boldsymbol{\gamma}) = \mathcal{N}^R(x_i;\mu_i,\Sigma_{ii})$. The first and second moments of a univariate RG density are well-known in closed form (i.e. Eqs. (2.45) and (2.46)) and are used in the R-SBL DA algorithm.

2.9 Acknowledgments

Chapter 2 is, in part, a reprint of the material published in two papers: (1) Alican Nalci, Igor Fedorov, Maher Al-Shoukairi, Thomas T. Liu, and Bhaskar D. Rao. "*Rectified Gaussian Scale Mixtures and the Sparse Non-negative Least Squares Problem.*" IEEE Transactions on Signal Processing 66, No. 12 (2018): 3124-3139, and (2) Alican Nalci, Bhaskar D. Rao, and Thomas T. Liu. "*Sparse Estimation of Quasi-Periodic Spatiotemporal Components in Resting-State FMRI,*" in 24th Annual Scientific Meeting of the ISMRM, 2016, p. 3824. I was the primary author in both papers and the research was supervised by B.D. Rao and T.T. Liu.

Chapter 3

Global Signal Regression Acts as a Temporal Downweighting Process in Resting-State fMRI

In resting-state functional MRI (rsfMRI), the correlation between blood oxygenation level dependent (BOLD) signals across different brain regions is used to estimate the functional connectivity of the brain. This approach has led to the identification of a number of resting-state networks, including the default mode network (DMN) and the task positive network (TPN). Global signal regression (GSR) is a widely used pre-processing step in rsfMRI that has been shown to improve the spatial specificity of the estimated resting-state networks. In GSR, a whole brain average time series, known as the global signal (GS), is regressed out of each voxel time series prior to the computation of the correlations. However, the use of GSR is controversial because it can introduce artifactual negative correlations. For example, it has been argued that anticorrelations observed between the DMN and TPN are primarily an artifact of GSR. Despite the concerns about GSR, there is currently no consensus regarding its use. In this paper, we introduce a new framework for understanding the effects of GSR. In particular, we show that the

main effects of GSR can be well approximated as a temporal downweighting process in which the data from time points with relatively large GS magnitudes are greatly attenuated while data from time points with relatively small GS magnitudes are largely unaffected. Furthermore, we show that a limiting case of this downweighting process in which data from time points with large GS magnitudes are censored can also approximate the effects of GSR. In other words, the correlation maps obtained after GSR show a high degree of spatial similarity (including the presence of anticorrelations between the DMN and TPN) with maps obtained using only the uncensored (i.e. retained) time points. Since the data from these retained time points are unaffected by the censoring process, this finding suggests that the observed anticorrelations inherently exist in the data from time points with small GS magnitudes and are not simply an artifact of GSR.

3.1 Introduction

In resting-state functional magnetic resonance imaging (rsfMRI), the correlation between the blood oxygenation level dependent (BOLD) signals from different brain regions is used to estimate the functional connectivity of the brain in the absence of an explicit task [150]. A set of regions that shows a high degree of mutual correlation is referred to as a resting-state network (RSN). Since the initial observation by [151] of resting-state correlations in the motor RSN, many other RSNs have been identified, including visual, auditory, and language networks [152,153,154], the default mode network (DMN) [49], and the task positive network (TPN) [50]. The DMN and the TPN consist of those brain regions that exhibit decreases and increases, respectively, in metabolic activity during the execution of attention demanding tasks [49]. In the resting-state it has been observed that signals in the DMN are negatively correlated with signals in the TPN, giving rise to the notion of the DMN and TPN as anti-correlated networks [50].

Many rsfMRI studies employ a pre-processing step known as global signal regression (GSR) in which a global mean time course is regressed out of each voxel time course prior to

the computation of the correlations. However, the use of GSR is controversial as it has been shown that the process can create artifactual negative correlations [51, 52, 155, 156, 157]. In particular, it has been been claimed that the observed anticorrelations between the DMN and TPN are largely a mathematical artifact introduced by GSR [52]. In response to these concerns, [51] acknowledged the validity of the mathematical argument put forth by [52], but argued that the characteristics of the correlation maps obtained with GSR could not be solely determined by the mathematical constraint. For example, they showed that negative correlations between the DMN and TPN were present even without the application of GSR, an observation supported by subsequent studies [158, 159, 160, 161]. To address some of the concerns regarding GSR, alternative approaches for either reducing global signal effects or determining when the application of GSR is valid have been proposed [158, 162, 163]. Related studies have demonstrated that there are significant neural contributions to the global signal [164, 165, 166] and have also examined the potential diagnostic value of the signal [67].

However, despite the growing understanding of the global signal and the potential limitations of GSR, a consensus regarding the use of GSR is lacking. Because the application of GSR greatly improves the spatial specificity of functional connectivity maps [51, 162], it is still widely used [167, 168] even in the face of the strong concerns that have been raised [169, 170]. From our perspective, this lack of consensus partly reflects the difficulty in thinking clearly about the process of GSR, even though it is a relatively straightforward and compact mathematical operation. In the prior work, mathematical proofs that describe the general characteristics of GSR (e.g. the sum of correlation values must be negative) or simplified models based on a few brain regions have been used to examine the effects of GSR [51, 52, 157]. While these approaches have provided useful perspectives on GSR, it has been difficult to leverage them to develop an intuitive understanding of the specific effects of GSR when it is applied to a typical rsfMRI dataset that has hundreds of time points and tens of thousands of voxels.

In this paper, we introduce a simple framework for understanding the effects of GSR.

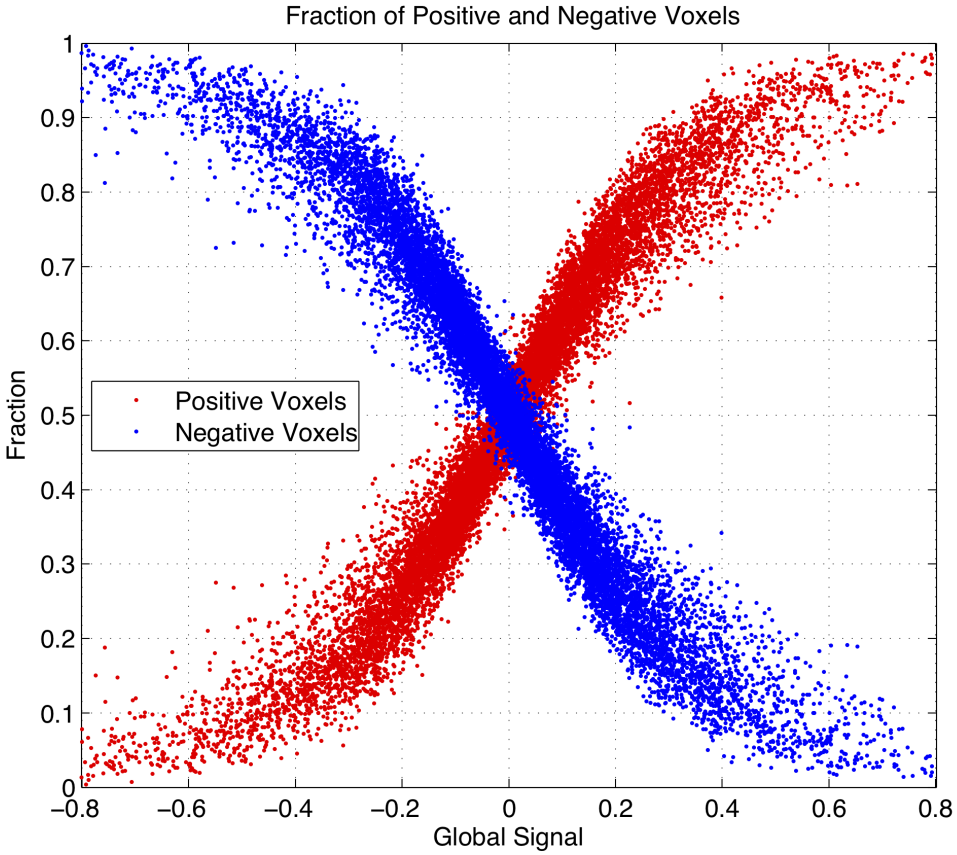


Figure 3.1: Fractions of positive (red) and negative (blue) voxels versus GS value for each time point in the dataset described under Methods. The GS magnitude is higher when a large fraction of the voxels have the same sign, and tends towards zero when there are roughly equal proportions of positive and negative signal values.

In particular, we show that the main effects of GSR can be well approximated by a temporal downweighting process in which the data from different time points are attenuated as a function of their global signal magnitude (i.e., absolute value), such that the time points with the highest global signal magnitude experience the greatest amount of attenuation. We also show that a limiting case of the downweighting process in which time points with a global signal magnitude greater than a specified threshold are censored (i.e. excluded) can provide a good approximation to GSR. Furthermore, we demonstrate that the ability to model the effects of GSR as a temporal downweighting or censoring process can provide insight into the controversy regarding the existence of anti-correlated networks.

3.2 Theory

3.2.1 The global signal as a time-varying measure of spatial homogeneity

As a starting point, it is useful to review the basic properties of the global signal (GS). In describing these properties we focus on providing a reasonable description of the average behavior that is observed empirically in the experimental data, and acknowledge that this "average" description is not intended to cover all possible cases.

The value of the GS at each time point is simply the average of the BOLD percent signal change values across all voxels in the brain. In considering the empirically observed properties of the GS, we find that the value of the GS will be positive when the majority of the voxels have a positive BOLD signal value and will be negative when the majority of voxels have a negative signal value. At time points where there are roughly equal proportions of positive and negative signal values, these signal components will tend to cancel out and the GS magnitude will tend towards zero. Thus, the GS magnitude will be higher at time points when there is a relatively high level of spatial homogeneity (e.g. most voxels have a positive BOLD signal) and lower for time points where the data exhibits spatial heterogeneity (e.g. a roughly equal mix of positive and negative values across the brain). These properties are demonstrated in Figure 3.1, where we plot the fraction of positive (red) and negative voxels (blue) versus the GS value for each timepoint in the dataset described under Methods. Large GS magnitudes (i.e large positive or negative values) correspond to zones where there is either a large fraction of either positive or negative voxels, respectively, whereas low GS magnitudes (i.e. small positive or negative values) correspond to zones where there are roughly equal fractions of positive and negative voxels. The GS magnitude can also be relatively low for time points when the BOLD values across voxels are uniformly small.

Another view of the basic properties of the GS is provided in Figure 3.2, where the BOLD images from a representative subject and slice are shown over 45 time points in panel

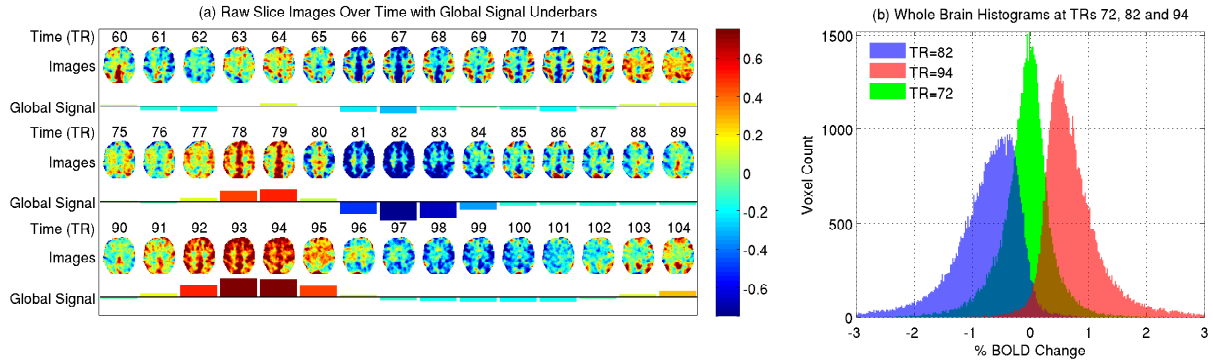


Figure 3.2: (a) Examples of brain images from a representative subject and slice with GS values represented by colored underbars for each image. A total of 45 consecutive time points are used to demonstrate a range of GS values. Large GS magnitudes correspond to brain images with greater spatial homogeneity, while small GS magnitudes correspond to images where there are similar proportions of negative and positive voxel values. (b) Histogram of whole brain voxel values corresponding to three selected time points with small (TR = 72) and large (TR = 82 and 94) GS magnitudes. The centers of mass of these histograms correspond to the GS values at their respective time points.

(a), along with colored bars below them indicating the value of the GS at each time point and histograms (in panel (b)) of the BOLD signal values at 3 selected time points. At time point 72, the image is relatively heterogeneous with a roughly equal mix of negative and positive values. The corresponding histogram is centered about zero and the GS magnitude is close to zero at this time point. In contrast, at time point 82 the image is relatively homogeneous with negative values, leading to a large negative value for the GS. The corresponding histogram clearly shows that the overwhelming majority of voxels have a negative value at this time point. At time point 94 the image is mostly positive with a large positive GS value. Taken together, the observations in Figures 3.1 and 3.2 suggest that the magnitude of the GS can be thought of as a time-varying measure of spatial homogeneity.

3.2.2 Global Signal Regression primarily affects time points where the GS magnitude is high

In this section we take a closer look at the process of global signal regression (GSR). We should note that the simple examples and arguments presented in this section are meant primarily to provide background and motivation for the experimental measures (described in Methods and Results) that empirically demonstrate the temporal downweighting effects of GSR.

Our basic observation is that GSR largely affects those time points where the magnitude of the GS is high. To see why this is the case, we first consider the toy example shown in the left column of Figure 3.3, where we have designed a set of three signals to highlight the properties of GSR. Each of the signals is constructed as the sum of six Gaussian waveforms of varying magnitudes and time shifts. In Figure 3.3 (a), the three signals are designed such that the Gaussian waveforms perfectly cancel out in the intervals centered about time points 30, 70, 90, and 110 and sum in a constructive fashion around time points 10 and 50. As a result, the resulting GS (computed as the mean of the three signals) in panel (d) has a large positive peak for the interval around time point 10 and a large negative peak for the interval around time point 50 (indicated by the thick black bars), but is otherwise equal to zero. Panel (g) shows the results of applying GSR to each of the original signals from panel (a). Because the GS consists solely of Gaussian waveforms centered about times points 10 and 50, the process of regression completely eliminates the signal components in these intervals. In contrast, regression has no effect on time points where the GS is equal to zero, so that the Gaussians centered about time points 30, 70, 90, and 110 are untouched by GSR.

In panel (b), we have slightly modified the signals from panel (a) so that there is no longer perfect cancellation of the Gaussian waveforms for the intervals centered about time points 30 and 70. As a result, the GS in panel (e) has small but non-zero values in these intervals. For this case, GSR still largely attenuates the waveforms at time points 10 and 50 with minimal attenuation

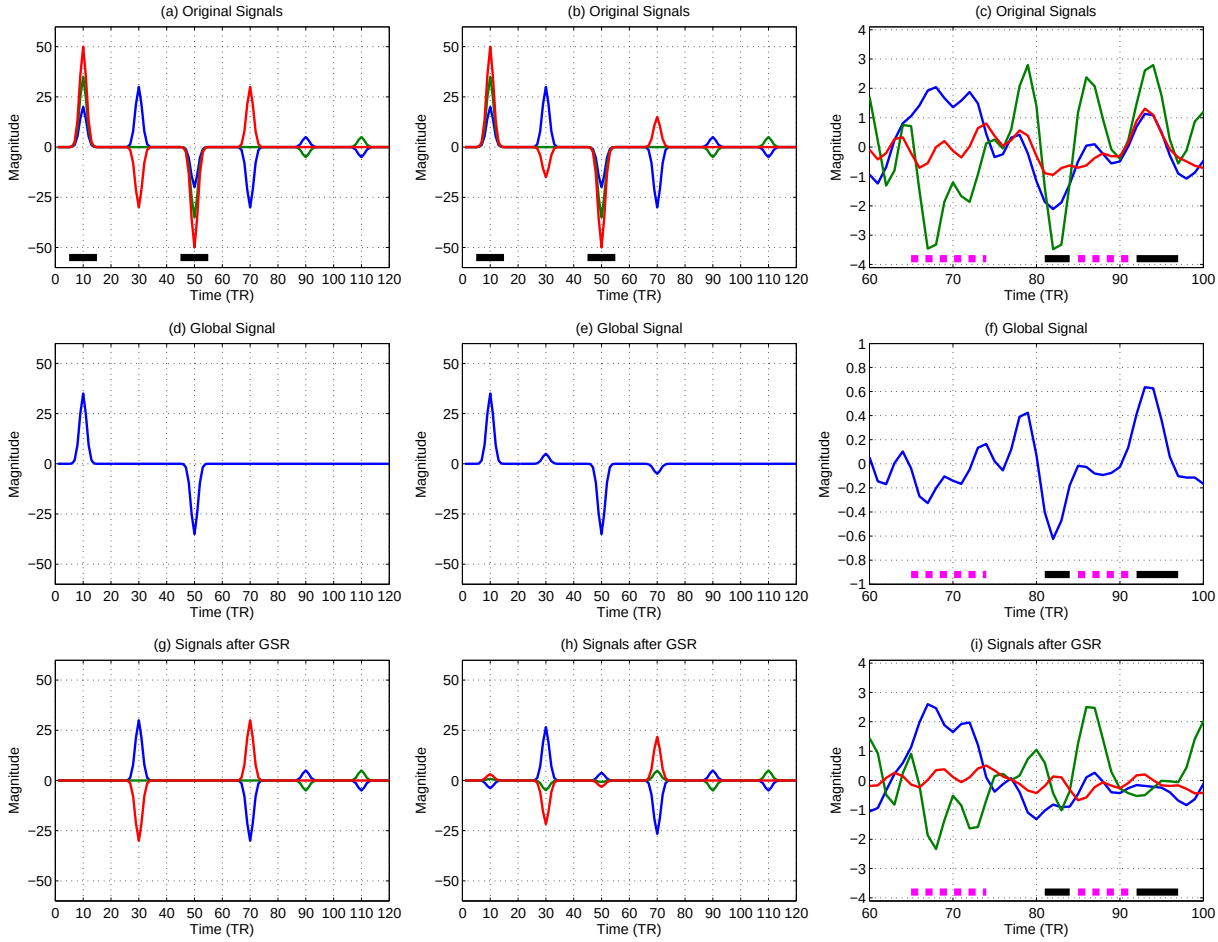


Figure 3.3: (a) Three voxel time series are constructed such that the Gaussian waveforms at time points 30, 70, 90 and 110 sum to zero, whereas the waveforms at time points 10 and 50 (indicated with thick black bars) sum constructively. This summation results in the large peaks at time points 10 and 50 in the GS shown in (d). Panel (g) shows that the original waveforms at time points 30, 70, 90 and 110 are preserved after GSR, whereas the waveforms at times points 10 and 50 are completely eliminated. Panels (b,e,h) demonstrate a more realistic case where the cancellation of signals at time points 30 and 70 is not complete and the GS has small values at these points. GSR greatly attenuates the waveforms at time points 10 and 50 and has a slight effect on the waveforms at time points 30 and 70. Panels (c,f,i) show three example BOLD time series where the GS in (f) is computed over the entire brain. The magenta and black lines indicates regions with low and high GS magnitudes, respectively. GSR greatly attenuates signals in the high GS regions and has a smaller effect on signals in the low GS regions.

of the waveforms at time points 30 and 70 (as shown in panel (h)). This reflects the fact that the computation of the regression coefficient in GSR is dominated by those time points with larger GS magnitudes, so that scaled versions of the GS will have a better fit to the voxel time series at these time points. From a mathematical point of view, those time points with larger GS magnitude have greater leverage on the regression [171, 172]. Denoting the global signal as the column vector $\mathbf{g}$, the leverage at each time point is given by the corresponding diagonal term of the projection matrix $\mathbf{g}(\mathbf{g}^T\mathbf{g})^{-1}\mathbf{g}^T$. Since $(\mathbf{g}^T\mathbf{g})^{-1}$ is simply a scalar, the leverage at the ith time point is proportional to the square g_i^2 of the corresponding global signal value. The concept of leverage is useful for understanding that time points with large magnitudes tend to be more influential in the regression, but the detailed effects of GSR require consideration (and therefore computation) of the entire projection matrix [172]. As a result, while we use theoretical concepts such as leverage to help motivate our approach, we need empirical measures to demonstrate the actual effects.

In Figure 3.3 (c), we show three BOLD time series from a representative subject (see Methods for details on the data). The GS for the dataset is shown in panel (f), where in contrast to the toy example, the GS is computed over all voxel time series (of which only 3 out of about 50,000 are shown in the plot). The dashed magenta and solid black lines correspond to time intervals in which the GS magnitude is either relatively small or large, respectively. After GSR (see panel (i)), the waveforms are greatly attenuated for those intervals where the GS magnitude is large and minimally affected for the intervals where the GS magnitude is small.

In summary, the examples shown above suggest that the main effect of GSR is to reduce the magnitude of waveforms in those time intervals in which the GS has a relatively large magnitude. In the appendix we present mathematical arguments that provides greater insight into this effect. Specifically, we introduce the concept of splitting the global signal $\mathbf{g}$ into two orthogonal components $\mathbf{g} = \mathbf{g}_H + \mathbf{g}_L$ where the subscripts H and L denote the sets of time points where the global signal is high and low, respectively. A voxel time series $\mathbf{x}$ can be similarly

decomposed into the sum $\mathbf{x} = \mathbf{x}_H + \mathbf{x}_L$ of two components. In addition, since GSR is a linear operation, the voxel time series after regression $\tilde{\mathbf{x}}$ can be written as the sum $\tilde{\mathbf{x}} = \tilde{\mathbf{x}}_H + \tilde{\mathbf{x}}_L$. In the appendix we show that $\tilde{\mathbf{x}} \approx \mathbf{x}_L$, such that to first order GSR can be viewed as attenuating the signal component $\mathbf{x}_H$ in temporal regions where the global signal is high while having a relatively small effect on the component $\mathbf{x}_L$ for intervals where the global signal is low. If the global signal is identical to zero for all time points in the set L, the approximation becomes an equality (i.e. $\tilde{\mathbf{x}} = \mathbf{x}_L$), consistent with the observations related to the toy example in the leftmost column of Figure 3.3.

As noted above, both the toy example and the mathematical perspectives are meant to provide background and insight that will be useful for understanding the extensive empirical findings that are presented in later sections. Our focus has been on presenting examples and arguments that help to explain the average behavior that is empirically observed, and it is certainly possible to come up with exceptions. For example, if the voxel time series is uncorrelated with the global signal, then GSR will have no effect on the time series and therefore there will also be no reductions in the GS magnitude. However, given the widespread correlation observed between the GS and voxel time series [173], uncorrelated voxels are the exception and do not drive the average behavior.

3.2.3 Effect of GSR on Seed-Based Correlations

We now consider the effects of GSR on the computation of the correlation coefficients between voxel time series. Without loss of generality, we will assume that the voxel time series have been normalized to be unit norm and zero mean. Then the correlation coefficient r between a seed voxel time series $\mathbf{s}$ and another voxel time series $\mathbf{x}$ is simply given by the inner product $r = \mathbf{s}^T\mathbf{x}$, where the time series are represented as column vectors. Using the orthogonal decomposition described above and in the Appendix, this may be rewritten as $r = \mathbf{s}_H^T\mathbf{x}_H + \mathbf{s}_L^T\mathbf{x}_L$. Similarly, the correlation coefficient after GSR (denoted as $\tilde{r}$) can be written as $\tilde{r} = \beta\tilde{\mathbf{s}}^T\tilde{\mathbf{x}} = \beta\left(\tilde{\mathbf{s}}_H^T\tilde{\mathbf{x}}_H + \tilde{\mathbf{s}}_L^T\tilde{\mathbf{x}}_L\right)$,

where the tilde notation indicates the values after GSR and the coefficient $\beta = (\|\tilde{\mathbf{s}}\|\|\tilde{\mathbf{x}}\|)^{-1}$ is needed to renormalize the signals. Using the approximations discussed in the prior section, the signals after GSR can be expressed as $\tilde{\mathbf{s}} \approx \mathbf{s}_L$ and $\tilde{\mathbf{x}} \approx \mathbf{x}_L$, so that the correlation coefficient can be approximated as $\tilde{r} \approx \beta \mathbf{s}_L^T \mathbf{x}_L$. In other words, the correlation coefficient after GSR is largely determined by those time points where the global signal has low magnitude. This argument can be extended to correlation maps, by defining a data matrix $\mathbf{X}$ consisting of individual voxel time series as columns which are normalized to have unit norm. The correlation map in vectorized form is then given by $\mathbf{R} = \mathbf{s}^T \mathbf{X}$, which can be approximated as $\tilde{\mathbf{R}} \approx \text{diag}(\beta)\, \mathbf{s}_L^T \mathbf{X}_L$ after GSR, where $\text{diag}(\beta)$ denotes the matrix with the vector β of renormalizing coefficients along the diagonal. This result suggests the following testable hypothesis: functional correlation maps obtained with GSR should resemble correlation maps obtained from a subset of the original dataset corresponding to time points where the global signal magnitude is low.

To further demonstrate the reasoning behind this hypothesis, we note that a correlation map can also be written as the sum of weighted images $\mathbf{R} = \sum_{t=1}^{T} s_t \mathbf{X}_t$ where s_t and $\mathbf{X}_t$ represent the unit norm seed signal value and the data image, respectively, acquired at time t. (Note $\mathbf{X}_t$ is the t'th row in $\mathbf{X}$). In Figure 3.4 we show examples of raw and weighted brain images (both before and after GSR) along with the corresponding correlation maps. The weighted images are obtained by multiplying the raw images by a seed voxel time series (represented by the colored bars) from the posterior cingulate cortex (PCC; see Methods). Due to space constraints, we only show representative images where the GS magnitude is very small (time points 66 to 68) or relatively large (time points 81 to 84 and 92 to 95). However, the correlation map is obtained from the sum of all weighted images (most of which are not shown). Because GSR reduces the magnitudes of both the PCC time series and the raw images at time points where the GS is large, the contribution of the weighted images at these time points to the overall correlation map is greatly diminished. As a result, the correlation map will be dominated by those time points where the GS is small in magnitude, i.e. $\tilde{\mathbf{R}} \approx \text{diag}(\beta)\, \mathbf{s}_L^T \mathbf{X}_L$ as noted above. This initial result suggests

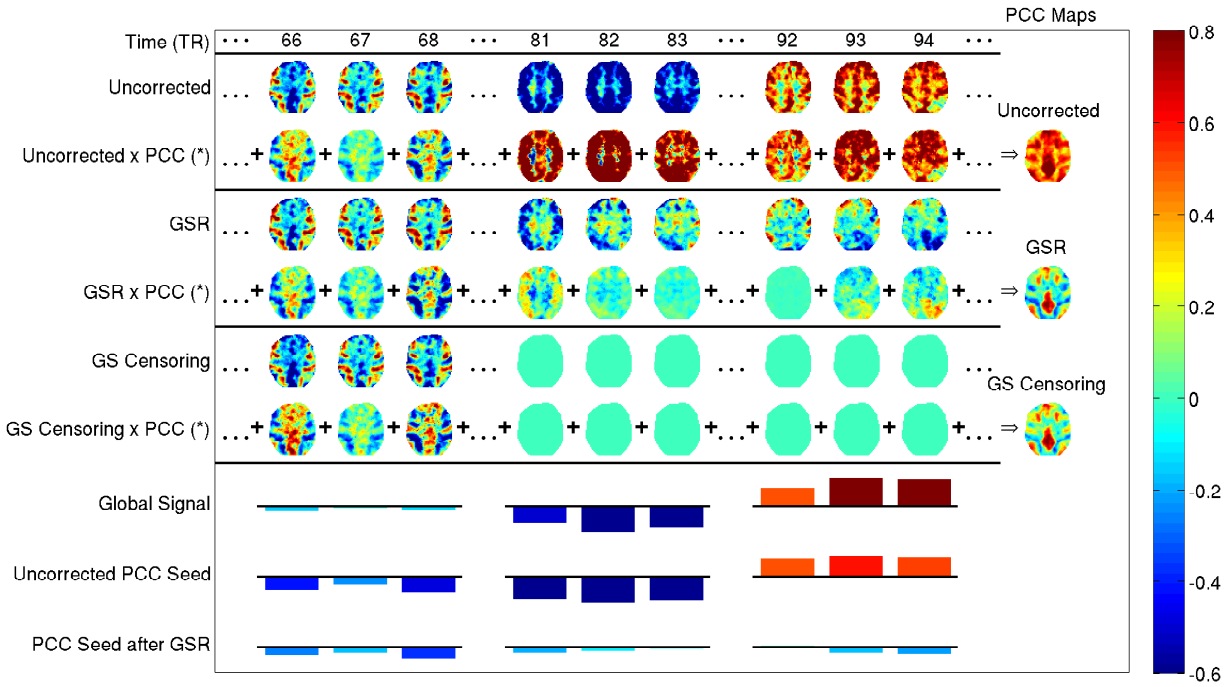


Figure 3.4: Construction of correlation maps for uncorrected data (rows 1 to 2), data after GSR (rows 3 to 4) and data with GS censoring (rows 5 to 6). Images are shown for representative time points where the GS magnitude is either very small (TR= 66-68) or large (TR = 81-83 and 92-94). In each group, the upper and lower rows show the images both prior to and after multiplication by the unit norm PCC seed signal (referred to as unweighted and weighted), respectively. Note that for the weighted images the voxel time series have been normalized to unit norm prior to multiplication by the unit norm PCC seed signal in order to be consistent with the mathematical description of the correlation sum provided in the text. The PCC correlation maps for each condition are obtained by summing all of the weighted images across the duration of a scan (note only 9 time points are shown here due to space limitations). For display purposes, the PCC weighted images in rows marked with (*) have been scaled to match the intensity scales of the images in the upper rows. The symbol ($\Rightarrow$) indicates that the summation of the images yields the correlation maps on the left, after taking into account the display scaling. The values of the GS and the PCC signal both prior to and after GSR are indicated by the colored bars at each time point, where the PCC-related signals in the bottom two rows have been divided by 2.25 for display purposes. Both the unweighted and weighted images are largely unaffected by GSR for time points where the GS magnitude is small, whereas the images at time points with large GS magnitude are greatly attenuated. As a result, the correlation map after GSR largely reflects the contribution of the weighted images from the time points with small GS magnitude. In the GS censoring approach, the unweighted images with large GS magnitudes are censored (e.g. multiplied by zero) while all other images are unaffected. The resulting correlation map is similar to that obtained with GSR.

that one might obtain maps similar to those with GSR by simply censoring those time points where the GS is large. An example correlation map obtained using this approach is shown in Figure 3.4.

As with the temporal downweighting arguments stated in the prior section, the examples and arguments provided in this section are meant to provide background and motivation for the experimental measures (described in Methods and Results) that empirically demonstrate the effect of GSR on correlation maps. In addition, it is important to note the fundamental difference between the temporal downweighting process, which is a time-varying modification of the signal magnitudes, and a simple scaling of the overall amplitude of the entire time series, which represents a time-invariant scaling of the signal magnitudes. To help make the distinction clear, we use term **magnitude** to refer to the absolute value of the GS at each time point and reserve the term **amplitude** to refer to the standard deviation of the GS over the course of the scan. Whereas the correlation coefficient is invariant to a simple scaling of the overall amplitude of the time series, it can be greatly altered by a temporal downweighting process. For example, censoring represents a limiting case of temporal downweighting and can have a significant effect on the correlation coefficient through the zeroing of data at the censored time points.

3.3 Methods

3.3.1 Subjects and Data Acquisition

We used data originally analyzed by [174] and downloaded from www.brainscape.org (dataset BS002). The data were acquired from 17 young adults (9 females) using a 3 T Siemens Allegra MR scanner. Subjects underwent 4 BOLD-EPI fixation runs (32 slices, TR=2.16 s, TE=25 ms, $4\times4\times4$ mm), each lasting 7 minutes (194 frames). Subjects were instructed to look at a cross-hair, and asked to remain still and awake. High-resolution T1-weighted anatomical images were also acquired for the purpose of anatomical registration (TR=2.1 s, TE=3.93 ms, flip angle=7

deg, $1 \times 1 \times 1.25$ mm).

3.3.2 Data Processing

Standard pre-processing steps were conducted with the AFNI software package [175]. The first 9 frames from each EPI run were discarded to minimize longitudinal relaxation effects, leaving 185 frames for further analysis. Images were slice-timing corrected and co-registered and the six head-motion parameter time series were retained. The resultant images were converted to coordinates of Talairach and Tournoux (TT), resampled to 3 mm cubic voxels, and spatially smoothed using a 6 mm full-width-at-half-maximum isotropic Gaussian kernel. Nuisance terms were removed from each voxel's time course using multiple linear regression. These were: (1) a constant term to model the temporal mean of each voxel, (2) a linear trend, and (3) six motion parameters obtained from the head motion correction algorithm. A temporal low pass filter was applied to the remaining time course using a cutoff frequency of 0.1 Hz [152, 176]. Each detrended and filtered voxel time series was converted into a percent change BOLD time series through division by its mean value. The GS was then calculated as the average of the percent change time series across all voxels within the brain for each scan. GSR was implemented by regressing the GS out of each voxel time series [177].

Correlation maps were formed by correlating a seed signal with the time-course from every voxel in the brain. The seed signal was computed as the average signal from a region selected with spheres of radius 9 mm (3 voxels) centered about the regions of interest. We obtained the following sphere centers by converting the MNI coordinates obtained from [178] to TT coordinates using the MNI to TT conversion algorithm [179]. For primary analysis we focused on the posterior cingulate cortex (PCC) with a seed region centered in TT coordinates [0,-51,26]. For the left and right auditory network (AUD) we used [-41,-26,14] and [41,-26,14]. For left and right motor networks (MOT) we used [-36,-25,57] and [36,-25,57], and for the medial prefrontal cortex we used [0,46,-7]. For MOT and AUD seeds we merged two spheres from the left and right

hemispheres of the brain to obtain a final region of interest for each seed. For right intraparietal sulcus we used the coordinates [27,-58,49] from [180] and for the white matter seed we defined the seed center as [31,-28,32]. Correlation maps were computed for (a) the pre-processed time series (referred to as uncorrected maps) and (b) the pre-processed time series after the application of GSR or one of the weighting schemes described below.

3.3.3 GSR Ratio: Characterizing the average effect of GSR

In Section 3.2.2 we made the observation that GSR largely attenuates the signals in time intervals where the GS is large and has a minimal effect in those intervals where the GS is small. To further characterize this effect, we calculated a GSR ratio $\mathbf{g}_r(t)$ that captures the average effect of GSR at each time point t. This is defined as

$$\mathbf{g}_r(t) = \left\langle \frac{\mathbf{v}_{i,a}(t)}{\mathbf{v}_{i,b}(t)} \right\rangle = \frac{1}{N}\sum_{i=1}^{N} \frac{\mathbf{v}_{i,a}(t)}{\mathbf{v}_{i,b}(t)} \tag{3.1}$$

where $\mathbf{v}_{i,b}(t)$ and $\mathbf{v}_{i,a}(t)$ are the values of the i'th voxel's time course before and after GSR, respectively, and N is the number of voxels. Because the ratio blows up for small values of $\mathbf{v}_{i,b}(t)$, outliers were removed prior to the computation of the mean. The threshold for outlier detection was defined as the median plus 2.5 times the median absolute deviation of the central 95% of the data, where initial censoring of the data in the extreme tails of the distribution was performed to avoid excessive skewness prior to the computation of the statistics [181, 182]. This process resulted in a threshold value of 2.25 that was applied to the magnitude of $\mathbf{g}_r(t)$.

By definition, the GSR ratio should be less than 1 when the average attenuation (across voxels) due to GSR is large and be close to 1 when the average effect of GSR is minimal. In Figure 3.5 (a,b) we show the GSR ratio plotted along with the magnitude of the GS from two representative scans. The GSR ratio exhibits local minima at time points where the GS magnitude has relatively large values. In contrast, when the GS magnitude has small values the GSR ratio is

close to 1, indicating minimal perturbation on average of the original voxel values.

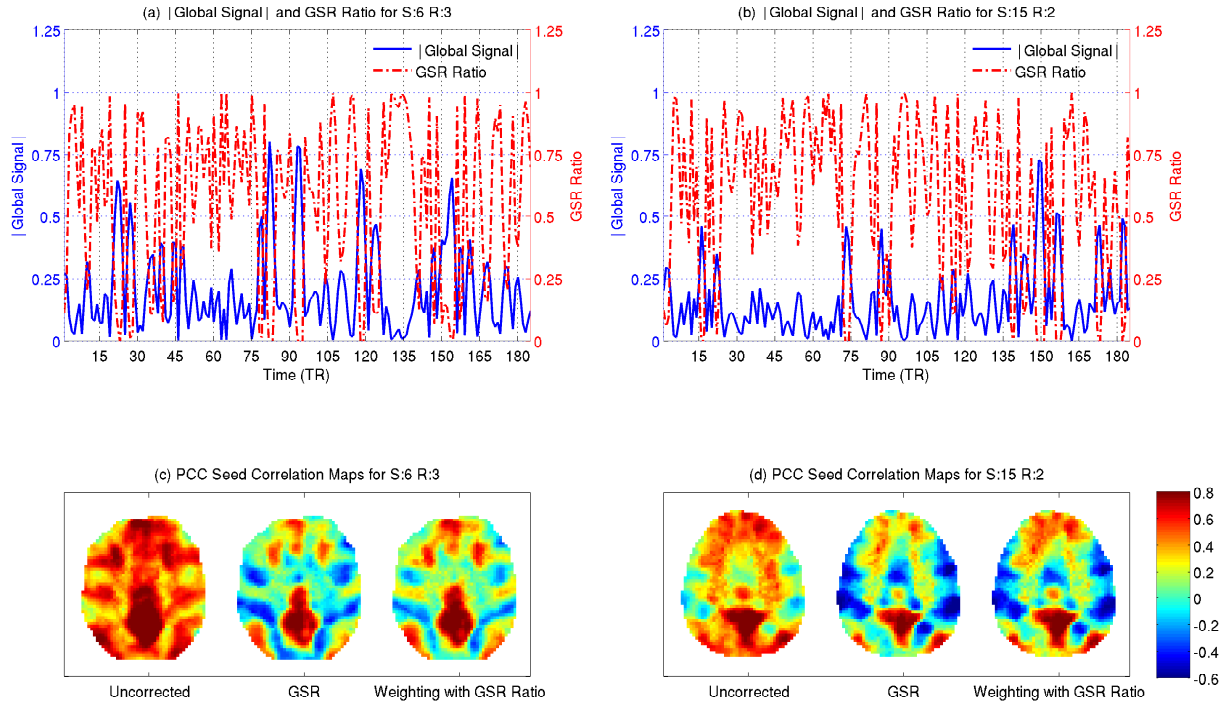


Figure 3.5: (a), (b) GSR Ratio time series (red dashed lines) and GS magnitudes (solid blue lines) from two representative scans. (c,d) PCC seed correlation maps obtained after: (1) preprocessing only, (2) application of GSR after preprocessing, and (3) GSR ratio weighting after pre-processing.

To further evaluate the properties of the GSR ratio, we multiplied each of the pre-processed voxel time series $\mathbf{v}_{i,b}(t)$ by the GSR ratio to obtain a new time series for each voxel of the form

$$\tilde{\mathbf{v}}_{i,GSRW}(t) = \left\langle \frac{\mathbf{v}_{i,a}(t)}{\mathbf{v}_{i,b}(t)} \right\rangle \cdot \mathbf{v}_{i,b}(t) = \mathbf{g}_r(t) \cdot \mathbf{v}_{i,b}(t) \tag{3.2}$$

where the subscript *GSRW* indicates that the time course is GSR ratio weighted. For time points where GSR has the same effect for all voxels, then $\left\langle \frac{\mathbf{v}_{i,a}(t)}{\mathbf{v}_{i,b}(t)} \right\rangle = \frac{\mathbf{v}_{i,a}(t)}{\mathbf{v}_{i,b}(t)}$ and multiplication by the GSR ratio acts in the same manner as GSR so that $\tilde{\mathbf{v}}_{i,GSRW}(t) = \mathbf{v}_{i,a}(t)$. On the other hand, for time points where the effect of GSR varies over voxels, multiplication by the GSR ratio only captures the average spatial effect of GSR so that the weighted values can only approximate the values obtained with GSR, such that $\tilde{\mathbf{v}}_{i,GSRW}(t) \approx \mathbf{v}_{i,a}(t)$.

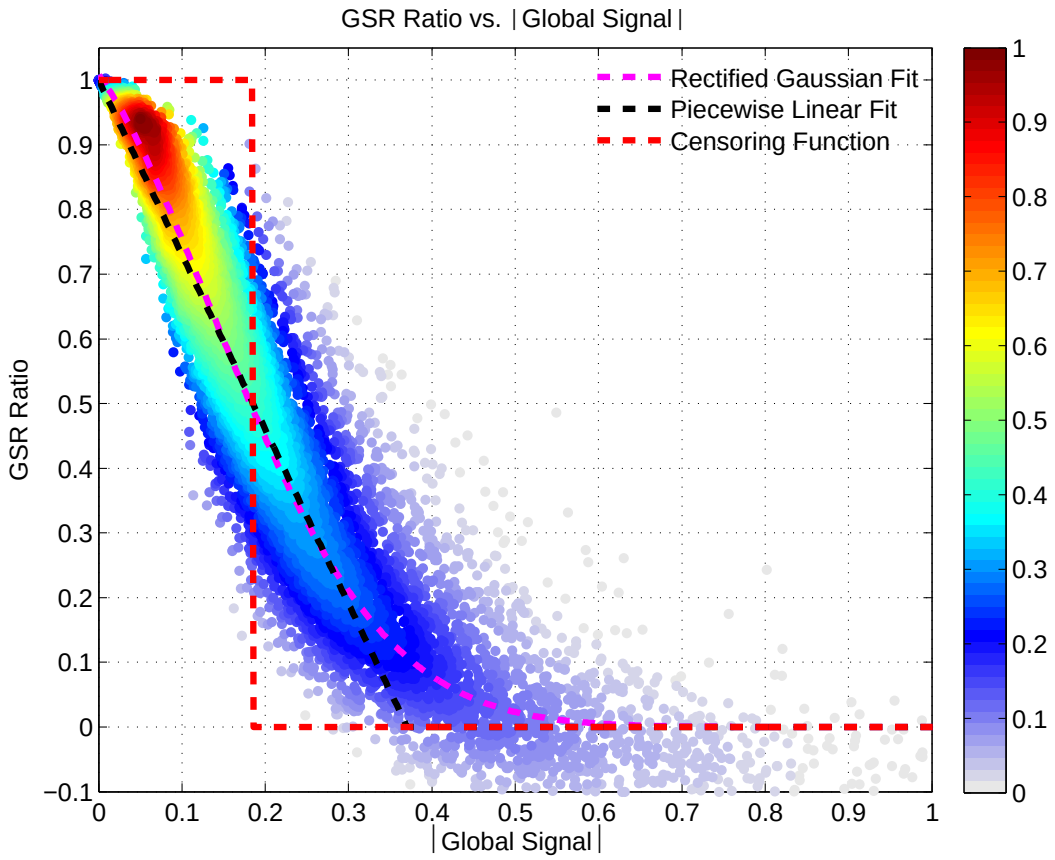


Figure 3.6: Density weighted scatter plot of GSR Ratios versus GS magnitude (% change) for all subjects and runs, i.e. a total of 12580 data points. Maximum data concentration is indicated by the red color and a density value of 1. Piecewise linear (black dashed line) and rectified Gaussian fits (magenta dashed line) are shown. In addition a censoring function with threshold $g_C = 0.5$ is shown (red dashed line).

The GSR ratio weighted time series $\tilde{\mathbf{v}}_{i,GSRW}(t)$ were used to compute an additional set of PCC correlation maps. In Figure 3.5 (c), (d), we compare these maps to PCC correlation maps obtained using the pre-processed voxel time series both before GSR (i.e. $\mathbf{v}_{i,b}(t)$) and after GSR (i.e. $\mathbf{v}_{i,a}(t)$). There is a strong similarity between the GSR and GSR Ratio weighted maps, suggesting that the main effects of GSR are reflected by its average attenuation (at each time point) of the voxel values. Differences in the maps reflect the fact that GSR performs a unique regression for each voxel time series, whereas the GSR ratio is constrained to use the same weighting function for each voxel time series. Additional examples of GSR ratio weighted maps are shown in the Results section.

3.3.4 Global Signal Weighting: An approximation for the GSR Ratio

In the previous section we presented preliminary evidence for a relation between the GSR ratio and the magnitude of the GS in two representative scans. We further evaluated the generality of this relation by computing the GS and GSR ratio for all scans. As shown in Figure 3.6, we found a strong inverse dependence of the GSR ratio on the GS magnitude. To quantify this dependence, we fit the data to a piece-wise linear model of the form

$$f\left(GS(t)\right) = \begin{cases} 1 - \alpha \mid GS(t) \mid & \text{for } \mid GS(t) \mid \leq g, \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

To calculate the optimal fit we used a robust regression method based on an iteratively reweighted least-squares (IRLS) algorithm with bisquare weights [183, 184] as implemented in the MATLAB Curve Fitting Toolbox. This resulted in parameter values of $\alpha = 2.7$ and $g = 0.37$ and a model fit (shown by the dashed black line) with $R^2 = 0.94$. We found that other models provided similar fits. For example, a rectified Gaussian model (shown by the dashed magenta line) provide a fit with $R^2 = 0.95$. The similarity in the performance of the fits reflects the fact that the main deviation between the models occurs in a region ($|GS| > 0.30\%$) where the density of data points is very low. Due to the simplicity of its form, we will use the piece-wise linear model for the remainder of this paper.

The explanatory power of the piece-wise linear model indicates that the average voxel-independent effect of GSR can be well approximated by a simple function of the GS. In other words, the average weighting effect of the GSR operation can be achieved without actually needing to perform the regression on each voxel time series. Instead, we can multiply each voxel time series by a 'GS weighting' function to obtain a new time series $\tilde{\mathbf{v}}_{i,GSW}(t) = f\left(GS(t)\right) \cdot \mathbf{v}_{i_b}(t)$ where the subscript *GSW* indicates that the time series is GS weighted with the piecewise linear function. Correlation maps computed using $\tilde{\mathbf{v}}_{i,GSW}(t)$ will be referred to as GS weighted maps.

3.3.5 Global Signal Censoring

When using the piece-wise linear function $f(GS(t))$ from Eq. (3.3) as a weighting function, the voxel time series values are scaled by a linearly decreasing function that gradually decreases to zero. This can be characterized as a soft censoring operation. We also considered a hard censoring function of the form

$$\mathbf{C}(GS(t)) = \begin{cases} 0 & \text{for } (1-\alpha|GS(t)|) \leq g_C, \\ 1 & \text{otherwise.} \end{cases} \tag{3.4}$$

This function censors those time points for which the predicted average weighting due to GSR is below a threshold value g_C. For example, a threshold value of $g_C = 0.5$ would censor those data points for which GSR would scale the data on average by a factor of 0.5 or less. This is equivalent to censoring those time points for which the GS magnitude is greater than 0.18% (i.e. $|GS| > (1-g_C)/\alpha$ using the value of $\alpha = 2.7$ estimated above). A plot of this censoring function is shown by the red dashed line in Figure 3.6.

The rationale behind the hard censoring approach relies in part on the preliminary observation (see Section 3.2.3) that the correlation maps after GSR are expected to be largely determined by those time points that are largely unaffected by GSR. The choice of g_C quantifies the degree to which a time point is considered to be "unaffected" by GSR. For example, for the threshold value of $g_C = 0.5$, those time points for which the average scaling due to GSR is predicted to range between 0.5 and 1.0 are considered to be unaffected and will not be censored. As the threshold is increased, the number of time points censored increases as fewer time points are considered to be unaffected by GSR (see Figure 3.13 in Results).

As with the previously discussed weighting functions, we can multiply each preprocessed voxel time series by the hard censoring function to obtain a new time series $\tilde{\mathbf{v}}_{i,GSC}(t) = \mathbf{C}(GS(t)) \cdot \mathbf{v}_{i_b}(t)$ where the subscript *GSC* indicates that the signal has been GS censored. Correlations maps

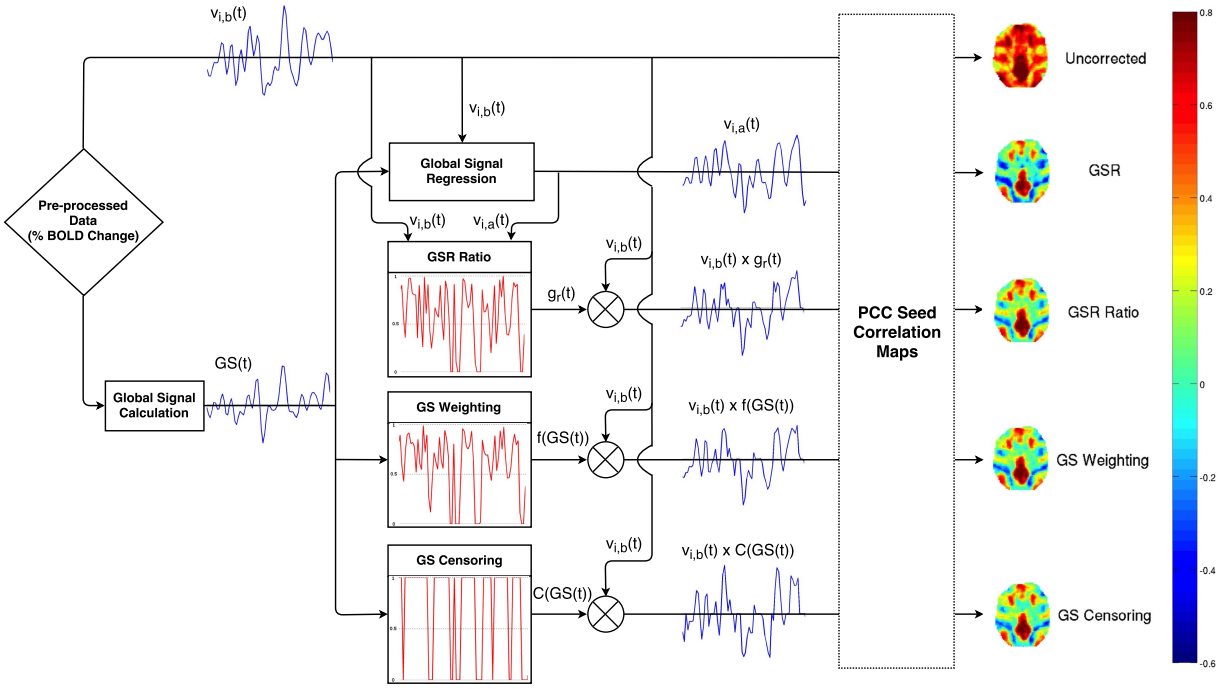


Figure 3.7: Diagram of the proposed GSR approximations and the data processing pipeline. The average effects of GSR can be approximated by multiplying the pre-processed voxel time courses $v_{i,b}(t)$ with GSR ratio, GS weighting or GS censoring functions (shown in red). The resulting time series are used to form PCC correlation maps shown on the righthand side. The GSR ratio approach uses all the voxel time courses prior to GSR ($v_{i,b}(t)$) and after GSR ($v_{i,a}(t)$) to determine a weighting function to approximate GSR. In contrast, the GS weighting and censoring operations use only the GS as an input to determine the weighting function.

computed using $\tilde{\mathbf{v}}_{i,GSC}(t)$ will be referred to as GS censored maps. Figure 3.7 summarizes the various weighting and censoring approaches and the signal paths used to compute the correlation maps shown in the Results sections.

3.3.6 Motion Censoring

As motion censoring approaches are sometimes used in the analysis of rsfMRI data, we wanted to compare the GS censoring approach with motion censoring. To do so, we implemented the motion censoring method described in [185], which is based on the frame-wise displacement (FD) and DVARS time series. An FD threshold value was computed for each scan as the mean plus one standard deviation of the FD time series, and an overall threshold for the sample was

then computed by averaging over the individual scan thresholds. This resulted in a threshold value of $T_{\mathrm{FD}} = 0.23$ mm. The DVARS threshold value was computed in the same fashion, yielding $T_{\mathrm{DVARS}} = 0.38\%$. Individual FD and DVARS masks were formed by thresholding the FD and DVARS time series with the pre-computed thresholds. The FD and DVARS masks were then dilated to include one frame prior to and one frame immediately after each of the time frames identified by thresholding. Finally, the intersection of the FD and DVARS masks was computed to form a combined motion censoring mask for each run. Correlation maps were then computed using those time points that survived the censoring process with the combined mask.

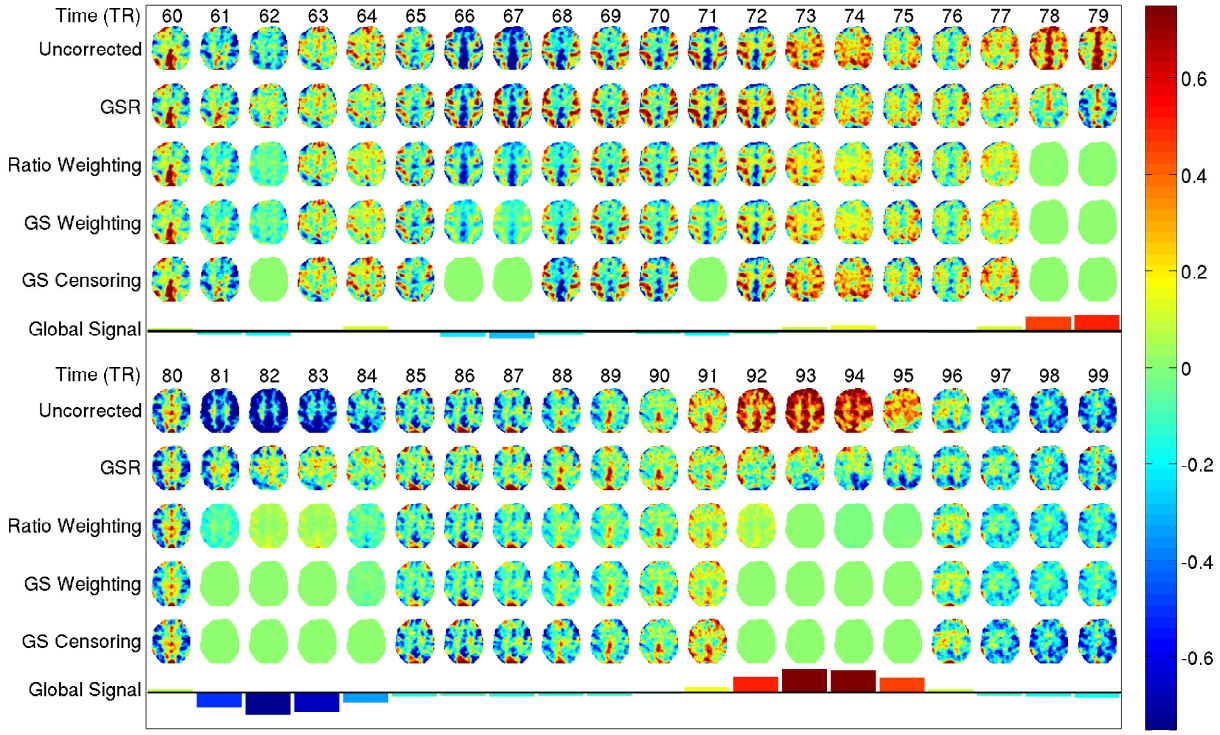


Figure 3.8: Images from a representative slice and scan obtained before GSR and after the application of GSR, GSR Ratio weighting, GS Weighting, and GS Censoring. The colored bars indicate the value of the GS at each time point.

3.3.7 Similarity Measures

We used cosine similarity to quantify the similarity of the correlation maps obtained with GSR and the various approaches [186, 187]. As cosine similarity preserves the means of the correlation maps, it is suitable for comparing approaches where visible differences in the associated maps reflect a strong dependence on the mean level (e.g. compare the first and second rows of Figure 3.9). We should note that cosine similarity would be equivalent to Pearson correlation if the mean were subtracted prior to the computation of the normalized inner product [187]. However, the mean subtraction procedure can introduce potential confounds related to post hoc centering of the correlation values, which biases correlation maps towards the GSR correlation map [51].

To assess the significance of the similarity between the maps obtained with GSR and each of the approaches on a per-scan basis, we constructed an empirical null distribution of the corresponding similarity measure for each scan. This was accomplished through random permutation (10,000 trials) of the temporal ordering of the calculated weights (i.e. GSR ratio, GS weight, or 0 or 1 for GS censoring). The cosine similarity was calculated between correlation maps obtained with GSR and the maps obtained after application of the permuted weights, and this process was repeated over all permutations in order to form a null distribution of similarity values. Note that for GS censoring the randomized approach is similar to the random censoring procedure used by [188]. In addition, we performed a paired t-test to assess whether there was a significant difference between the similarity of the correlation maps obtained with GSR and the GS weighting approach versus those obtained with GSR and the GS censoring approach.

3.4 Results

The application of the various weighting approaches to data from a representative scan is demonstrated in Figure 3.8, where the raw images and global signal from Figure 3.2 are shown

again for comparison. In addition, the images after the application of GSR, GSR ratio weighting, GS weighting, and GS censoring (with threshold $g_C = 0.5$) are shown. For time points where the GS magnitude is low (such as time point 75), the images are largely unaffected by GSR. In contrast, for time points where the GS magnitude is high (such as time points 82 and 93), GSR results in an average decrease in the magnitude of the images. These average effects are reflected in the uniform downweighting of the images with either GSR ratio weighting or GS weighting. Note that the attenuation introduced by GSR can vary across voxels whereas the attenuation

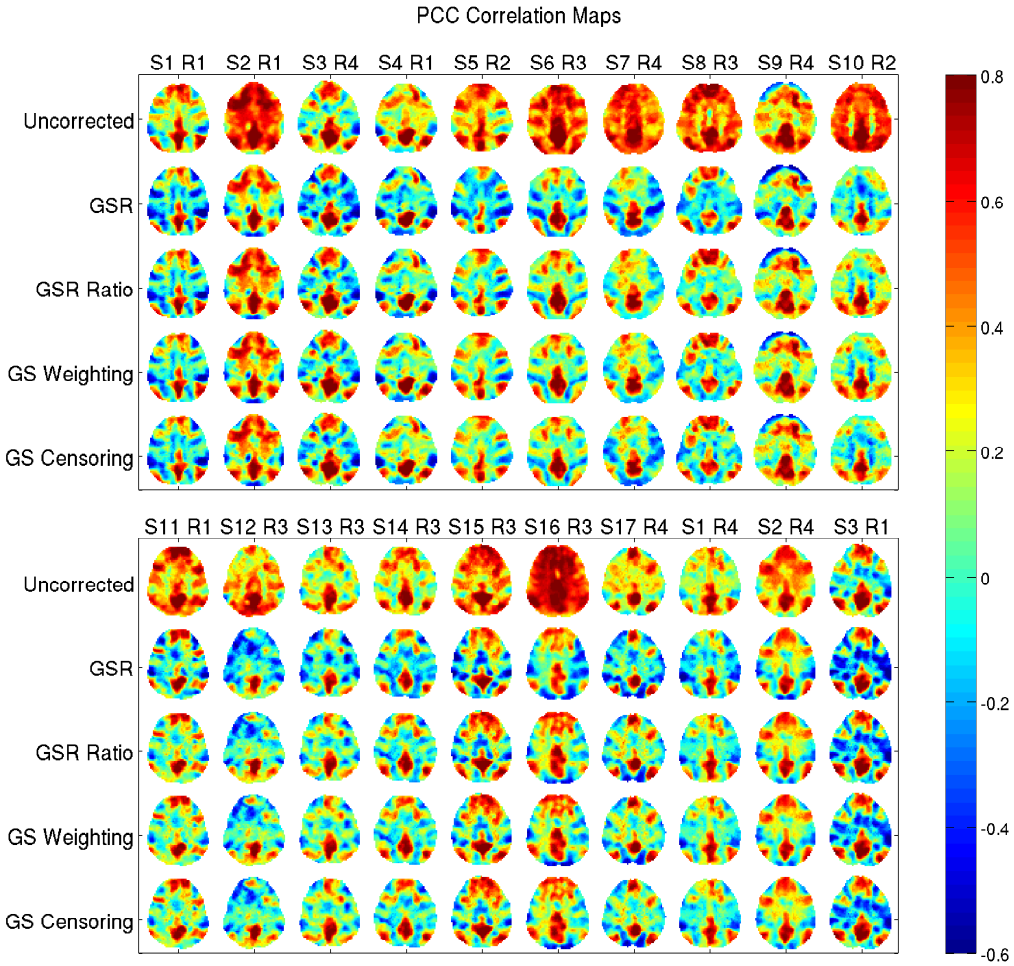


Figure 3.9: PCC seed correlation maps obtained before GSR, after GSR, and after application of GSR ratio weighting, GS weighting, and GS censoring. Maps are shown for 20 different scans, with at least one scan from each subject. For GS censoring a threshold of $g_C = 0.50$ was used.

due to GSR ratio weighting and GS weighting is by definition the same across voxels, as these approaches are designed to capture the average effect of GSR across all voxels. In the bottom row, the GS censoring approach censors (i.e. multiplies by zero) images at those time points where there is significant attenuation introduced by GSR. With the specific threshold used in this example, the censored time points occur for images where the average attenuation is large (i.e. scaling factor less than $g_C = 0.5$) in the GS weighted images.

Correlation maps obtained with the PCC seed and using the proposed approximation approaches are shown in Figure 3.9 for a representative slice from each of 20 different scans. Maps obtained with the uncorrected data (i.e. after preprocessing but before GSR) are shown in the top row. From a qualitative viewpoint, the correlation maps obtained after GSR are very similar to those obtained after GSR ratio, GS weighting and GS censoring. This suggests that the main effects of GSR on the correlation maps are well approximated by using voxel-independent weighting and censoring approaches. Maps using other seed voxels are provided in Supplementary Figures 7.1 to 7.4.

To quantitatively assess the similarity of the PCC correlation maps obtained with GSR and the various approaches, we tested the similarity values against the empirical null distributions

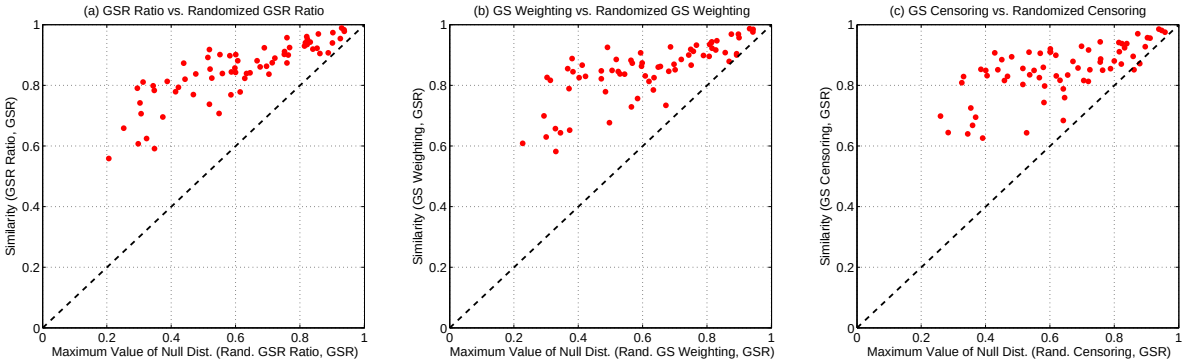


Figure 3.10: The similarity of the maps obtained with the proposed approaches to the GSR versus the maximum similarity in the respective null distribution for each scan (derived from randomized weighting or censoring). In (a) the absolute difference in the means was 0.2326 and effect size was $d = 1.46$, in (b) the difference in the means was 0.2318 and effect size was $d = 1.50$ and in (c) the difference in the means was 0.2191 and the effect size was $d = 1.49$.

obtained with randomized weighting or censoring, as described in the Methods section. This was done for each scan. For the similarities between the maps obtained with GSR and either the GSR ratio or GS weighting approaches, the empirical p-values were in the range $p < 1 \times 10^{-4}$ (i.e. less than the smallest empirical p-value that could be assessed with 10,000 random trials). For the similarities between maps obtained with GSR and GS censoring, the empirical p-values were in the range $p < 3 \times 10^{-4}$. These results indicate that the observed similarities were significantly different from those that could be obtained with randomized weighting or censoring of the data. To provide a complementary view, we show in Figure 3.10 the actual similarity values between the various approaches versus the maximum similarity values obtained from the respective null distributions. These maximum values from the null distributions represent highly unlikely values ($p = 1 \times 10^{-4}$). The absolute differences in the means of these maximum null distribution values and the actual similarity values ranged from 0.2191 to 0.2326, and the effect sizes (Cohen's d)

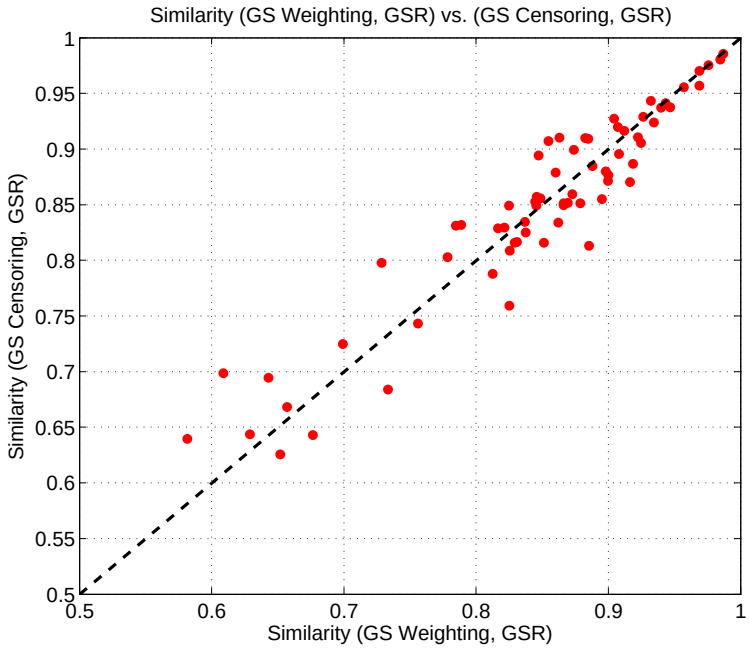


Figure 3.11: Comparison of similarity values between GS censoring and GSR maps versus similarities between GS weighting and GSR maps. There was not a significant difference ($p = 0.92$) between the values.

ranged from $d = 1.46$ to $d = 1.50$. These results indicate that the effects of randomized weighting and censoring greatly differ from the proposed approaches and the similarity of the correlation maps obtained with these methods to GSR maps cannot be explained by randomized weighting and censoring of the data.

In Figure 3.11 we compare the similarity between the GSR and GS weighted maps with the similarity between the GSR and GS censored maps. The similarity values were not significantly different ($p = 0.92$), with an absolute difference in similarity means of 0.0003 and effect size $d = 0.0040$. This result suggests that the hard censoring approach (with weights of either 0 or 1) can capture the important spatial features of the GSR maps about as well as the soft censoring approach that has linearly varying weights. The advantage of the hard censoring approach is that the maps obtained can be simply interpreted as estimates of the functional connectivity for the subset of time points that survive the censoring operation.

As motion censoring (or scrubbing) is finding increasing use in the analysis of rsfMRI data, we compared the effects of motion censoring and GS censoring. First we looked at the percent overlap of time points that were censored by both approaches. This was calculated by

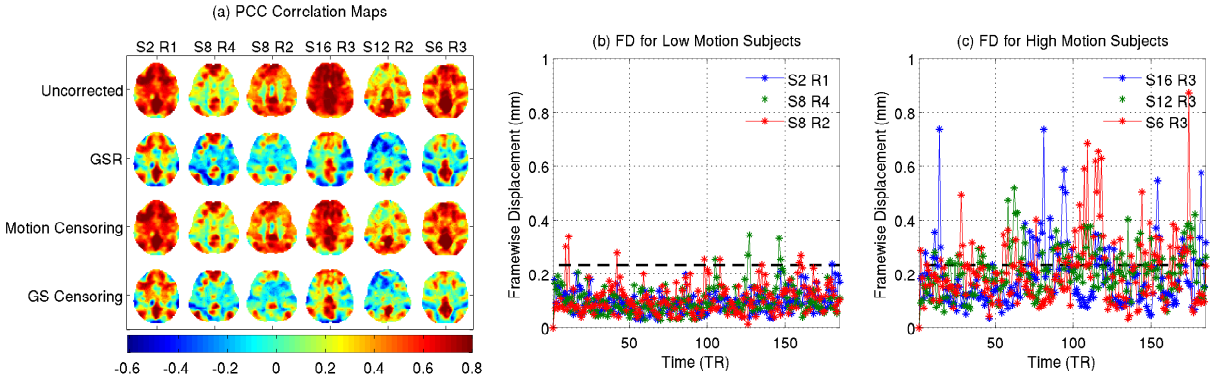


Figure 3.12: (a) PCC correlation maps for three low head motion runs (S2 R1, S8 R4, S8 R2) and three high motion runs (S16 R3, S12 R2, S6 R3). The maps shown were obtained before GSR and after the application of GSR, motion censoring, and GS censoring. FD time series are shown for (b) low motion and (c) high motion subjects, with the censoring threshold shown by the dashed black lines. The differences between the motion censoring and GS censoring maps indicate that the two approaches are fundamentally different.

dividing the number of time points that were censored by both approaches (i.e. the intersection) by the total number of time points that were censored by either approach (i.e. the union) and then multiplying the ratio by 100. We found that there was relatively low overlap between time points censored by the two approaches, with a median percent overlap of 5.67%. We then investigated temporal censoring based on individual FD and DVARS masks instead of their intersection. Application of the FD only and DVARS only masks resulted in median percent overlaps of 12.73% and 24.83%, respectively. The FD only censoring had a much smaller overlap as compared to DVARS, suggesting that the results obtained with the intersection mask were largely determined by the FD mask. This most likely reflects a closer link between the GS and DVARS time courses (i.e. root-mean square of the temporal derivative of the brain images), as compared to the link between the GS and FD time courses.

Next we compared PCC correlation maps obtained using the two different censoring

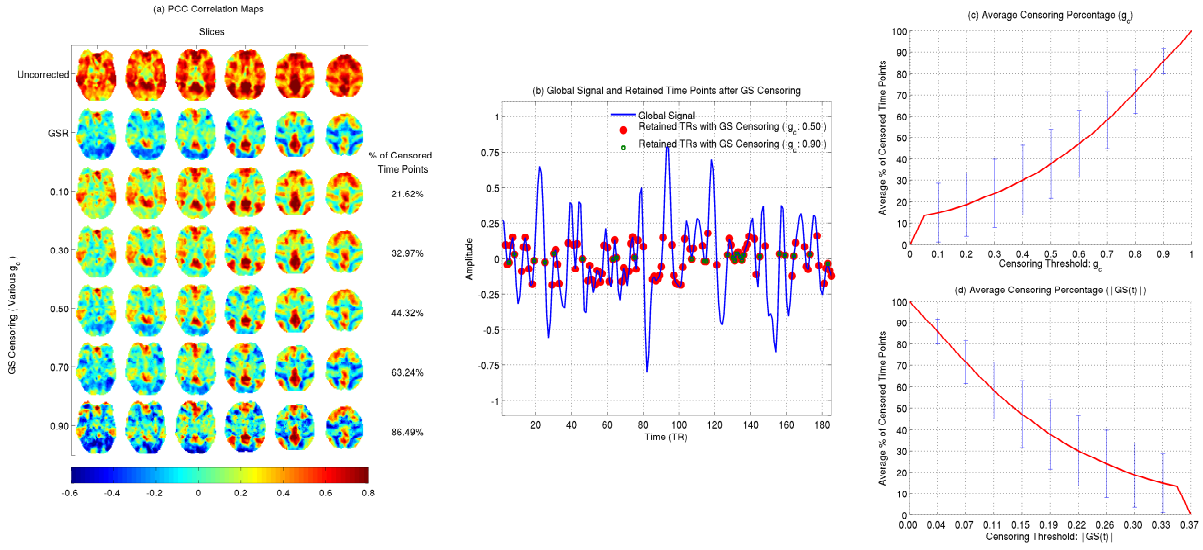


Figure 3.13: (a) PCC maps from a representative subject obtained before GSR and after the application of GSR and GS censoring approaches using different thresholds. (b) GS of the same subject with time points retained after censoring with $g_C = 0.50$ (red dots) and $g_C = 0.90$ (green circles). (c) Average percentage of censored time points versus censoring threshold g_C, with standard deviation across all scans indicated by the bars. (d) Average percentage of censored time points versus percent change GS magnitude threshold, where the threshold is computed using $|GS| = (1 - g_C)/\alpha$ with $\alpha = 2.7$.

approaches. Figure 3.12 shows PCC correlation maps for scans from 6 representative subjects with three low head-motion scans (S2 R1, S8 R4, S8 R2) and three moderate to high head-motion runs (S16 R3, S12 R2, S6 R3) that were selected according to the RMS values of their respective FD time series. FD time series for low-motion and high-motion runs are shown in panels (b) and (c), respectively, with the censoring threshold for FD indicated with dashed black lines. Low motion runs incur little or no motion censoring and the resultant motion-censored scans are similar or identical to the uncorrected maps. On the other hand, high motion runs incur a large degree of motion censoring, resulting in some clean-up of the correlation maps. For both low motion and high motion runs, the maps obtained with GS censoring are similar to those obtained with GSR but are markedly different from those obtained with motion censoring.

To provide additional insight into the GS censoring approach, we show in Figure 3.13(a) a representative subject's PCC correlation maps obtained prior to GSR and after the application of either GSR or GS Censoring using different thresholds ranging from $g_C = 0.10$ to $g_C = 0.90$. As we increase the the threshold value, the number of time points censored increases, as indicated by the percentage values listed next to each row. For a threshold value of $g_C = 0.10$, only time points with a GSR ratio below 0.10 and a relatively high GS magnitude ($|GS| > 0.33\%$) are censored, corresponding to about 22% of the total number of points. At this level of censoring, the DMN and TPN are more clearly delineated as compared to the uncorrected maps, but the value of the correlations in the TPN are more positive than those observed in the after GSR map. As we increase the value of g_C and censor an increasing number of time points, the correlation values in the TPN become more negative and the maps become more similar to those obtained with GSR. In Figure 3.13(b) we indicate the location of the censored time points for threshold levels of $g_C = 0.50$ (red dots) and $g_C = 0.90$ (green circles). At a threshold value of $g_C = 0.90$ only about 14% of the time points are retained and yet the PCC correlation maps are similar to those obtained with GSR. In Figure 3.13(c), we show that the average percentage (across all scans) of censored time points increases with the threshold g_C, consistent with the qualitative behavior

discussed above. Because of the approximate relation between GSR ratio and GS magnitude (see Figure 3.6 and Eq. (3.3)), we can associate each threshold value g_C with a GS magnitude threshold $|GS| = (1 - g_C)/\alpha$ where the value $\alpha = 2.7$ is obtained from the fit to the data. Only those time points with GS magnitude higher than the GS magnitude threshold are censored. As the GS threshold is increased, the average percentage of censored points decreases, as shown in Figure 3.13(d).

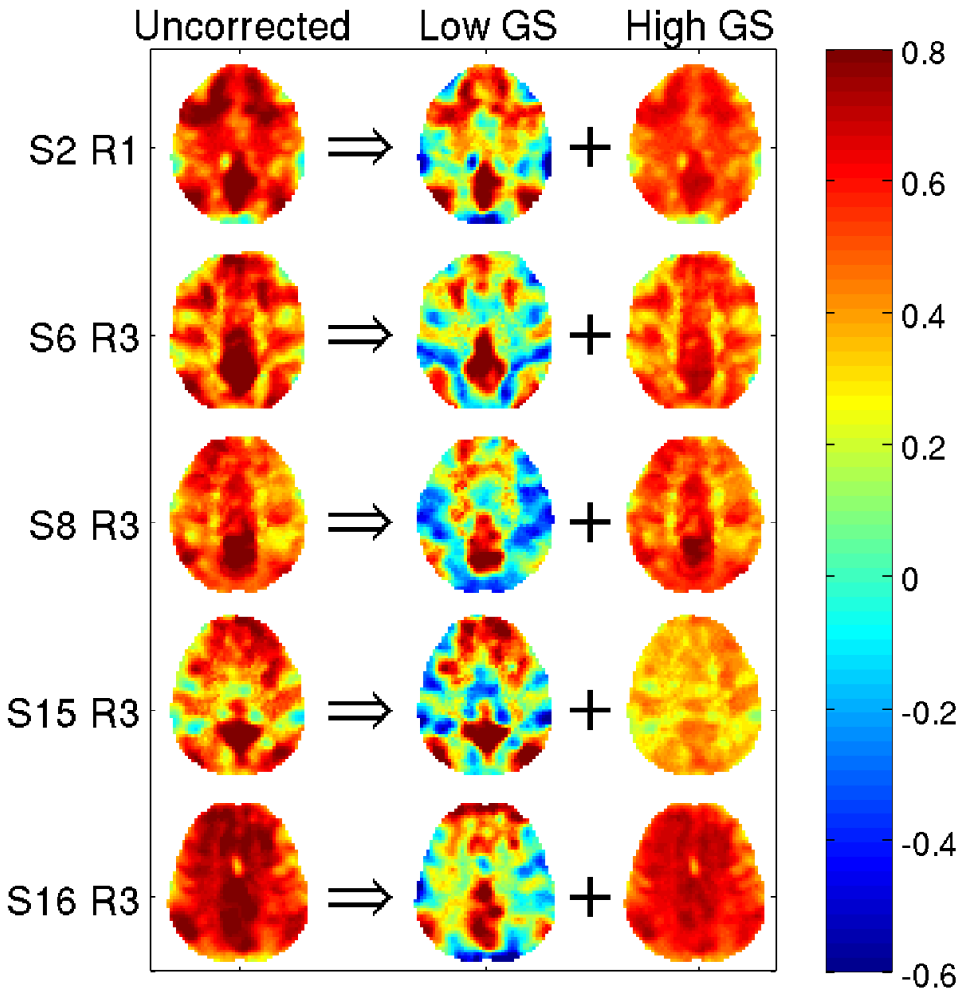


Figure 3.14: PCC correlation maps expressed as the sum of a map corresponding to retained points with low GS and a map based on the censored points with high GS. For display purposes, the low GS map were scaled to better delineate the features in the map. The symbol ($\Rightarrow$) indicates that the correlation maps is equal to the sum of the low GS and high GS maps, after taking into account the display scaling of the low GS maps.

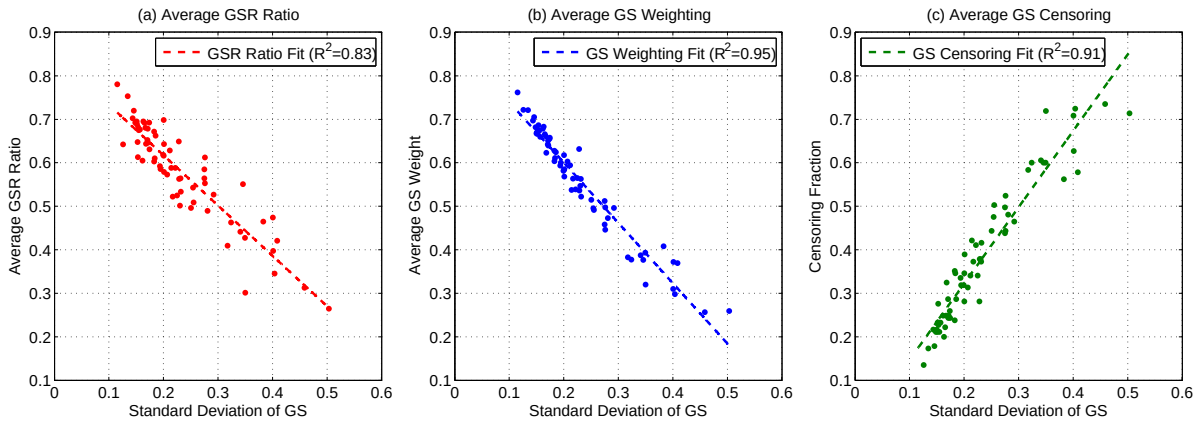


Figure 3.15: (a) The average GSR ratio for each scan (computed as the mean of GS ratios across all time points within a scan) versus the standard deviation of the GS across the scan. (b) The average of temporal weights obtained with the GS weighting approach for each scan versus the corresponding GS standard deviation. (c) The average fraction of censored time points versus the GS standard deviation for each scan.

The GS censoring approach implements a temporal partitioning of the data into a subset of retained points (with relatively low GS magnitude) and a subset of censored points (with relatively high GS magnitude), where the exact delineation between low and high magnitudes is determined by the threshold. As noted previously in Section 3.2.3, a correlation map can be decomposed into a sum of weighted images. Using this observation, we can decompose correlation maps into the sum of a map corresponding to the subset of retained points (these are the GS censored maps already shown) and a map corresponding to the subset of censored points. Figure 3.14 shows examples of this decomposition applied to PCC correlation maps with a censoring threshold of $g_C = 0.5$. As expected, the high GS subset maps exhibit a high degree of spatial homogeneity while the low GS subset maps reveal the spatial structure of the anti-correlated networks.

By averaging the GSR ratio across the duration of a scan we can obtain a value that represents the average temporal downweighting due to GSR. In Figure 3.15(a) we show that the average GSR ratio exhibits an inverse relation with the standard deviation of the GS. Thus, scans with greater GS fluctuations undergo more downweighting (on average) by GSR. In Figure 3.15(b) we show that a similar inverse dependence on GS standard deviation holds for the average

GS weighting across the scan (blue dots). The average fraction of time points censored by GS censoring (green dots) increases with GS standard deviation (shown in Figure 3.15(c)), consistent with the associated increase in downweighting shown in panels (a) and (b) (i.e. more censoring for lower GS ratio and GS weighting values). The implications of these relations are discussed further in the following sections.

3.5 Discussion

3.5.1 A new framework for understanding GSR

We have shown that the average effects of GSR can be well approximated by a temporal downweighting of the voxel time series, where the weighting factor varies with time but is uniform across space. The weighting factor decreases with the magnitude of the global signal, so that time points with large global signal magnitudes are greatly attenuated whereas those time points with small global signal magnitudes are largely unaffected. We also introduced GS censoring as a limiting case of the downweighting approach, in which the weighting factor is equal to zero for time points where the GS magnitude exceeds a specified threshold and is equal to one otherwise.

In the prior studies and debates concerning the use of GSR, a combination of mathematical arguments, empirical findings, and simulations have been used to examine the strengths and limitations of the approach [51, 52, 157, 162]. However, the process of GSR has remained somewhat mysterious as it is not easy to visualize the process of regression with a global mean signal in a high-dimensional signal space. The use of differing perspectives in prior studies has complicated the development of a unified view of GSR and made it difficult to resolve the ongoing debate. By showing that GSR can be well approximated by a temporal downweighting operation, we have introduced a simple way of viewing GSR that facilitates a more intuitive understanding of its effects. Instead of having to visualize how regression affects each individual voxel time series, we can simply consider consider how the data from each time point is downweighted

prior to the computation of functional correlations. In the limit of GS censoring, the data from each time point is either included in (weighting of 1.0) or excluded (weighting of 0.0) from the computation. The similarity in the maps obtained with GSR and GS censoring suggests that the first order effects of GSR are well approximated by censoring a large fraction (i.e. greater than 50% or more) of the data.

3.5.2 Anti-correlated Networks

With respect to the highly debated anti-correlated networks, our work suggests that the negative correlations between the DMN and the TPN are not simply an artifact of GSR. Specifically, the application of the GS censoring approach shows that negative correlations can be observed when computing the functional correlation for a temporal subset of the data. Within this subset, there is no mathematical constraint that forces the existence of negative correlations, because the images retained in the subset are not modified by the censoring operation. In other words, the negative correlations are inherent in the data when considering time points with low GS magnitudes.

Our results suggest that the presence of anti-correlated networks is often obscured by the contributions of images at those time points where the GS magnitude is high (see for example Figure 3.4). Because these peaks in GS magnitude occur when there is a high degree of spatial homogeneity in the images (e.g. either largely positive or negative), the inclusion of these images gives rise to spatial homogeneity in the correlation maps. By downweighting the contribution of these images, GSR and the various associated approaches presented in this paper leads to a reduction in the spatial homogeneity of the correlation maps.

Prior work has shown that the application of GSR [51, 52] forces the sum of correlation values to be negative, thus potentially introducing artifactual negative correlations. The present work does not contradict these prior findings, but rather presents a new way of looking at the effects of GSR. In essence, the process of GSR can be viewed as approximately partitioning the

data into two sets – a set of images in which the GS magnitude is relatively high and a set in which the magnitude is relatively low. GSR greatly attenuates the images in the high magnitude set while having a smaller effect on the low images in the low magnitude set. When viewed from this perspective, GSR helps to reveal the inherent negative correlations that exists in the low magnitude set by reducing the contributions of the images in the high magnitude set, which have a high degree of spatial homogeneity.

3.5.3 Differences in Global Signal Magnitude

What differentiates images with high GS magnitude from those with low GS magnitude? While the origins of the GS are not completely understood, there is growing evidence that a significant component of the GS is due to neural fluctuations [164, 165, 166]. In particular, [165] found that the average amplitude of the global signal (computed as its standard deviation over the course of a scan) exhibited an inverse relation to EEG measures of vigilance, with higher amplitudes corresponding to lower states of vigilance. In a recent preliminary study, [189] demonstrated a negative correlation between the GS and EEG vigilance time courses over the course of a scan. Thus, at time points where the GS signal exhibited positive peaks, the EEG vigilance was found to be low, whereas negative peaks in the GS corresponded to high EEG vigilance values. [190] found evidence for a similar negative correlation between the GS time course and LFP measures of arousal in non-human primates. Furthermore, a recent study identified characteristic events in global electro-cortical activity that were related to drops in arousal [191]. These temporally distinct events may contribute to the appearance of high magnitude events observed in the GS. The link between global activity and dynamic changes in arousal finds further potential support in the observations of [192], who used pupilometry to measure arousal states in mice. They found that increases in arousal were associated with decreases in an optical measure of global hemodynamic activity.

While the neurobiological mechanisms linking vigilance to global fluctuations are still

not fully understood, it is thought that the global nature of the fluctuations may reflect widespread projections of various arousal systems onto the cortex [193,194]. Although further work is needed, the existing evidence suggests that images with high GS magnitudes (i.e. uniformly positive or negative) are associated with temporal peaks and valleys in the state of vigilance or arousal. Because GSR and its variants downweights these images, it is likely that these approaches are minimizing the contribution of vigilance fluctuations to the resulting correlation maps. In addition, the results in Figure 3.15 demonstrate that scans with higher GS standard deviation undergo a higher degree of downweighting or censoring, suggesting that GSR and its variants may have a greater effect on scans with low mean vigilance levels. Further studies are needed to better understand the origins of vigilance-related signal components in fMRI and to assess the effects of GSR (and its variants) on these components.

As shown in Figures 3.2, 3.4, and 3.8, images from time points with low GS magnitudes can contain spatial patterns resembling the DMN and TPN. For example at time point 68 in Figure 3.4, the GS is relatively small and the image values in the DMN and TPN are largely negative and positive, respectively. The opposite signs in these two networks (that are both fairly spread out across space) results in a cancelation of voxel values when computing the global average, consistent with the low GS magnitude. When multiplied by the value of the PCC signal (which is negative because it is in the DMN), the resulting weighted images have positive values in the DMN and negative values in the TPN. The inclusion of these weighted images in the correlation sum will tend to enhance the presence of anti-correlations in the resulting correlation map. Thus, the presence of anti-correlated DMN and TPN networks after GS censoring largely reflects the fact that the retained images already exhibit this spatial relationship.

3.5.4 Related Approaches

Our findings with the GS censoring approach complement prior studies, which demonstrated that key features of resting-state functional connectivity maps could be obtained using

a fraction (e.g. 5 to 15%) of the original time points [195, 196]. In particular, [196] found that an average image constructed from a set of timepoints showing the highest (top 15%) PCC signal values exhibited a pattern that almost perfectly matched the group average PCC correlation map, which exhibited anti-correlation between the DMN and TPN. They further decomposed the selected images into a sum of co-activation patterns (CAPs). It is likely that the CAP time points identified by [196] partially overlap with the time points that would be retained with a GS censoring approach, since it is these time points that exhibit the DMN and TPN spatial patterns in the individual images. The partial overlap reflects the fact that the CAP time points occur for positive peaks in the PCC signal, whereas the GS censoring approach will pick out time points that can have either positive or negative peaks in the PCC signal. As support for this conjecture, we note that [196] found that GSR did not affect the spatial patterns of the CAPs. This is consistent with our finding that GSR has a minimal effect on those time points for which the GS magnitude is low. In addition, in Supplementary Figure 7.6, we show that the average of images from those time points that are retained by GS censoring and have a positive PCC signal shows a strong similarity with the PCC correlation maps obtained after GSR. Further work to elucidate the connection between the GS censoring and CAP approaches would be of interest.

At first glance the downweighting approaches presented here may be considered to be similar to the frame-to-frame intensity stabilization that has been used in prior studies (also referred to as global signal normalization) [51]. In the stabilization approach, all computations, including calculation of the global signal, are performed prior to the removal of the temporal means of the voxel time courses. At each time point, the image data are divided by the global signal and then a constant term (typically equal to 1.0) is subtracted from the normalized images prior to the computation of inter-voxel correlations. The apparent similarity between the stabilization and downweighting approaches lies in the fact that they both apply a uniform scaling to all voxels at a given time point. A key difference is that the global signal used in the stabilization approach has a large positive mean while the global signal signal in GSR and its variants has zero mean. Taking

into account this difference, it can be shown that the stabilization process is equivalent to simply subtracting out the zero mean global signal from each demeaned voxel time series (see Appendix B). Thus, intensity stabilization is more accurately viewed as a global mean subtraction approach (also known as global signal subtraction) as opposed to a downweighting or scaling approach. In comparing intensity stabilization to GSR, the key difference is that GSR finds the optimal fit (through regression) between the voxel time series and the GS prior to removal of an appropriately scaled version of the GS, whereas intensity normalization simply subtracts the GS from each voxel time series, without any voxel-specific scaling. Supplementary Figure 7.5 provides further examples of the fundamental difference between stabilization, GSR, and downweighting.

In this paper we have focused on the effects of GSR on seed-based correlation analysis. Independent components analysis (ICA) is another popular approach for the analysis of resting-state fMRI data, and some ICA implementations use global signal subtraction as a pre-processing step [197]. As discussed above and in Appendix B, when the voxel time series is approximately equal to the GS (i.e. voxel-specific scaling is approximately 1.0), then GSR and global signal subtraction would be expected to have similar downweighting effects. However, the validity of this approximation is likely to vary across voxels and scans. In addition, some ICA implementations do not use global signal subtraction. [197] have reported differences in ICA results obtained with and without global signal subtraction, with slightly better performance observed with subtraction. Additional work would be useful for better understanding the effect of the global signal on ICA-based analyses, including an examination of the effects of GS downweighting and censoring.

In the Results section, we showed that the GS censoring approach differs from the motion censoring methods that have been adopted by some rsfMRI studies, with a relatively small overlap between the censored points. As it is the average of all brain signals, the GS captures the contributions of many potential signal sources, such as motion-related artifacts and the neuronal contributions previously discussed. In computing the GS, we used a "clean" version of the GS in which nuisance terms, such as motion covariates, were regressed out as part of the preprocessing

pipeline [162]. This step minimizes but does not entirely eliminate the contributions of motion to the GS. We also performed our analyses with a version of the GS in which motion covariates were not removed, and still found clear differences in the effects of GS censoring and motion censoring (see Supplementary Figure 7.7). Nevertheless, further work is needed to better understand the relation between GS and motion censoring.

In prior work, our group presented evidence suggesting that the GS could be viewed as an additive confound [160, 162]. The current work provides further support for this view, as it presents a useful temporal partitioning of the GS into sections with low and high signal magnitudes. From this perspective, the additive effect of the GS is primarily confined to the time points with high GS magnitudes, and it is these time points that are largely attenuated by GSR and the other approaches proposed here. It is also important to note that the orthogonality between the temporal segments with low and high signal magnitudes is fundamentally different from the approximate orthogonality between the GS and other principal components discussed previously in [162, 198]

3.5.5 Future Steps

While our results provide insight into the effects of GSR, the question of how best to use GSR or one of the variants proposed in this paper will require further discussion and investigation. Much of the prior concern regarding the use of GSR arose out of the potential introduction of negative correlations. The current work partly mitigates this concern because it provides a means (GS censoring) of defining a temporal subset of the original data in which the negative correlations already exist. Researchers can use this approach to examine group differences in negative correlations within these temporal subsets. For example, it would be instructive to use the censoring approach to compare anticorrelations in studies where global signal differences have been found (e.g. schizophrenic versus healthy controls [67]).

In addition, by comparing the correlation maps obtained before and after GSR with

those obtained after GS censoring, researchers can estimate the extent to which the functional connectivity maps are dominated by the GS signal in the censored time points. In comparing maps across groups or conditions, the proportion of points that are censored could prove to be a useful metric. Given the prior work relating average GS amplitude to vigilance [165] and the results shown in Figure 3.15, it is likely that scans in subjects with a lower level of vigilance and higher GS amplitudes will experience a greater degree of censoring.

In this work, we introduced the concept of GS censoring using a threshold based upon a piecewise linear function of the GS, as this approach was useful for making a link between GSR and GS censoring. In general, there is no need to use a function of the GS and instead one can censor time points using a threshold on the GS magnitude (see for example Figure 5). However, the choice of threshold may depend on the specifics of the fMRI acquisition (such as echo time and magnetic field strength), as these parameters can affect the range of GS magnitudes. Further work is needed to best determine how to best generalize the censoring approach to a wide range of studies. In addition, there is a direct loss in the temporal degrees of freedom when using GS censoring, and the implications of this loss on between-subject and between-group comparisons will need to be considered. Furthermore, our observation that GSR can be approximated as a temporal downweighting process suggests that there is an effective loss of temporal degrees of freedom when using GSR. Additional work is needed to quantify this loss and assess its impact on subsequent analyses.

As shown in Figure 3.13, the choice of the threshold has a direct effect on the extent and magnitude of negative correlations, with the anti-correlations becoming more pronounced as a greater number of time points are censored. However, this observation does not tell us how best to select a threshold level. Is a threshold that reveals more anti-correlations preferred to one that yields less pronounced anti-correlations? To answer this question we will need a deeper understanding of the mechanisms that give rise to large GS magnitudes, including whether or not these mechanisms can be viewed as introducing an additive signal confound that obscures the

presence of anti-correlations in the censored data points. If this is the case, it may be possible to devise algorithms that use the data in the retained time points to estimate the underlying data in the censored time points. Future studies are needed to address these issues.

3.6 Appendix

3.6.1 Approximating the Effects of GSR

In this appendix we provide a mathematical perspective of the basic observations presented in Section 3.2.2. Our goal is to provide background and motivation for the empirical findings (see Methods and Results) demonstrating that GSR largely attenuates signals for time intervals where the global signal magnitude is high. To begin, we express the global signal as the sum of two non-overlapping parts in time:

$$\mathbf{g} = \mathbf{1}_H\mathbf{g} + \mathbf{1}_L\mathbf{g} = \mathbf{g}_H + \mathbf{g}_L, \tag{3.5}$$

where $\mathbf{1}_H$ is the indicator function for the set H of points surrounding large values in the global signal and $\mathbf{1}_L$ is the indicator function for the set L of remaining points corresponding to relatively small values of the global signal.

In the example shown in Figure 3.3, the set H corresponds to the time points surrounding the peaks of the global signal and is indicated by the thick black horizontal bars. Similarly, we can also partition an arbitrary voxel time series in the same way $\mathbf{x} = \mathbf{x}_H + \mathbf{x}_L$. Note that by construction, $H \cap L = 0$ and the time series from different sets are orthogonal, such that $\mathbf{g}_H^T\mathbf{g}_L = 0$, $\mathbf{x}_H^T\mathbf{x}_L = 0$, $\mathbf{g}_H^T\mathbf{x}_L = 0$, and $\mathbf{g}_L^T\mathbf{x}_H = 0$. Note that without loss of generality, we assume that both $\mathbf{g}$ and $\mathbf{x}$ are zero-mean percent normalized time series (i.e. both represent percent BOLD signal changes).

Using the decomposition above, we rewrite the voxel time series after GSR as:

$$\begin{aligned}
\tilde{\mathbf{x}} &= \mathbf{x} - \mathbf{g}\left(\mathbf{g}^T\mathbf{g}\right)^{-1}\mathbf{g}^T\mathbf{x} && (3.6)\\
&= \mathbf{x}_H + \mathbf{x}_L - \frac{\left(\mathbf{g}_H + \mathbf{g}_L\right)\left(\mathbf{g}_H + \mathbf{g}_L\right)^T\left(\mathbf{x}_H + \mathbf{x}_L\right)}{\left(\mathbf{g}_H^T\mathbf{g}_H + \mathbf{g}_L^T\mathbf{g}_L\right)} && (3.7)\\
&= \left(\mathbf{x}_H - \frac{\mathbf{g}_H\left(\mathbf{g}_H + \mathbf{g}_L\right)^T\left(\mathbf{x}_H + \mathbf{x}_L\right)}{\left(\mathbf{g}_H^T\mathbf{g}_H + \mathbf{g}_L^T\mathbf{g}_L\right)}\right) + \left(\mathbf{x}_L - \frac{\mathbf{g}_L\left(\mathbf{g}_H + \mathbf{g}_L\right)^T\left(\mathbf{x}_H + \mathbf{x}_L\right)}{\left(\mathbf{g}_H^T\mathbf{g}_H + \mathbf{g}_L^T\mathbf{g}_L\right)}\right) && (3.8)\\
&= \underbrace{\left(\mathbf{x}_H - \frac{\mathbf{g}_H\left(\mathbf{g}_H^T\mathbf{x}_H + \mathbf{g}_L^T\mathbf{x}_L\right)}{\left(\mathbf{g}_H^T\mathbf{g}_H + \mathbf{g}_L^T\mathbf{g}_L\right)}\right)}_{\tilde{\mathbf{x}}_H} + \underbrace{\left(\mathbf{x}_L - \frac{\mathbf{g}_L\left(\mathbf{g}_H^T\mathbf{x}_H + \mathbf{g}_L^T\mathbf{x}_L\right)}{\left(\mathbf{g}_H^T\mathbf{g}_H + \mathbf{g}_L^T\mathbf{g}_L\right)}\right)}_{\tilde{\mathbf{x}}_L}, && (3.9)
\end{aligned}$$

where we have used the orthogonality relations described above. The first part $\tilde{\mathbf{x}}_H$ in the expression corresponds to the part of $\tilde{\mathbf{x}}$ with points in the set H while the second term $\tilde{\mathbf{x}}_L$ corresponds to points in the complementary set L.

As noted in Section 3.2.1, large peaks in the global signal occur when the majority of voxel waveforms sum in a constructive fashion. As an approximation, we represent the voxel waveforms surrounding these peaks as a scaled version of the global signal $\mathbf{x}_H \approx \alpha\mathbf{g}_H$. Although we expect that for most voxels $\alpha > 0$ because of the widespread positive correlation that exists empirically between the global signal and the voxel time series [173], it is important to note that the derivations below do not depend on the sign of α. Using this approximation, we rewrite the first term $\tilde{\mathbf{x}}_H$ in Eq. (3.9) as

$$\begin{aligned}
\tilde{\mathbf{x}}_H &\approx \alpha\mathbf{g}_H - \frac{\mathbf{g}_H\left(\alpha\mathbf{g}_H^T\mathbf{g}_H + \mathbf{g}_L^T\mathbf{x}_L\right)}{\left(\mathbf{g}_H^T\mathbf{g}_H + \mathbf{g}_L^T\mathbf{g}_L\right)} && (3.10)\\
&= \alpha\mathbf{g}_H\left(1 - \frac{\mathbf{g}_H^T\mathbf{g}_H + \left(\mathbf{g}_L^T\mathbf{x}_L\right)/\alpha}{\mathbf{g}_H^T\mathbf{g}_H + \mathbf{g}_L^T\mathbf{g}_L}\right) && (3.11)\\
&\approx \alpha\mathbf{g}_H\left(1 - \frac{\mathbf{g}_H^T\mathbf{g}_H}{\mathbf{g}_H^T\mathbf{g}_H + \mathbf{g}_L^T\mathbf{g}_L}\right) && (3.12)
\end{aligned}$$

where the last approximation reflects our expectation that $\mathbf{g}_H^T\mathbf{g}_H \gg \mathbf{g}_L^T\mathbf{x}_L$, due to the small

magnitude of the signal components in $\mathbf{g}_L$ and the fact that the cancellation of signals across voxels will tend to make $\mathbf{g}_L$ orthogonal to the voxel time courses for time points within the set L.

In the event that $\mathbf{g}_L$ is equal to zero, it follows directly from Eq. (3.11) that $\tilde{\mathbf{x}}_H = 0$. In addition, this will lead to $\tilde{\mathbf{x}}_L = \mathbf{x}_L$ in Eq. (3.9) , so that the remaining signal is simply $\tilde{\mathbf{x}} = \mathbf{x}_L$. In other words, GSR will eliminate the signal at time points where the global signal is large but will not affect the original voxel values $\mathbf{x}_L$ for those time points where the global signal is zero. This is the case for the example shown on the lefthand column of Figure 3.3, where the peaks in the original voxel time series that coincide with the large peaks in the global signal are completely eliminated by GSR, whereas those signals that correspond to segments where the global signal is zero (due to complete cancelation of the signals across voxels) are unaffected by GSR.

In the more realistic case where the cancellation of signals is not complete, the magnitudes of the signal components in $\mathbf{g}_L$ are still expected to be smaller than the magnitudes in $\mathbf{g}_H$, but $\mathbf{g}_L$ will differ from zero for some points in L. An example of this case is shown in the middle column of Figure 3.3, where the cancellation of the signal components centered about time points 30 and 70 is not complete, leading to the appearance of small magnitude peaks in the global signal around these time points. The attenuation term $\left(1 - \frac{\mathbf{g}_H^T \mathbf{g}_H}{\mathbf{g}_H^T \mathbf{g}_H + \mathbf{g}_L^T \mathbf{g}_L}\right)$ in Eq. (3.12) can be rewritten as the ratio $\frac{\mathbf{g}_L^T \mathbf{g}_L}{\mathbf{g}_H^T \mathbf{g}_H + \mathbf{g}_L^T \mathbf{g}_L}$ of the energy (i.e. squared norm) of the smaller magnitudes signals in $\mathbf{g}_L$ to the overall energy of $\mathbf{g}$. Based on the partitioning of the signal, this ratio is expected to be much smaller than 1. As a result, the signal components in $\mathbf{x}_H$ will still be greatly reduced by GSR as shown in panels (h) and (i) of Figure 3.3.

We can approximate $\tilde{\mathbf{x}}_L$ in Eq. (3.9) as

$$\tilde{\mathbf{x}}_L \approx \left(\mathbf{x}_L - \frac{\mathbf{g}_L \left(\alpha \mathbf{g}_H^T \mathbf{g}_H + \mathbf{g}_L^T \mathbf{x}_L\right)}{\left(\mathbf{g}_H^T \mathbf{g}_H + \mathbf{g}_L^T \mathbf{g}_L\right)}\right) \tag{3.13}$$

$$\approx \mathbf{x}_L - \mathbf{g}_L \frac{\alpha \mathbf{g}_H^T \mathbf{g}_H}{\mathbf{g}_H^T \mathbf{g}_H + \mathbf{g}_L^T \mathbf{g}_L} \tag{3.14}$$

$$\approx \mathbf{x}_L - \alpha \mathbf{g}_L \tag{3.15}$$

where we have used the approximations already noted above ($\mathbf{x}_H \approx \alpha \mathbf{g}_H$, $\mathbf{g}_H^T \mathbf{g}_H \gg \mathbf{g}_L^T \mathbf{g}_L$, and $\mathbf{g}_H^T \mathbf{g}_H \gg \mathbf{g}_L^T \mathbf{x}_L$). The perturbation $\alpha \mathbf{g}_L$ in the final expression is proportional to the global signal. Since this occurs in time intervals where the global signal magnitude is assumed to be small (by definition), the magnitude of the perturbation is also expected to be small. This is consistent with the minimal perturbation of the signals at time points 30 and 70 in panel (h) of Figure 3.3 and in the intervals indicated by the magenta dashed lines in panel (i) of the same figure. As discussed above, when $\mathbf{g}_L = 0$, the waveforms are completely unaffected and $\tilde{\mathbf{x}}_L = \mathbf{x}_L$ (i.e. time points 30, 70, 90, and 110 in panel (g) and time points 90 and 110 in panel (h).

3.6.2 Intensity Stabilization Approach

In this approach, computations are performed prior to the removal of the voxel-wise temporal means or percent normalization. Thus, we can define the unnormalized global signal $\mathbf{g}_U$ as the sum $\mathbf{g}_U = \bar{g}\mathbf{1}_M + \tilde{\mathbf{g}}$ of a mean term $\bar{g}$ and a zero-mean fluctuation term $\tilde{\mathbf{g}}$, where $\mathbf{1}_M$ indicates a $M \times 1$ column vector of ones. An unnormalized voxel time series is similarly defined as $\mathbf{x}_U = \bar{x}\mathbf{1}_M + \tilde{\mathbf{x}}$. The process of intensity stabilization (or GS normalization) is formally defined as

$$\left(\text{diag}\left(\mathbf{g}_U\right)\right)^{-1} \mathbf{x}_U - \mathbf{1}_M \tag{3.16}$$

where $\text{diag}\left(\mathbf{g}_U\right)$ denotes the matrix with $\mathbf{g}_U$ along the diagonal [51].

To proceed, let $\tilde{g}[i]$ and $\tilde{x}[i]$ denote the values of $\tilde{\mathbf{g}}$ and $\tilde{\mathbf{x}}$ at the ith time point. Then the

normalized values are

$$\begin{aligned}
\frac{\bar{x}+\tilde{x}[i]}{\bar{g}+\tilde{g}[i]}-1 &= \frac{\tilde{x}[i]-\tilde{g}[i]+\bar{x}-\bar{g}}{\bar{g}+\tilde{g}[i]} && (3.17)\\
&= \frac{\tilde{x}[i]-\tilde{g}[i]}{\bar{g}+\tilde{g[i]}} && (3.18)\\
&\approx \frac{\tilde{x}[i]-\tilde{g}[i]}{\bar{g}}\left(1-\frac{\tilde{g}[i]}{\bar{g}}\right) && (3.19)\\
&\approx \frac{\tilde{x}[i]-\tilde{g}[i]}{\bar{g}} && (3.20)\\
&= \frac{\tilde{x}[i]}{\bar{x}}-\frac{\tilde{g}[i]}{\bar{g}} && (3.21)
\end{aligned}$$

where we have made use of the fact that the magnitude of the fluctuations in fMRI GS time series are typically only a few percent at most of the overall mean, so that $\bar{g} \gg \tilde{g}[i]$. In addition we assumed that $\bar{x} = \bar{g}$, where this last relation holds because we can always scale the data such that the means of all the voxels are the same. Thus, to first order the process of GS normalization is equivalent to simply taking the difference between the percent change voxel time series $\tilde{x}[i]/\bar{x}$ and the percent change global signal $\tilde{g}[i]/\bar{g}$. Using vector notation, the approximation to GS normalization is expressed as the difference $\mathbf{x}-\mathbf{g}$ of the percent normalized voxel time series and global signals. Thus, GS normalization and GS subtraction are nearly identical methods.

The GS normalization (and subtraction) approach and GSR are expected to give similar results when

$$\mathbf{x}-\mathbf{g} \approx \mathbf{x}-\mathbf{g}\left(\mathbf{g}^T\mathbf{g}\right)^{-1}\mathbf{g}^T\mathbf{x} \qquad (3.22)$$

This will occur whenever the fit coefficient $\alpha = \left(\mathbf{g}^T\mathbf{g}\right)^{-1}\mathbf{g}^T\mathbf{x}$ between the voxel time series and the GS is approximately equal to 1.0.

3.7 Acknowledgments

Chapter 3 is, in full, a reprint of the material as it appears in: Alican Nalci, Bhaskar D. Rao, and Thomas T. Liu. "*Global Signal Regression Acts as a Temporal Downweighting Process in Resting-state fMRI.*" NeuroImage 152 (2017): 602-618. I was the primary author of this paper and T.T. Liu supervised the research. The work was partially supported by NIH grant R21MH112155 and a UC San Diego Frontiers of Innovation Scholars Program (FISP) Project Fellowship.

Chapter 4

Nuisance Effects and the Limitations of Nuisance Regression in Dynamic Functional Connectivity fMRI

In resting-state fMRI, dynamic functional connectivity (DFC) measures are used to characterize temporal changes in the brain's intrinsic functional connectivity. A widely used approach for DFC estimation is the computation of the sliding window correlation between blood oxygenation level dependent (BOLD) signals from different brain regions. Although the source of temporal fluctuations in DFC estimates remains largely unknown, there is growing evidence that they may reflect dynamic shifts between functional brain networks. At the same time, recent findings suggest that DFC estimates might be prone to the influence of nuisance factors such as the physiological modulation of the BOLD signal. Therefore, nuisance regression is used in many DFC studies to regress out the effects of nuisance terms prior to the computation of DFC estimates. In this work we examined the relationship between seed-specific sliding window correlation-based DFC estimates and nuisance factors. We found that DFC estimates were significantly correlated with temporal fluctuations in the magnitude (norm) of various nuisance regressors.

Strong correlations between the DFC estimates and nuisance regressor norms were found even when the underlying correlations between the nuisance and fMRI time courses were relatively small. We then show that nuisance regression does not necessarily eliminate the relationship between DFC estimates and nuisance norms, with significant correlations observed between the DFC estimates and nuisance norms even after nuisance regression. We present theoretical bounds on the difference between DFC estimates obtained before and after nuisance regression and relate these bounds to limitations in the efficacy of nuisance regression with regards to DFC estimates.

4.1 Introduction

In resting-state functional magnetic resonance imaging (fMRI), the correlation between the BOLD time courses from different brain regions is used to estimate the functional connectivity (FC) of the brain in the absence of an explicit task. This approach has revealed a number of resting-state networks, where each network consists of brain regions that exhibit a high degree of mutual correlation [50]. For the most part, FC estimates have assumed temporal stationarity of the underlying BOLD time courses and are obtained by computing correlations over the entire scan duration. As the resulting estimates represent a *temporally averaged* measure of brain connectivity over a typical scan duration of 5-10 minutes, they can miss important dynamic temporal changes in FC [55, 56, 57].

An increasing number of studies have focused on the dynamics of functional brain connectivity by considering dynamic FC (DFC) measures that are computed over time scales typically on the order of tens of seconds and thus much shorter than the scan duration [55, 56, 57]. Approaches for estimating DFC include sliding window correlation method [56, 199], time-frequency methods such as wavelet transform coherence [144, 200], and probabilistic inference methods such as hidden Markov modeling [201]. To date, the sliding window correlation approach is the most widely used DFC estimation method [56]. In this approach, the correlations

between BOLD time courses from different brain regions are computed over sliding windows with durations typically ranging from 30-60 seconds [57]. Regardless of the analysis method, non-neural processes that affect the BOLD time series can also contaminate the DFC estimates [56, 57, 202]. These confounds are often referred to as *nuisance terms* and include the effects of motion, cardiac and respiratory activity, and fluctuations in arterial CO_2 concentration. [56] noted that non-neuronal sources that introduce spatial correlations into the time series can also give rise to spurious dynamics in FC measures. Recently, [203] reported that temporal fluctuations in network degree were related to fluctuations in both heart rate and end-tidal CO_2. [204] found that temporal fluctuations in FC were related to temporal variations in a global measure of average BOLD signal magnitude.

A common step in most fMRI analyses is the use of nuisance regression to minimize the contributions of nuisance terms in BOLD time courses [205, 206]. Nuisance regressors include cardiac and respiratory activity derived time courses [63, 64], head motion parameters, Legendre polynomials to model scanner drift, signals from white-matter (WM) and cerebrospinal fluid (CSF) regions, and a whole brain global signal (GS) [78]. Although nuisance regression is widely employed prior to the computation of DFC estimates [202], efforts to examine its efficacy with regards to DFC estimates have been limited. [56] noted that residual nuisance effects "inevitably remain" in the BOLD time series and must be considered in the interpretation of DFC measures. In support of this view, [203] found that nuisance regression diminished but did not completely remove the relationship between network degree and measures of cardiac and respiratory activity. On the other hand, [207] reported that global signal regression (GSR) had a spatially heterogeneous impact on DFC estimates, but did not assess whether GSR removed GS contributions from the DFC estimates.

In this paper, we take a closer look at (1) the role of nuisance terms in correlation-based DFC estimates and (2) the efficacy of nuisance regression for DFC studies. In particular, we use two independent datasets to examine the relationship between seed-specific sliding window

correlation-based DFC estimates and the norms of various nuisance regressors (WM and CSF signals, GS, cardiac and respiratory measurements, and head motion). We then assess the effect of nuisance regression on this relation, considering both regression applied to the entire scan and regression applied on a sliding window basis. To interpret the empirical findings, we derive mathematical expressions to describe the effect of regression on DFC estimates and compare the experimental results with the theoretical predictions. Preliminary versions of this work have been presented in [208, 209].

4.2 Methods

4.2.1 Datasets

In this work, we analyzed two datasets in order to show the generality of our results and to experiment with different types of nuisance measurements. First, to understand the effect of nuisance measurements derived directly from the MRI images such as the GS, WM and CSF, and head motion (HM) time courses on the DFC estimates, we used a publicly available dataset originally analyzed by [174], which we will refer to as the BS002 dataset. Second, to understand the effect of physiological measurements such as changes in the respiration and cardiac rate on the DFC estimates, we used the dataset analyzed by [160], which we will refer to as the CFMRI dataset.

BS002 data were acquired from 17 young adults (9 females) using a 3 T Siemens Allegra MR scanner. Each subject underwent 4 BOLD-EPI fixation runs (32 slices, TR=2.16 s, TE=25 ms, $4\times4\times4$ mm), each lasting 7 minutes (194 frames). The subjects were instructed to look at a cross-hair and asked to remain still and awake. High-resolution T1-weighted anatomical images were acquired for the purpose of anatomical registration (TR=2.1 s, TE=3.93 ms, flip angle=7 deg, $1\times1\times1.25$ mm).

CFMRI data were acquired from 10 healthy volunteers (4 males and 6 females) using a 3

Tesla GE Discovery MR750 system. Each subject underwent four separate resting-state scans (5 minutes per scan with either their eyes open or closed). High resolution anatomical data were collected using a magnetization prepared 3D fast spoiled gradient (FSPGR) sequence (TI=600 ms, TE=3.1 ms, flip angle = 8 degrees, slice thickness = 1 mm, FOV = 25.6 cm, matrix size = $256\times256\times176$). Whole brain BOLD data were acquired over thirty axial slices using an echo planar imaging (EPI) sequence (flip angle = 70 degrees, slice thickness = 4 mm, slice gap = 1 mm, FOV = 24 cm, TE = 30 ms, TR = 1.8 s, matrix size = 64×64). Cardiac pulse and respiratory data were monitored using a pulse oximeter (InVivo) and a respiratory effort transducer (BIOPAC), respectively. The pulse oximeter was placed on each subject's index finger while the respiratory effort belt was placed around the subject's abdomen. Physiological data were sampled at 40 Hz using a multi-channel data acquisition board (National Instruments).

4.2.2 Preprocessing steps for the BS002 dataset

Standard pre-processing steps were conducted with the AFNI software package [175]. The initial 9 frames from each EPI run were discarded to minimize longitudinal relaxation effects. Images were then slice-time corrected and co-registered, and the 6 head motion parameter time series were retained. The resultant images were converted to coordinates of Talairach and Tournoux (TT), resampled to 3 mm cubic voxels, and spatially smoothed using a 6 mm full-width-at-half-maximum isotropic Gaussian kernel. The 1^{st} and 2^{nd} order Legendre polynomials (a constant term to model the temporal mean and a linear trend) were projected out from each voxel's time course. Each voxel time series was then converted into a percent change BOLD time series through demeaning and division by its mean value.

To analyze the effect of nuisance regression on DFC estimates, we used both (1) full linear regression in which regression is performed over the whole scan duration and (2) block regression (also referred to as sliding window regression) in which regression is performed for each window separately (more details on block regression are provided in Section 4.2.6). For each type of

regression, we performed a separate regression on the data using one of the following types of regressor: (1) the set of all 6 head motion parameters, (2) nuisance signals from WM and CSF regions, and (3) the global signal (GS), which was calculated as the average of the percent change time series across all voxels within the brain.

For DFC analysis, we used seed signals derived from the posterior cingulate cortex (PCC), intraparietal sulcus (IPS), frontal eye fields (FEF), and motor network (MOT). These seed signals were obtained by averaging time series selected over spheres of radius 6 mm (2 voxels) centered about their corresponding TT coordinates [162]. The sphere centers were obtained by converting the MNI coordinates from [178] to TT coordinates [179]. For the PCC, left MOT, and right MOT seeds we used the coordinates [0,-51,26], [-36,-22,52] and [37,-21,52], respectively. A combined MOT seed was obtained by using the left and right MOT coordinates to define two spheres and by merging them. For the IPS and FEF seeds, we used the coordinates [27,-58,49] and [24,-13,51] from [180]. Finally, for the WM and CSF nuisance signals, we defined the sphere centers as [12,-36,27] and [9,-9,15], respectively.

4.2.3 Preprocessing steps for the CFMRI dataset

We used FSL and AFNI packages to preprocess the CFMRI dataset [210, 211]. First, high-resolution anatomical data were skull stripped and segmentation was applied to estimate WM, gray matter, and CSF partial volume fractions. Images were then slice-time corrected and co-registered and the 6 head motion parameter time series were retained. The anatomical data were then aligned to the functionals. Each subject's native space data were transferred to TT coordinates and spatially smoothed using a 6 mm full-width-at-half-maximum isotropic Gaussian kernel. A binary brain mask was created for each subject using the transferred data and was eroded by 2 voxels along the brain edges to eliminate possible inclusion of non-brain areas [212]. The first 6 frames (10.8 seconds) from each functional data were discarded to minimize longitudinal relaxation effects. For each run, the 1^{st} and 2^{nd} order Legendre polynomials and the 6 head

motion parameters were projected out to obtain the 'baseline' CFMRI data.

We performed RVHRCOR by simultaneously projecting out the physiological nuisance regressors derived from cardiac and respiratory signals using both full linear regression and block regression [63, 64]. Specifically, a respiratory variation (RV) signal was computed as the standard deviation of the respiratory signal using a 7.2 second sliding window and was then convolved with the respiration response function (RRF) to obtain the respiration regressor (RVf). Similarly, a heart rate (HR) signal was computed as the inverse of average peak to peak time interval between two consecutive heartbeats over a 7.2 second sliding window. The HR was then convolved with a cardiac response function (CRF) to obtain the cardiac regressor (HRf). In this process, we discarded 4 scans out of 40 due to the poor quality of physiological signals, leaving a total of 36 scans for analysis. For DFC analysis, we used the same seed coordinates as mentioned for the BS002 dataset.

4.2.4 Calculation of the DFC estimates: Sliding window correlations

The DFC estimates were obtained by computing the sliding window correlations between pairs of seed signals selected from the PCC, IPS, FEF, and MOT regions. Specifically, we computed the sliding window correlations between the following seed pairs: PCC and IPS, PCC and FEF, PCC and MOT, IPS and FEF, IPS and MOT, and FEF and MOT. Denoting x_1 and x_2 as two seed time series and $x_{1,k}$ and $x_{2,k}$ as the windows taken from those signals (where k denotes the window index), we computed the window correlation value as $r_k = \frac{x_{1,k}^T x_{2,k}}{|x_{1,k}||x_{2,k}|} = \cos\theta_k$. Here, θ_k is the observed angle between those time courses and $|.|$ is the vector norm. Note also that here and throughout the paper the time courses are represented as column vectors. This process was repeated by temporally shifting the window index until all sliding window correlations were carried out over the entire scan duration. The final aggregated set of correlation values $r_{DFC} = [r_1, r_2, \cdots, r_T]$ was used as an estimate for the DFC.

In our primary analysis, to maximize the number of available DFC samples per scan,

the sliding window duration in this paper was fixed at 30 TRs and a window shift of 1 TR was used [56, 207]. This duration corresponds to a window length of 63 seconds for the BS002 dataset and 54 seconds for the CFMRI dataset. These window lengths lie within the range of durations typically used in the existing DFC literature [57]. In supplementary analyses, we considered two additional window sizes: a longer window of duration of 100 seconds, which corresponds to 48 TRs for the BS002 dataset and 56 TRs for the CFMRI dataset, and a shorter window size of 40 seconds, which corresponds to 20 TRs for the BS002 dataset and 22 TRs for the CFMRI dataset.

4.2.5 Calculation of the nuisance metrics: Nuisance norms

To examine the relationship between the nuisance measurements and the DFC estimates, we used the norm of the regressor as our nuisance metric. Note that the norm of a regressor is simply the geometric length of the column vector used to represent the regressor time course. In Section 4.4 we present some toy examples and geometric arguments to show how DFC estimates can be related to the norms of various nuisance terms. The theoretically inclined reader may choose to read that section first before continuing.

For a single regressor, the computation of the windowed norms was done by first demeaning the nuisance regressor for each window. Denoting n_k as the demeaned nuisance measurement in the kth window, the window norm was computed as $|n_k| = \sqrt{\sum_{t=t_k}^{t_k+T_D} n_k(t)^2}$, where t indicates time in TR units and T_D is the window duration minus 1 (e.g. $T_D = 29$ for a window duration of 30 TRs). When multiple regressors were used in the regression (e.g. for HM), we computed the total norm as $|n_k|_{\text{Total}} = \sqrt{\sum_{i=1}^{N_R} |n_{k,i}|^2}$, where i denotes the index for different regressors and N_R is the number of regressors. In the text we use the term norm and the notation $|n_k|$ to denote both $|n_k|$ and $|n_k|_{\text{Total}}$. The norm was computed for all windows by temporally shifting the window index. The final set $\{|n_1|, |n_2|, ..., |n_T|\}$ of nuisance norms comprised the nuisance norm time course. To examine the relationship between the nuisance norm and DFC estimates, we computed the correlation coefficient between the DFC estimates and the nuisance norm time courses.

4.2.6 Analysis of nuisance regression techniques

We applied two types of linear regression, which we refer to as full regression and block regression. In full regression, the nuisance terms were projected out of the voxel time series over the entire scan duration. For example, for a single regressor, the clean time course after full regression was obtained as $\tilde{x} = x - n(n^T n)^{-1} n^T x = x - bn$, where n is the column vector representing the nuisance regressor and $b = (n^T n)^{-1} n^T x$ is the scalar fit coefficient for the entire scan duration. In block regression, the nuisance measurement was projected out of each window separately, such that the clean time course for the kth window was $\tilde{x}_k = x_k - n_k(n_k^T n_k)^{-1} n_k^T x_k = x_k - b_k n_k$, where $b_k = (n_k^T n_k)^{-1} n_k^T x_k$ is the window-specific scalar fit coefficient. Note that all time courses (seed and nuisance signals) were demeaned prior to the regression process. The signals were demeaned across the full scan duration when performing full regression and demeaned for each window independently when performing block regression.

Note that we will refer to the "clean" DFC estimates obtained after the application of regression as Post FullReg DFC and Post BlockReg DFC estimates for full and block regression, respectively. To simplify the presentation, we will also use the shorter term Post DFC to refer to Post FullReg DFC, since full regression is the method that has typically been used in the literature.

4.2.7 Significance testing of the relationship between DFC estimates and nuisance norms

To assess the statistical significance of the relationship between the DFC estimates and nuisance norms, we used an autoregressive (AR) bootstrapping procedure based on the work of [144, 213, 214]. As our null hypothesis was that the DFC estimates and nuisance norms were not linearly related, we fit (on a per scan basis) separate AR processes of model order q to the DFC estimates and nuisance norm time courses both before and after full/block regression. The model order q was determined according to the Bayesian information criterion (BIC). Using the

estimated AR coefficients from each scan, we generated 10,000 surrogate time series for both the DFC estimates and nuisance norms and then computed the associated surrogate correlation coefficients.

To create a null distribution for the assessment of significance across the study sample, we computed the absolute value of each of the surrogate correlation coefficients in the sample and took the mean of the absolute values across the sample (i.e. mean across all scans, seed pairs, and regressor types). The resulting null distribution consisted of 10,000 surrogate mean absolute correlation values. We used this null distribution to assess the significance of the sample mean absolute correlation value.

As a secondary descriptive analysis, we also assessed significance on a per-scan basis. We formed a null distribution of correlation values for each scan by correlating the surrogate nuisance norms with the surrogate DFC estimates. We used the resulting null distribution to compute the p-value associated with the correlation coefficient computed from each scan's measured data. We then calculated the percentage of scans that showed significant correlations between the DFC estimates and nuisance norms for significance thresholds of 0.05 and 0.10.

4.3 Results

In this section, we first show that the DFC estimates obtained between pairs of seed time courses can be significantly correlated with the norms of various nuisance measurements. We demonstrate that strong correlations between the DFC estimates and nuisance norms exist even when the correlations between the nuisance and seed time courses are small. We then show that performing nuisance regression prior to the computation of the DFC estimates does not necessarily eliminate the presence of strong and significant correlations between the nuisance norms and DFC estimates.

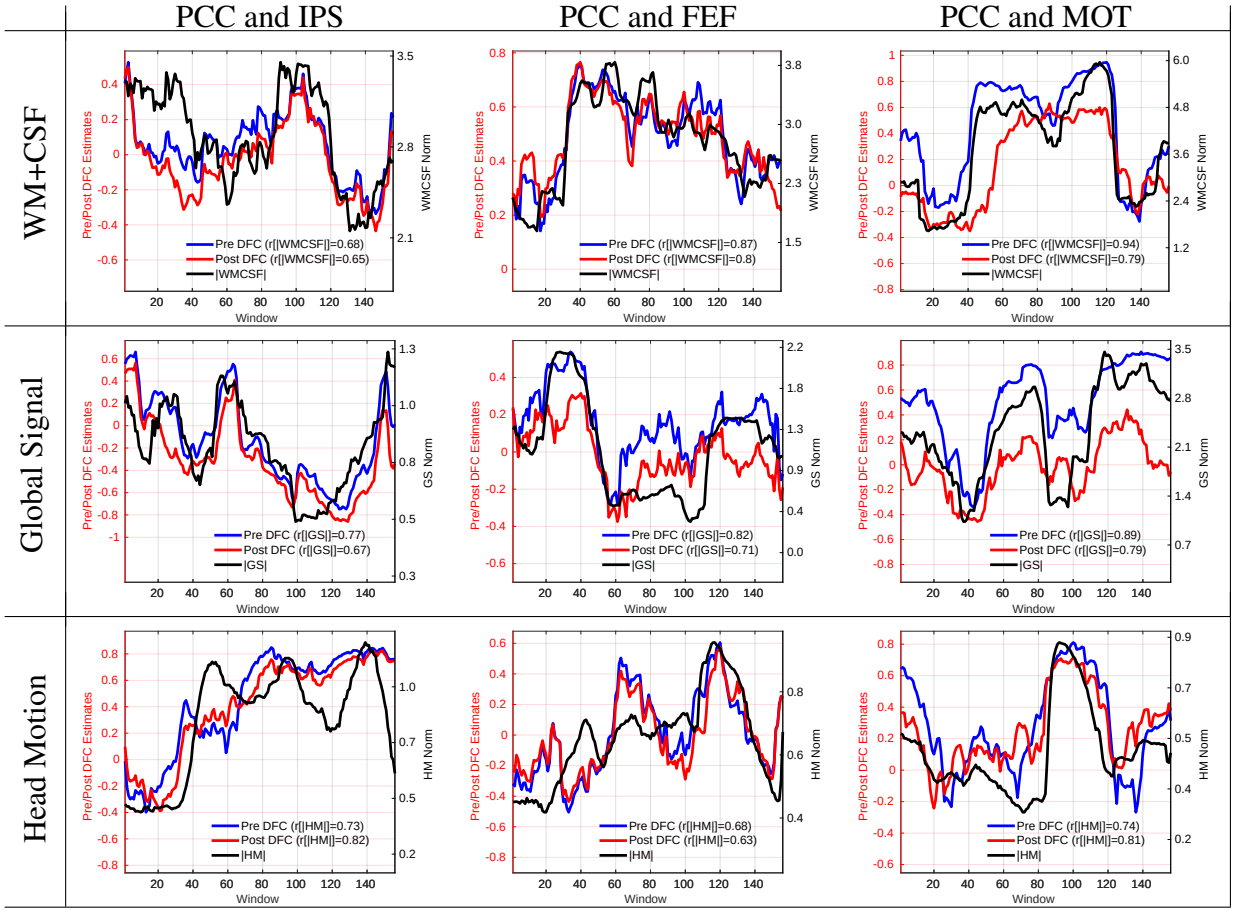


Figure 4.1: 9 example scans that demonstrate strong and significant ($p < 0.02$) correlations between the nuisance norms and DFC estimates. The type of nuisance regressor is indicated by the row label. The seed pair for the DFC estimate is indicated by the column label. The solid blue line in each panel shows the DFC estimate prior to nuisance regression (labeled as Pre DFC on the legend) and the solid black line shows the corresponding nuisance norm. Correlation values between the DFC estimates and nuisance norms are indicated in the legend labels. The correlations between the nuisance norms and the Pre DFC estimates varied from $r = 0.68$ to $r = 0.94$ with a mean value of $r = 0.79$. The DFC estimates after performing full linear regression are shown with solid red lines (labeled as Post DFC on the legend). The correlations between the nuisance norms and the Post DFC estimates ranged from $r = 0.63$ to $r = 0.82$ with a mean value of $r = 0.74$. See Figure 4.3 and Supplementary Figure 7.8 for examples of weak and moderate correlations.

4.3.1 Examples of correlations between DFC estimates and nuisance norms

In Figure 4.1, we present examples of scans from the BS002 dataset in which significant positive correlations between the DFC estimates and nuisance norms were observed. The column labels indicate the seed region pair (e.g. PCC and IPS, PCC and FEF, and PCC and MOT) and the row labels indicate the type of nuisance norm. The solid blue line in each panel shows the DFC estimate before nuisance regression (labeled as Pre DFC) and the solid black line shows the respective nuisance norm. In these scans, the correlations between the nuisance norms and the Pre DFC estimates ranged from $r = 0.68$ to $r = 0.94$ with a mean correlation value of $r = 0.79$, corresponding to a range of explained variances from 46% to 88% over the example scans.

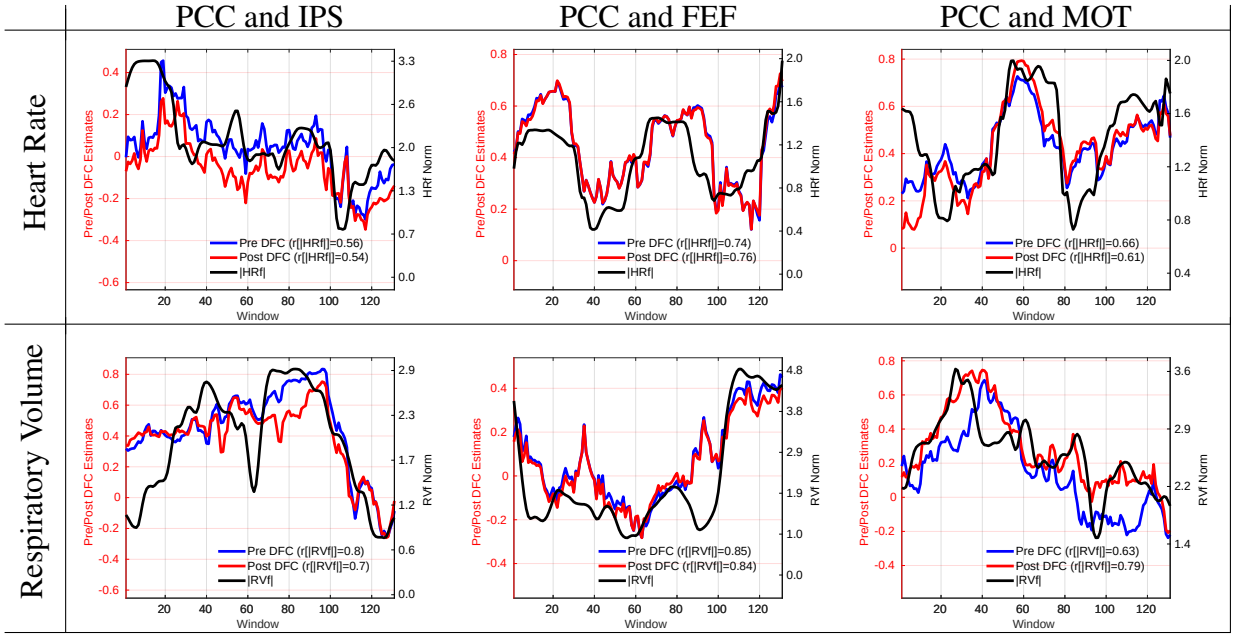


Figure 4.2: 6 example scans showing that the DFC estimates are significantly ($p < 0.03$) correlated with the norms of physiological nuisance terms. Row labels indicate the type of physiological nuisance term considered and column labels indicate the seed pair. Correlation values between the DFC estimates and nuisance norms are indicated in the legend labels. The correlations between the Pre DFC estimates and the nuisance norms ranged from $r = 0.56$ to $r = 0.85$ with a mean correlation of $r = 0.71$. The DFC estimates obtained after performing RVHRCOR (Post DFC) are similar to the Pre DFC estimates. The correlations between the Post DFC and the nuisance norms ranged from $r = 0.54$ to $r = 0.84$ with a mean correlation of $r = 0.71$. See Figure 4.3 and Supplementary Figure 7.8 for examples of weak and moderate correlations.

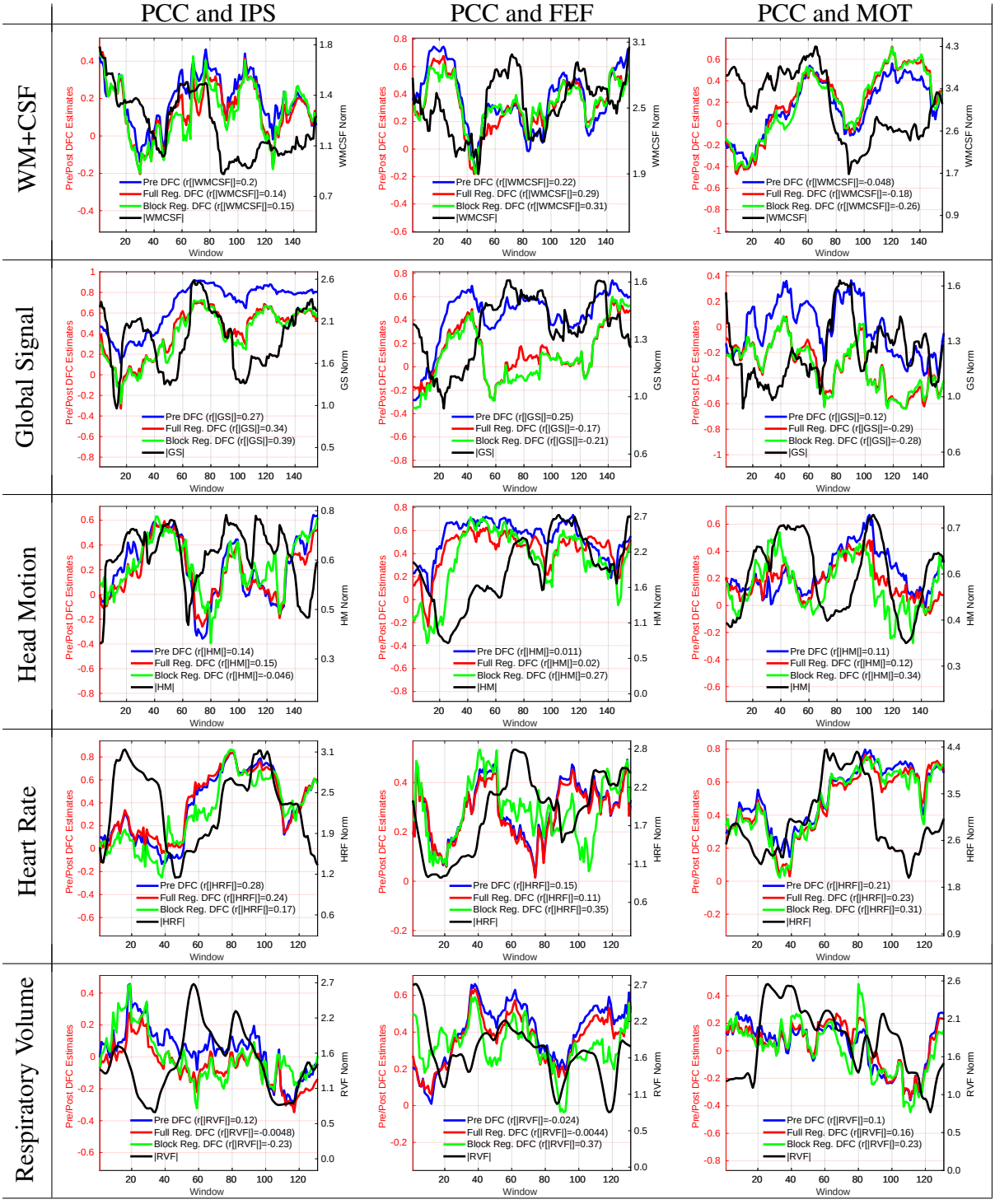


Figure 4.3: 15 example scans that demonstrate weaker correlation values ($r < 0.30$) between the nuisance norms (black) and the Pre DFC estimates (blue). The type of nuisance regressor is indicated by the row label. The seed pair for the DFC estimate is indicated by the column label. Even though the observed correlations are weaker in this Figure, there is some visual similarity between the nuisance norms and DFC estimates. Post DFC estimates with full (red) and block (green) regression are also shown.

The nuisance norms in these scans were significantly correlated with the DFC estimates with per-scan p-values across the example scans of $p < 0.0075$ for WM+CSF, $p < 0.0009$ for GS and $p < 0.0237$ for HM regressors.

In Figure 4.2, we show 6 example scans from the CFMRI dataset demonstrating significant positive correlations between DFC estimates and physiological nuisance norms. The first row shows 3 scans using the HRf norm as the nuisance measure and the second row shows 3 scans using the RVf norm. Each column shows a different seed pair for the DFC estimates. The correlations between the Pre DFC estimates and the nuisance norms in these scans ranged from $r = 0.56$ to $r = 0.85$ with a mean correlation of $r = 0.71$ corresponding to a range of explained

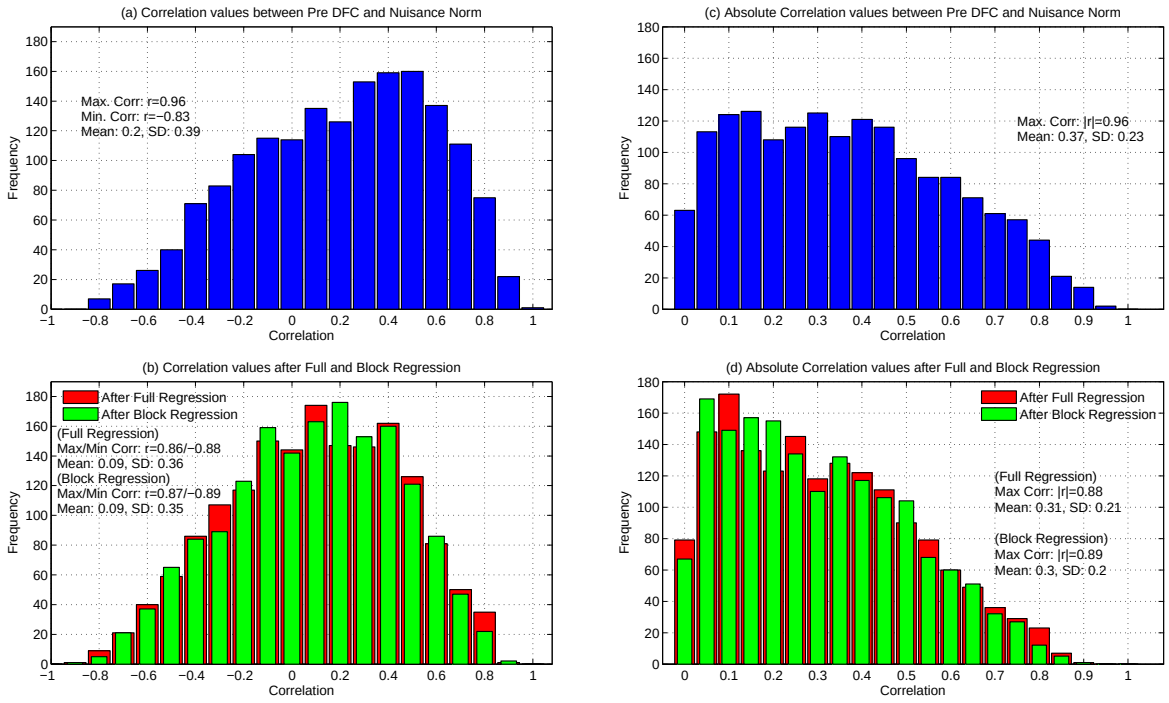


Figure 4.4: (a) Histogram of correlation values obtained between the Pre DFC estimates and the nuisance norm across all scans, seed pairs, and nuisance norms. Distribution of the correlation values ranged from $r = -0.83$ to $r = 0.96$. (b) After regression, the correlation values ranged from a minimum of $r = -0.88$ to a maximum of $r = 0.86$ for full regression, and from a minimum of $r = -0.89$ to a maximum of $r = 0.87$ for block regression. (c) Histogram of absolute correlation values between the nuisance norm and Pre DFC estimates with a sample mean of $|r| = 0.37$. (d) Absolute correlation values after full and block regression with sample means of $|r| = 0.31$ for full and $|r| = 0.30$ for block regression.

variances from 31% to 72%. The RVf and HRf norms were significantly correlated with the DFC estimates with per-scan p-values of $p < 0.0345$ for RVf and $p < 0.0355$ for the HRf regressors.

In Figure 4.3 and Supplementary Figure 7.8, we show additional examples for each regressor and seed pair in cases of weak to medium positive correlations observed between the DFC estimates and nuisance norms. Note that all qualitative examples shown in this paper are from different scans but can be from the same subject.

In Figure 4.4a, we show the histogram of correlations between the nuisance norms and the DFC estimates across all scans, seed pairs, and nuisance norms. The histogram includes a total of $1,656$ correlation values with $1,224$ belonging to the BS002 dataset (68 scans $\times$ 6 seed pairs $\times$ 3 types of nuisance norm) and the remaining 432 from the CFMRI dataset (36 scans $\times$ 6 seed pairs $\times$ 2 types of nuisance norm). The correlations ranged from a negative value of $r = -0.83$ to a positive value of $r = 0.96$, with a skewed distribution in which 68% of the correlations were positive and the remaining 32% were negative. In Supplementary Figure 7.9, we present the correlations from Figure 4.4a separated into the values for individual regressors. Figure 4.4c shows the histogram of absolute correlation values, with a sample mean value of $|r| = 0.37$. The significance of the sample mean absolute correlation value is assessed in the following section.

In Figure 4.5, we show 3 example scans to demonstrate the existence of significant

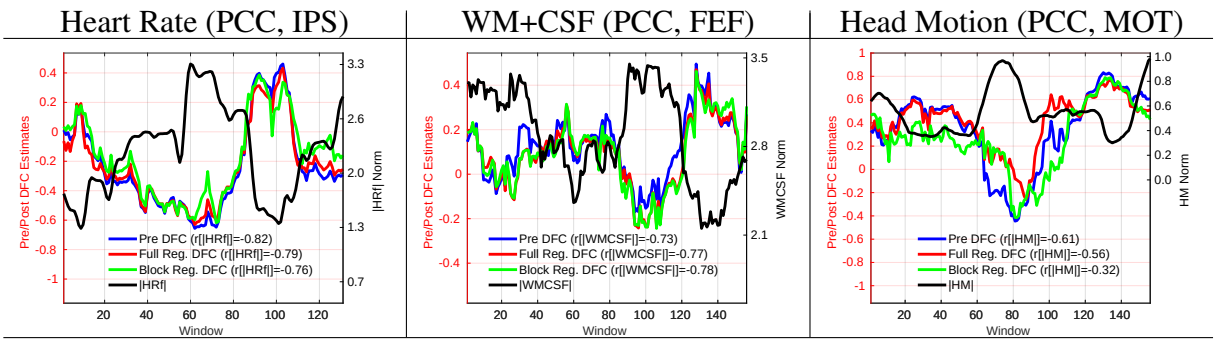


Figure 4.5: 3 scans showing examples of strong anti-correlations between the nuisance norms (black) and the DFC estimates both before (blue) and after full (red) and block (green) regression . The type of nuisance regressor and the seed region pair are indicated in the column labels. The correlation values between the DFC estimates and the nuisance norms are indicated in the legend labels.

anti-correlations between the Pre DFC estimates and various regressor norms. The type of nuisance regressor and the seed region pair are indicated in the column labels. For these scans, the correlation values were strongly negative ranging from $r = -0.82$ to $r = -0.61$ with an increase in nuisance norm corresponding to a decrease in the DFC estimate. The nuisance norms in these scans were significantly anti-correlated with the DFC estimates with per-scan p-values of $p < 0.001$ for HRf, $p < 0.0024$ for WM+CSF and $p < 0.0226$ for the HM regressors.

As an exploratory analysis, we also assessed whether the state of the eyes (open or closed) in the CFMRI dataset affected the relationship between the DFC estimates and the norms of the RVf and HRf time courses. We found no significant difference ($p = 0.95$, two-tailed paired t-test) in the correlations obtained between eyes open and eyes closed conditions.

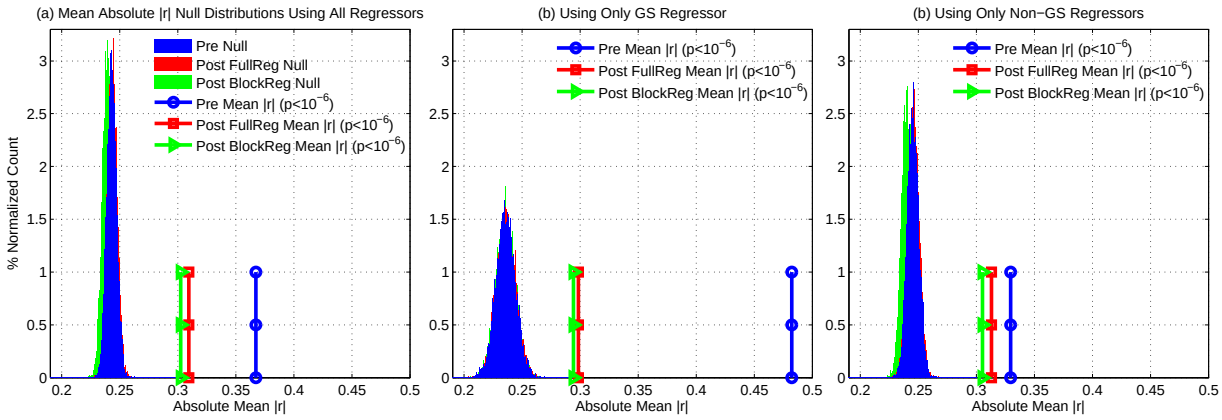


Figure 4.6: (a) Empirical null distributions for mean absolute correlation values when considering all regressors both before regression (blue) and after full regression (red) or block regression (green). The sample mean absolute correlation values for these conditions are shown by the circles, squares, and triangles, and are equal to the values previously reported in 4.4c-d. The sample mean absolute correlation values were found to be significant ($p < 10^{-6}$) both before and after full or block regression. (b) Empirical null distributions and sample mean absolute correlation values when considering only the GS regressor. There was a marked reduction in the sample mean absolute correlation values after regression. However, the sample mean absolute correlation values were still found to be significant ($p < 10^{-6}$) both before and after nuisance regression. (c) Empirical null distributions and sample mean absolute correlation values for the Non-GS regressors. Sample mean absolute correlation values were found to be significant ($p < 10^{-6}$) both before and after nuisance regression.

4.3.2 Assessing significance across the sample

We now consider whether the DFC estimates are significantly correlated with nuisance norms across the study sample. As demonstrated in the previous section, both positive and negative correlations can exist between the DFC estimates and nuisance norms. Taking this into account, we focused on the absolute value of the correlation values and computed the mean of the absolute correlation values across the sample. As described in the Methods section, we assessed the significance of the mean absolute value using an empirical null distribution. Figure 4.6a shows the empirical null distribution and the sample mean absolute value, which was found to be significant with $p < 10^{-6}$. Note that the sample mean absolute value indicated here is simply the mean of the histogram of the absolute correlation values shown in Figure 4.4c. In panels (b) and (c), we show the null distributions and sample values when considering complementary subsets of the sample where the global signal (GS) is either included or not included, respectively, with associated p-values of $p < 10^{-6}$ for both subsets. In Supplementary Figure 7.10, similar plots are provided for analyses with window lengths of approximately 40 and 100 seconds.

To complement our primary analysis, we also considered the extent to which the correlation coefficients were significant at the per-scan level. Using the scan-specific null distributions (see Methods), we found that 24% of the correlations between the nuisance norms and Pre DFC estimates were significant at the $p < 0.05$ level and 30% of the correlation values were significant at the $p < 0.10$ level. In supplementary material, Table 7.1 summarizes these findings and provides results for analysis window lengths of 40s and 100s. The results for a shorter window length of 40s were fairly similar to that of our primary results (approximate window length of 60s). However, we observed a slightly higher percentage of significance scans at the longer window length of 100s.

4.3.3 Strong correlations between DFC estimates and nuisance norms exist even when the underlying correlations between the nuisance and seed time courses are small

In this section, we ask the question: can strong correlations between DFC estimates and nuisance norms occur even when the nuisance time course are only weakly correlated with the seed time courses? Figure 4.7a shows the correlations between the DFC estimates and the nuisance norms plotted against the root-mean-squared (RMS) correlations between the underlying nuisance and seed time courses. Each point corresponds to a specific scan, nuisance term, and

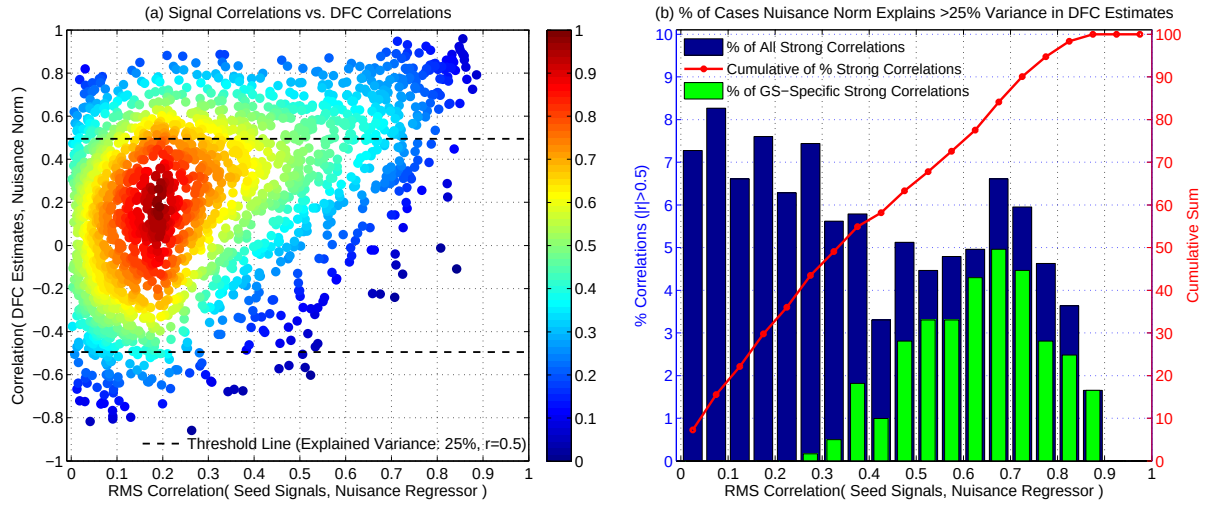


Figure 4.7: (a) Correlation between the DFC estimate and nuisance norm versus the RMS correlation between the raw nuisance signal and the two seed time courses. Correlation thresholds are indicated by the black dashed lines such that points that lie above and below these lines correspond to cases in which the DFC estimates are considered to be 'strongly correlated' ($|r| > 0.5$ with empirical $p \leq 0.09$) with the nuisance norms such that nuisance norms explain more than 25% of the total variance in the DFC estimates. The relative density of correlations (maximum density is normalized to 1.0) is indicated by the color map on the right-hand side. The relative density is computed by summing the total number of data points in smaller sub-grids and normalizing by the total number of points. (b) The blue bars show the percentage (%) of cases in which the nuisance norms explain more than 25% of the total variance in the DFC estimates as a function of the RMS correlation between the raw nuisance and the two seed time courses. The red dotted line (dots are located at bin centers) shows the cumulative sum of the percentage of strong correlations as a function of the RMS correlations. The green bars show the percentage (%) of strong correlations that are accounted for by the GS regressor.

seed region pair. For each point, the RMS correlation is computed by correlating the nuisance time course with each of the two seed time courses (from the same scan) and then calculating the RMS value (e.g. $r_{RMS} = \sqrt{(r^2_{\text{Seed1}} + r^2_{\text{Seed2}})/2}$, where r_{Seed1} and r_{Seed2} are the correlation coefficients between the nuisance time course and the first and second seed signals, respectively).

We first observe that strong correlations (points above and below the dashed lines in Figure 4.7a) between the DFC estimates and nuisance norms exist even when the RMS correlation between the nuisance and seed time courses is close to zero. We defined strong correlations between the DFC estimates and nuisance norm as those correlations for which $|r| > 0.5$, corresponding to an explained variance of greater than 25%. This threshold corresponded to an empirical $p = 0.09$ as assessed with the null distribution obtained by correlating the surrogate time series across all scans and realizations. Note that the threshold we have used merely serves as a convenient way to delineate strong correlations and its exact value is not critical. Conclusions similar to those stated here and below would be obtained even if we increased or decreased the threshold.

For further analysis, we binned the points according to the RMS correlation values using a bin width of 0.05. For each RMS correlation bin, we counted the number of points in which the nuisance norm explained more than 25% variance in the DFC estimates (i.e. number of scans above the threshold) and then divided this count by the total number of points across all bins. In Figure 4.7b, the blue bars indicate the percentage of strong correlations between the DFC estimates and nuisance norms as a function of the binned RMS correlation values. The green bars indicate the percentage of strong correlations associated with the GS regressor. The cumulative percentage of strong correlations versus RMS correlation value is shown by the red dotted line.

We found that a large percentage of the strong correlations between the DFC estimates and nuisance norms occurred for fairly small RMS correlation values between the raw nuisance time course and seed signals. For example, 22% of the all strong correlations between the DFC estimates and nuisance norms occurred for RMS values in the interval $r = 0$ to $r = 0.15$. Note

that in this interval, the nuisance terms account for less than 2.2% of the average variance of the seed time courses. Excluding the GS, 33% of the strong correlations between the DFC estimates and Non-GS nuisance norms occurred in the same interval. If we expand the interval $r = 0$ to $r = 0.25$, then 36% of the strong correlations lie within this expanded interval, corresponding to nuisance terms that account for less than 6.25% of the average variance of the seed time courses. Excluding the GS, 54% of the strong correlations lie within the same interval. In other words, for 36% of the all cases and 54% of the Non-GS cases examined, the nuisance norm explains more than 25% of the variance in the DFC estimates even though the raw nuisance time courses explain less than 6.25% of the average variance of the seed time courses.

Overall, we have the rather surprising observation that strong correlations between the DFC estimates and nuisance norms can exist even when the underlying nuisance terms account for only a small fraction of the variance of the seed time courses. We present a plausible explanation for this observation in Section 4.4.3 and examine its implications for nuisance regression later in the text.

Finally, the green bars in Figure 4.7b show the percentage of strong correlations between the DFC estimates and nuisance norm due to the GS regressor. This shows that the RMS correlation values between the GS and seed signals are relatively higher as compared to the Non-GS regressors. This is expected since the GS is computed as the average of all voxels time courses and is expected to correlate well with the seed time courses. On the other hand, the Non-GS regressors (difference between the blue and green bars) have fairly small RMS correlation values with the seed time courses even though their norms can be strongly correlated with the DFC estimates.

4.3.4 Regression does not eliminate the relation between the DFC estimates and nuisance norms

In this section, we show that full regression does not necessarily eliminate the relationship between the DFC estimates and the nuisance norms. Revisiting Figures 4.1 and 4.2, our focus now is on the solid red lines (Post DFC), which show the DFC estimates obtained after projecting out the relevant nuisance regressors. The fluctuations in the Pre DFC and Post DFC estimates in both figures are similar to each other and there is very little difference in their relationships to the nuisance norm. In the scans shown in both figures, the mean correlation value between the Post DFC estimates and the nuisance norms was $r = 0.73$, which is very close to the mean correlation value of $r = 0.76$ between the Pre DFC estimates and the nuisance norms. For these example scans, the Post DFC estimates were still significantly correlated with the nuisance norms after nuisance regression for the WM+CSF, GS and HM regressors ($p < 0.02$) and also for the RVf and HRf regressors ($p < 0.04$). Additional examples of Post DFC estimates are provided in Figures 4.3 and 4.5 and Supplementary Figure 7.8.

Viewed across the entire sample, the Post DFC correlation values shown in Figure 4.4b ranged from a minimum of $r = -0.88$ to a maximum $r = 0.86$ with a skewed distribution in which 60% of the correlations were positive and the remaining 40% were negative. The correlation distributions were similar to the Pre DFC histograms in Figure 4.4a,b with a cosine similarity values of $S = 0.77$ and $S = 0.87$ for the correlation and absolute correlation values, respectively.

Over the sample, the mean absolute correlation $|r|$ was found to be significant with $p < 10^{-6}$ (see Figure 4.6a). Moreover, we found that 14% of the correlations between the nuisance norms and Post DFC estimates were still significant at the $p < 0.05$ level and 22% of the correlation values were significant at the $p < 0.10$ level.

To further demonstrate the relationship between the Pre and Post DFC estimates, in Figure 4.8a we plotted the correlations between the Post DFC estimates and nuisance norms versus the

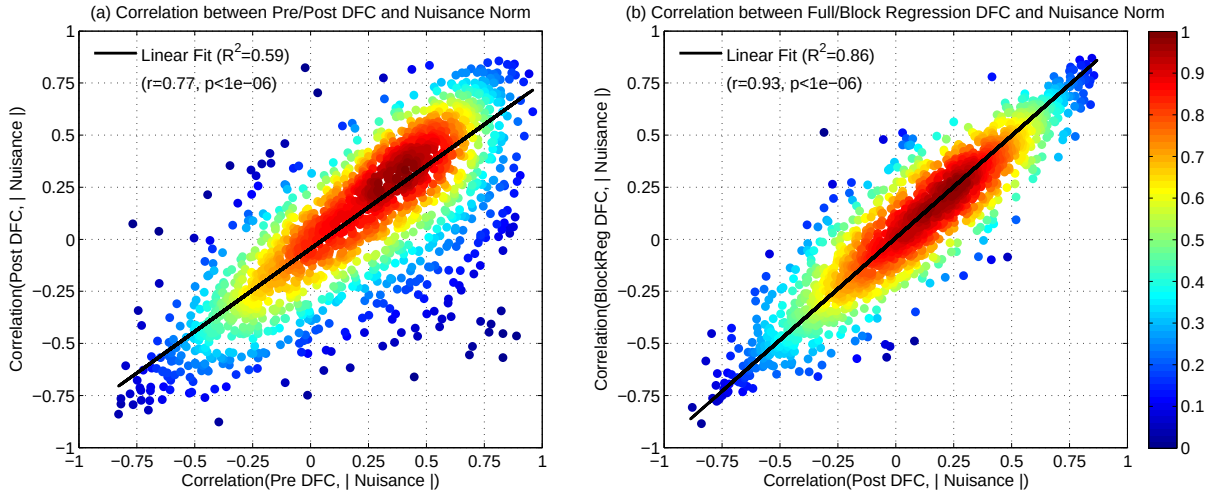


Figure 4.8: (a) The correlations between the Post DFC estimates and the nuisance norms (y-axis) versus the correlations between the Pre DFC estimates and the nuisance norms (x-axis). Post DFC estimates were still largely correlated with the nuisance norms and there was a strong linear relationship when compared to the correlations between the Pre DFC and nuisance norms (linear fit shown with black solid line, $R^2 = 0.59$, $r = 0.77$, $p < 10^{-6}$). The relative density of correlations (normalized to 1.0) is indicated by the color map on the right hand side. (b) The correlations between the DFC estimates and the nuisance norms after block regression (y-axis) and after full regression (x-axis). There was a strong linear relationship between the correlation populations (fit shown with black solid line, $R^2 = 0.86$, $r = 0.93$, $p < 10^{-6}$). The correlation distributions for full and block regression were not significantly different (paired two-tailed t-test $p = 0.86$). The effect size ($d = 0.0017$) and the absolute difference in correlation population means (0.0006) were negligibly small.

corresponding correlations between the Pre DFC estimates and the nuisance norms across all scans, nuisance time courses, and seed pairs. The relative density of correlation values is indicated with the color bar on the right hand side. There was a significant linear relationship between the two correlation populations ($r = 0.77$, $p < 10^{-6}$) indicating that the Post DFC estimates largely retain the correlation with the nuisance norms that is observed in the Pre DFC estimates.

4.3.5 Block regression is similarly ineffective in removing nuisance effects from the DFC estimates

As an alternative nuisance removal approach, we performed block regression in which nuisance measurements were projected out from each window separately. In Figure 4.9, we superimposed plots of the DFC estimates after performing block regression (green lines; referred to as Post BlockReg DFC) on the plots previously shown in Figure 4.1 for the BS002 dataset. The correlations between the Post BlockReg DFC estimates and nuisance norms ranged from $r = 0.63$ to $r = 0.86$ with a mean correlation of $r = 0.70$ and per-scan $p < 0.03$. For the CFMRI dataset,

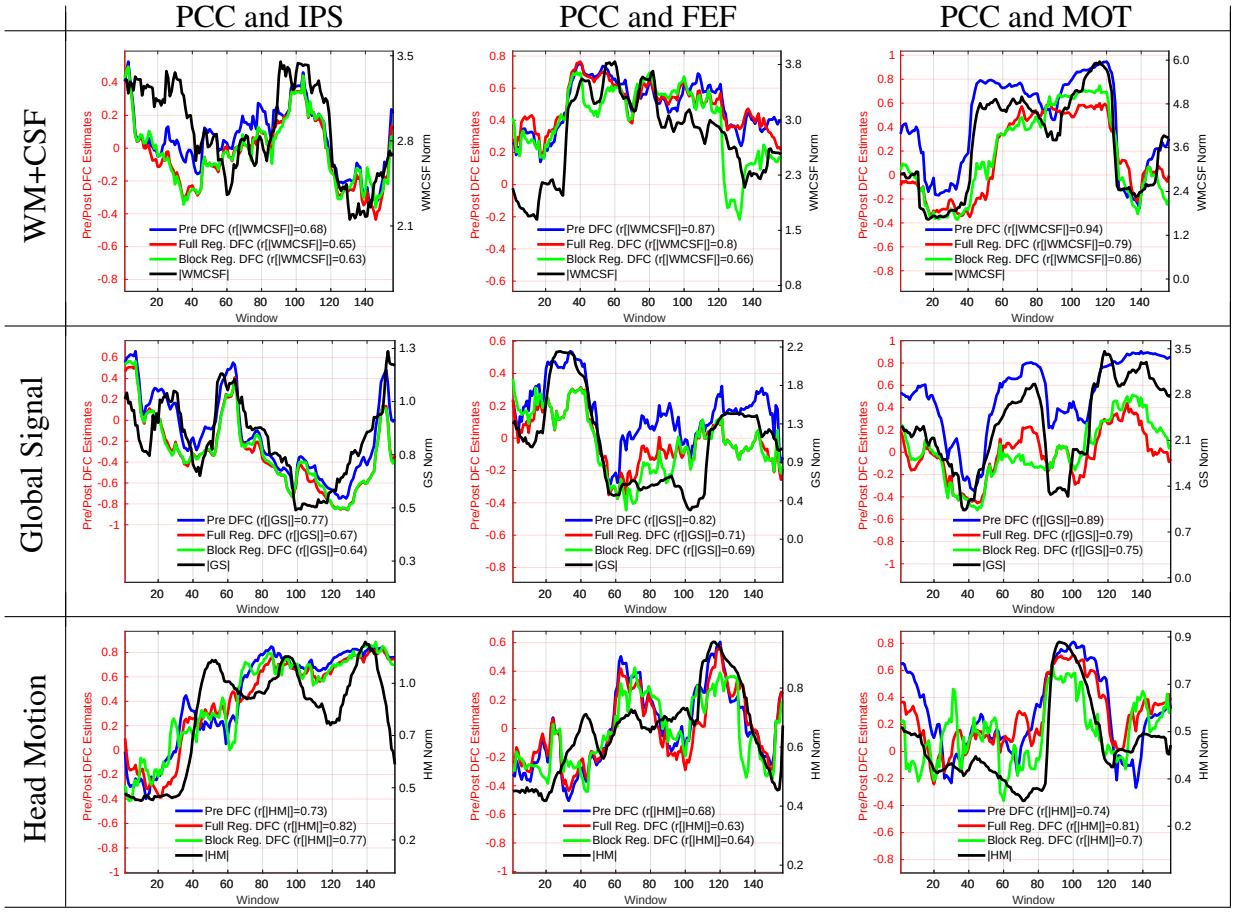


Figure 4.9: Same 9 representative scans from Figure 4.1 with the DFC estimates after block regression (solid green line) superimposed. The block regression DFC estimates were significantly correlated ($p < 0.03$) with the respective nuisance norms with a mean correlation value of $r = 0.70$.

the corresponding qualitative figures for block regression results are given in Supplementary Figure 7.11. Additional examples are provided in Figures 4.3 and 4.5 and Supplementary Figure 7.8.

Viewed across the entire sample, the Post BlockReg DFC correlation values shown in Figure 4.4b ranged from a minimum of $r = -0.89$ to a maximum $r = 0.87$. The correlation distributions were similar to the Post FullReg DFC histograms with cosine similarity values of $S = 0.93$ and $S = 0.95$ for the correlation and absolute correlation values, respectively.

Over the sample, the mean absolute correlation $|r|$ between the Post BlockReg DFC estimates and nuisance norms across all scans was significant with $p < 10^{-6}$ (see Figure 4.6a). We also found that 14% of the correlations between the nuisance norms and Post BlockReg DFC estimates were significant at the $p < 0.05$ level and 22% of the correlation values were significant at the $p < 0.10$ level.

Lastly, in Figure 4.8b, we plot the correlations between the Post BlockReg DFC estimates and the nuisance norms versus the correlations between the Post FullReg DFC estimates and the nuisance norms. A linear fit between the two correlation populations (shown with solid black line) revealed a significant linear relationship ($R^2 = 0.86$, $r = 0.93$, $p < 10^{-6}$). In addition, the correlation distributions for full and block regression were not significantly different from each other ($p = 0.86$, paired two-tailed t-test). The effect size ($d = 0.0017$) and the absolute difference in correlation population means (0.0006) were negligibly small. Thus, with respect to the relationship between the DFC estimates and the nuisance norms, block and full regression have nearly identical effects.

In the results presented so far, we have looked at the effects of nuisance regression when using different groups of regressors separately (e.g. WM and CSF, RVf and HRf, GS, and the 6 HM regressors). Note that multiple regression was used when there were more than one regressor in a group. In Supplementary Figures 7.12 and 7.13, we show that similar results are obtained when performing multiple regression with (a) WM and CSF grouped with the 6 HM regressors

and (b) GS grouped with WM, CSF, and the 6 HM regressors.

4.4 Interpretation

4.4.1 Nuisance effects on correlation estimates

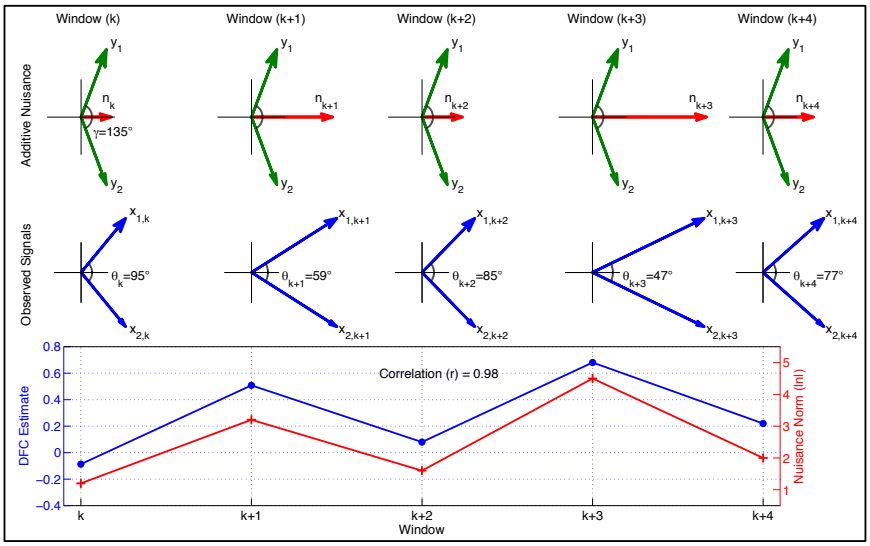
In Section 4.3.1, we showed that DFC estimates can be related to the norms of various nuisance terms. Here, we aim to provide an intuitive understanding of how this relationship might arise. We use simple toy examples to demonstrate the key principles and to establish concepts that will be further developed in the Theory section. In the toy examples, we represent time series as vectors in a low-dimensional (2D or 3D) space, such that the correlation between time series is simply the cosine of the angle between the vectors. For all of the examples, we will assume that there is a set of two underlying vectors with a fixed angle across time windows, corresponding to an idealized case in which the windowed correlation between two time series is fixed across time. Then, we examine what happens when a nuisance term is added to the underlying vectors. Note that due to considerations of simplicity and mathematical tractability, we restrict our presentation to the case of a single nuisance regressor in both this section and the following Theory section.

4.4.2 2D Examples

Positive correlations (Aligned Case)

We begin with a simple 2D toy example in which an additive nuisance term points in the same general direction as the underlying vectors. The underlying vectors are depicted with green vectors and denoted as y_1 and y_2 in Figure 4.10a for 5 consecutive windows (k through $k+4$). The angle between the vectors is fixed at an angle of 135°, such that the correlation is $r = -0.70$ for each window. We consider an additive nuisance term pointing along the horizontal axis with a norm that varies across windows (indicated by the red vectors and denoted as n_k through n_{k+4}).

(a) Aligned Case



(b) Anti-aligned Case

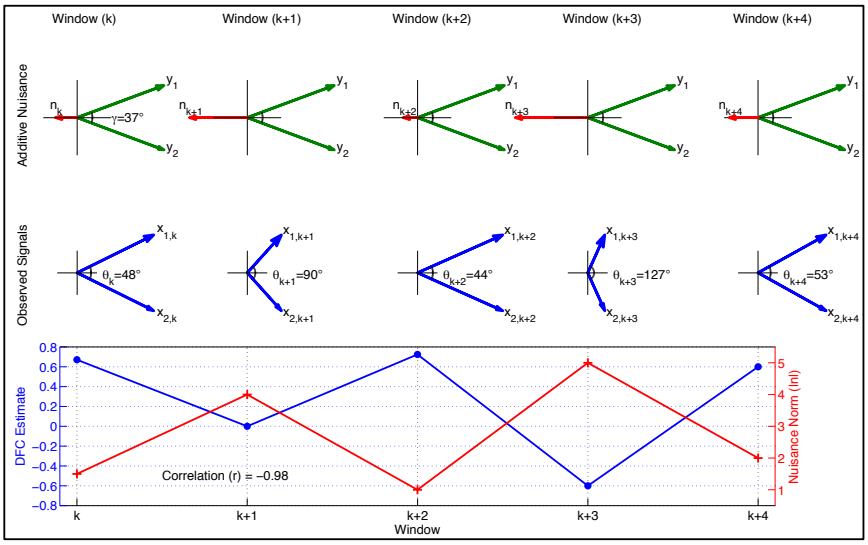


Figure 4.10: The relationship between the DFC estimates and the nuisance norms under an additive nuisance model. We present two cases indicated as the (a) aligned case and the (b) anti-aligned case. In the first row in (a), we simulate a set of fixed underlying vectors y_1 and y_2 across 5 windows (green arrows). The nuisance vectors n_k have different norms in each window (red arrows). The observed time courses $x_{1,k}$ and $x_{2,k}$ in the second row (shown with blue arrows) are the sum of the corresponding underlying vectors (y_1 and y_2) and the nuisance vector n_k. In the last row in (a), the simulated DFC estimates ($\cos\theta_k$, solid blue line) follows the norm of the nuisance term ($|n_k|$, solid red line). The aligned case shows that an increase in the norm of a nuisance vector that points in the direction of the underlying signals decreases the inner angle between the observed time courses and increases the value of DFC estimates. The anti-aligned case in (b) shows that an increase in the norm of a nuisance vector, that points in a direction opposite to that of the underlying signals, can decrease the DFC estimate and lead to anti-correlation between the nuisance norm and the DFC estimate.

In this example, the nuisance vectors n_k are aligned with the underlying vectors y_1 and y_2, meaning that the inner products ($y_1^T n_k > 0$ and $y_2^T n_k > 0$) are positive. The sum of the underlying vectors and additive nuisance terms yields the observed time series as shown by the blue vectors in the second row of Figure 4.10a and denoted as $x_{1,k}$ and $x_{2,k}$. The angle between the observed vectors (shown with θ_k) varies with the nuisance norm, with smaller angles observed for larger nuisance norms. In the last row of Figure 4.10, we plot the correlation between the observed signals (cosine of the angle between observed signal vectors) and the nuisance norm across windows. There is a strong relationship ($r = 0.98$) between the window correlation values and the nuisance norm. This simple example shows how variations in the norm of a nuisance term may induce variations in the windowed correlation estimates that are highly correlated with the norm of the nuisance term.

Anti-correlations (Anti-aligned Case)

As noted in Section 4.3.1, the DFC estimates can sometimes be anti-correlated with the nuisance norms. To see how this might arise, we consider the 'anti-aligned' case shown in Figure 4.10b. In contrast to the aligned case discussed above, the nuisance vector points in a direction opposite to the average direction of the underlying signals, such that the corresponding dot products ($y_1^T n_k < 0$ and $y_2^T n_k < 0$) are negative. In this example, the correlation between y_1 and y_2 is fixed at a value of $r = \cos 37^\circ = 0.80$.

The addition of the anti-aligned nuisance vector leads to cancellation of vector components along the horizontal axis. As a result, the angle between the observed vectors $x_{1,k}$ and $x_{2,k}$ in the second row tends to increase as the nuisance norm increases. Taking the cosine of the angle between the observed vectors to obtain the correlation estimates, we find that the correlation values are anti-correlated ($r = -0.98$) with the nuisance norm, as shown in the third row of Figure 4.10b. This simple example shows that an additive nuisance term can induce variations in the windowed correlation estimates that are anti-correlated with the nuisance norm.

We should note that if the nuisance norm becomes extremely large for the anti-aligned case, the observed vectors can be dominated by the nuisance term in a manner that can cause the nuisance norm to be positively correlated with the resulting DFC estimates (see the first three windows n_k through n_{k+2} in Supplementary Figure 7.14 for an illustration). It is possible that this mechanism may contribute to the positive skew observed in the correlation histograms in Figure 4.4a-b. However, in the absence of ground truth it is not currently possible to make any definitive conclusions in this regard. In general, the extent to which a nuisance term induces positive or negative correlations between the correlation estimates and the nuisance norm depends on both the angle between the underlying vectors and the relative direction and magnitude of the nuisance vector.

Finally, in Supplementary Figure 7.14 we present a last 2D example in which the relationship between the DFC estimates and nuisance norms exhibits both positive and negative correlations such that the overall observed correlation is $r = 0$. The nuisance norm is extremely large in the first three windows (k through $k+2$) and dominates over the underlying vectors such that the nuisance norm is positively correlated with the resulting DFC estimates. The nuisance norm then becomes much smaller in the last two windows ($k+3$ and $k+4$) so that the nuisance term is no longer able to dominate over the underlying vectors and the nuisance norm is anti-correlated with the resulting DFC estimates. Because of this varying relationship between the DFC estimates and nuisance norms, the observed correlation across all windows is equal to $r = 0$ even though there is a clear effect of the nuisance term on the observed vectors. In general, the relationship between DFC estimates and nuisance norms will exhibit a highly variable and complex behavior across different scans, such that there are many other possible scenarios (in addition to the simple one presented here) that could be used to explain the observation of small correlations between DFC estimates and nuisance norms.

4.4.3 Extension to 3D with addition of an orthogonal nuisance component

In the 2D examples discussed so far, the nuisance term was completely within the 2D plane spanned by the observed vectors. We now expand the example to 3D by including an additional nuisance component that is orthogonal to the 2D plane, such that the overall nuisance term is the sum $n = n_I + n_O$ of an in-plane component n_I that lies in the subspace spanned by the observed vectors x_1 and x_2 and an orthogonal component n_O that is orthogonal to the subspace.

In Figure 4.11, we have constructed the observed signals (blue vectors in the first row) using the fixed vectors and in-plane nuisance components previously used in the 2D example of Figure 4.10a. We then modify the nuisance terms by adding orthogonal components, such that the overall nuisance terms (red vectors) have both in-plane and orthogonal components. We construct the orthogonal component in each window such that its norm is proportional to the norm of the in-plane component (i.e. $|n_{O,k}| \propto |n_{I,k}|$). Here, it is useful to define a metric $|n_O|^2/(|n_I|^2 + |n_O|^2) = |n_O|^2/|n|^2$ that reflects the relative fraction of nuisance energy that lies in the orthogonal component. The orthogonal nuisance fraction is equal to 1.0 when the nuisance component is completely orthogonal to the observed signal subspace and is equal to 0.0 when the nuisance component lies within the subspace. For this example, the orthogonal nuisance fraction (indicated by the black dots in the fourth row) is relatively large and fairly constant with a mean value of 0.94.

The resulting 3D example has several important properties. First, the Pre DFC estimates (blue dots in the third row) exhibit a strong correlation ($r = 0.98$) with the nuisance norm (red dots), similar to that previously shown for the 2D example in Figure 4.10a. This is because (1) the observed signals are the same in the 2D and 3D examples and (2) by construction the nuisance norm of the 3D nuisance term scales with the norm of the in-plane component.

Second, because the orthogonal nuisance fraction is relatively large, the underlying correlations between the nuisance terms and the observed vectors (i.e. cosine of the angle between the red and blue vectors) are relatively small as shown by the red and blue curves in the fourth

row. Thus, this example is consistent with the empirical findings discussed in Section 4.3.3: a strong correlation between the Pre DFC estimates and the nuisance norms can exist even when the correlation between the nuisance and seed signals is small and the orthogonal fraction is high.

Third, linear regression has a minimal effect in this example, such that the signal vectors

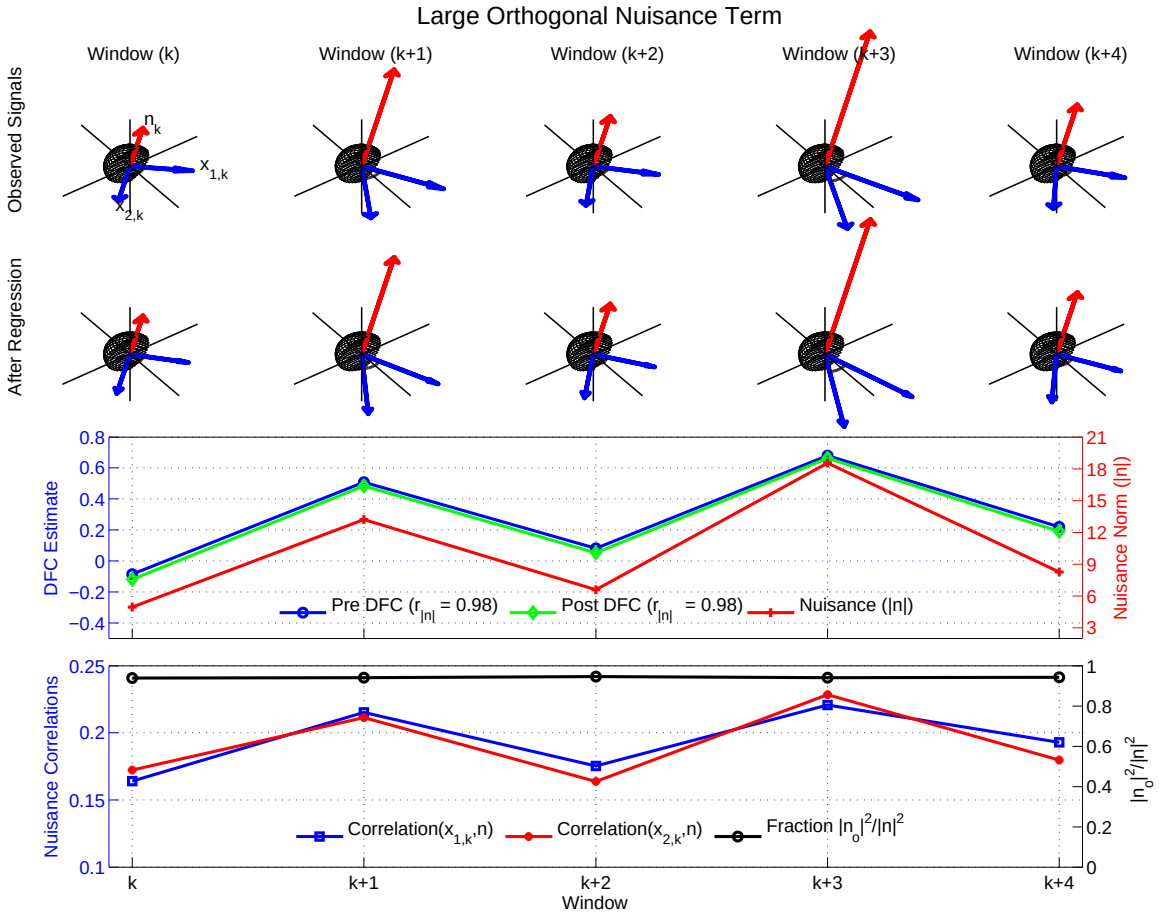


Figure 4.11: Linear regression can fail to eliminate the relationship between the nuisance norm and the DFC estimates due to a large degree of orthogonality between the nuisance measurement and fMRI time courses. The observed signals (blue vectors) were obtained using nuisance terms (red vectors) for which the in-plane nuisance components are the same as those previously used in Figure 4.10a. The Pre DFC estimate (blue line) is highly correlated ($r = 0.98$) with the nuisance norm (red line) . The nuisance term also has a large orthogonal nuisance fraction (indicated by the black dots in the fourth row) with a mean value of 0.94. Therefore, linear regression has a minimal effect and Post DFC estimates (green line) are also highly correlated ($r = 0.98$) with the nuisance norm. The correlations between the nuisance term and the observed vectors (red and blue lines in bottom row) are relatively weak. Note that for display purposes, the orthogonal axes (vertical) in the top 2 rows are compressed ($\times 0.7$) with respect to the in-plane axes.

after regression (blue vectors in the second row) are similar to the original observed signals (blue vectors in the first row). Note that to simplify the presentation, we have used block regression to obtain the Post DFC estimates. The resultant Post DFC estimates (shown with solid green line and diamond markers in the third row) are also highly correlated ($r = 0.98$) with the nuisance norm. This is consistent with the empirical findings from Sections 4.3.4 and 4.3.5 that nuisance regression can have a minimal effect on the relation between the DFC estimates and the nuisance norms.

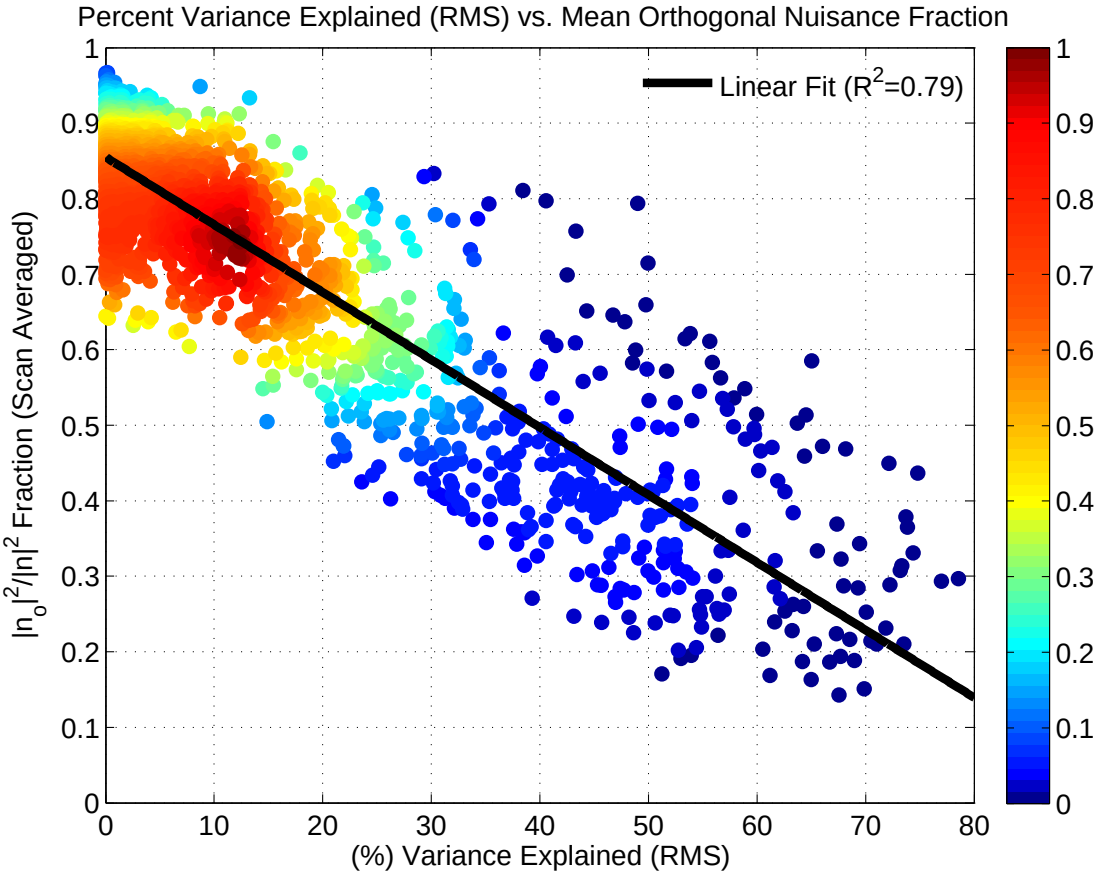


Figure 4.12: The scan-averaged orthogonal nuisance fraction $|n_O|^2/|n|^2$ versus the percent variance explained by the raw nuisance regressors (squared RMS correlation values from Figure 4.7). As shown by the black least squares line, the orthogonal fraction decreases in a linear fashion ($R^2 = 0.79$) with increasing percent variance. The relative density of values (maximum density is normalized to 1.0) is indicated by the color map on the right-hand side.

To aid in relating the orthogonal nuisance fraction to the experimental results, Figure 4.12 plots the orthogonal fraction values (averaged over each scan) versus the average percent variance in the seed time courses that is explained by the nuisance regressors (computed as the square of the RMS correlation values previously shown in Figure 4.7). As the average percent variance explained increases, the fraction of the nuisance term energy that is orthogonal to the plane spanned by the observed time series decreases. As noted in Section 4.3.3, 36% of the strong correlations between the DFC estimates and nuisance norms occurred when the percent variance explained was less than 6.25%. Using the linear fit shown in Figure 4.12, this percentage corresponds to orthogonal nuisance fractions greater than 0.80, consistent with our use of a large orthogonal fraction in the example.

4.4.4 Regression effects depend on the orthogonal nuisance fraction

In this section, we present a 3D example to demonstrate how the effects of regression depend on the orthogonal nuisance fraction. In each of the 5 cases shown in Figure 4.13 the nuisance vector n has different orientations with respect to the 2D subspace spanned by the observed signals x_1 and x_2. The orthogonal fraction starts at 1.0 for Case 1 and then decreases to 0.0 for Case 5, with intermediate values for the other cases. The correlation between the observed time courses is fixed at $r = 0.78$ for all of the cases. In the middle column, we show the vectors (denoted as $\tilde{x}_1$ and $\tilde{x}_2$) after nuisance regression.

When the orthogonal fraction is 1.0 (Case 1), the nuisance vector is completely orthogonal to the space spanned by the observed signals, and therefore the regression coefficient between the nuisance vector and each observed signal is zero. As a result, linear regression has no effect, and the post regression signals are identical to the observed signals. Consequently, for Case 1 the post regression correlation value is equal to the pre regression correlation value, as shown in the rightmost plot in Figure 4.13 as the intersection between the red dot and the horizontal green line representing $r = 0.78$.

When the orthogonal fraction is 0.0 (Case 5), the nuisance vector lies completely in the plane spanned by the observed time courses. After regression, the vectors $\tilde{x}_1$ and $\tilde{x}_2$ must point in opposite directions in order to achieve orthogonality to the in-plane nuisance term. As a result, the correlation between the post regression time courses will be -1. (Note that if the nuisance term n is in-plane but lies outside of the inner angle formed by x_1 and x_2 then regression will force $\tilde{x}_1$ and $\tilde{x}_2$ to point in the same direction and the correlation will be forced to be $+1$. In the

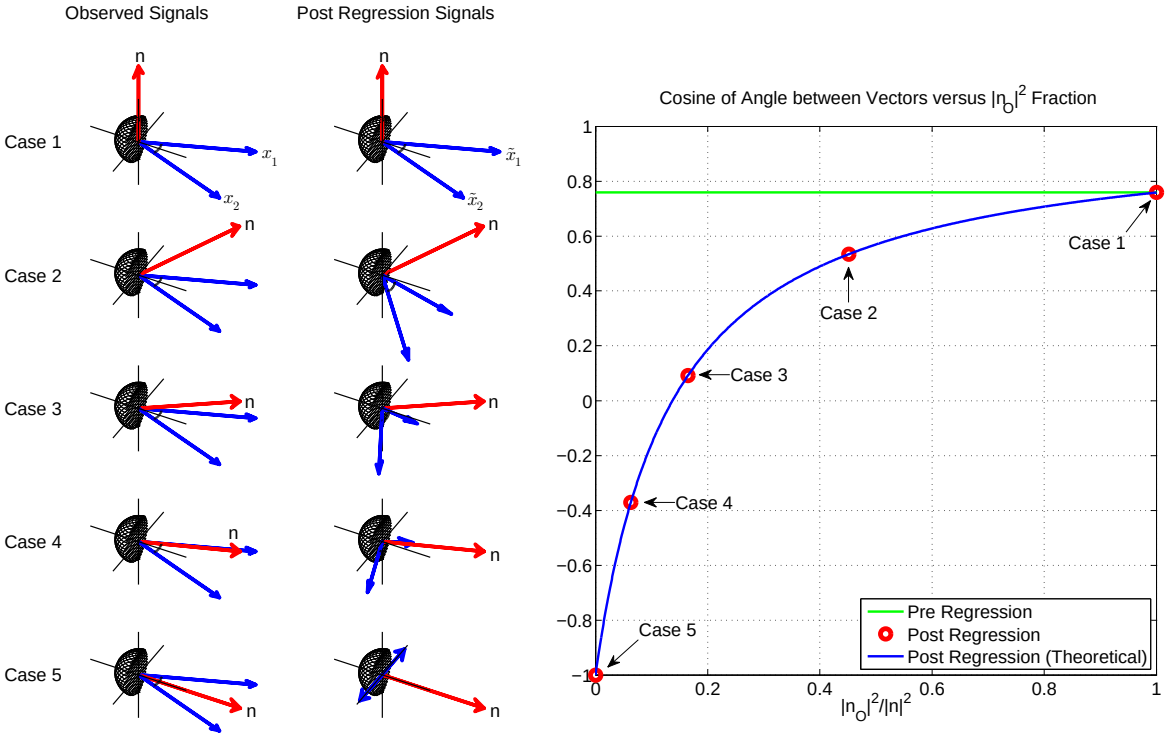


Figure 4.13: The effect of linear regression on cosine of angle between two vectors. We illustrate 5 cases for the same observed fMRI time courses with a fixed correlation value ($r = 0.78$; as shown with the green solid line) and a variable nuisance component. In Case 1, the nuisance vector is orthogonal to the observed vectors and the orthogonal nuisance fraction ($|n_O|^2/|n|^2$) is 1.0, in which case regression has no effect on the pre regression correlation values. Thus, the post regression correlation value is also $r = 0.78$ as shown with the red marker. In Case 5, nuisance vector lies within the observed vectors with $|n_O|^2/|n|^2 = 0$ and post regression correlation value is -1.0 regardless of the pre regression correlation value. For Cases 2, 3, and 4, the fraction $|n_O|^2/|n|^2$ takes on intermediate values, such that the difference between the pre and post values increases as the orthogonal fraction decreases. The blue line represents the theoretical values obtained from Eq. (4.2).

Appendix, we refer to this as the complementary case).

For the remaining cases, the post regression correlation values vary as a function of the orthogonal nuisance fraction as shown by the red circles on the right hand side of Figure 4.13. The green line in the plot indicates the pre-regression signal correlation values ($r = 0.78$), which are independent of the fraction. Cases with a larger orthogonal nuisance fraction have post regression correlation values that tend towards the pre regression correlation value of $r = 0.78$, whereas cases with a smaller orthogonal nuisance fraction have post regression correlation values that tend towards -1. The blue line in the plot shows the theoretical relation between the post regression correlation values and the orthogonal nuisance fraction. This relation is discussed in greater detail in the Theory section.

Overall, we see that when the orthogonal nuisance fraction is relatively high (e.g. greater than 0.5), the difference between the pre and post regression DFC estimates (distance between blue and green lines) will be relatively small. The exact bound on this difference (over all possible pre-regression correlation values) is provided in the Theory section, where it is also shown that most nuisance regressors lie within the high orthogonal fraction regime and therefore exhibit a small difference between the pre and post regression DFC estimates. The primary exception is the global signal, which provides the motivation for our next example.

4.4.5 DFC estimates after regression with smaller orthogonal nuisance term

In this subsection, we consider a second example in which the orthogonal nuisance fraction is relatively modest with a mean value of 0.47 across windows, as shown with the black line with circles in the fourth row of Figure 4.14. As discussed further in the Theory section, this value for the orthogonal nuisance fraction is consistent with what is observed for the global signal nuisance term. In this case, linear regression has a noticeable effect as can be seen by the difference between the original signal vectors (blue vectors in the top row) and the post regression

vectors (blue vectors in the second row). Regression moves the post regression correlation values away from the pre regression correlation value towards -1.0, similar to Cases 2 and 3 in Figure 4.13. However, since the orthogonal nuisance fraction is relatively constant across windows, the difference between the Pre and Post DFC estimates is also fairly constant. As shown in the third row in Figure 4.14 the Post DFC estimates can be approximated as a shifted version of the Pre

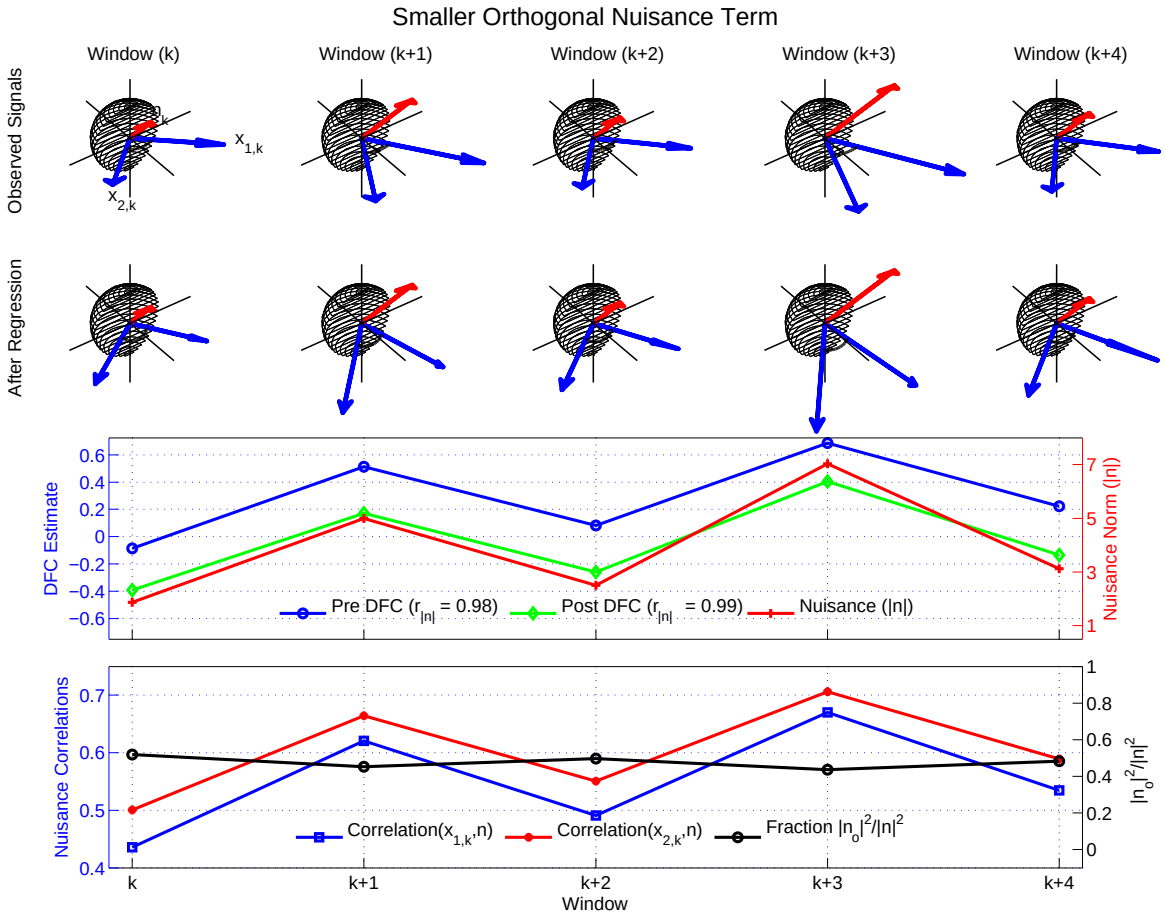


Figure 4.14: Toy example illustrating how linear regression can fail to eliminate the relationship between the nuisance norms and the DFC estimates when the nuisance orthogonality is weak. Similar to the first toy example in Figure 4.11, the in-plane nuisance components are the same as those previously used in Figure 4.10a, however, the orthogonal part of the nuisance term in this example is much smaller as compared to Figure 4.11. Therefore, linear regression has a noticeable effect on the DFC estimates. Since the orthogonal nuisance fraction is fairly constant across windows, the difference between the Pre and Post DFC estimates is also fairly constant. As the correlation coefficient is invariant to constant offsets, the correlation between the nuisance norm and Pre DFC estimate ($r = 0.98$) is essentially the same as the correlation between the nuisance norm and the Post DFC estimate ($r = 0.99$).

DFC estimates. Because correlation is invariant to constant offsets, the correlation ($r = 0.99$) between the Post DFC estimates with the nuisance norm is essentially the same as the correlation ($r = 0.98$) between the Pre DFC estimates and the nuisance norm. The construction of the toy example is consistent with our overall observation of a fairly constant difference between the Pre and Post DFC estimates obtained after GS regression. A real example scan with slightly smaller but confined orthogonal nuisance fraction is shown in the second row of Figure 4.15 for GS regression. In Section 4.5.3 we provide additional empirical and theoretical results regarding this effect.

The correlations between the observed signals and the nuisance terms are shown with the blue square and red dotted lines in the fourth row of Figure 4.14, with values ranging from $r = 0.44$ to $r = 0.70$. Consistent with the smaller orthogonal fraction in this example, these correlations are higher than those observed in the example shown in Figure 4.11. The range of correlation values used in this example is consistent with the range of RMS correlations that is empirically observed for scans, in which the DFC estimates and GS norm are significantly correlated. Specifically, as shown by the green bars in Figure 4.7b, the empirical correlations range from $r = 0.25$ to $r = 0.89$ with a mean of $r = 0.59$.

In concluding the Interpretation section, it is important to note that the toy examples considered here are designed to provide a basic level of intuition that can be helpful for understanding both the empirical findings in the Results section and the theoretical expressions which will be presented in the Theory section. While the examples shown here demonstrate behavior similar to that observed in the experimental data, they are by no means exhaustive and alternative examples might be useful to consider in future work.

4.5 Theory

4.5.1 The DFC estimate after block regression

In this section, we provide the expression for the DFC estimate after block regression. The detailed steps of this derivation can be found in 4.7.1. Dropping the window index subscript k for simplicity, the correlation coefficient after block regression for a single window is given by:

$$\tilde{r} = \frac{\tilde{x}_1^T \tilde{x}_2}{|\tilde{x}_1||\tilde{x}_2|} = \frac{x_1^T x_2 - \frac{|n_I|^2}{|n|^2} x_1^T P_{n_I} x_2}{|x_1||x_2| \sqrt{1 - \frac{|n_I|^2}{|n|^2} \frac{x_1^T P_{n_I} x_1}{|x_1|^2}} \sqrt{1 - \frac{|n_I|^2}{|n|^2} \frac{x_2^T P_{n_I} x_2}{|x_2|^2}}} \tag{4.1}$$

where x_1 and x_2 are the time course column vectors prior to regression, $\tilde{x}_1$ and $\tilde{x}_2$ are the time course column vectors after regression, n and n_I are the nuisance term column vector and the in-plane component (also a column vector), respectively, and P_{n_I} denotes the projection matrix onto the in-plane nuisance component.

This can rewritten as

$$\tilde{r} = \lambda r + O \tag{4.2}$$

where $r = x_1^T x_2 / (|x_1||x_2|)$ is the correlation coefficient before regression and the scaling λ and offset O terms are defined as

$$\lambda = \frac{1}{\sqrt{1 - \frac{|n_I|^2}{|n|^2} \frac{x_1^T P_{n_I}^T x_1}{|x_1|^2}} \sqrt{1 - \frac{|n_I|^2}{|n|^2} \frac{x_2^T P_{n_I} x_2}{|x_2|^2}}} \tag{4.3}$$

$$O = -\lambda \frac{|n_I|^2}{|n|^2} x_1^T P_{n_I} x_2 / (|x_1||x_2|) . \tag{4.4}$$

As the orthogonal component n_O becomes arbitrarily large compared to the in-plane component

n_I, then the orthogonal fraction $\frac{|n_O|^2}{|n|^2} \to 1$ and the term $\frac{|n_I|^2}{|n|^2} = \left(1 - \frac{|n_O|^2}{|n|^2}\right) \to 0$. In this case, the scaling term $\lambda \to 1$ and the offset term $O \to 0$, such that the correlation coefficient after regression approaches the pre-regression value $\tilde{r} \to r$. This corresponds to Case 1 in Figure 4.13.

On the other hand, when the orthogonal component becomes arbitrarily small and the orthogonal nuisance fraction $\frac{|n_O|^2}{|n|^2} \to 0$, it can be shown that $\tilde{r}$ approaches either -1.0 or 1.0 (see Eq. (4.29) in the Appendix). When the terminal value is -1.0, this corresponds to Case 5 in Figure 4.13. For intermediate values $0 < \frac{|n_O|^2}{|n|^2} < 1$ of the orthogonal nuisance fraction, the correlation coefficient $\tilde{r}$ takes on values between -1 and r, corresponding to Cases 2 through 4.

4.5.2 A mathematical bound on the change in DFC using block regression

Here, we consider the difference $\Delta\text{DFC} = \tilde{r} - r$ between the correlation coefficients obtained before and after block regression. In 4.7.2, we show that this quantity is bounded as follows:

$$|\Delta\text{DFC}| \leq 2\left(\frac{1 - \sqrt{|n_O|^2/|n|^2}}{1 + \sqrt{|n_O|^2/|n|^2}}\right) \tag{4.5}$$

As an example of this bound, in the second column of Figure 4.15 we plot ΔDFC versus the orthogonal nuisance fraction $\frac{|n_O|^2}{|n|^2}$ for WM regression (first row) and GS regression (second row) applied separately to two representative scans. Consistent with the discussion in the previous sections, when the orthogonal nuisance fraction $\frac{|n_O|^2}{|n|^2}$ is close to 1.0, the theoretical bound (black dashed line) approaches 0.0. Consequently, the post regression correlation coefficients for the WM signal in the first column of Figure 4.15 are constrained to be close to the pre regression coefficients.

On the other hand, as the orthogonal nuisance fraction $\frac{|n_O|^2}{|n|^2}$ approaches zero, the theoretical bound relaxes and approaches ± 2. This corresponds to the case where the post regression correlation coefficient approaches either -1.0 or 1.0. Since the pre regression correlation coef-

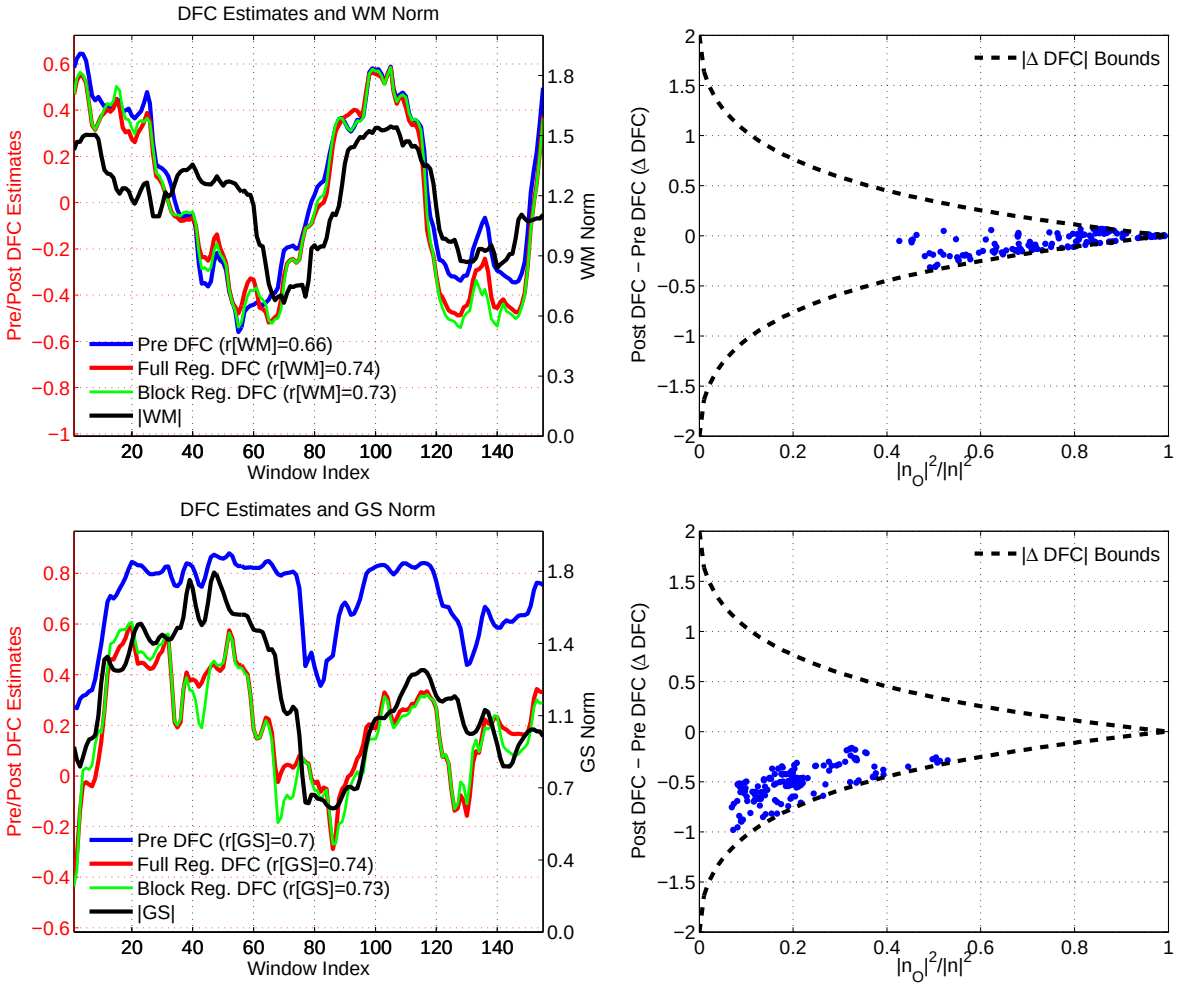


Figure 4.15: Two representative scans that demonstrate the theoretical bounds on ΔDFC. In the first column, we show the DFC estimates obtained between the PCC and IPS seeds. In these scans the DFC estimates show a large degree of correlation with the norm of the WM time course ($r = 0.66$) in the first row, and with the norm of the GS time course ($r = 0.7$) in the second row. Performing full or block regression does not reduce these correlations and the correlations are larger than 0.7 after each regression technique. In the second column, we show the ΔDFC values versus the orthogonal nuisance fraction for the WM and GS regressors. Each point in these plots corresponds to the orthogonal nuisance fraction and ΔDFC values in a specific window. We also superimpose the theoretical bounds. For the WM nuisance term, the effects of regression are limited by the tight bound imposed by the large orthogonal nuisance fractions. For the GS nuisance term, the bounds are more relaxed due to the smaller orthogonal nuisance fraction. However, the ΔDFC values are clustered around -0.5 and the difference between the Pre DFC and Post DFC values in the first column is also fairly constant. Thus, the Post DFC values remain correlated with the nuisance norms after GS regression.

ficient is bounded between -1.0 and 1.0, the maximum absolute difference in coefficients is 2, consistent with the theoretical bound on $|\Delta\text{DFC}|$. The post regression correlation coefficients for the GS signal in the first column of Figure 4.15 have a noticeable negative offset when compared to the pre regression coefficients.

To demonstrate the general validity of the bound, in Figure 4.16, we plot ΔDFC versus the fraction $\frac{|n_O|^2}{|n|^2}$ for six different nuisance regressors using the data from all scans and different seed pairs. For head motion (HM), the nuisance regressor was defined as the first principal component of the 6 motion regressors. All of the empirical DFC differences are found to lie within the theoretical bounds. For WM, CSF, HM, RVf, and HRf regressors, the mean of the orthogonal fractions ranged between 0.78 and 0.83, reflecting the fact that for most of the data windows there is a large orthogonal fraction and a fairly tight bound on ΔDFC such that mean difference between the pre and post regression values is small, ranging from -0.049 to -0.0065. In contrast, for the GS regressor, the mean orthogonal fraction is 0.47, and so the data for most of the windows lie in a range where the bounds on ΔDFC are not as tight and the mean difference between pre and post regression values is -0.26. The smaller orthogonal nuisance fraction observed for GS reflects the fact that it is derived as the mean of all the voxel time courses in the brain and therefore is expected to exhibit a greater similarity (and hence a greater in-plane component) with the seed voxel time courses, as compared to the other regressors.

4.5.3 Approximate Constant Offset observed in ΔDFC for GS regression

In both the previous section and Section 4.4.5, we noted that GS regression resulted in a relatively constant difference ΔDFC between the Pre and Post DFC values. In addition, we can see from the upper left-hand panel in Figure 4.16 that the per-window ΔDFC values with GS regression are clustered close to the lower bound.

To further investigate this effect, we computed the mean and standard deviation of the ΔDFC values for each scan. We also computed the mean lower theoretical bound for each scan

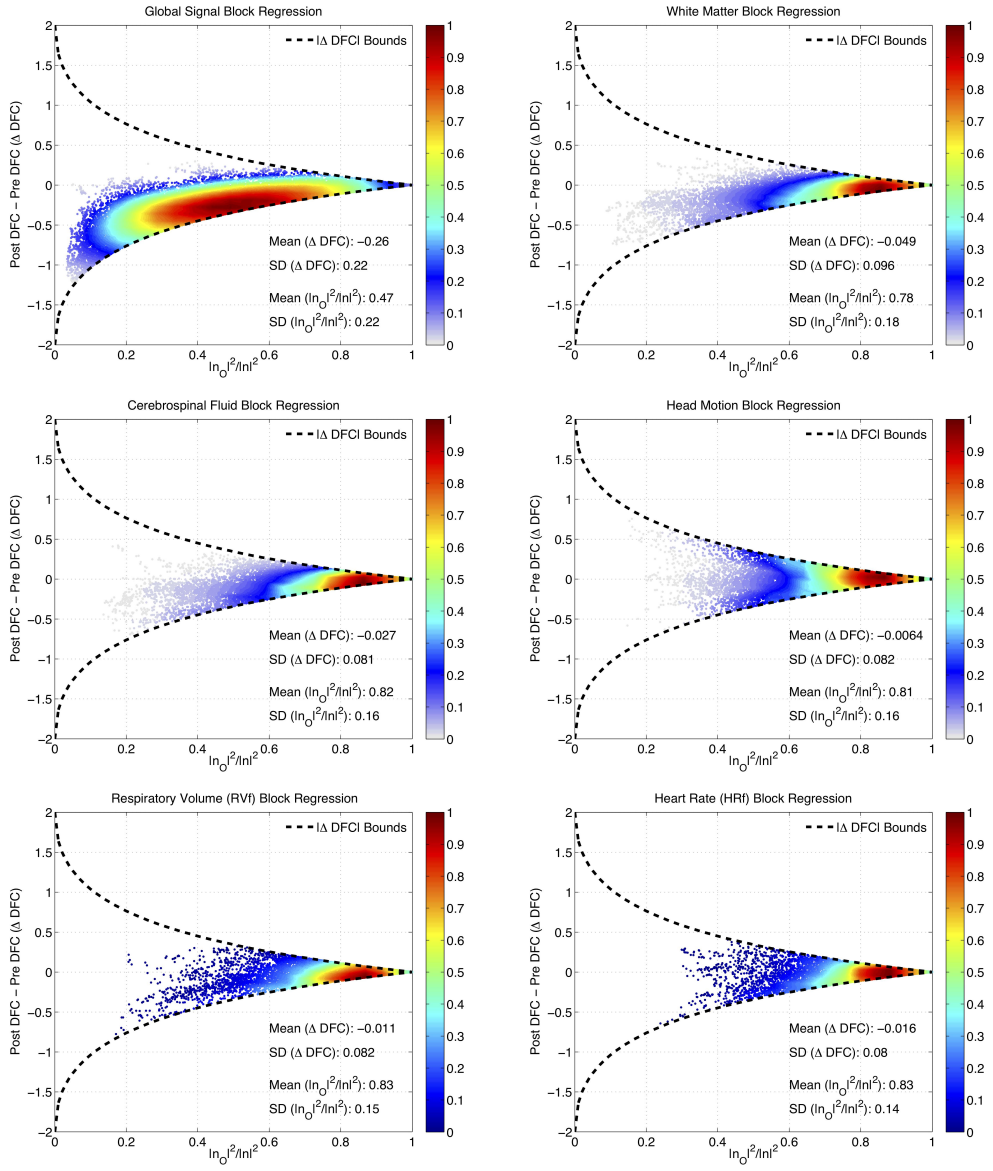


Figure 4.16: ΔDFC versus the fraction $\frac{|n_O|^2}{|n|^2}$ for six different nuisance regressors (from left to right and top to bottom: GS, WM, CSF, HM, RVf, and HRf) using the data from all scans. Each point in a given plot corresponds to a single window. The theoretical bounds for block regression are shown with the dashed black lines. The empirical ΔDFC values were found to lie within the theoretical bounds. For WM, CSF, HM, RVf and HRf regressors, the mean of the orthogonal fractions ranged between 0.78 and 0.83. In this region, the bounds are fairly narrow such that the mean differences between the pre and post regression values were small, ranging from -0.049 to -0.0065. For GS, the mean orthogonal fraction was 0.47 and the bounds on ΔDFC were not as tight. However, the the ΔDFC values were clustered about a mean value of -0.26 and closely followed the lower bound. Note that the relative density of data points (maximum density is normalized to 1.0) is indicated by the color map on the right-hand side of each plot.

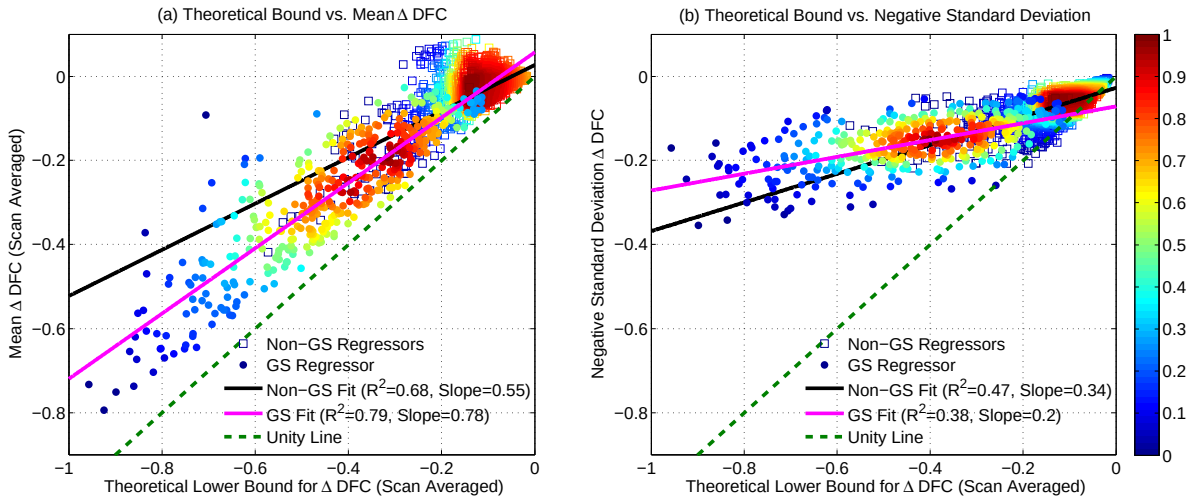


Figure 4.17: (a) Mean ΔDFC versus scan-averaged theoretical bound. Data for GS and non-GS regressors are indicated by the circle and square markers, respectively. For GS regression, the linear fit (magenta line, $R^2 = 0.79$, Slope $= 0.78$) is close to the line of unity (dashed green line). For non-GS regressors, the linear fit is indicated by the black line and most values occur at small ΔDFC values. (b) Negative Standard Deviation (NSD) of ΔDFC (over each scan) versus the mean theoretical bound. For GS regression, both the slope of the best fit line (magenta line, $R^2 = 0.38$, Slope $= 0.2$) and the magnitude of the NSD values are considerably lower than the slope and magnitudes observed for the mean ΔDFC values. The relative density of data points (maximum density is normalized to 1.0) is computed for GS and non-GS data separately and indicated by the color map on the right-hand side.

by computing the lower bound for each window and then averaging over windows. Figure 4.17a plots the mean ΔDFC values versus the mean theoretical bounds, with the data for GS and non-GS regressors indicated by the circle and square markers, respectively. The line of unity is indicated by the dashed green line. Due to the additive property of inequality, the mean bound is strict so that all points must lie above this line. A linear fit to the GS data is shown with the magenta line ($R^2 = 0.79$, Slope $= 0.78$), which is fairly close to the line of unity. This indicates that the mean ΔDFC values after GS regression approximately follow the mean theoretical bound over a large range of ΔDFC values. In contrast, the mean ΔDFC values shown for the non-GS regressors are relatively small, consistent with the large orthogonal nuisance fractions observed for these regressors in Figure 4.16.

In Figure 4.17b, we plot the negative standard deviation (NSD) of ΔDFC for each scan

versus the average bound, where the standard deviation is negated for display purposes. Note that in contrast to the mean ΔDFC values, there is no requirement that these points lie above the line of unity (dashed green line). The slope of the linear fit for GS (Slope = 0.2, magenta line) is much smaller than the slope of 0.79 observed for mean ΔDFC in Figure 4.17a. In addition, the magnitudes of the SD values are significantly smaller than the magnitudes of the mean ΔDFC values ($p < 10^{-6}$, paired t-test). Thus, the primary effect of GS regression is to induce a negative offset in the mean of the DFC estimates accompanied by a relatively smaller change in the DFC fluctuations about the mean. As a result, the Post DFC estimates after GS regression can be approximated to first order as a shifted version of the Pre DFC estimates and will therefore largely retain the correlation with the nuisance norm. An example of this limitation was previously shown in the GSR example in the second row in Figure 4.15.

4.5.4 DFC estimates after full regression

In the Results section we have shown that effects of block regression and full regression on the DFC estimates were very similar. Here, we present the expression for the DFC estimate after full regression as a modified version of the expression obtained for block regression. The main difference between the two approaches is that the regression fit coefficients for full regression are computed from the entire time series data, whereas the fit coefficients for block regression are computed using only the data in the window of interest. However, it important to note that both approaches subtract out a scaled version of the nuisance term in the window of interest.

As derived in 4.7.3, the correlation coefficient after full regression for the kth window can

be written as:

$$\frac{\hat{x}_{1,k}^T \hat{x}_{2,k}}{|\hat{x}_{1,k}||\hat{x}_{2,k}|} = \frac{x_{1,k}^T x_{2,k} - [1-\gamma_{1,k}\gamma_{2,k}]\frac{|n_{I,k}|^2}{|n_k|^2} x_{1,k}^T P_{n_{I,k}} x_{2,k}}{|x_{1,k}||x_{2,k}|\sqrt{1-[1-\gamma_{1,k}^2]\frac{|n_{I,k}|^2}{|n_k|^2}\frac{x_{1,k}^T P_{n_{I,k}} x_{1,k}}{|x_{1,k}|^2}}\sqrt{1-[1-\gamma_{2,k}^2]\frac{|n_{I,k}|^2}{|n_k|^2}\frac{x_{2,k}^T P_{n_{I,k}} x_{2,k}}{|x_{2,k}|^2}}} \tag{4.6}$$

where the hat notation is used to refer to time courses after full regression, and $\gamma_{1,k}$ and $\gamma_{2,k}$ are scalar correction terms that account for the difference between full and block regression. When these correction terms are zero, the expressions for block and full regression are identical.

The correction terms $\gamma_{1,k}$ and $\gamma_{2,k}$ are empirically determined and vary with the specific features of the signal and nuisance time series. Thus, in contrast to the block regression, it is challenging to derive theoretical bounds on ΔDFC for the full regression case. However, as shown in Figure 4.18, the empirical values of ΔDFC for full regression are largely within the theoretical bounds obtained for block regression. This is consistent with the empirical similarity of the DFC estimates obtained after block and full regression.

The correction term for a time course x_1 is given in 4.7.3 as $\gamma_{1,k} = \frac{\beta_{1,F} - \beta_{1,k}}{\beta_{1,k}}$, where $\beta_{1,F}$ is the regression fit coefficient using full regression and $\beta_{1,k}$ is the per-window fit coefficient. As defined, the correction term $\gamma_{1,k}$ blows up for very small $\beta_{1,k}$ values. However, in practice this term is actually multiplied by $\beta_{1,k}$ (see Eq. (4.45)), so that it is sufficient to consider the difference between the fit coefficients $m_{1,k} = \beta_{1,F} - \beta_{1,k}$. Similarly, for the second time series x_2, we may consider the difference term $m_{2,k}$. Note that as $m_{1,k}$, $m_{2,k}$, and the product term $m_{1,k}m_{2,k}$ go to zero, then $\gamma_{1,k}$, $\gamma_{2,k}$, and the product $\gamma_{1,k}\gamma_{2,k}$ all approach zero, and full regression and block regression become identical operations. To provide examples of the behavior of these differences, Supplementary Figure 7.15 shows that $m_{1,k}$ and $m_{1,k}$ as well as the product $m_{1,k}m_{2,k}$ are centered around zero for the GS and WM regressors.

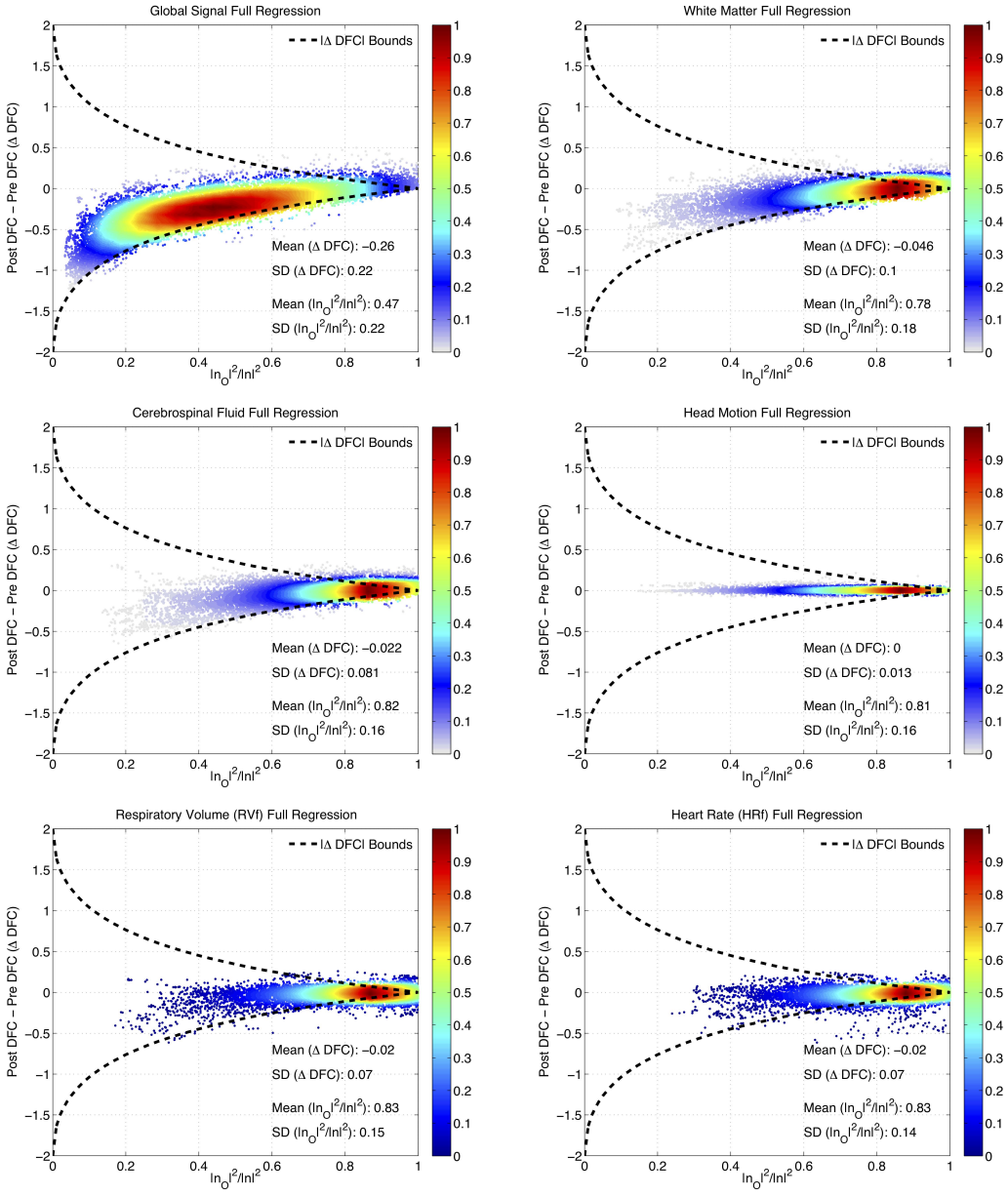


Figure 4.18: The empirical ΔDFC versus the $\frac{|n_O|^2}{|n|^2}$ fraction for the same six nuisance regressors used in Figure 4.16. This time full regression was performed to obtain the empirical values but the same bounds are taken from the block regression derivation. The empirical values of ΔDFC for full regression are largely within the theoretical bounds obtained for block regression, consistent with the empirical similarity of the DFC estimates obtained after block and full regression. Note that the relative density of data points (maximum density is normalized to 1.0) is indicated by the color map on the right-hand side of each plot.

4.6 Discussion

4.6.1 Summary

We have shown that sliding window correlation DFC estimates can be strongly and significantly correlated with the sliding window norms of various nuisance measurements. This relationship between the DFC estimates and the nuisance norms can exist even when the correlations between the underlying nuisance and seed time courses are relatively weak. Moreover, we found that significant correlations between the DFC estimates and nuisance norms can persist even after performing nuisance regression. We derived mathematical expressions to describe the effects and limitations of nuisance regression on DFC estimates and demonstrated that the empirical results lie within the theoretically predicted bounds.

Based on our empirical and theoretical findings, we identified two main mechanisms for the inefficacy of nuisance regression. First, as shown in Figure 4.7, the DFC estimates can be strongly correlated with the nuisance norms even when the underlying correlation between the nuisance terms and the seed signals is relatively low. As a result, for most cases a large fraction of the nuisance term is orthogonal to the subspace spanned by the seed signals. This greatly reduces the efficacy of nuisance regression, such that the difference between the Pre and Post DFC values is relatively small and the relation between the DFC estimates and nuisance norms is largely unaffected. We observed this major limitation of nuisance regression particularly for WM, CSF, RVf, HRf, and HM regressors as demonstrated in Figure 4.16.

The second mechanism applies primarily to the GS nuisance term which has a smaller orthogonal nuisance fraction. The reduced orthogonal nuisance fraction reflects the fact that the GS is computed as the average of the BOLD time courses across the brain and will therefore tend to have a higher correlation with the seed time courses. As shown in Figure 4.17, the mean ΔDFC values observed for GS regression closely follow the mean theoretical bound, which approaches -1 as the orthogonal fraction approaches zero. In contrast, the standard deviations of

the ΔDFC values exhibit a much weaker dependence on the orthogonal fraction. As a result, the Post DFC estimates after GS regression can be approximated to first order as a shifted version of the Pre DFC estimates, as illustrated in the second row of Figure 4.15 for a representative subject. Because the Post DFC estimates largely retain the fluctuations in the Pre DFC estimates, they will also retain the correlation with the nuisance norm. Although this mechanism is most often observed for the GS nuisance term, it can sometimes be observed for WM and CSF nuisance terms (e.g. upper righthand panel in Figure 4.9) in which partial volume effects with gray matter can make the WM and CSF nuisance terms to behave more like the GS.

4.6.2 Nuisance effects in DFC studies

It has been previously noted that nuisance effects might account for a significant portion of the fluctuations in DFC estimates [56, 57]. The effects may be especially pronounced for sliding window DFC estimates since transient nuisance effects can greatly alter the correlation estimates within a short temporal window [56]. [144] looked at the inter-subject correlation between measures of motion and DFC variability, but did not find a significant relation. They also found that intra-subject correlations between sequences of motion-based and DFCs-based deviations were not significant. However, they did not consider the intra-subject correlation between DFC estimates and motion nuisance norms, as was done in this study.

[203] found that dynamic fluctuations in network degree were significantly correlated with the spectral power of end-tidal CO_2 and heart rate measurements. Since the network degree calculated in that study was based on the summation of DFC magnitudes, these findings are roughly consistent with our observation of a significant correlation between the DFC estimates and the norms of the RVf and HRf regressors.

[215] used multivariate kurtosis to assess deviations from constant covariance and found that the removal of high-motion frames reduced the observed kurtosis. They concluded that a significant portion of the resting state functional connectivity dynamics could be attributable to

subject motion.

[204] reported that measures of instantaneous BOLD variance (averaged over the entire brain) were significantly correlated with instantaneous average correlation (over all pairwise correlations). These results are consistent with our findings of significant correlations between DFC estimates and the GS norm. A distinction is that the GS norm is the magnitude of an average signal whereas the instantaneous BOLD variance was calculated as the average across variances from different regions. However, prior work has shown that these two measures are highly related [160, 162].

A key finding of this paper is that DFC estimates can be significantly correlated with the nuisance norms even when the underlying nuisance time courses are largely orthogonal to the observed seed time courses. As discussed in the Interpretation section, one plausible scenario that can give rise to this effect is the presence of an additive nuisance term consisting of time-varying in-plane and orthogonal terms, where the magnitudes of the two terms are roughly in sync and the magnitude of the orthogonal term is much larger than that of the in-plane term. Under this scenario, the time-varying in-plane component can cause DFC fluctuations in the observed vectors (since it affects the inner-angle between the observed vectors) even when the correlation of the putative underlying signal vectors (which are not observed) remains constant. Because of the assumed relation between the in-plane and orthogonal terms, the resulting DFC fluctuations are correlated with the overall nuisance norm. In addition, because the in-plane term is much smaller than the orthogonal term, the overall nuisance term exhibits a weak correlation with the observed vectors. It is important to note that while this plausible scenario provides some insight into the empirical observations and helps to motivate the theoretical findings, it is by no means intended to serve as a "model" of the data. The modeling of nuisance effects in fMRI is still an area of active investigation [202, 205] and future work will be needed to more fully characterize the impact of nuisance terms on DFC estimates. Such efforts are likely to require a consideration of the non-linear and non-stationary aspects of the underlying signals [56, 216].

4.6.3 Efficacy of Nuisance regression in DFC studies

Although nuisance regression is widely performed in DFC studies [56, 57], its effects on DFC estimates have received relatively little attention. [203] reported that regression reduced but did not eliminate the correlation between dynamic measures of network degree and measures of heart rate and respiratory spectral power. These findings are aligned with our results showing that regression has a limited effect on the correlation between DFC estimates and HRf and RVf nuisance norms.

In recent work, [207] found that global signal regression (GSR) had a maximal impact on DFC estimates in temporal windows where the GS mean absolute magnitude was large and a lesser impact in windows where the magnitude was small. They interpreted their findings using the framework introduced in [54], where it was shown that GSR can be approximated as a temporal down-weighting process. The authors noted that the main effect of GSR was a spatially heterogenous negative shift in the sliding window correlation values, an observation consistent with the predominantly negative ΔDFC values found in this study (see the upper righthand panel in Figure 4.16). However, in contrast to our study, the authors did not consider the relationship between the DFC estimates and the GS norm.

There is related work for static FC studies regarding the inefficacy of nuisance regression [64, 173, 185, 217]. For instance, [173] observed that frame-wise displacement (FD), a summary measure for head motion, remained highly related with resting-state fMRI time courses even after HM regression. As we have discussed in detail, the efficacy of nuisance regression increases as the orthogonal nuisance fraction decreases. Due to the inverse relationship between the orthogonal fraction and the percent variance explained (see Figure 4.12), this means that the efficacy of regression will decrease as the amount of variance that can be explained by the nuisance regressors decreases. [64] found that RVf and HRf regressors together explained only 15.8% of the average total variance in voxels which showed a significant correlation between the BOLD time series and nuisance regressors, and the RVf regressor explained only 11.7% of the average total variance

when used on its own. This is roughly in-line with the results shown in Figure 4.7 where the percent variance values observed for non-GS regressors were largely below 20%, with 54% of the strong correlations between the DFC estimates and non-GS nuisance norms occurring when the percent variance was less than 6.25%.

4.6.4 Other Approaches

In this work, we used the sliding window correlation approach, which is widely used in DFC studies [56, 57, 199] both as a primary analysis approach and as an intermediate analysis step (e.g. used to generate DFC estimates that are then further analyzed with k-means clustering or principal components analysis). Other approaches include time-frequency methods such as wavelet transform coherence [144, 200] and probabilistic methods such as hidden Markov modeling [201]. Regardless of the approach, nuisance regression is a standard preprocessing step using nuisance regressors based on either independent measures (e.g. physiological measurements) or data-driven measures (e.g. GS, WM+CSF, or independent component analysis (ICA) components). Our findings regarding the limited efficacy of nuisance regression for sliding-window DFC estimates suggest that caution must also be exercised when interpreting DFC estimates obtained with other approaches. Nevertheless, future work to assess the impact of nuisance terms and the efficacy of nuisance regression when applied to additional DFC approaches would be of great interest.

In this study, we have examined the use of both physiological (respiratory and cardiac) and data-driven (motion, GS, and WM+CSF) nuisance regressors, which are all widely used in the field [206]. There has also been growing interest in the use of ICA-based approaches for denoising of fMRI data. Recent work has pointed out that both traditional nuisance regressor (excluding GS) and ICA-based approaches are limited in their ability to eliminate spatially widespread effects and suggest the continued need for some type of global signal regression [173, 206, 218, 219]. In the case of ICA-based approaches, the limitations most likely reflect the spatial independence criteria

that is inherent to the ICA algorithm. Given this limitation it is unlikely that ICA-based approaches alone can eliminate the relation between DFC estimates and nuisance norms. Nevertheless, the development and comparison of nuisance regressor and ICA-based approaches continues to be an active area of research, and future work to assess the efficacy of ICA-based approaches with regards to the relation between DFC estimates and nuisance norms would be of great interest. In addition, as methods for the removal of nuisance terms continue to evolve, it will be useful to evaluate whether these future approaches can better attenuate the relation between DFC estimates and nuisance norms.

For our examination of the relation between DFC estimates and nuisance norms, we used a seed-based approach in which the average time series from four different seed regions were used. The utilization of the seed based approach with two independent datasets serves as a solid approach for an initial and detailed characterization of the relation between DFC estimates and nuisance norms and the limited efficacy of nuisance regression. An extension of the current work to further characterize these effects using other brain regions and networks would be of interest. This could entail the use of cortical parcellations or ICA-based network components as described in [56, 57, 199, 215]. The fact that the various potential approaches all utilize some type of spatial weighting to derive region-based or network-based time courses suggests that they may yield similar results with regards to DFC estimates and nuisance norms. Nonetheless, further work in this area would be useful to determine if there are any substantial differences.

4.6.5 Nuisance Norm Regression

Given the inefficacy of nuisance regression in reducing the correlations between the DFC estimates and the nuisance norms, it is reasonable to consider alternative approaches. A potential solution is to compute the sliding window norm for each regressor and then project these out from the DFC estimates using linear regression. This procedure is described in 4.7.4 and referred to as *nuisance norm regression* (NNR). This approach differs from traditional nuisance regression

techniques which regress out nuisance measures directly from the fMRI time courses. Instead, NNR acts on the correlation coefficients. Further work is needed to characterize the potential advantages and disadvantages of this approach.

4.6.6 Vigilance Effects

There is a growing appreciation that variations in vigilance may account for a considerable portion of the dynamic fluctuations in resting-state fMRI data [190, 220, 221]. In addition, the state of vigilance can have an effect on physiological measures, such as the recently reported dependence of pulse oximetry amplitudes on the level of wakefulness [222]. Our preliminary analysis in Section 4.3.1 did not reveal an effect of the state of the eyes (open or closed) on the correlations between the DFC estimates and nuisance norms. Nevertheless, future work to assess how the relation between DFC estimates and nuisance norms might depend on the state of vigilance would be of great interest.

4.6.7 Implications for static FC estimates

In this paper, we have considered the relation between nuisance norms and FC estimates obtained across different temporal windows. In DFC studies, the temporal windows under consideration are typically on the order of 40 to 100 second windows. However, the framework for interpreting the observed effects is applicable to arbitrary window lengths. It is therefore of interest to consider whether similar effects are observed for static FC estimates where the window length is equal to the scan duration. In this case both the static FC estimates and nuisance norms are computed over the full duration of each scan and we look to see whether static FC estimates and nuisance norms are related across scans.

In Supplementary Figures 7.16 and 7.17 , we present a preliminary analysis showing the existence of significant correlations (both before and after nuisance regressions) between static

FC estimates and nuisance norms across the 68 scans in the BS002 dataset. These preliminary findings are roughly consistent with recent work looking at the correlation between static FC measures and motion-related metrics [167, 206, 223]. In the preliminary example, the relation between static FC estimates and nuisance norms is examined using different combinations of regressors, as detailed in the figure captions. Similar to the prior work, the inclusion of additional regressors led to a greater reduction (with nuisance regression) in the observed correlations, as compared to the use of only motion regressors.

This preliminary work suggests that the framework we have presented may be useful for providing deeper insights into the effects of nuisance terms and regression in static FC studies. Further work that considers the proposed framework within the context of the growing body of related studies on the effects of nuisance terms in static FC studies would be of interest.

4.6.8 Conclusion

We have provided a detailed examination of nuisance effects and regression in DFC measures. Our findings both confirm and significantly extend the limited prior work in this area. In particular, we have shown that DFC estimates can be strongly correlated with nuisance norms even when there is only a weak correlation between the nuisance and seed signals. We have demonstrated that although nuisance regression can sometimes reduce the correlation between DFC measures and nuisance norms (e.g. when the GS is used as regressor), significant correlations can persist after regression with either GS or non-GS regressors. Furthermore, we provided a mathematical framework to describe the effects of nuisance regression and showed that the experimental findings are in agreement with the theoretical predictions. The current mathematical framework considers a single nuisance regressor and provides valuable insights into the experimental findings. It can be used to approximate the multiple regressor case by taking the first principal component of the regressors, as was done for the HM regressors in Figure 4.16. While our preliminary efforts indicate that an extension of the framework to readily

handle multiple regressors is not straightforward, future work in this area would be of interest.

It is important to note that there is an ongoing discussion in the literature regarding whether DFC estimates are largely due to sampling variability versus true changes in FC (e.g. fluctuations in wakefulness) [215, 221, 224, 225, 226]. The present work does not make any claims as to whether DFC estimates reflect dynamics changes in brain functional connectivity as opposed to sampling variability or other artifactual factors. Instead, the focus is on presenting an empirical finding of a relation between DFC estimates and nuisance norms. This represents an additional factor that should be considered in the interpretation of DFC estimates.

This work highlights a potential confound in the interpretation of DFC studies, which typically make the implicit assumption that nuisance effects are largely minimized in the pre-processing stage. Our findings suggest that the interpretation of DFC measures should be expanded to consider potential correlations with the nuisance norms. If these correlations are not adequately considered, differences in DFC estimates (e.g. between groups or treatment conditions) may be incorrectly interpreted as reflecting meaningful effects when in fact they may be largely attributable to differences in nuisance activity. Because nuisance effects such as subject motion, respiration, and cardiac activity originate in brain networks that control these functions [205, 227], it will be important to distinguish between spatially widespread effects that may be largely considered as measurement confounds and more localized effects that may be viewed as meaningful reflections of links between DFC measures and physiological activity. Future DFC studies would benefit from the development of analysis methods that more effectively take into account the origins of the nuisance effects.

4.7 Appendix

4.7.1 Block regression on windowed time series

In this Appendix section we derive expressions that make explicit the dependence of the correlation coefficient after block nuisance regression on the orthogonal and in-plane components of the nuisance term. We then use these expressions in 4.7.2 to derive bounds on the difference between the pre and post regression DFC estimates.

To link our derivations with the geometric concepts presented in Section 4.4, we represent the time courses as column vectors and use projection matrices to represent the process of regression. We note that alternate derivations are certainly possible. For example, it is possible to start with the Yule expression for partial correlation [228] to derive Eq. (4.17). However, such an approach would not be as straightforward since the Yule expression does not explicitly model the dependence on the orthogonal and in-plane components of the nuisance term.

We begin by presenting the notation for the time series, where each time course is represented as a column vector. For the kth window, let $x_{1,k}$ and $x_{2,k}$ be a pair of windowed fMRI time series, with $\tilde{x}_{1,k}$ and $\tilde{x}_{2,k}$ corresponding to the time courses after block regression using n_k as the nuisance regressor. Since each window is treated independently in block regression, we can simplify the derivations by dropping the window index k for now. We also assume without loss of generality that the time series have zero mean.

For x_1 we can write the time course after regression as:

$$\tilde{x}_1 = x_1 - n(n^T n)^{-1} n^T x_1 = x_1 - P_n x_1, \tag{4.7}$$

where $P_n = n(n^T n)^{-1} n^T$ is the projection matrix onto n, and a similar expression holds for x_2. For the derivations that follow, we decompose the nuisance regressor into in-plane and orthogonal components such that $n = n_I + n_O$ where n_I is the component of n that lies in the subspace spanned

by x_1 and x_2, and n_O is the orthogonal complement.

The squared norm for the time series after regression is:

$$\tilde{x}_1^T \tilde{x}_1 = (x_1 - P_n x_1)^T (x_1 - P_n x_1) \tag{4.8}$$

$$= x_1^T x_1 - x_1^T P_n x_1 \tag{4.9}$$

$$= x_1^T x_1 - \frac{x_1^T (n_I + n_O)(n_I + n_O)^T x_1}{|n|^2} \tag{4.10}$$

$$= x_1^T x_1 - \frac{x_1^T n_I n_I^T x_1}{|n|^2} \tag{4.11}$$

$$= x_1^T x_1 - \frac{|n_I|^2}{|n|^2} x_1^T P_{n_I} x_1, \tag{4.12}$$

where $|n|^2 = n_I^T n_I + n_O^T n_O$, $P_{n_I} = n_I (n_I^T n_I)^{-1} n_I^T$, and we have made use of the relation $n_O^T x_1 = 0$ and the symmetry of the projection matrices. The corresponding norm can then be written as:

$$|\tilde{x}_1| = \sqrt{x_1^T x_1 - \frac{|n_I|^2}{|n|^2} x_1^T P_{n_I} x_1} \tag{4.13}$$

$$= |x_1| \sqrt{1 - \frac{|n_I|^2}{|n|^2} \frac{x_1^T P_{n_I} x_1}{|x_1|^2}}. \tag{4.14}$$

A similar derivation holds for $|\tilde{x}_2|$.

To compute the correlation coefficient after regression, we start with the dot product between $\tilde{x}_1$ and $\tilde{x}_2$ and use the orthogonal decomposition to write:

$$\tilde{x}_1^T \tilde{x}_2 = (x_1 - P_n x_1)^T (x_2 - P_n x_2) \tag{4.15}$$

$$= x_1^T x_2 - \frac{|n_I|^2}{|n|^2} x_1^T P_{n_I} x_2, \tag{4.16}$$

where the omitted steps in the derivation are similar to those shown above for the derivation of the norm. Normalizing the dot product by the appropriate norms yields the correlation coefficient

after regression as:

$$\tilde{r} = \frac{\tilde{x}_1^T \tilde{x}_2}{|\tilde{x}_1||\tilde{x}_2|} = \frac{x_1^T x_2 - \frac{|n_I|^2}{|n|^2} x_1^T P_{n_I} x_2}{|x_1||x_2| \sqrt{1 - \frac{|n_I|^2}{|n|^2} \frac{x_1^T P_{n_I} x_1}{|x_1|^2}} \sqrt{1 - \frac{|n_I|^2}{|n|^2} \frac{x_2^T P_{n_I} x_2}{|x_2|^2}}}. \tag{4.17}$$

This can be rewritten in the following form:

$$\tilde{r} = \lambda r + O, \tag{4.18}$$

where $r = \frac{x_1^T x_2}{|x_1||x_2|}$ is the correlation coefficient prior to regression, and the scaling λ and offset O terms are defined as:

$$\lambda = \frac{1}{\sqrt{1 - \frac{|n_I|^2}{|n|^2} \frac{x_1^T P_{n_I} x_1}{|x_1|^2}} \sqrt{1 - \frac{|n_I|^2}{|n|^2} \frac{x_2^T P_{n_I} x_2}{|x_2|^2}}} \tag{4.19}$$

$$O = \frac{-\lambda \frac{|n_I|^2}{|n|^2} x_1^T P_{n_I} x_2}{|x_1||x_2|}. \tag{4.20}$$

Note that if the orthogonal nuisance component n_O is large compared to the in-plane component n_I, the ratio $\frac{|n_I|^2}{|n|^2} = 1 - \frac{|n_O|^2}{|n_I|^2 + |n_O|^2}$ approaches 0, such that the scaling term λ approaches 1, and the offset term O goes to 0. As a result, the correlation coefficients before and after regression will be approximately equal $\tilde{r} \approx r$.

To gain further insight, it is useful to rewrite Eq. (4.17) using trigonometric functions. First, note that the correlation coefficient prior to regression can be also written as $r = \cos\theta$, where θ is the angle between x_1 and x_2. Then without loss of generality, we can write $\theta = \theta_1 + \theta_2$ where θ_1 is the angle between n_I and x_1 and θ_2 is the angle between n_I and x_2, where both θ_1 and θ_2 are assumed to be non-negative. This corresponds to the case where the in-plane term n_I "lies"

between the vectors x_1 and x_2. Note it also possible for the in-plane term to lie outside of the vectors, such that $\theta = \pm(\theta_1 - \theta_2)$. We refer to this second case as the complementary case.

With the stated assumptions, we rewrite the post regression DFC in Eq. (4.17) as:

$$\tilde{r} = \frac{\tilde{x}_1^T \tilde{x}_2}{|\tilde{x}_1||\tilde{x}_2|} = \frac{\dfrac{x_1^T x_2}{|x_1||x_2|} - \dfrac{|n_I|^2}{|n|^2}\dfrac{x_1^T P_{n_I} x_2}{|x_1||x_2|}}{\sqrt{1 - \dfrac{|n_I|^2}{|n|^2}\dfrac{x_1^T P_{n_I} x_1}{|x_1|^2}}\sqrt{1 - \dfrac{|n_I|^2}{|n|^2}\dfrac{x_2^T P_{n_I} x_2}{|x_2|^2}}} \tag{4.21}$$

$$= \frac{\cos\theta - \dfrac{|n_I|^2}{|n|^2}\cos\theta_1\cos\theta_2}{\sqrt{1 - \dfrac{|n_I|^2}{|n|^2}\cos^2\theta_1}\sqrt{1 - \dfrac{|n_I|^2}{|n|^2}\cos^2\theta_2}} \tag{4.22}$$

$$= \frac{\cos\theta - \dfrac{|n_I|^2}{|n|^2}(\cos\theta \pm \sin\theta_1\sin\theta_2)}{\sqrt{1 - \dfrac{|n_I|^2}{|n|^2}\cos^2\theta_1}\sqrt{1 - \dfrac{|n_I|^2}{|n|^2}\cos^2\theta_2}} \tag{4.23}$$

$$= \frac{\left(1 - \dfrac{|n_I|^2}{|n|^2}\right)\cos\theta \mp \dfrac{|n_I|^2}{|n|^2}\sin\theta_1\sin\theta_2}{\sqrt{1 - \dfrac{|n_I|^2}{|n|^2}\cos^2\theta_1}\sqrt{1 - \dfrac{|n_I|^2}{|n|^2}\cos^2\theta_2}}, \tag{4.24}$$

where we have used the identities:

$$x_1^T P_{n_I} x_2 = \frac{x_1^T n_I n_I^T x_2}{|n_I|^2} = |x_1||x_2|\cos\theta_1\cos\theta_2, \tag{4.25}$$

$$x_1^T P_{n_I} x_1 = |x_1|^2\cos^2\theta_1, \tag{4.26}$$

$$x_2^T P_{n_I} x_2 = |x_2|^2\cos^2\theta_2. \tag{4.27}$$

Also note that in Eq. (4.23) we have used the trigonometric identity $\cos\theta_1\cos\theta_2 = \cos\theta \pm \sin\theta_1\sin\theta_2$. If $\theta = \theta_1 + \theta_2$, the sign of $\mp$ in Eq. (4.24) is a minus $(-)$, and in the complementary case for $\theta = \pm(\theta_1 - \theta_2)$ (i.e. when n_I lies outside of x_1 and x_2) the sign is a $+$.

Finally, by making the substitution $\frac{|n_I|^2}{|n|^2} = 1 - \frac{|n_O|^2}{|n|^2}$, the correlation coefficient after regression is:

$$\tilde{r} = \frac{\frac{|n_O|^2}{|n|^2}\cos\theta \mp \left(1 - \frac{|n_O|^2}{|n|^2}\right)\sin\theta_1 \sin\theta_2}{\sqrt{\sin^2\theta_1 + \frac{|n_O|^2}{|n|^2}\cos^2\theta_1}\sqrt{\sin^2\theta_2 + \frac{|n_O|^2}{|n|^2}\cos^2\theta_2}}. \tag{4.28}$$

If the orthogonal component becomes small and the orthogonal fraction $\frac{|n_O|^2}{|n|^2} \to 0$, then $\tilde{r}$ reduces to:

$$\tilde{r} = \pm\frac{\sin\theta_1 \sin\theta_2}{|\sin\theta_1||\sin\theta_2|} = \pm 1, \tag{4.29}$$

where the minus sign applies to the case where the nuisance vector lies between the observed vectors and the plus sign applies to the complementary case where the nuisance vector lies outside the vectors. In terms of the underlying fMRI data, when the orthogonal component is zero, the nuisance term completely lies within the hyperplane spanned by the two voxel time series (e.g. Case 5 in Figure 4.13, which corresponds to the conditions that give rise to a minus sign). As a result, when the nuisance term is projected out, the resulting time series are forced to be completely anti-correlated, as represented by the post regression vectors pointing in opposite directions in Case 5.

4.7.2 Derivation of limits on difference in DFC estimates

Here we derive the analytical bound on the difference $\Delta\text{DFC} = \tilde{r} - r$, between the pre and post regression DFC estimates. First, we provide derivations for the case when $\theta = \theta_1 + \theta_2$ (we

will consider the complementary case $\theta = \pm(\theta_1 - \theta_2)$ later). In this case, the ΔDFC is:

$$\Delta\text{DFC} = \frac{\frac{|n_O|^2}{|n|^2}\cos(\theta_1+\theta_2) - \left(1 - \frac{|n_O|^2}{|n|^2}\right)\sin\theta_1\sin\theta_2}{\sqrt{\sin^2\theta_1 + \frac{|n_O|^2}{|n|^2}\cos^2\theta_1}\sqrt{\sin^2\theta_2 + \frac{|n_O|^2}{|n|^2}\cos^2\theta_2}} - \cos(\theta_1+\theta_2). \tag{4.30}$$

Note that since n_I lies between the observed vectors x_1 and x_2, nuisance regression will increase the angle (and decrease the correlation value) between the vectors x_1 and x_2, and we will have $\tilde{r} < r$. As a result, we are looking for a lower bound on ΔDFC as a function of $\frac{|n_O|^2}{|n|^2}$.

It can be shown empirically that the expression for ΔDFC is minimized when $\theta_1 = \theta_2$.

Using this, we simplify the original expression for ΔDFC in Eq. (4.30) as:

$$\Delta\text{DFC} = \frac{\frac{|n_O|^2}{|n|^2}\cos 2\theta_1 - \left(1 - \frac{|n_O|^2}{|n|^2}\right)\sin^2\theta_1}{\sin^2\theta_1 + \frac{|n_O|^2}{|n|^2}\cos^2\theta_1} - \cos 2\theta_1 \tag{4.31}$$

$$= \frac{\left(-\sin^2\theta_1 - \frac{|n_O|^2}{|n|^2}\cos^2\theta_1 + \frac{|n_O|^2}{|n|^2}\right)\cos 2\theta_1 + \left(\frac{|n_O|^2}{|n|^2} - 1\right)\sin^2\theta_1}{\sin^2\theta_1 + \frac{|n_O|^2}{|n|^2}\cos^2\theta_1} \tag{4.32}$$

$$= \frac{\left(\frac{|n_O|^2}{|n|^2}\sin^2\theta_1 - \sin^2\theta_1\right)\cos 2\theta_1 + \left(\frac{|n_O|^2}{|n|^2} - 1\right)\sin^2\theta_1}{\sin^2\theta_1 + \frac{|n_O|^2}{|n|^2}\cos^2\theta_1} \tag{4.33}$$

$$= \frac{\left(\frac{|n_O|^2}{|n|^2} - 1\right)\sin^2\theta_1(2\cos^2\theta_1 - 1) + \left(\frac{|n_O|^2}{|n|^2} - 1\right)\sin^2\theta_1}{\sin^2\theta_1 + \frac{|n_O|^2}{|n|^2}\cos^2\theta_1} \tag{4.34}$$

$$= \frac{\left(\frac{|n_O|^2}{|n|^2} - 1\right)\sin^2 2\theta_1}{(1-\cos 2\theta_1) + \frac{|n_O|^2}{|n|^2}(1+\cos 2\theta_1)} \tag{4.35}$$

$$= \frac{\left(\frac{|n_O|^2}{|n|^2} - 1\right)\sin^2 2\theta_1}{1 + \frac{|n_O|^2}{|n|^2} + \left(\frac{|n_O|^2}{|n|^2} - 1\right)\cos 2\theta_1}, \tag{4.36}$$

where in the numerator of Eq. (4.35), we used the trigonometric identity: $2\sin^2\theta_1\cos^2\theta_1 = \sin^2 2\theta_1/2$, and in the denominator of the same equation, we used: $\sin^2\theta_1 = (1-\cos 2\theta_1)/2$ and

$\cos^2\theta_1 = (1+\cos 2\theta_1)/2$. We then take the partial derivative with respect to θ_1 as:

$$\frac{\partial}{\partial\theta_1}[\Delta\text{DFC}] = \frac{\partial}{\partial\theta_1}\left[\frac{\left(\frac{|n_O|^2}{|n|^2}-1\right)\sin^2 2\theta_1}{1+\frac{|n_O|^2}{|n|^2}+\left(\frac{|n_O|^2}{|n|^2}-1\right)\cos 2\theta_1}\right] \tag{4.37}$$

$$= 2\sin 2\theta_1\left(1-4\frac{\frac{|n_O|^2}{|n|^2}}{\left(1+\frac{|n_O|^2}{|n|^2}+\left(\frac{|n_O|^2}{|n|^2}-1\right)\cos 2\theta_1\right)^2}\right). \tag{4.38}$$

Now solving for the non-trivial roots of the derivative yields $\theta_1^* = \pm tan^{-1}\left(\sqrt{\frac{|n_O|}{|n|}}\right)$. Finally, by using this value in the ΔDFC expression we find the lower bound as:

$$\Delta\text{DFC}_{LB} = -\left(\frac{4}{1+|n_O|/|n|}-2\right) = -2\left(\frac{|n|-|n_O|}{|n|+|n_O|}\right). \tag{4.39}$$

For the complementary case: $\theta = \pm(\theta_1-\theta_2)$, in which n_I is positioned outside of x_1 and x_2, linear regression will decrease the angle between the vectors x_1 and x_2 and the correlation value after regression will increase yielding $\Delta\text{DFC} \geq 0$. Therefore, we are seeking an upper bound. It can be shown empirically that the expression for ΔDFC is maximized when $\theta_1+\theta_2 = \pi$. Following a derivation similar to the one used for the lower bound, the upper bound is obtained as:

$$\Delta\text{DFC}_{UB} = +2\left(\frac{|n|-|n_O|}{|n|+|n_O|}\right). \tag{4.40}$$

4.7.3 Derivation of DFC estimate after full regression

In this section, we show that the correlation coefficients obtained after full regression can be expressed as a modified version of the correlation coefficients obtained with block regression.

First, we note that for the kth window the residual after block regression can be written as:

$$\tilde{x}_{1,k} = x_{1,k} - P_{n_k} x_{1,k} \tag{4.41}$$

$$= x_{1,k} - \beta_{1,k} n_k, \tag{4.42}$$

where $P_{n_k} = n_k (n_k^T n_k)^{-1} n_k$ is the projection matrix onto the windowed nuisance time series n_k, and $\beta_{1,k} = (n_k^T n_k)^{-1} n_k^T x_{1,k}$ is the corresponding scalar fit coefficient computed for the windowed time series $x_{1,k}$.

For full regression the residual can be written as:

$$\hat{x}_{1,k} = x_{1,k} - \beta_{1,F} n_k, \tag{4.43}$$

where $\beta_{1,F} = (n^T n)^{-1} n^T x_1$ is the scalar fit coefficient computed using the entire time series x_1 and the full nuisance time series n.

Noting that the expressions in Eq. (4.42) and Eq. (4.43) differ only in the scalar fit coefficients $\beta_{1,F}$ and $\beta_{1,k}$, we can rewrite Eq. (4.43) as:

$$\hat{x}_{1,k} = x_{1,k} - \left(\beta_{1,k} + m_{1,k}\right) n_k \tag{4.44}$$

$$= x_{1,k} - \left(1 + \frac{m_{1,k}}{\beta_{1,k}}\right) \beta_{1,k} n_k \tag{4.45}$$

$$= x_{1,k} - (1 + \gamma_{1,k}) P_{n_k} x_{1,k}, \tag{4.46}$$

where $m_{1,k} = \beta_{1,F} - \beta_{1,k}$ is the difference in fit coefficients and $\gamma_{1,k} = \frac{m_{1,k}}{\beta_{1,k}}$ is a correction term that captures the difference between full and block regression. Note that if $\gamma_{1,k} = 0$ then full regression and block regression have identical effects for the kth window. The expression for $\hat{x}_{2,k}$ has the same form as Eq. (4.46).

To compute the correlation coefficient between $\hat{x}_{1,k}$ and $\hat{x}_{2,k}$, we first note that the dot

product can be written as:

$$\hat{x}_{1,k}^T \hat{x}_{2,k} = (x_{1,k} - (1+\gamma_{1,k})P_{n_k} x_{1,k})^T (x_{2,k} - (1+\gamma_{2,k})P_{n_k} x_{2,k}) \tag{4.47}$$

$$= x_{1,k}^T x_{2,k} - [1-\gamma_{1,k}\gamma_{2,k}] x_{1,k}^T P_{n_k} x_{2,k} \tag{4.48}$$

$$= x_{1,k}^T x_{2,k} - [1-\gamma_{1,k}\gamma_{2,k}] \frac{|n_{I,k}|^2}{|n_k|^2} x_{1,k}^T P_{n_{I,k}} x_{2,k}, \tag{4.49}$$

which is similar in form to the dot product expression previously derived in Eq. (4.16) for block regression with the addition of a correction term $[1-\gamma_{1,k}\gamma_{2,k}]$ and the window index k. The norms $|\hat{x}_{1,k}|$ and $|\hat{x}_{1,k}|$ can be derived using similar modifications.

Incorporating all the correction terms, we can write the correlation coefficient after full regression as:

$$\frac{\hat{x}_{1,k}^T \hat{x}_{2,k}}{|\hat{x}_{1,k}||\hat{x}_{2,k}|} = \frac{x_{1,k}^T x_{2,k} - [1-\gamma_{1,k}\gamma_{2,k}] \frac{|n_{I,k}|^2}{|n_k|^2} x_{1,k}^T P_{n_{I,k}} x_{2,k}}{|x_{1,k}||x_{2,k}| \sqrt{1-[1-\gamma_{1,k}^2] \frac{|n_{I,k}|^2}{|n_k|^2} \frac{x_{1,k}^T P_{n_{I,k}} x_{1,k}}{|x_{1,k}|^2}} \sqrt{1-[1-\gamma_{2,k}^2] \frac{|n_{I,k}|^2}{|n_k|^2} \frac{x_{2,k}^T P_{n_{I,k}} x_{2,k}}{|x_{2,k}|^2}}}. \tag{4.50}$$

Note that if $\gamma_{1,k} = \gamma_{2,k} = 0$ and we drop the window index k, then this expression is identical to the expression for block regression previously shown in Eq. (4.17).

4.7.4 Nuisance Norm Regression (NNR)

Nuisance norm regression aims to remove the nuisance norm-related variance from the DFC estimates. In this technique, nuisance removal can be performed in the relevant DFC metric space in addition to traditional nuisance removal on raw time series. For sliding window correlations obtained between pairs of ROIs, we define a sliding window correlation vector $R = [r_1, r_2, \cdots, r_k, \cdots, r_M]^T$, where k is the window index, M is the total number of windows, and r_k is the pairwise correlation value for the kth window. Similarly, we define a nuisance norm

vector (or matrix for multiple nuisance terms) $N = [|n_1|, |n_2|, \cdots, |n_k|, \cdots, |n_M|]^T$, where $|n_k|$ is the norm of the nuisance regressor for the kth window. Then the effects of a nuisance norm can be removed through linear regression to obtain a "clean" DFC estimate $\tilde{R} = R - N(N^T N)^{-1} N^T R$. In the case of multiple nuisance terms, the nuisance norm matrix can be expanded to include additional norm vectors. For example, let N_{WM}, N_{CSF}, N_{RVf}, N_{HRf}, and N_{HM} be the sliding window norm vectors for the WM, CSF, RVf, HRf, and HM (total norm) regressors, respectively. Then the corresponding nuisance norm matrix is $N = [N_{WM}, N_{CSF}, N_{RVf}, N_{HRf}, N_{HM}]$. Potential variations include expansion of the nuisance norm matrix to include additional terms, such as the squared norms or temporal derivatives of the sliding window norms.

4.8 Acknowledgments

Chapter 4 is, in full, a reprint of the material as it appears in: Alican Nalci, Bhaskar D. Rao, and Thomas T. Liu. "*Nuisance Effects and the Limitations of Nuisance Regression in Dynamic Functional Connectivity fMRI.*" NeuroImage 184 (2019): 1005-1031. I was the primary author of this paper and T.T. Liu supervised the research. The work was partially supported by NIH grant R21MH112155.

Chapter 5

Nuisance Effects in Inter-scan Functional Connectivity Estimates Before and After Nuisance Regression

In resting-state functional MRI (fMRI), the correlation between blood-oxygenation-level-dependent (BOLD) signals across brain regions is used to estimate the functional connectivity (FC) of the brain. FC estimates are prone to the influence of nuisance factors including scanner-related artifacts and physiological modulations of the BOLD signal. Nuisance regression is widely performed to reduce the effect of nuisance factors on FC estimates on a per-scan basis. However, a dedicated analysis of nuisance effects on the variability of FC metrics across a collection of scans has been lacking. This work investigates the effects of nuisance factors on the variability of FC estimates across a collection of scans both before and after nuisance regression. Inter-scan variations in FC estimates are shown to be significantly correlated with the geometric norms of various nuisance terms, including head motion measurements, signals derived from white-matter and cerebrospinal regions and the whole-brain global signal (GS) both before and after nuisance regression. It is shown that GS regression (GSR) can introduce negative GS norm fluctuations into

the inter-scan FC estimates. This work shows that caution must be exercised when interpreting inter-scan FC measures across scans both before and after nuisance regression.

5.1 Introduction

Resting-state functional magnetic resonance imaging (fMRI) is a widely used method that aims to characterize the functional organization of the brain at rest [150, 229]. The blood-oxygenation-level-dependent (BOLD) signal reflects metabolic changes in the brain that result from neuronal activity [62, 151]. The correlation between BOLD signals across different brain regions is computed to estimate the functional connectivity (FC) of the brain [49, 50].

It is well-known that the BOLD signal is prone to the influence of various nuisance confounds including thermal noise, scanner drift, head motion, and physiological activity such as changes in respiration and heart rate [60, 61, 62, 63, 64]. If these confounds are not removed from the BOLD signal prior to analysis, they can lead to an increase in the number of false positives and negatives, causing erroneous interpretations of the fMRI results [65].

Nuisance regression (NR) is widely performed to improve the spatial specificity of FC estimates on a per-scan basis. This involves projecting out a combination of nuisance measurements from the BOLD data prior to the computation of FC estimates. Nuisance measurements typically include but are not limited to head motion (HM) measurements, signals from the white-matter (WM) and cerebrospinal fluid (CSF) regions, cardiac and respiratory activity derived time courses, and the whole-brain global signal (GS) [63, 78, 159, 205, 230].

Despite the fact that NR is adopted with the assumption that it removes nuisance confounds from the FC estimates, it has been previously shown that it can be quiet ineffective in reducing the effects of nuisance confounds [60, 185, 231]. For example, HM regression has been shown to be a largely ineffective approach for reducing HM confounds in FC estimates even after projecting out 12 motion regressors [185, 232, 233].

More recently, in [231], we have shown that *dynamic* FC (DFC) estimates were related to the norms of various nuisance regressors such as the HM, WM+CSF, GS, heart rate, and respiratory derived time courses. We have shown that NR was largely ineffective in removing nuisance effects from DFC estimates with significant relations between the nuisance norms and DFC estimates remaining even after NR. We presented a theoretical basis explaining why NR was ineffective and found that the effects of nuisance norms on the DFC estimates were significant even when the raw nuisance signals and BOLD signals were largely unrelated.

The fundamental difference between DFC and FC studies is the temporal duration over which the FC estimates (correlations between BOLD signals) are computed. In DFC studies the temporal window is typically on the order of 30-60 seconds, where as in static FC studies the duration is the whole scan duration which is typically several minutes long [56, 57].

Although the effects of nuisance terms and efficacy of NR have been investigated on a per-scan basis [78], efforts to examine nuisance effects with regards to variations in FC estimates across scans have been rather limited. A better understanding of these effects is critical considering the increasing use of fMRI to examine the differences in FC measures between important disease populations and healthy controls [234]. Some important studies include the investigation of FC metrics in Alzheimer's disease [66, 67, 68], Parkinson's disease [69], depression [70], schizoprenia [71], dementia [235], and amyotrophic lateral sclerosis (ALS) [236, 237].

In this work, we first investigate the effects of nuisance terms on the variability of FC estimates across different scans. Specifically, we compute the norm of the HM measurements, WM+CSF time courses, and the GS for each scan. We show the existence of significant correlations across scans between the FC estimates and nuisance norms. We find nuisance regression using non-GS regressors to be largely ineffective in reducing the correlations between FC estimates and nuisance norms. We show that though GSR is partially effective in reducing the relation between GS norm and FC estimates, a considerable portion of the GS norm-related variance remains in the FC estimates, and strong GS norm-related fluctuations can be injected

into the FC estimates.

Our work significantly extends the preliminary results in [231]. We provide a more extensive analysis of various nuisance effects on the FC estimates before and after NR. We generalize the theory developed in [231] to static FC measures and confirm the validity of the theoretical limitation of nuisance regression for correlation-based static FC estimates. We introduce *nuisance contamination maps* which illustrate the spatial distribution and extent of correlations between nuisance norms and FC estimates across scans. We also provide a detailed analysis of the limited efficacy of GSR and show how GSR can introduce GS norm-related fluctuations in the FC estimates across scans.

5.2 Methods

5.2.1 Data

We used a publicly available dataset originally analyzed by [174]. The data were acquired from 17 young adults using a 3T Siemens Allegra MR scanner. Each subject underwent 4 BOLD echo-planar imaging (EPI) scans (32 slices, TR=2.16 s, TE=25 ms, $4\times4\times4$ mm) each lasting 7 minutes (194 frames). The subjects were instructed to look at a cross-hair and asked to remain still and awake. High-resolution T1-weighted anatomical images were acquired for the purpose of anatomical registration (TR=2.1 s, TE=3.93 ms, flip angle=7 deg, $1\times1\times1.25$ mm).

Standard pre-processing steps were conducted with the AFNI software package [175]. The initial 9 frames from each EPI run were discarded to minimize longitudinal relaxation effects. Images were then slice-time corrected and co-registered, and the 6 head motion parameter time series were retained. The images were converted to Talairach and Tournoux (TT) coordinates, resampled to 3 mm cubic voxels, and spatially smoothed using a 6 mm full-width-at-half-maximum isotropic Gaussian kernel. The 1^{st} and 2^{nd} order Legendre polynomials (a constant term to model the temporal mean and a linear trend) were projected out from each voxel's time

course. Each voxel time series was then converted into a percent change BOLD time series through demeaning and division by its mean value. This version of the data will be referred to as "uncorrected" data in this paper.

We used seed signals derived from the posterior cingulate cortex (PCC), intraparietal sulcus (IPS), frontal eye fields (FEF), auditory (AUD) and motor (MOT) networks. These seed signals were obtained by averaging time series selected over spheres of radius 6 mm (2 voxels) centered about their corresponding TT coordinates [162]. The sphere centers were obtained by converting the MNI coordinates from [178] to TT coordinates [179]. For the PCC, IPS, FEF and MPF seeds we used the coordinates [0,-51,26], [32,-51,41], [24,-13,51], and [6,32,28], respectively. For the left MOT, and right MOT seeds we used the coordinates [-36,-22,52] and [37,-12,52], respectively. A combined MOT seed was obtained by using the left and right MOT coordinates to define two spheres and by merging the spheres. A combined AUD seed was obtained by using the left and right AUD coordinates [-41,-26,14] and [41,-26,14], respectively. Finally, for the WM and CSF nuisance signals we defined the sphere centers as [31,-28,32] and [-15,-28,21], respectively.

5.2.2 Inter-scan variations in FC estimates

To investigate the variations in FC estimates across scans, we computed the Pearson correlation between a seed signal and every other voxel in the brain for each scan. Denoting a pair of zero mean percent change BOLD signals from the PCC and IPS regions with x_1 and x_2 in vector notation, the FC estimate for the kth scan was obtained by computing $\mathrm{FC}_k = (x_1^T x_2)/(|x_1||x_2|)$, where $|.|$ denotes the ℓ_2 norm and k is the scan index. For each seed-voxel pair, we computed the FC estimates across all scans and concatenated them to form a vector of FC estimates: $\mathrm{FC}_{\mathrm{Vec}} = \{\mathrm{FC}_1, \mathrm{FC}_2, ..., \mathrm{FC}_K\}$ where $K = 68$ is the total number of scans. This vector will be referred to as the inter-scan FC estimates or simply as FC estimates. We obtained a separate vector $\mathrm{FC}_{\mathrm{Vec}}$ for each seed voxel pair in the brain (i.e. for a single seed, we have N vectors where

N is the number of voxels).

5.2.3 Nuisance regressions

To investigate the effects of nuisance regression on FC estimates, we performed 4 separate nuisance regressions on the uncorrected data. This was done prior to the computation of FC estimates. Nuisance regressions involved projecting out (1) 6 HM parameters, (2) 6 HM parameters combined with the signals from the WM and CSF regions, (3) the GS time course, and (4) HM, WM, CSF signals combined with the GS. The global signal (GS) was obtained as the average of all (percent) change BOLD time courses across the whole brain volume. For each nuisance regression, the vector of inter-scan FC estimates prior to nuisance regression will be referred to as "Pre FC" estimates and after regression as "Post FC" estimates.

5.2.4 Norm as a nuisance metric on FC estimates across scans

To measure the effect of nuisance terms on the FC estimates across scans we adopted the approach in [231]. For GS regression, we first computed the ℓ_2 norm of the GS time course for each scan. Denoting the GS time course for a scan k with n_k, we computed the ℓ_2 norm as $|n_k| = \sqrt{\sum_{t=1}^{T} n_k^2(t)}$, where t indexes over time and T is the total number of time points. We then concatenated the GS norms across different scans to obtain a vector of GS norms as $|n|_{\text{Vec}} = \{|n_1|, |n_2|, ..., |n_K|\}$.

For multiple nuisance regressions (e.g. HM+WM+CSF) we obtained a total norm of all the regressors involved by computing $|n_k| = \sqrt{\sum_{i=1}^{N_R} \sum_{t=1}^{T} n_{i,k}^2(t)}$, where $n_{i,k}$ is a single regressor time course, i is the index over multiple regressors and N_R is the total number of regressors. Finally, we concatenated the corresponding total nuisance norms across different scans to obtain a nuisance norm vector as $|n|_{\text{Vec}} = \{|n_1|, |n_2|, ..., |n_K|\}$.

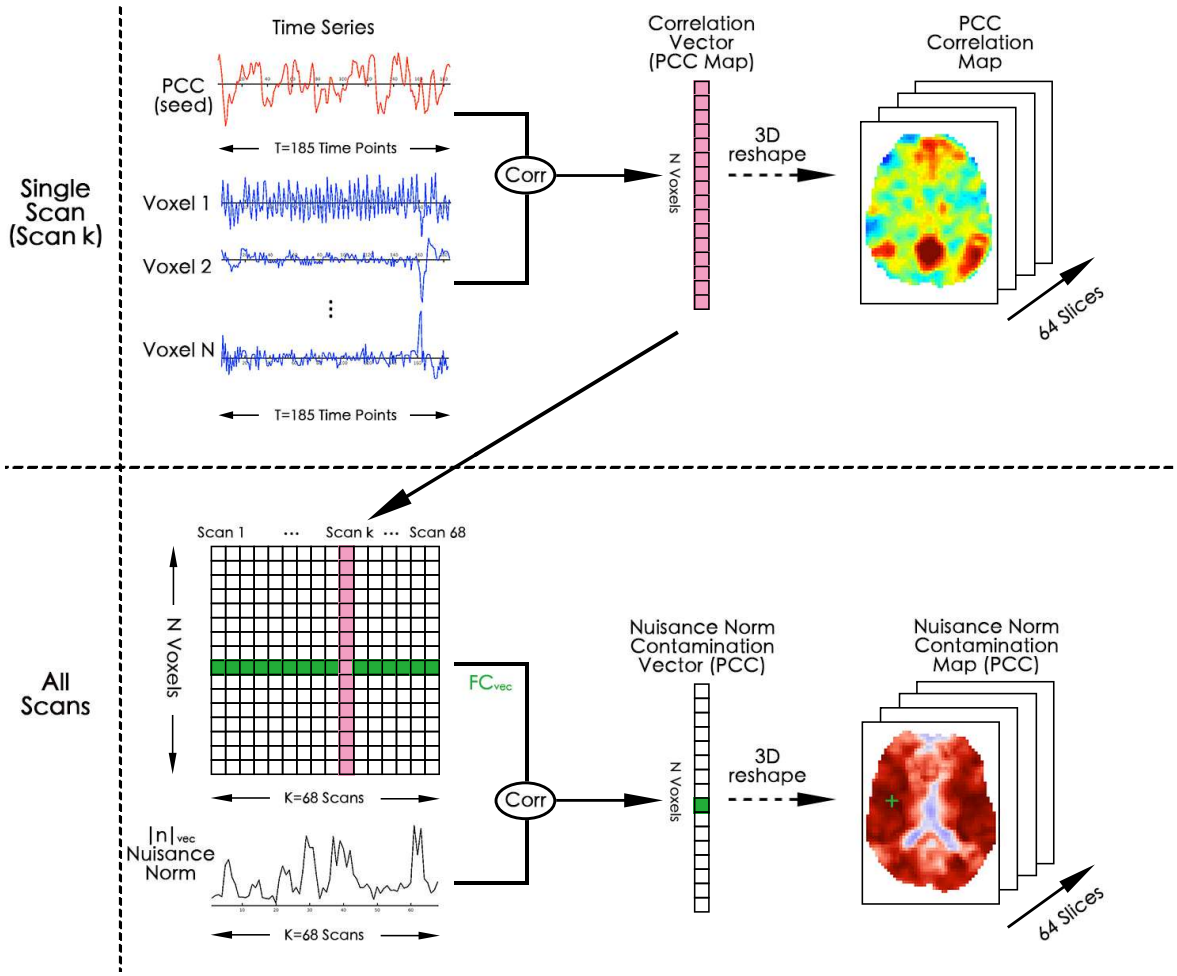


Figure 5.1: Diagram illustrating how to obtain nuisance contamination maps. These maps visualize the spatial distribution of the correlations between inter-scan FC estimates (FC_{Vec}) and nuisance norms ($|n|_{Vec}$). In the first row, a seed signal (e.g. PCC seed) is correlated with signals from other voxels in the brain to form a vector (vertical vector with red color) of FC estimates. This vector represents the canonical seed-based PCC correlation map. This step is repeated across all 68 scans and the resulting PCC maps are concatenated column-wise to form a matrix of FC estimates across scans. This matrix is shown on the bottom left-hand side. The individual rows of the matrix correspond to inter-scan variations in the FC between the PCC seed and a single voxel (an example is shown with green color). Each column is the PCC map for a single scan. The individual rows of this matrix are then correlated with the nuisance norm (shown at the bottom left) to obtain the nuisance contamination values as a 1D vector. An entry of the nuisance contamination vector corresponds to the correlation between the nuisance norm and the FC vector from a single seed-voxel pair. This is illustrated with the dark green (+) symbol on the contamination map. The nuisance contamination vector can be reshaped into a 3D volume to investigate regions of nuisance contamination across the brain.

5.2.5 Nuisance Contamination Maps: nuisance contamination of FC estimates across scans

We quantify the nuisance contamination in inter-scan FC estimates by correlating the nuisance norm vectors $|n|_{\text{Vec}}$ with the vector of FC estimates FC_{Vec} for each seed-voxel pair. This approach is illustrated in Figure 5.1. In the top row, we first computed the correlations between a seed signal (e.g. PCC seed shown with red color) and the time series from every other voxel (lines with blue color) to form a seed-based correlation map for each scan (represented as a $N \times 1$ column vector with red color). We then repeated this for all 68 scans and concatenated the resulting seed-based correlation vectors (FC maps) to form a $(N \times 68)$ matrix as shown in the left hand side in the second row. Each row of this matrix corresponds to the FC estimates vector for a single seed voxel pair (an example row is shown with green color). We then computed the correlations between the FC estimates rows and nuisance norm vector (time series with black color) to form a nuisance contamination vector which can be reshaped into a 3D nuisance contamination map. The green colored square in the nuisance contamination vector corresponds to a single correlation coefficient obtained between the nuisance norm and the FC vector from a single seed voxel pair. This is also depicted on the nuisance contamination map with the green $(+)$ symbol.

We obtained nuisance contamination maps both before and after each regression and for different seed signals including the PCC, IPS, FEF, MOT, and AUD seeds. Note that these maps are **not** functional connectivity maps, but instead quantify the relations between seed-based FC estimates and nuisance norms across different scans.

5.2.6 Theoretical bound on ΔFC

In [231] we presented a theoretical expression for the difference $\Delta\text{DFC} = (\text{Post DFC} - \text{Pre DFC})$ between the dynamic FC (DFC) estimates obtained before and after nuisance regression

in seed correlation-based DFC studies. This mathematical theory applies for static FC studies as well by noting that a temporal sliding window in DFC analysis can be replaced by the whole scan duration. Thus, the following theoretical bounds apply for the difference between Pre FC and Post FC estimates obtained before and after nuisance regression:

$$|\Delta\text{FC}| = |\text{Post FC} - \text{Pre FC}| \leq 2\left(\frac{1-\sqrt{|n_O|^2/|n|^2}}{1+\sqrt{|n_O|^2/|n|^2}}\right). \tag{5.1}$$

Here, n is a single nuisance regressor time course represented in vector notation. The nuisance regressor can be decomposed as $n = n_I + n_O$, where n_I is an in-plane component that lies in the subspace spanned by a single seed-voxel pair x_1 and x_2 and n_O is the component orthogonal to this subspace.

The orthogonal nuisance fraction $0 \leq \frac{|n_O|^2}{|n|^2} \leq 1$ reflects the nuisance energy that lies in the orthogonal subspace and serves as a measure of orthogonality between the nuisance regressor and the seed-voxel pair (e.g. x_1 and x_2). If n_O becomes arbitrarily large then the fraction $\frac{|n_O|^2}{|n|^2} \to 1$ and $|\Delta\text{FC}| \to 0$. An example of this bound is provided in Figure 5.5 where a large orthogonal nuisance fraction for the HM, WM, CSF and HM+WM+CSF regressors impose a strict bound on $|\Delta\text{FC}|$ values forcing them to cluster close to 0.

Note that an exact value for the orthogonal nuisance fraction $\frac{|n_O|^2}{|n|^2}$ can be obtained when using a single regressor such as the GS. In the case of multiple regressors, an estimate of the orthogonal nuisance fraction and $|\Delta\text{FC}|$ can be obtained by using the first principal component (PC) of the multiple regressors as in [231]. This simple approximation enables us to understand the an approximate relation between $|\Delta\text{FC}|$ an the orthogonal nuisance fraction. When we analyze multiple regressors, we will provide the approximate orthogonal nuisance fraction values and will also show that regression with the first PC is a good approximation to performing multiple regression.

5.2.7 Significance testing of the relation between FC variations and nuisance norms across scans

We assessed the statistical significance of the relation between the FC estimates and nuisance norms across scans using non-parametric null testing. As the ordering of scans is not important, we formed null distributions by randomly permuting the scan ordering of FC estimates for each seed voxel pair and nuisance norm over 10,000 trials. We then correlated the resulting surrogate FC estimates with nuisance norms and obtained 10,000 null correlation values both before and after nuisance regression. We used the null distributions to assess the statistical significance of the correlations between the non-permuted FC estimates and nuisance norms.

5.3 Results

In this section, we show that variations in FC estimates across multiple scans are significantly correlated with geometric norms of various nuisance terms. We demonstrate that a considerable portion of the FC estimates still remain significantly correlated with nuisance norms even after nuisance regression. We make use of the theoretical findings in [231] to show that the inefficacy of nuisance regression for non-GS regressors such as HM, WM and CSF is largely due to the large orthogonality between nuisance regressors and the BOLD data within each scan. We further show that GSR can introduce negative GS norm fluctuations into the FC estimates.

In the first row of Figure 5.2 we show 3 examples of the relation between nuisance norms and FC estimates for 3 seed pairs: PCC&AUD, MPF&AUD and PCC&IPS. The column labels indicate both the type of nuisance norm and the specific seed-pair. The FC estimates in each column (blue lines, labeled as Pre FC) are significantly ($p < 10^{-3}$) correlated with various nuisance norms (black lines) before nuisance regression. The correlations obtained between the Pre FC estimates and HM, HM+WM+CSF, and GS norms in the first, second, and third columns are $r = 0.56$, $r = 0.58$, and $r = 0.82$, respectively. After nuisance regression, the Post FC

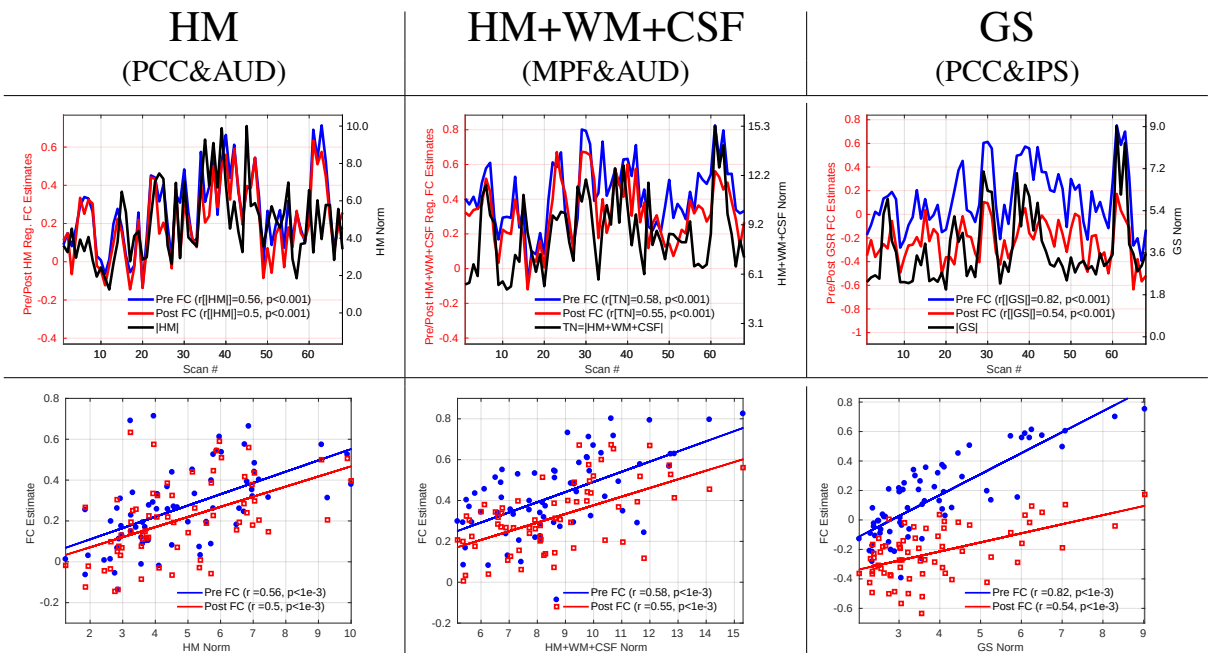


Figure 5.2: FC estimates obtained from 3 example seed pairs are significantly correlated ($p < 10^{-3}$) across scans with various nuisance norms before (blue lines) and after (red lines) nuisance regression. The type of nuisance norm and seed pair are indicated by the column labels. In the first row the FC estimate values are indicated on the left y-axis in each column, nuisance norm values are indicated on the right y-axis, and scan numbers are indicated by the x-axis. Before nuisance regression, the correlations between the Pre FC estimates and nuisance norms were $r = 0.56$, $r = 0.58$, and $r = 0.82$ for the HM, HM+WM+CSF and GS norms, respectively. After nuisance regression, the correlations between the Post FC estimates and nuisance norms were $r = 0.50$, $r = 0.55$, and $r = 0.54$ for the HM, HM+WM+CSF and GS norms, respectively. The first row serves as a nice visual demonstration of the similarity between the fluctuations in nuisance norms and FC estimates. The second row shows the relation between the FC estimates and nuisance norms using scatter-plots.

estimates are still significantly correlated with the nuisance norms with correlations of $r = 0.50$, $r = 0.55$, and $r = 0.54$ observed for HM, HM+WM+CSF, and GS norms, respectively. The second row shows the same relations using scatter-plots where the FC estimates are shown in the y-axes and respective nuisance norms are shown in the x-axes.

The results presented below generalize the relation between various nuisance norms and PCC-based FC estimates to include all PCC seed-voxel pairs. We provide the results for other seeds in the supplementary material and main text below.

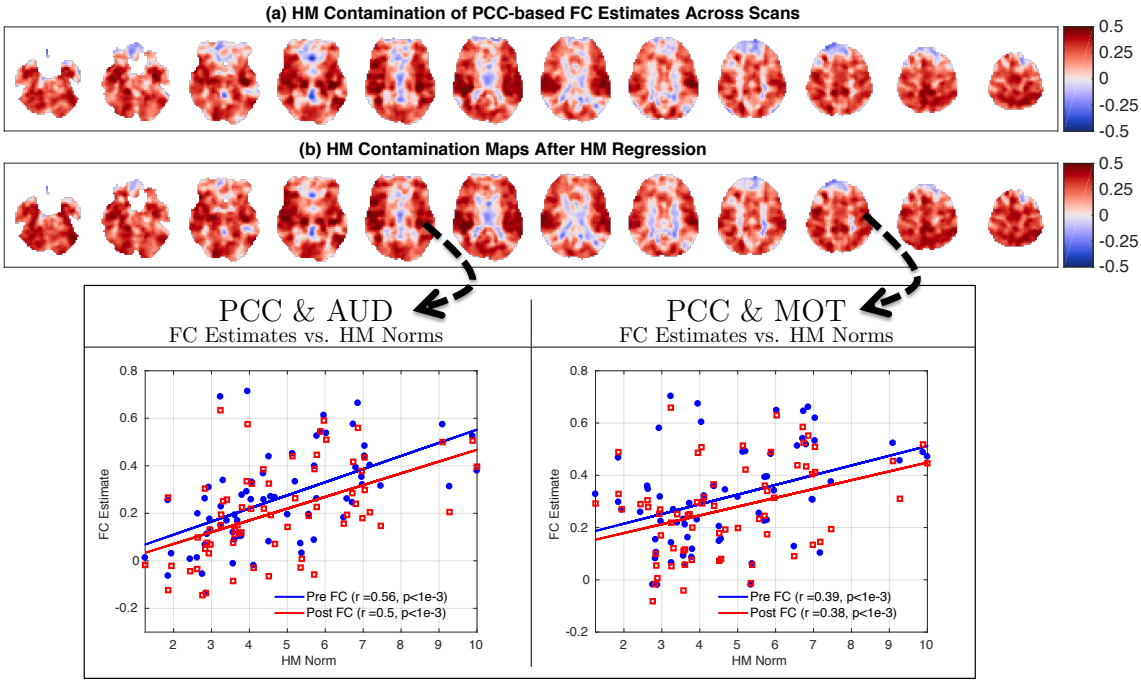


Figure 5.3: HM contamination maps obtained by correlating the HM norm with FC estimates before (a) and after (b) HM regression. These contamination maps in (a) and (b) are fairly similar to each other (cosine similarity $S = 0.98$) and both show widespread correlations between the HM norm and FC estimates across scans. This indicates that HM regression is largely ineffective in removing the relation between the HM norm and FC estimates. We show two example seed-pairs (PCC & AUD and PCC & MOT) at the bottom to illustrate the relation between HM norms and FC estimates for two different regions.

5.3.1 HM regression

In Figure 5.3a,b we show the PCC-based HM contamination maps before and after HM regression. These maps are very similar to each other (cosine similarity $S = 0.98$) and show widespread correlations between the HM norm and FC estimates across scans both before and after HM regression.

In the scatter plot shown in Figure 5.4 we plot the correlations obtained between the Post FC estimates and HM norm versus the correlations obtained between the Pre FC estimates and HM norm. The sideways histogram along the y-axis shows the distribution of correlation values obtained between the Post FC estimates and HM norm, which ranged from $r = -0.36$ to $r = 0.60$ with a mean of 0.27. The histogram along the x-axis at the bottom shows the correlations obtained

Table 5.1: Summary of the relationship between the FC estimates and nuisance norms computed for each seed-voxel pair before and after nuisance regressions. The seed signal and regression state (e.g. before and after) are indicated in the first and second columns, respectively. For each nuisance regression we report the % seed-voxel pairs for which there was a significant correlation between the FC estimates and nuisance norms. We also provide the mean and standard deviation of the percent variance explained in the FC estimates by the nuisance norms over the respective subsets of significant seed-voxel pairs.

		HM		HM+WM+CSF		GS		HM+WM+CSF+GS	
Seed	State	%Brain	%Var mean±SD	%Brain	%Var mean±SD	%Brain	%Var mean±SD	%Brain	%Var mean±SD
PCC	Pre FC	45	11±4	88	27±12	96	37±15	91	31±14
	Post FC	51	11±4	56	13±6	27	11±5	26	11±5
MOT	Pre FC	45	13±5	96	26±9	99	38±13	98	30±10
	Post FC	49	13±5	74	15±7	30	13±6	20	11±5
AUD	Pre FC	60	13±5	98	28±10	99	40±13	98	33±12
	Post FC	63	14±6	75	16±8	25	11±5	22	10±4
FEF	Pre FC	43	11±4	100	32±9	99	49±11	99	38±10
	Post FC	49	11±4	82	17±8	24	11±5	23	11±5
IPS	Pre FC	40	11±4	97	25±9	99	40±13	98	29±10
	Post FC	43	11±4	63	13±6	20	12±5	23	11±5

between the Pre FC estimates and HM norm, which ranged from $r = -0.34$ to $r = 0.61$ with a mean of 0.26.

These two correlation distributions were strongly related to each other ($r = 0.94$, $p < 10^{-3}$). The linear fit (blue line) between these correlation distributions (Slope= 0.91, Offset= 0.04) was very close to the line of unity (yellow dashed line). This indicates that HM regression had a very limited effect in removing the correlations between the Pre FC estimates and HM norms, and the correlations with the HM norm largely persisted after HM regression.

As shown in Table 5.1, 45% of the PCC-based Pre FC estimates were significantly correlated ($p < 0.05$) with HM norms. We further computed the mean and standard deviation of the percent variance explained by the HM norm for those significant 45% Pre FC estimates. The HM norm explained an average of 11%±4% (mean+SD) of the variance in those significant Pre FC estimates. After HM regression, 51% of the Post FC estimates were significantly correlated with the HM norm. The HM norm explained an average of 11%±4% (mean+SD) of the variance in those Post FC estimates, which is the same percent variance explained prior to HM regression.

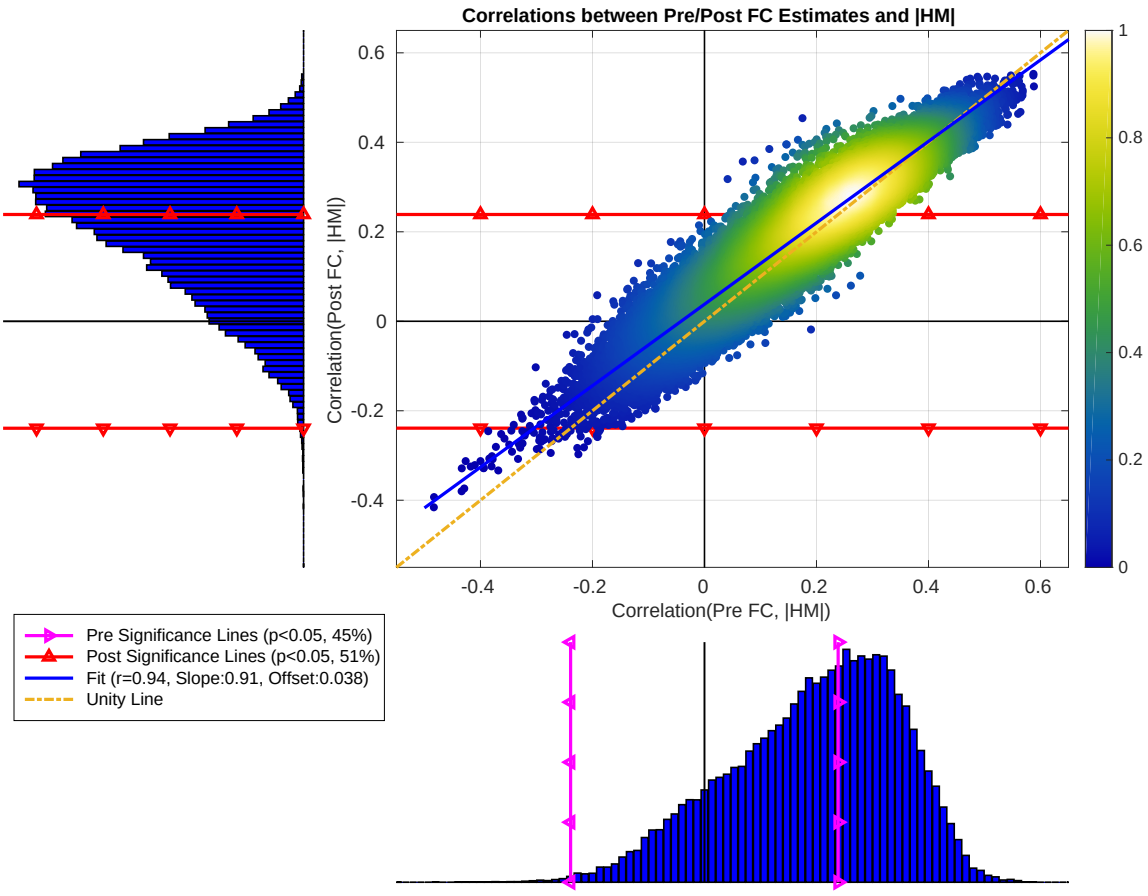


Figure 5.4: The correlations between the Post FC estimates and HM norm after regression versus the correlations obtained between Pre FC estimates and HM norm. These correlation distributions were significantly related ($r = 0.94$, $p < 10^{-6}$) to each other. The linear fit (blue line, Slope= 0.91 and Offset= 0.04) between the two correlation distributions was fairly close to the line of unity (dashed yellow line). At the bottom histogram, the correlations between the Pre FC estimates and HM norms ranged from $r = -0.34$ to $r = 0.61$ with mean 0.26. We superimposed the significance lines ($p < 0.05$) on this histogram using magenta lines with triangles (labeled as Pre Significance). Based on the significance line, 45% of the Pre FC estimates were significantly correlated with HM norms. In the sideways histogram to the left, the correlations between the Post FC estimates and HM norms ranged from $r = -0.36$ to $r = 0.60$ with mean 0.27, where significance lines are shown with red lines with triangles. 51% of the Post FC estimates were significantly correlated with HM norms after HM regression.

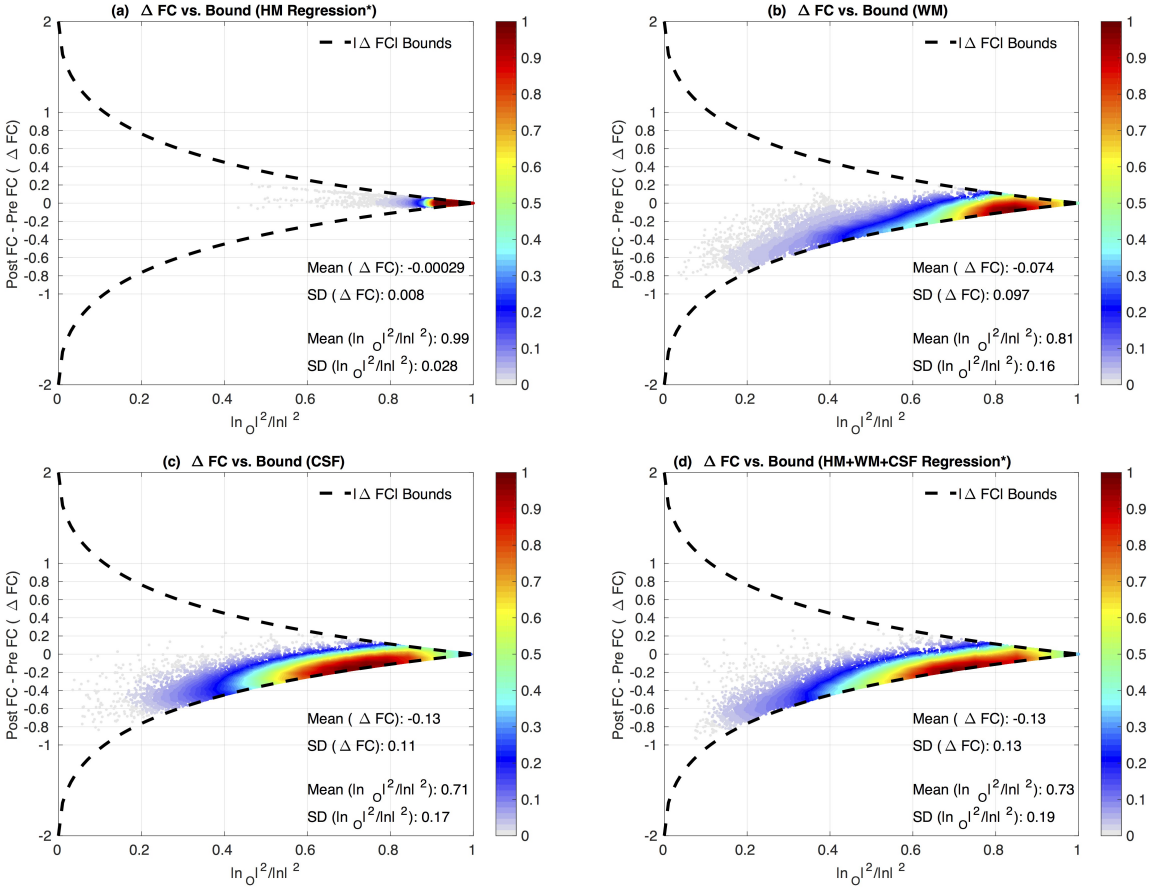


Figure 5.5: ΔFC versus orthogonal nuisance fraction $\frac{|n_O|^2}{|n|^2}$ for HM ($^*1^{\text{st}}$ PC), WM, CSF and HM+WM+CSF ($^*1^{\text{st}}$ PC) regressor groups. Each point in these plots represent the values for a single scan and a single seed-voxel pair. In (a) we plot the ΔFC versus the orthogonal nuisance fraction for HM regression. We see that the points are strictly clustered on the right hand side around a mean orthogonal nuisance fraction value of 0.99. This imposes an extremely tight bound on the ΔFC values before and after regression and HM regression resulting in a negligible effect on the FC estimates (mean $\Delta\text{FC} = 0$). In (b), (c) and (d) we plot the ΔFC versus the orthogonal nuisance fraction for the WM, CSF, and 1^{st} PC of HM+WM+CSF regressors, respectively. These regressors are largely orthogonal to BOLD data with $\frac{|n_O|^2}{|n|^2} > 0.70$ which results in fairly tight bounds on the ΔFC values. Thus, effects of these regressor groups on the FC estimates are very limited with $|\Delta\text{FC}| \leq 0.13$.

Ideally, HM regression should fully remove the HM norm variance from the Post FC estimates. However, from Figure 5.4 we see that the Post FC estimates are still correlated with HM norm nearly as much as the Pre FC estimates are correlated with HM norm. In Figure 5.5a we show the theoretical limitation of HM regression. Each point in this plot shows the ΔFC (difference between the Pre FC and Post FC estimates) versus the orthogonal nuisance fraction $|n_O|^2/|n|^2$ for a single seed-voxel pair in a single scan. The HM regressors are largely orthogonal to most seed voxel pairs across scans with a mean orthogonal fraction of $|n_O|^2/|n|^2 = 0.99$. Due to the large orthogonal fraction, the theoretical bounds force the ΔFC values to be clustered around a mean value of 0. Since HM regression cannot really alter the Pre FC estimates, the slope of the linear fit in Figure 5.4 remains close to the line of unity.

Note that as in [231] we used the first principal component (PC) across all head motion measurements for the ΔFC plot in Figure 5.5a. We provide the ΔFC plot after performing multiple HM regression in Supplementary Figure 7.18, which shows that ΔFC values are still largely clustered around zero (mean $\Delta FC = -0.037$), roughly within the tight theoretical bounds.

We present the HM contamination maps and their respective correlation distributions before and after HM regression for other seeds (e.g. MOT, AUD, FEF, and IPS) in Supplementary Figures 7.20 to 7.23. These were similar to the PCC-based results discussed above. Additionally, Table 5.1 summarizes the significance results of the relation between nuisance norms and FC estimates for other seeds.

5.3.2 HM+WM+CSF regression

In Figure 5.6a,b we show the HM+WM+CSF contamination maps both before and after nuisance regression. These maps show widespread correlations between the nuisance norm and FC estimates across scans. In panel (b) there is a visible reduction in the positive correlation values (i.e. red regions) after regression with a slight increase in anti-correlations (i.e. blue regions start to appear).

In Figure 5.6c the correlations between the Pre FC estimates and nuisance norm ranged from $r = -0.26$ to $r = 0.81$ with mean 0.46. After nuisance regression, the correlation between the Post FC estimates and nuisance norm ranged from $r = -0.42$ to $r = 0.66$ with mean 0.23. We found a strong linear relation between the two correlation distributions ($r = 0.82$, $p < 10^{-3}$) with a slight increase in significant anti-correlations residing below the lower red significance line. The linear fit between the two correlation distributions was close to the line of unity with a large slope (0.83) and a small negative offset (-0.15).

As noted in Table 5.1, 88% of the Pre FC estimates were significantly correlated ($p < 0.05$) with the nuisance norm. The nuisance norm explained an average of 27%$\pm$12% (mean+SD) of the variance in those significant correlations. After regression, 56% of the Post FC estimates were still significantly correlated with the nuisance norm, and the nuisance norm explained an average of 13%$\pm$6% (mean+SD) of the variance in those Post FC estimates.

In Figure 5.5b,c we show the ΔFC plots for the WM and CSF regressors. We found that the mean orthogonal nuisance fraction for the WM and CSF regressors were still relatively large with $|n_O|^2/|n|^2 = 0.81$ and $|n_O|^2/|n|^2 = 0.71$, respectively. Due to the theoretical bounds associated with the large orthogonal fractions, the ΔFC points were clustered close to 0 with mean ΔFC values of -0.08 and -0.13 for the WM and CSF regressors, respectively. In Figure 5.5d we show the ΔFC plot for the HM+WM+CSF regression using the first PC of those regressors. This revealed a mean orthogonal fraction of $|n_O|^2/|n|^2 = 0.73$ (similar to individual WM and CSF regressors) with mean ΔFC around -0.13 and ΔFC values bounded below by the theoretical bound. In Supplementary Figure 7.19, we show the 'true' ΔFC values (e.g. not approximated with the 1$^{\text{st}}$ PC) were within the theoretical bounds with a mean $\Delta\text{FC} = -0.17$, similar to Figure 5.5d.

To summarize, although HM+WM+CSF regression partially reduced the correlations between the nuisance norm and Pre FC estimates (as compared to HM regression alone), a large fraction of seed-voxel pairs (56%) exhibited significant correlations between the nuisance

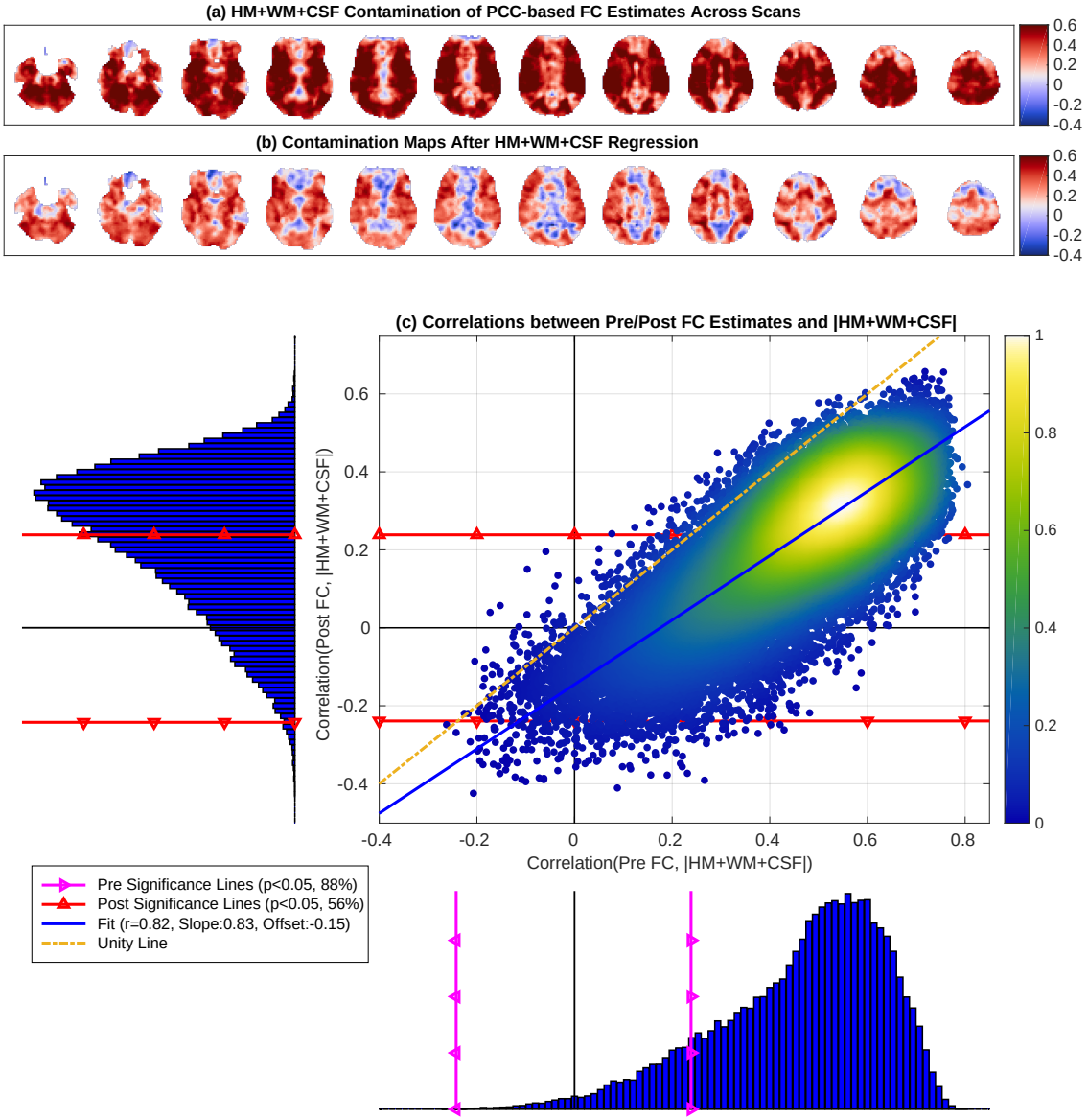


Figure 5.6: Nuisance contamination maps obtained by correlating the total HM+WM+CSF norm with the FC estimates before (a) and after (b) nuisance regression. In (c) the correlations between the Pre FC estimates and the nuisance norm (x-axis) ranged from $= -0.26$ to $r = 0.81$ with mean 0.46. The correlations between the Post FC estimates and nuisance norm (y-axis) ranged from $= -0.42$ to $r = 0.66$ with mean 0.23. There was a strong linear relation between the two correlation distributions (linear fit $r = 0.82$, $p < 10^{-6}$). The linear fit between the two correlation distributions was close to the line of unity with a large slope (0.83) and a small negative offset (-0.15).

norm and Post FC estimates. The limited efficacy of nuisance regression reflects the fact that the theoretical bounds on ΔFC estimates in Figure 5.5d are still very tight since the orthogonal nuisance fraction is large and close to 1.0.

We present the HM+WM+CSF contamination maps and their respective correlation distributions before and after nuisance regression for other seeds (e.g. MOT, AUD, FEF, and IPS) in Supplementary Figures 7.24 to 7.27. These were similar to the PCC-based results discussed above. Additionally, Table 5.1 summarizes the relations between nuisance norms and FC estimates for other seeds.

5.3.3 GS regression

In Figure 5.7a,b we show GS contamination maps both before and after GSR. In panel (a) we observe strong positive correlations (red regions) between the GS norm and Pre FC estimates. The distribution of these correlations is shown in Figure 5.8, with values ranging from $r = -0.22$ to $r = 0.87$ with mean 0.57. Across the sample, 96% of these correlations were significant ($p < 0.05$) and 99% were positive with a strong left skew $S = -0.96$. The GS contamination map after GSR is given in Figure 5.7b. This map shows brain regions consisting of both positive (red regions) and anti-correlations (blue regions) between the GS norms and Post FC estimates.

In Figure 5.7b we observe that GSR introduces anti-correlations between the GS norm and FC estimates which were not present prior to GSR in panel (a). An example seed pair is provided at the bottom of Figure 5.7 as a scatter plot. The FC estimates obtained between the PCC&MPF seeds on the left-hand side are positively correlated with the GS norm before GSR and anti-correlated with the GS norm after GSR. As shown along the y-axis of Figure 5.8 the correlations between the GS norm and Post FC estimates ranged from $r = -0.61$ to $r = 0.66$ centered around a mean 0 with standard deviation 0.21. 49% of these correlations were positive (remaining 51% were negative) and 27% were significant ($p < 0.05$).

In Figure 5.8 we found a significant linear relationship ($r = 0.59, p < 10^{-6}$) between

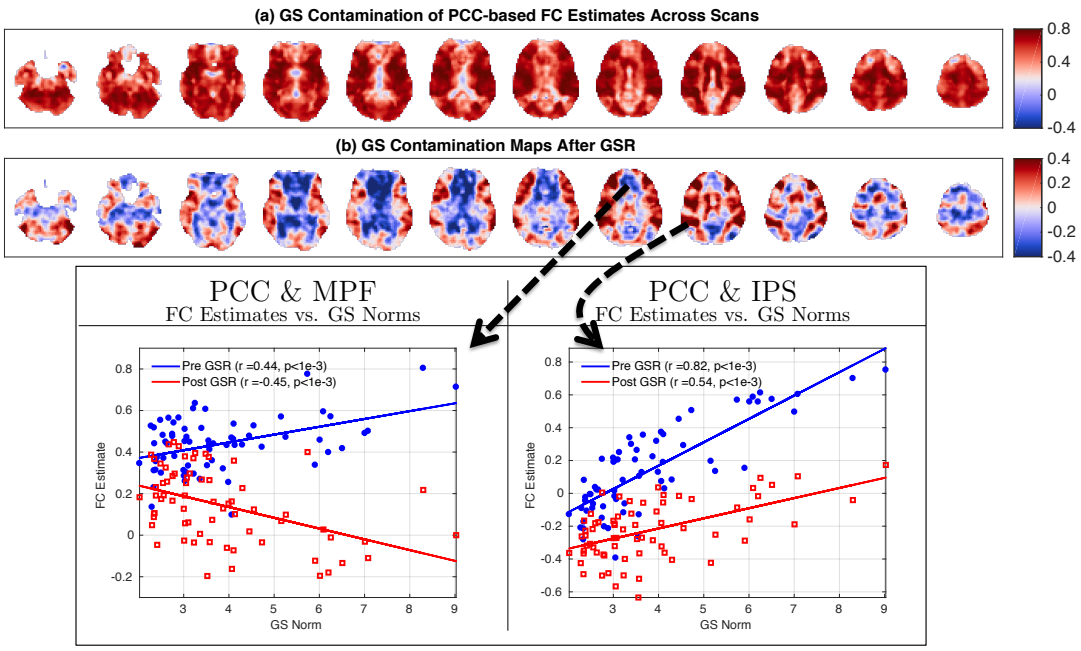


Figure 5.7: GS contamination maps obtained both (a) before and (b) after GSR. In panel (a) the contamination maps show positive and significant correlations (99% positive and 96% significant) between Pre FC estimates and GS norm. In (b) the contamination map after GSR involves both positive (49%) and negative (51%) correlations between the GS norms and Post FC estimates. 27% of these correlations were significant ($p < 0.05$). In the seed-pair examples shown at the bottom, the FC estimates between the PCC & MPF across scans are positively correlated with GS norms before GSR ($r = 0.44$) but become anti-correlated after GSR ($r = -0.45$). The FC estimates between PCC & MPF seed pair are positively correlated with the GS norm both before and after GSR with correlation values of $r = 0.82$ and $r = 0.54$, respectively.

the correlations obtained between the Post FC estimates and GS norm versus those correlations obtained before GSR. The linear fit (shown with the blue line) is fairly parallel to the line of unity (green dashed line) with a slope of 0.77 and has a very large negative offset -0.44. The negative offset indicates that GSR produces a strong negative shift in the correlations obtained between the Post FC estimates and GS norms when compared to those correlations obtained before GSR. This results in significant anti-correlations ($p < 0.05$) between the GS norms and Post FC estimates. The significant anti-correlations between the GS norms and Post FC estimates are shown with the points residing below the bottom significance line (red line with triangles) within the dark blue zone in the scatter plot. The points residing above the top significance line in the dark red

zone are significant positive correlations remaining after GSR. These correlation values remain significant and positive despite the negative offset.

These results indicate that effects GS norm can largely remain in the Post FC estimates after GSR. GSR can result in "residual" positive correlations between the GS norm and Post FC estimates and can also "introduce" significant anti-correlations between the GS norm and Post FC estimates due to the negative shift in correlation values. To understand why GSR results in both positive and negative correlations between the GS norm and Post FC estimates we start by defining two zones as illustrated in Figure 5.8 with dark red and dark blue zones labeled by *Zone I* and *Zone II*.

Zone I (GSR undercorrects): GSR is limited in removing GS norm from FC estimates

In Figure 5.8, red background colors (both light and dark colors) show those correlations between GS norm and FC estimates that are **positive** both before and after GSR. In this case, "GSR undercorrects" and is unable to fully remove the positive correlations between the GS norm and FC estimates.

We define Zone I (dark red background) to include the **significant** positive correlations ($p < 0.05$) between GS norm and FC estimates remaining after GSR. In Figure 5.9a we plot the average FC estimates (across Zone I seed-voxel pairs) versus the GS norm before GSR (blue line and dots) and after GSR (red line and diamonds), where each point corresponds to a single scan. The correlation between the average Pre FC and GS norm is $r = 0.87$. The correlation between the average Post FC and GS norm after GSR is still large $r = 0.77$. This indicates that GSR is unable to fully remove the GS norm effects from the Pre FC estimates.

The effect of GSR is fully characterized by the average ΔFC values (black squares) shown in Figure 5.9a. The average ΔFC is anti-correlated with the GS norm ($r = -0.81$) where the slope of the linear fit (black line) is -0.068. The slope (0.11) of the linear fit (blue line) between the Pre FC estimates and GS norm has a greater magnitude than the slope (-0.068) for ΔFC. Since

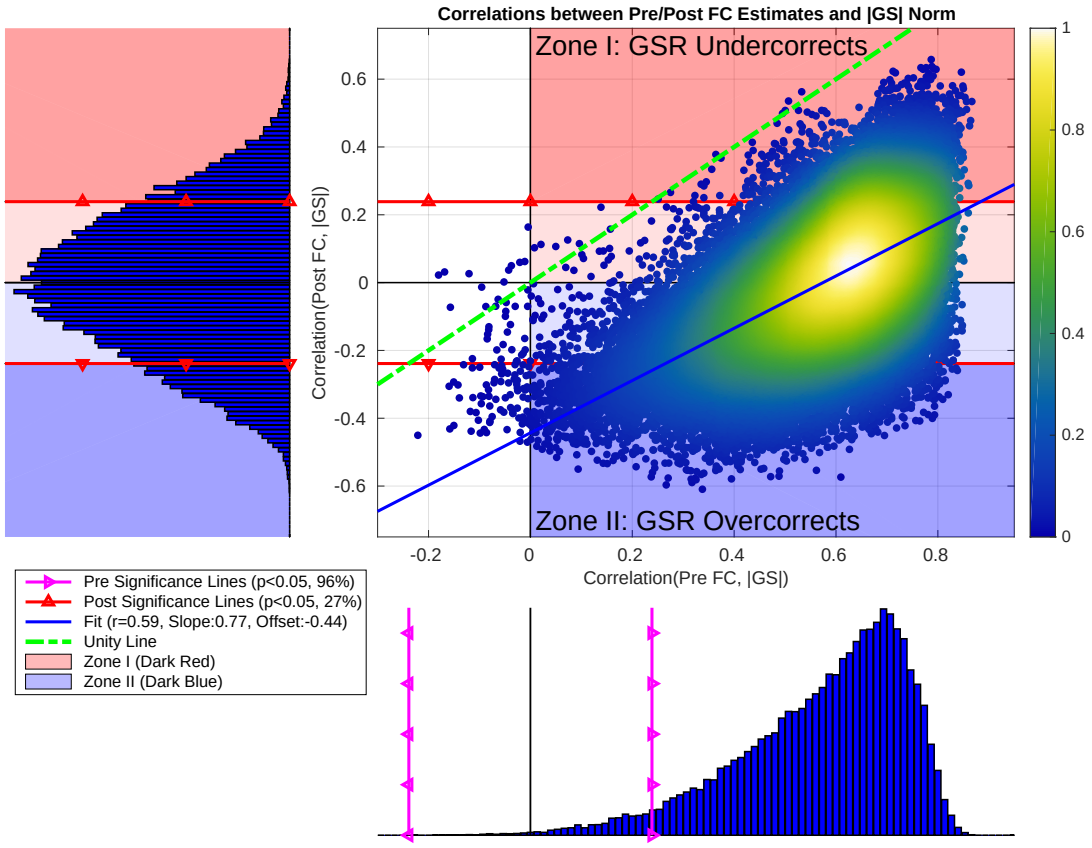


Figure 5.8: In the scatter plot, we show the correlations between the Post FC estimates and GS norms plotted versus the correlations obtained between the Pre FC estimates and GS norms. We found a significant linear relationship between the two correlation distributions ($r = 0.59, p < 10^{-6}$). The linear fit (blue line) had a slope of 0.77 and a very large negative offset -0.44. In the bottom histogram, the correlations between the Pre FC estimates and GS norms ranged from $r = -0.22$ to $r = 0.87$ with mean 0.57 and standard deviation 0.16. 99% of the correlations were positive with a strong left skew (Skewness $= -0.96$) and 96% of the correlations were significant ($p < 0.05$) (correlations residing to the right of magenta line with triangles). In the sideways histogram on the left, the correlations between the Post FC estimates and GS norm ranged from $r = -0.61$ to $r = 0.66$ with mean 0 and standard deviation 0.21, where significance lines ($p < 0.05$) are shown with red lines with triangles. We found 27% of the Post FC estimates to be significantly correlated with HM norms after regression. The labels on the scatter plot ("Zone I" (red zone) and "Zone II" (blue zone)) refer to two main cases where GSR fails to remove correlations between the GS norm and Post FC estimates.

Post FC = Pre FC + ΔFC, this difference in the slope magnitudes results in a positive slope for the relation between Post FC estimates and GS norm (linear fit shown with red line).

In Figure 5.10a we plot the ΔFC values versus orthogonal nuisance fraction for Zone I, where each point represents the values obtained for a single seed-voxel pair in a given single scan. The ΔFC values are strongly related to the orthogonal nuisance fraction with $r = 0.93$ ($p < 10^{-3}$), where the linear fit shown with the green line is almost tangent to the lower theoretical bound. We observe that ΔFC values are clustered around the lower theoretical bound with mean value

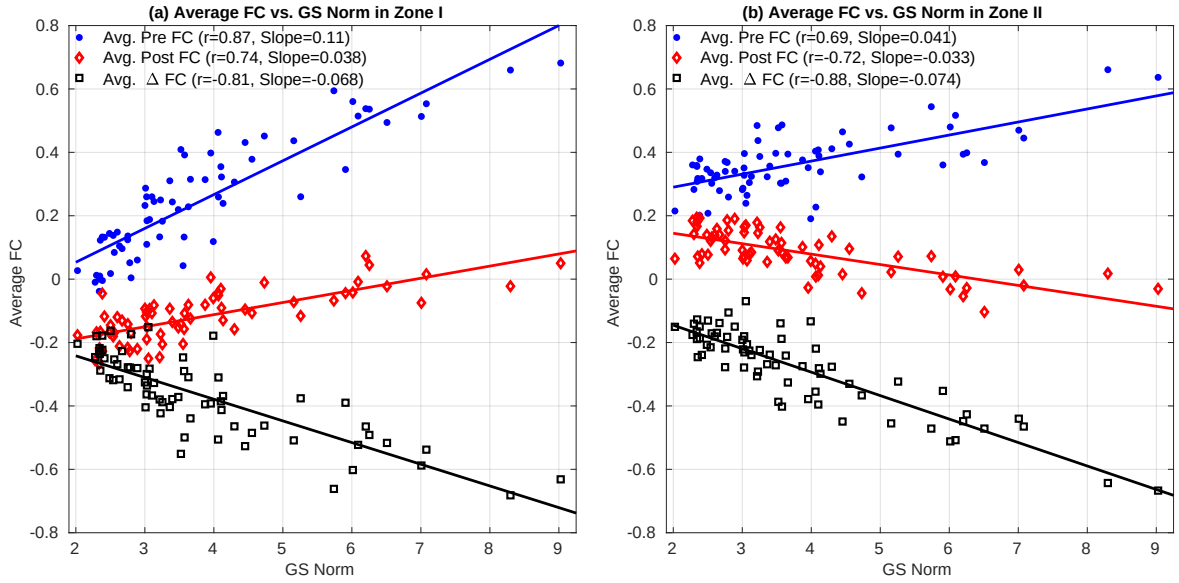


Figure 5.9: Average FC estimates and average ΔFC versus GS norm in Zone I (a) and Zone II (b). In panel (a) GSR adds the average ΔFC (shown with black color) to the average Pre FC estimates (blue color). The resulting average Post FC estimates (red color) are still positively correlated with the GS norm $r = 0.74$. This is because the slope for average ΔFC (-0.068) has a smaller magnitude than the slope for average FC estimates ($+0.11$). Thus, the slope for average Post FC estimates remains positive $0.038 \approx +0.11 - 0.068$. In panel (b) the average Pre FC estimates for Zone II are significantly larger than those in Zone I in panel (a) $p < 10^{-3}$ (paired two-tailed t-test) with a smaller standard deviation (0.09 in Zone II as compared to the 0.19 in Zone I). Thus, the slope for average Pre FC estimates in Zone II has smaller magnitude $+0.041$ as compared to Zone I. The slope (-0.074) of average ΔFC in (b) is similar to the slope (-0.068) of average ΔFC in (a). Since the average Post FC estimates are simply the sum of average Pre FC points and average ΔFC points, the average Pre FC estimates are dominated by the larger ΔFC effect and the relation between the average Post FC estimates and GS norm exhibits a negative correlation ($r = -0.72$, $p < 10^{-3}$).

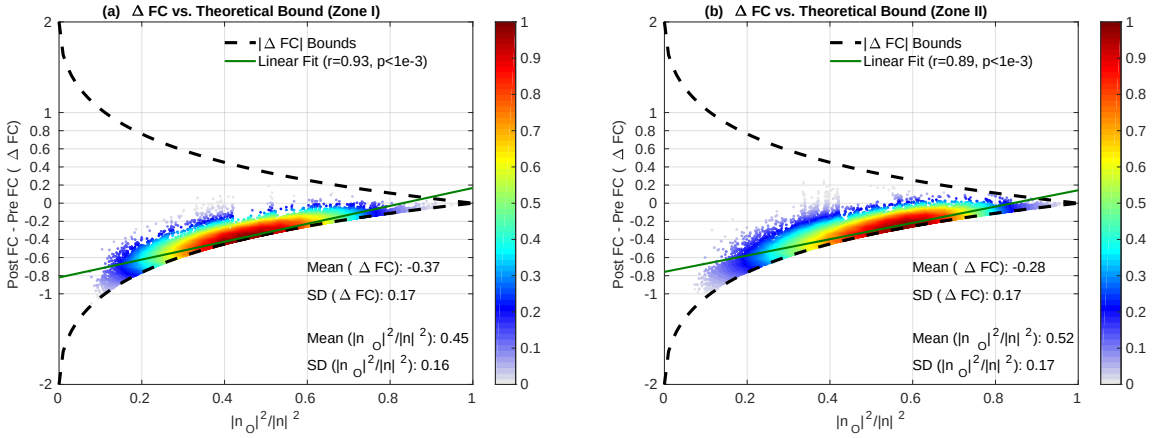


Figure 5.10: ΔFC versus orthogonal nuisance fraction $\frac{|n_O|^2}{|n|^2}$ for GS regression in (a) Zone I and (b) Zone II. In both zones, ΔFC values are clustered close to the lower theoretical bound with mean ΔFC -0.37 for Zone I and -0.28 for Zone II with similar mean orthogonal nuisance fractions of 0.45 and 0.52. Moreover, the standard deviation for ΔFC is 0.17 which is much smaller than the mean ΔFC values. The linear fits shown with solid green lines show that ΔFC values are strongly correlated ($p < 10^{-3}$) with the orthogonal nuisance fractions with $r = 0.93$ and $r = 0.89$ for Zone I and Zone II, respectively. Both linear fits are extremely close to the lower theoretical bound.

of -0.37 with a relatively smaller standard deviation -0.17. This is consistent with the average ΔFC values in Figure 5.9a. The bound plot reveals that effects of GSR on the FC estimates in Zone I are bounded below by the theoretical curve such that the magnitudes of ΔFC cannot exceed the lower bound. As a result, the magnitude of ΔFC in Figure 5.9a cannot increase as rapidly as the Pre FC estimates.

Zone II (GSR overcorrects): GSR introduces anti-correlation between GS norm and FC estimates

The light and dark blue backgrounds in Figure 5.8 show those correlations between GS norm and FC estimates that are **positive** before GSR and are **negative** after GSR. In this case, "GSR overcorrects" by first removing the positive correlations between GS norm and Pre FC estimates and then introducing anti-correlations between the GS norms and Post FC estimates. The brain regions which are red in Figure 5.7a but are blue in Figure 5.7b correspond to this case.

These regions broadly include the default mode network, medial-prefrontal cortex and thalamus regions, when using the PCC as the seed signal.

In Figure 5.8 we define Zone II (dark blue background) to include **significant** negative correlations ($p < 0.05$) between GS norm and FC estimates after GSR. In Figure 5.9b we plot the average FC estimates versus the GS norm for Zone II. The slope (0.041) of the relation between Pre FC estimates and GS norm in Zone II in Figure 5.9b is weaker in magnitude as compared to the slope (-0.074) of the relation between the ΔFC and GS norm. Since Post FC = Pre FC + ΔFC, a stronger negative relation between ΔFC and GS norm dominates over the weaker relation between Pre FC estimate and GS norm and the average Post FC estimates become significantly anti-correlated with the GS norm ($r = -0.72$, $p < 10^{-3}$).

The average Pre FC estimates in Zone II in Figure 5.8b are significantly greater than those in Zone I in panel (a) $p < 10^{-3}$ (paired two-tailed t-test). This means that BOLD signals residing in Zone II exhibit greater intrinsic similarity (in average) to the seed signal as compared to those in Zone I. To verify this, in Figure 5.11a we plot the correlations between the seed signal (PCC) and the average BOLD signal in Zone II versus the correlations between the seed signal and the average BOLD signal in Zone I. The average BOLD signal in Zone II was significantly ($p < 10^{-3}$ paired two-tailed t-test) more correlated with the seed signal as compared to Zone I.

In Figure 5.11b we plot the correlations between the GS and average BOLD signal in Zone II versus the correlations between the GS time course and average BOLD signal in Zone I. We found no significant difference ($p = 0.45$ paired two-tailed t-test) between those correlation values for Zone I and Zone II. This means that the GS regressor is similar to the raw time courses in both zones. This is consistent with the fact that the average ΔFC in both zones were fairly similar to each other ($r = 0.94$) in Figure 5.9, with similar linear fits (green lines) between ΔFC and orthogonal nuisance fractions ($r = 0.93$ for Zone I and $r = 0.89$ for Zone II).

The difference in Zone I and Zone II behavior is caused by the fact (1) GSR imposed a similar ΔFC in both zones as shown in Figure 5.10a,b. This was largely because that the GS time

course were similarly related to the average BOLD signals in Zone I and Zone II (Figure 5.11b) and thus the orthogonal nuisance fraction fractions between the GS and seed-voxel pairs were similar in Figure 5.10 for both zones. (2) the underlying differences in FC estimates in Zone I and Zone II were significantly different as shown in Figure 5.9. FC estimates in Zone II were significantly greater as compared to those in Zone I and exhibited a weaker dependency on the GS norm (slope 0.041, $r = 0.69$) in Figure 5.9b as compared to the Zone I (slope 0.11, $r = 0.87$). GSR fails to model this difference in intrinsic similarities and the relation (slope) between FC estimates and GS norm.

Figure 5.11a shows that brain regions that are intrinsically similar to the seed signal are more likely to involve Zone II effect. Brain regions that suffer either Zone I limitation (remaining positive correlation between the GS norm and Post FC estimates) or Zone II limitation (an

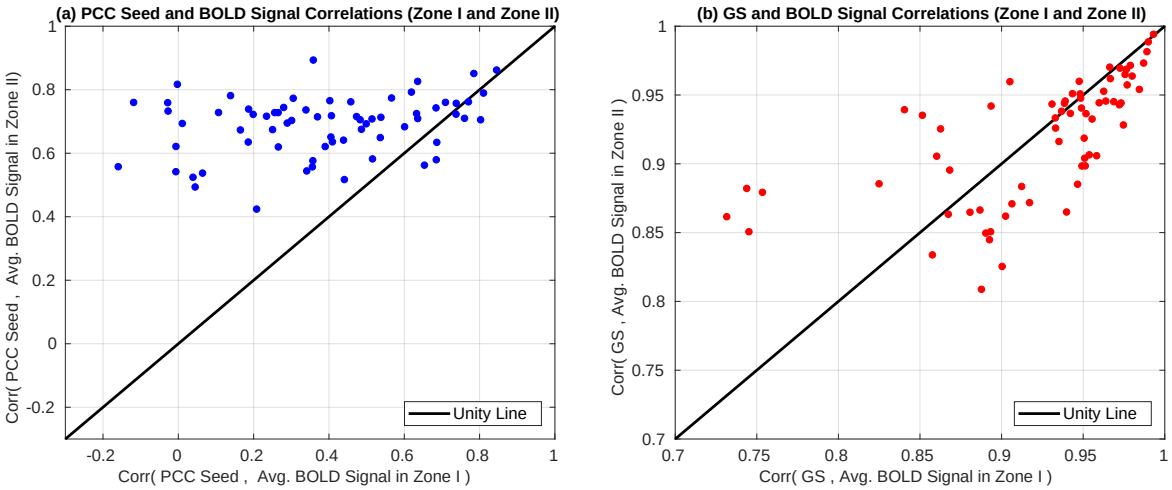


Figure 5.11: In (a) we show the correlations obtained between the PCC seed and the average BOLD signal in Zone II versus the correlations obtained between the PCC seed and the average BOLD signal in Zone I. The average BOLD signal in Zone II was significantly ($p < 10^{-3}$ paired two-tailed t-test) more correlated with the PCC seed signal as compared to Zone I. This means that brain regions that are more similar to the seed signal are more likely to have the Zone II limitation of GSR. In (b) we show the correlations between the GS the average BOLD signal in Zone II versus the correlations between the GS and the average BOLD signal in Zone I. We found no significant difference ($p = 0.45$ paired two-tailed t-test) between Zone I and Zone II correlations. This indicates that in average, the difference in Zone I and Zone II effects are not dependent on the relation between the GS and the underlying BOLD time courses in those zones.

introduced negative correlation between the GS norm and Post FC estimates) depends on the seed signal used. These regions can be determined by looking at the GS contamination maps after GSR for other seeds as provided in the Supplementary Figures 7.28 to 7.31. Table 5.1 generalizes the significance results of the relation between the GS norm and FC estimates.

5.4 Discussion

We have shown that inter-scan variations in FC estimates can be strongly and significantly correlated with the geometric norms of various nuisance measurements. We found that the relationship between the FC estimates and nuisance norms can persist even after performing multiple nuisance regression. We used the mathematical framework developed in [231] to describe the limitations of nuisance regression in cleaning inter-scan FC estimates and demonstrated that the empirical results were largely within the theoretically predicted bounds.

For non-GS regressions, we found nuisance regression to be largely ineffective. This was because non-GS regressors were largely orthogonal to BOLD data (as shown in Figure 5.5 with large values for orthogonal nuisance fractions). The large orthogonality imposed tight bounds on the difference between the FC estimates obtained before and after nuisance regression, which limited the removal of nuisance norms from the FC estimates. This limitation is largely consistent with our previous findings for dynamic FC measures in [231].

We introduced *nuisance contamination maps* to visualize the effects of nuisance norms on the FC estimates both before and after nuisance regression. These maps serve as a useful tool in analyzing the spatial location and extent of nuisance contamination present in seed-based FC estimates across scans. In Figures 5.3 and 5.6, the FC estimates obtained between the PCC seed and BOLD signals from other brain regions involved strong correlations with the HM and HM+WM+CSF norms both before and after nuisance regression. As summarized in Table 5.1, depending on the seed signal, 40-99% of the seed-voxel pairs in the brain exhibited significant

correlations with nuisance norms before nuisance regression. After nuisance regression, FC estimates from 40-80% of the seed-voxel pairs still exhibited significant correlations with nuisance norms. This reveals that non-GS nuisance regressions were largely ineffective in cleaning the nuisance norms from FC estimates across scans.

We found that GSR removed a large portion of the GS norm fluctuations from the FC estimates across scans. Unfortunately, depending on the seed signal, a considerable portion (20-30%) of the FC estimates still remained significantly correlated with the GS norm. We provided a detailed analysis of the residual GS norm effects after GSR. We found that GSR can 'undercorrect' FC estimates by leaving a positive GS norm residual in the FC estimates (Zone I limitation) or it can overcorrect by introducing negative GS norm fluctuations (Zone II limitation). If GSR is not performed then a majority portion (>95%) of the FC estimates significantly depend on the GS norm. This means that differences in GS norms will largely determine the variations in FC metrics across scans. On the other hand, if GSR is performed then the dependency of the FC estimates to the GS norm will be weaker, however there is a risk of introducing negative correlations between the GS norm and inter-scan FC estimates.

The difference in Zone I and Zone II behavior is caused by the fact (1) GSR imposed a similar ΔFC in both zones as shown in Figure 5.10a,b, and (2) the underlying differences in FC estimates in Zone I and Zone II were significantly different as shown in Figure 5.9. The former is due to the fact that GSR does not take into account the underlying differences in FC estimates since GS is similarly correlated with the underlying BOLD signals in both zones as shown in Figure 5.11b.

Lastly, GS contamination maps in Figure 5.7b resembled FC maps appearing after GSR [50, 53]. However the contamination maps are fundamentally different from FC maps as contamination maps show the relation between the GS norm and FC estimates. So why do they look similar to each other? In Figure 5.11a we have shown that the average BOLD signal in Zone II exhibited a greater intrinsic similarity to the seed signal as compared to Zone I. Moreover,

in Figure 5.9b for Zone II, GSR resulted in significant negative correlations between the FC estimates. This shows that brain regions that are functionally more similar to each other (e.g. involving greater correlation values) are more likely to involve a negative correlation between their FC estimates and GS norms after GSR. An example was provided in Figure 5.7b where a negative GS norm is introduced to the default-mode network (DMN) regions (blue regions).

5.4.1 Studies Investigating FC Across Scans

A multitude of fMRI studies have investigated the differences in FC measures between disease populations and healthy controls. Examples of such studies include investigations on Alzheimer's disease [66, 67, 68], Parkinson's disease [69], depression [70], schizoprenia [71], dementia [235], and amyotrophic lateral sclerosis (ALS) [236, 237]. These studies have all used a form of nuisance regression to remove possible effects of nuisance terms on their analyses.

Our results strongly indicate that these studies can be confounded by various nuisance effects both before and after nuisance regression. Therefore future FC studies should consider the possible relations between nuisance terms (norm) and FC estimates. To provide an example, in [69] the authors observed an increased subthalamic nucleus-motor cortex FC in Parkinson's disease as compared to healthy controls [69]. They made an attempt to analyze the effect of HM on their group analysis by averaging (across scans) the HM parameters in the Parkinson's and healthy control groups and did not find a significant difference in the averaged HM parameters. However, they did not consider the effects of HM directly on the FC estimates. We have shown that HM norms can be significantly positively correlated with the FC estimates across scans both before and after HM regression. It is highly possible that FC estimates will be larger in with Parkinson's group due to increased motion as compared to healthy controls, and this is regardless whether or not HM regression is performed. We strongly believe that any future study that compares FC estimates across subject groups should also consider the potential link between FC estimates and nuisance norms.

5.4.2 Nuisance Effects in FC Studies

Understanding the relationship between nuisance terms and the BOLD signal has been a challenging task in fMRI [61, 64, 65, 159, 167, 185, 217, 238, 239, 240, 241, 242]. Performance of nuisance removal techniques is usually characterized by their effects on raw BOLD signals on a per-scan basis e.g. percentage of variance removed from raw BOLD signals. However, investigating the effects of nuisance terms and the efficacy of nuisance regression is often neglected, especially when considering a collection of scans.

fMRI studies rely on linear regression to remove nuisance effects from BOLD data by either using direct nuisance measurements [61, 64, 159, 167, 185, 240] or using a set of nuisance regressors derived by data-driven methods [65, 239, 241]. Unfortunately, so far none of these methods have become a gold-standard approach for cleaning fMRI data [167]. We believe this is largely because the underlying relations between nuisance effects and BOLD data are more complicated. As we have shown in this work and in [231] a simple general linear model (GLM) has theoretical limitations in cleaning nuisance effects from FC metrics.

Our results indicate that as long as a regressor exhibits a moderate to large orthogonality to the BOLD data then nuisance regression will suffer from theoretical limitations. This is regardless of whether a regressor is obtained using data-driven methods or measured directly. Therefore, we believe that a proper analysis of nuisance effects on the fMRI data should also consider the relation between nuisance norm on FC measures.

5.4.3 Nuisance Norm Regression

The present study shows the inefficacy of nuisance regression in removing the relation between nuisance norms and FC estimates across scans. Given this limitation, a reasonable approach is to perform nuisance norm regression (NNR) on the inter-scan FC estimates, as originally proposed in [231] for cleaning sliding-window dynamic FC estimates. This involves

projecting out the norms of nuisance measures from the inter-scan FC estimates. Though this approach would ensure orthogonality between the nuisance norms and inter-scan FC estimates it might not be suitable for cleaning FC estimates.

From a mathematical point of view, one potential issue with NNR is the leverage effects [171, 172]. In computing the regression fit coefficient between the nuisance norms and inter-scan FC estimates, scans with larger nuisance norms will have greater leverage on the regression effect. This means that a scan with a large nuisance norm value will determine the amount of nuisance norm removed from other scans. Therefore, although NNR might help reduce the effects of nuisance norms on FC estimates across scans, we believe NNR should not be used in studies that investigate FC estimates on a per-scan basis. Future work and new approaches are of interest to understand how to best remove nuisance norm effects from the inter-scan FC estimates.

5.4.4 Vigilance Effects

There is growing evidence that vigilance fluctuations are responsible for a considerable portion of the resting-state fMRI data [78, 160, 165, 243, 244]. Vigilance fluctuations can account for approximately 10-20% variance in the whole-brain GS [78]. Our results in Table 5.1 reveal that inter-scan FC estimates strongly depend on the differences in GS norms across scans. It is also known that the global signal takes on large values (either positive or negative) when there are large deviations in vigilance levels [78, 244]. This implies that scans exhibiting larger fluctuations in their vigilance levels (this will correspond to a larger GS norm) can also have larger values in their FC estimates. Therefore, future investigations into the effect of vigilance on FC estimates will be very useful.

5.4.5 Anti-correlated Networks

Prior work has demonstrated that GSR places a constraint on the sum of seed-based correlations (FC estimates) for each scan [50, 52]. GSR increases the spatial extent and strength of anti-correlations present in fMRI data, especially between the default-mode (DMN) and task-positive networks (TPN) [51, 53, 157]. It was also shown that anti-correlations between the DMN and TPN were inherently present in a temporal subset of fMRI data [54]. Still use of GSR remains a highly debated issue. We found that GSR largely removed the GS norm-related fluctuations from the FC estimates, however about 20-30% of the FC estimates after GSR remained significantly correlated with GS norm.

To understand the behavior of GSR on inter-scan FC estimates, we divided the brain into two spatially non-overlapping regions (or zones). In Zone I, GS norm remained significantly positively correlated with inter-scan FC estimates even after performing GSR. In Zone II, significant negative correlations between the GS norm and inter-scan FC estimates were introduced.

For the PCC seed, the TPN region resided in Zone I and DMN region resided in Zone II. In both zones, GSR produced a large negative offset (mean -0.37) in the FC estimates. As a result of this negative offset, the mean FC estimates in Zone I in Figure 5.9a became predominantly negative, however the FC estimates in Figure 5.9b still remained largely positive despite the negative offset. This negative offset in the correlation values obtained after GSR is consistent with [52]. Despite the negative offset, we revealed that GSR also highlighted the differences in the underlying similarity between Zone I and Zone II (e.g. between DMN and TPN).

5.5 Conclusion

We provided a detailed examination of nuisance effects and the efficacy of nuisance regression on the variability of FC estimates across scans. We have shown that the norms of various nuisance terms can be strongly and significantly correlated with the variations in inter-scan

FC estimates both before and after nuisance regression. We found that non-GS regressions (HM, HM+WM+CSF) were largely ineffective in reducing the correlations between nuisance norms and FC estimates. We showed that though GSR removed a large fraction of the GS norm fluctuations from the FC estimates, a considerable portion $(20-30\%)$ of the FC estimates still remained significantly correlated with the GS norm. We identified two distinct effects of GSR and have shown that the process can introduce significant negative correlations between the GS norm and FC estimates.

This work stresses an important issue in the interpretation of FC measures. Most FC studies implicitly assume that the effect of nuisance terms on FC estimates are minimized by nuisance regression in the pre-processing stage. Our findings strongly suggest that the interpretation of FC measures should also consider the effect of nuisance terms (e.g. nuisance norms) on the relationship between BOLD signals (e.g. correlation-based FC measures). If the relationship between FC estimates and nuisance norms are not considered, the differences in FC estimates may be incorrectly interpreted as meaningful effects, when in fact they may be largely due to differences in nuisance activity. This is especially true in FC studies comparing FC measures between disease groups or treatment conditions.

As we have shown, linear regression-based nuisance removal approaches have important theoretical limitations in cleaning the FC estimates. Therefore, future FC studies will greatly benefit from newer nuisance removal approaches that avoid the potential issues highlighted in this work.

5.6 Acknowledgments

Chapter 5 is, in full, a reprint of material in submission under the title "*Nuisance Effects in Inter-scan Functional Connectivity Estimates Before and After Nuisance Regression.*" I was the primary author of this paper. This work was partially supported by NIH grant R21MH112155

and a UC San Diego Frontiers of Innovation Scholars Program (FISP) Project Fellowship.

Chapter 6

Conclusion

This dissertation focused on the problems of sparse signal recovery (SSR) and linear regression-based nuisance removal in functional MRI. In the first part of the dissertation, we developed a Bayesian algorithm called Rectified Sparse Bayesian Learning (R-SBL) to recover non-negative sparse solutions in under-determined systems of linear equations. R-SBL was able to outperform the state-of-the-art non-negative SSR approaches in a variety of experiments. R-SBL was particularly robust to the structure of the dictionary in coherent, low-rank, ill-conditioned or non-negative settings.

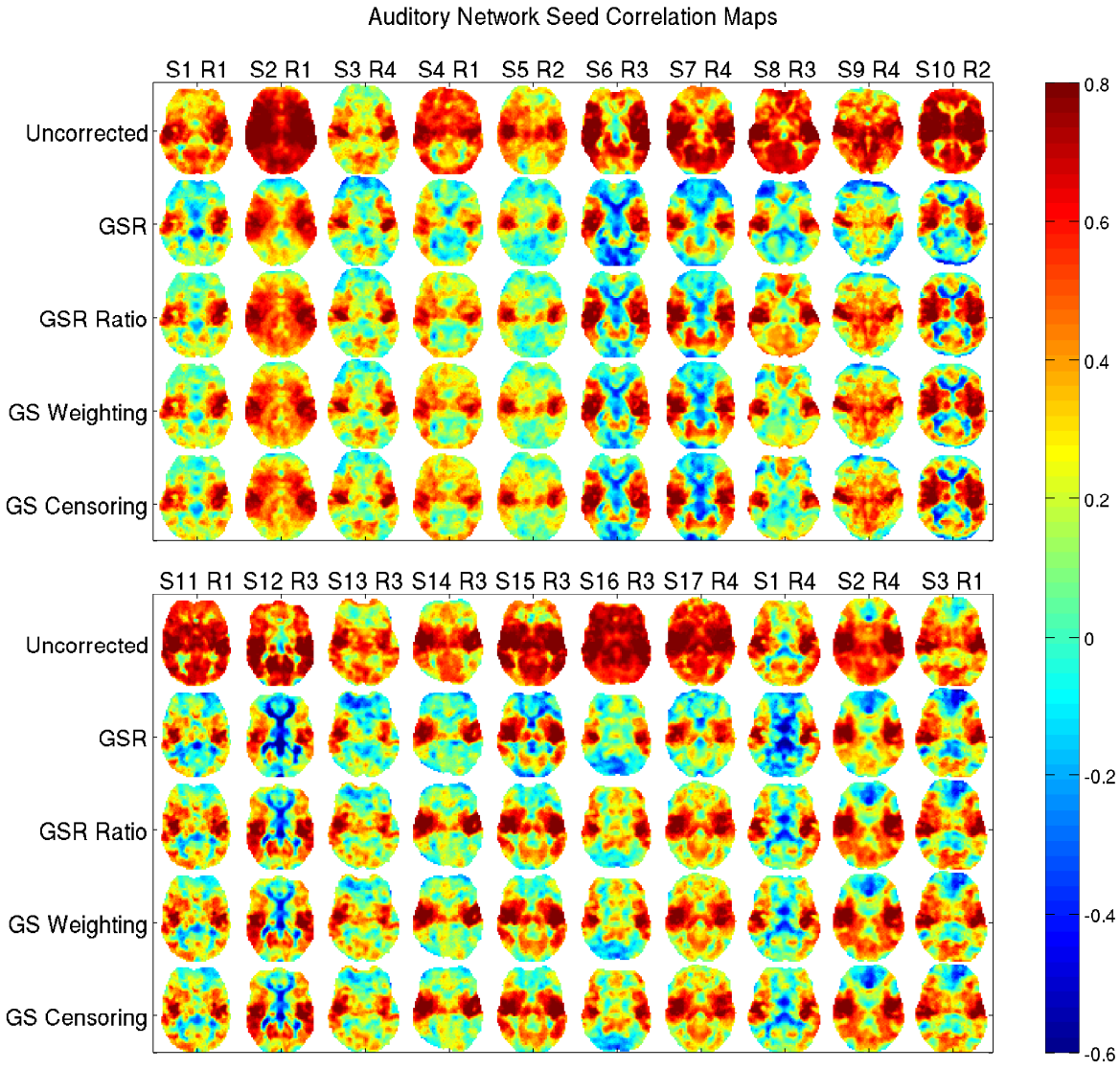
In the second part of the dissertation, we analyzed the effects of nuisance regression on the raw fMRI data and functional connectivity estimates (correlation values) computed across fMRI data. In Chapter 3, we proposed temporal downweighting and censoring approximations to model the effects of global signal regression (GSR). These approximations revealed that the anti-correlated functional brain networks in the brain were intrinsic to the fMRI data and were not an artifact introduced by GSR.

In Chapters 4 and 5, we have shown that functional connectivity (FC) estimates obtained between pairs of fMRI signals across different brain regions were significantly correlated with the norms of various nuisance terms in both static and dynamic functional connectivity (FC) studies.

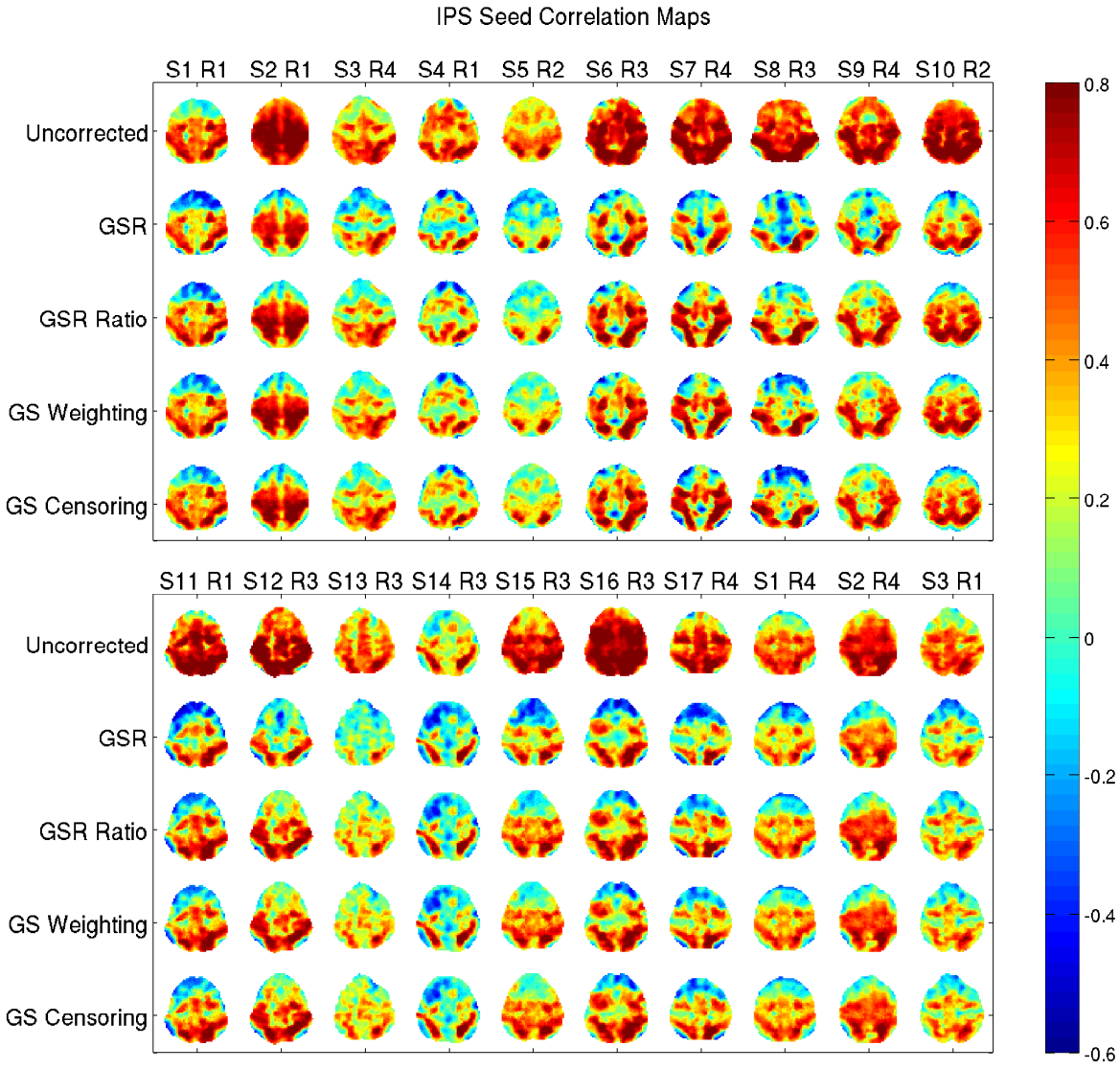
We found that linear regression methods were ineffective in removing the significant correlations observed between the nuisance norms and FC estimates. We derived a mathematical bound of linear regression on the difference between correlation coefficients obtained before and after regression. This mathematical bound restricted nuisance regression in removing the nuisance norm effects from FC estimates.

Chapter 7

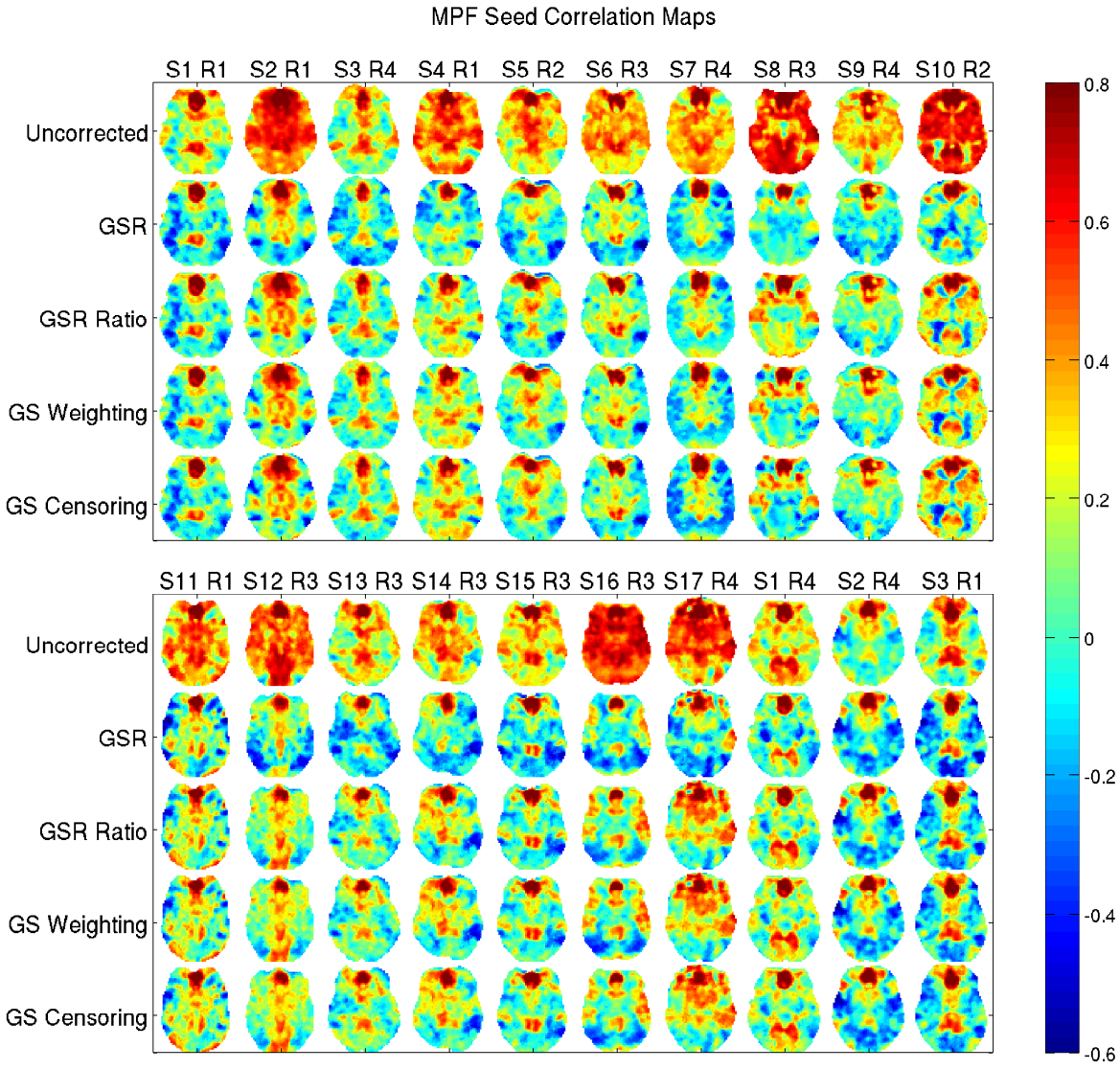
Supplementary Material



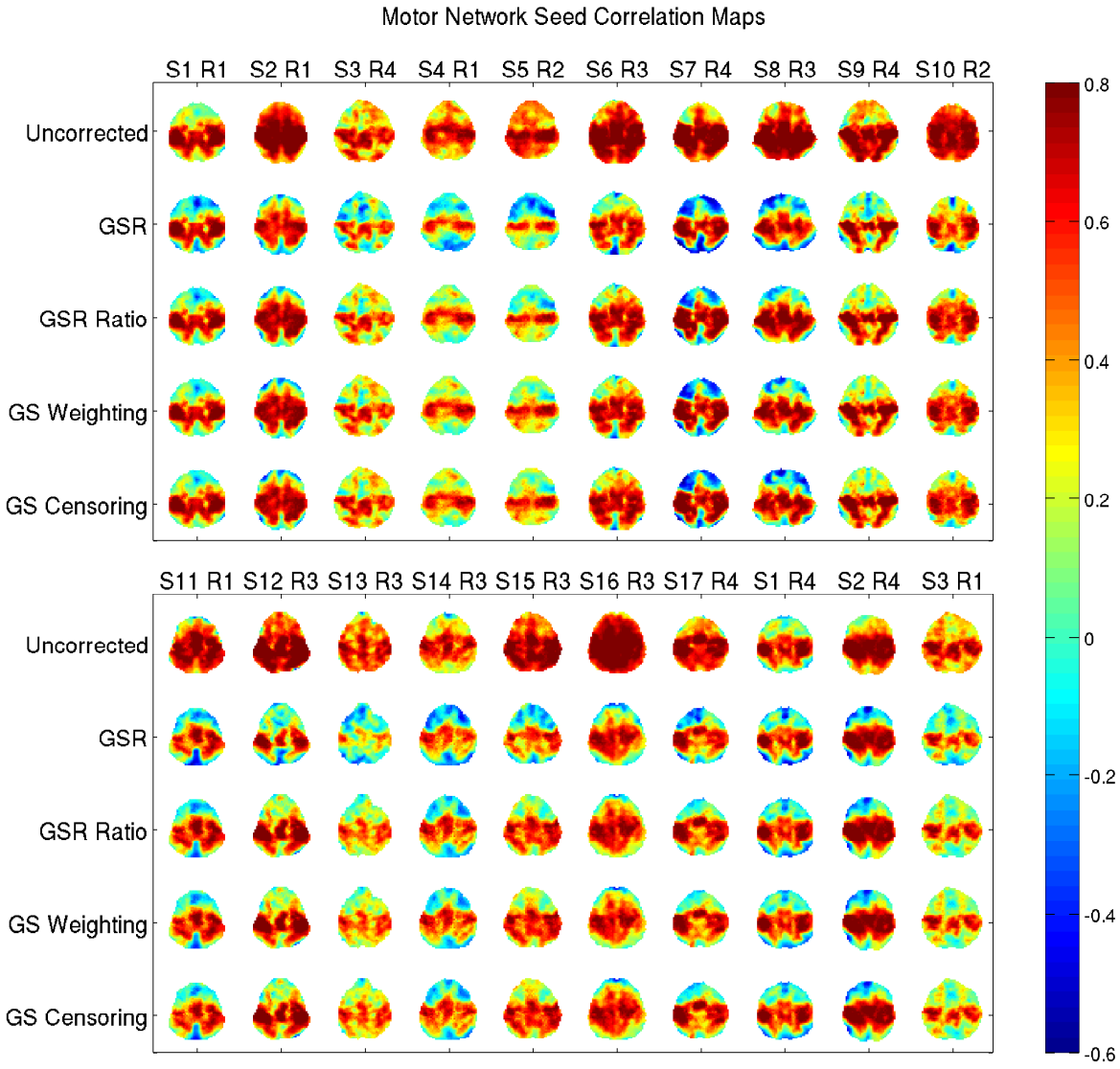
Supplementary Figure 7.1: Auditory network seed correlation maps obtained before GSR, after GSR, and after application of GSR ratio weighting, GS weighting, and GS censoring. For GS censoring a threshold of $g_C = 0.50$ was used.



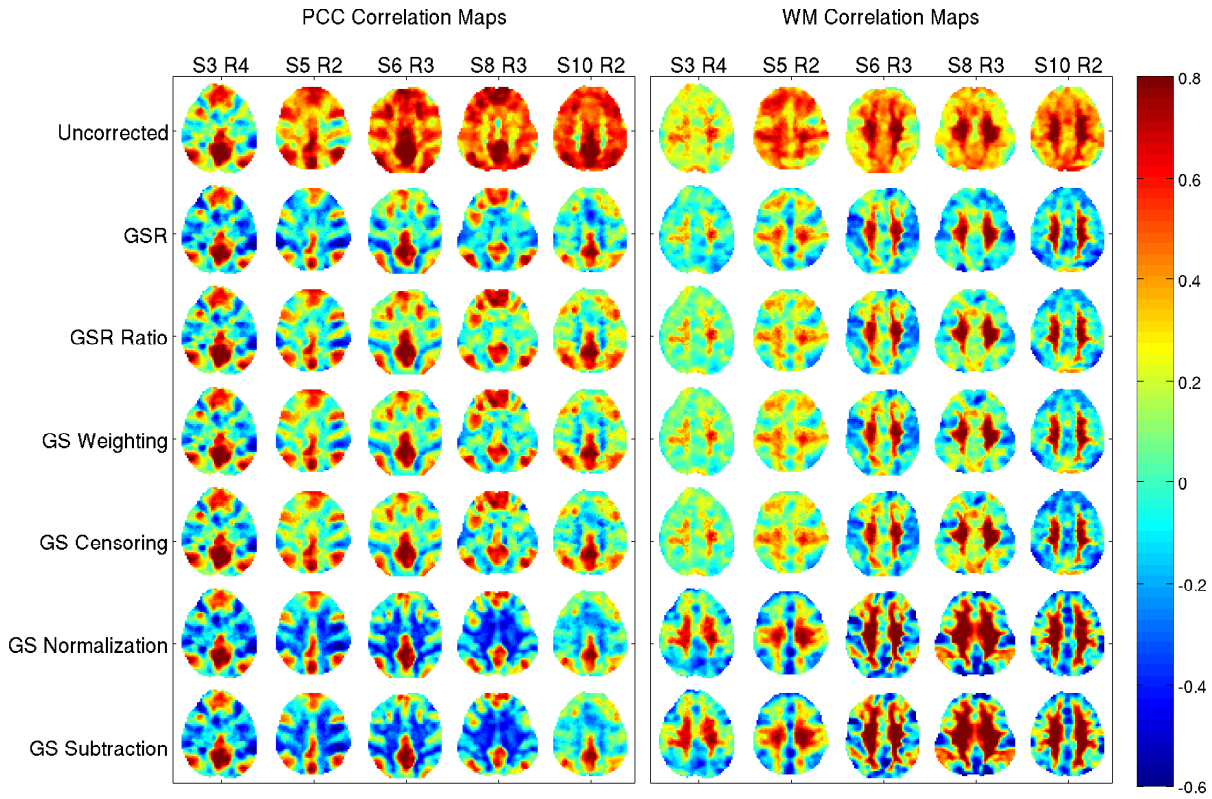
Supplementary Figure 7.2: IPS seed correlation maps obtained before GSR, after GSR, and after application of GSR ratio weighting, GS weighting, and GS censoring. For GS censoring a threshold of $g_C = 0.50$ was used.



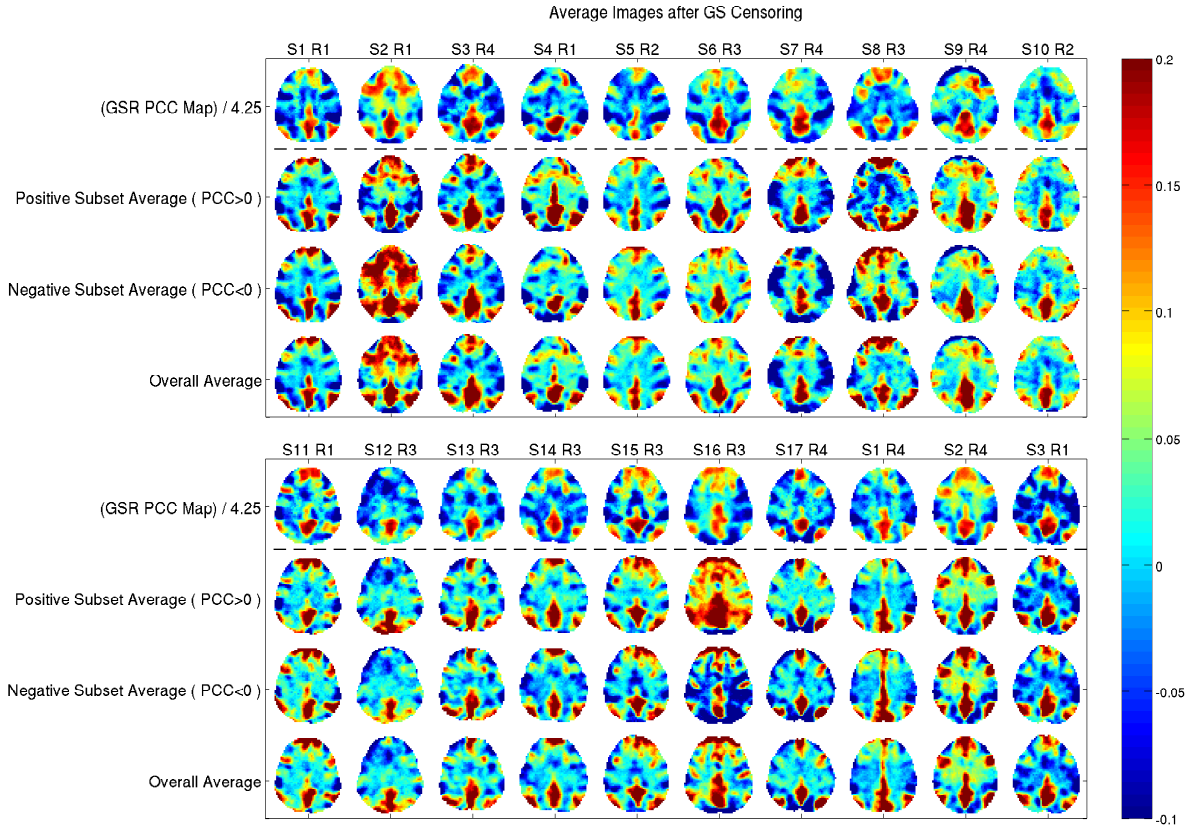
Supplementary Figure 7.3: MPF seed correlation maps obtained before GSR, after GSR, and after application of GSR ratio weighting, GS weighting, and GS censoring. For GS censoring a threshold of $g_C = 0.50$ was used.



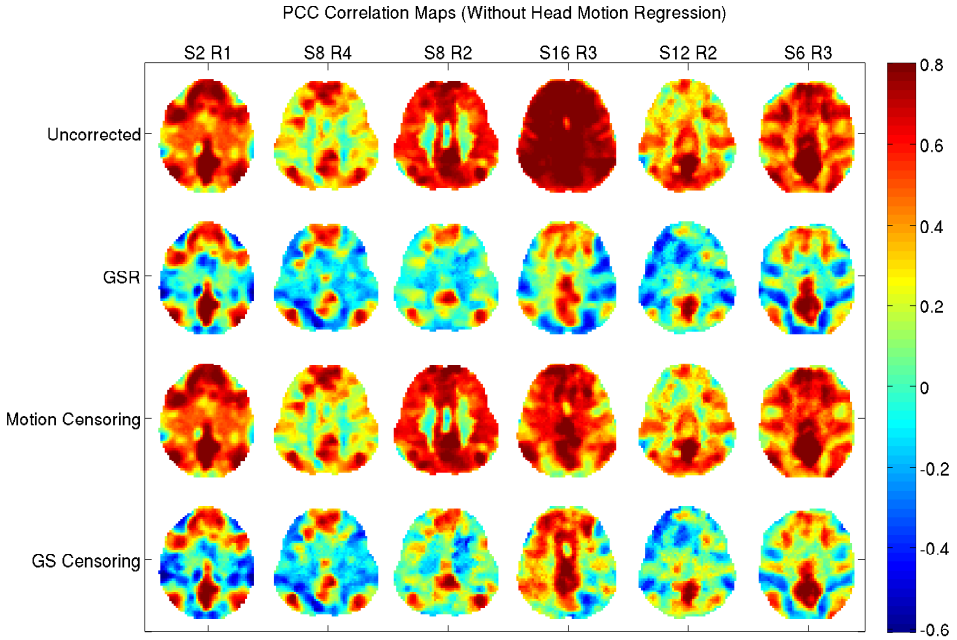
Supplementary Figure 7.4: Motor network seed correlation maps obtained before GSR, after GSR, and after application of GSR ratio weighting, GS weighting, and GS censoring. For GS censoring a threshold of $g_C = 0.50$ was used.



Supplementary Figure 7.5: Posterior-cingulate cortex (PCC) and white-matter (WM) seed correlation maps obtained prior to GSR and after the application of GSR, GSR ratio weighting, GS weighting, GS censoring, GS normalization (intensity stabilization), and GS subtraction. Consistent with the approximation shown in Appendix B, GS normalization and GS subtraction yield nearly identical maps for all scans. For some scans, the GS normalization and subtraction maps show significant differences with the maps obtained after GSR and the other GS weighting approaches, indicating clear differences in the approaches. These differences are greater when using a WM seed, where the fit between the WM seed voxel time series and the GS is expected to be lower, such that subtraction of the GS is not a good approximation for subtraction of an optimally scaled version of the GS. When the GS normalization and GSR approaches provide similar PCC correlation maps, there is a reasonably good fit between the PCC and GS time series, such that subtraction of the GS is a good approximation to subtraction of a scaled version of the GS (see also Appendix B)



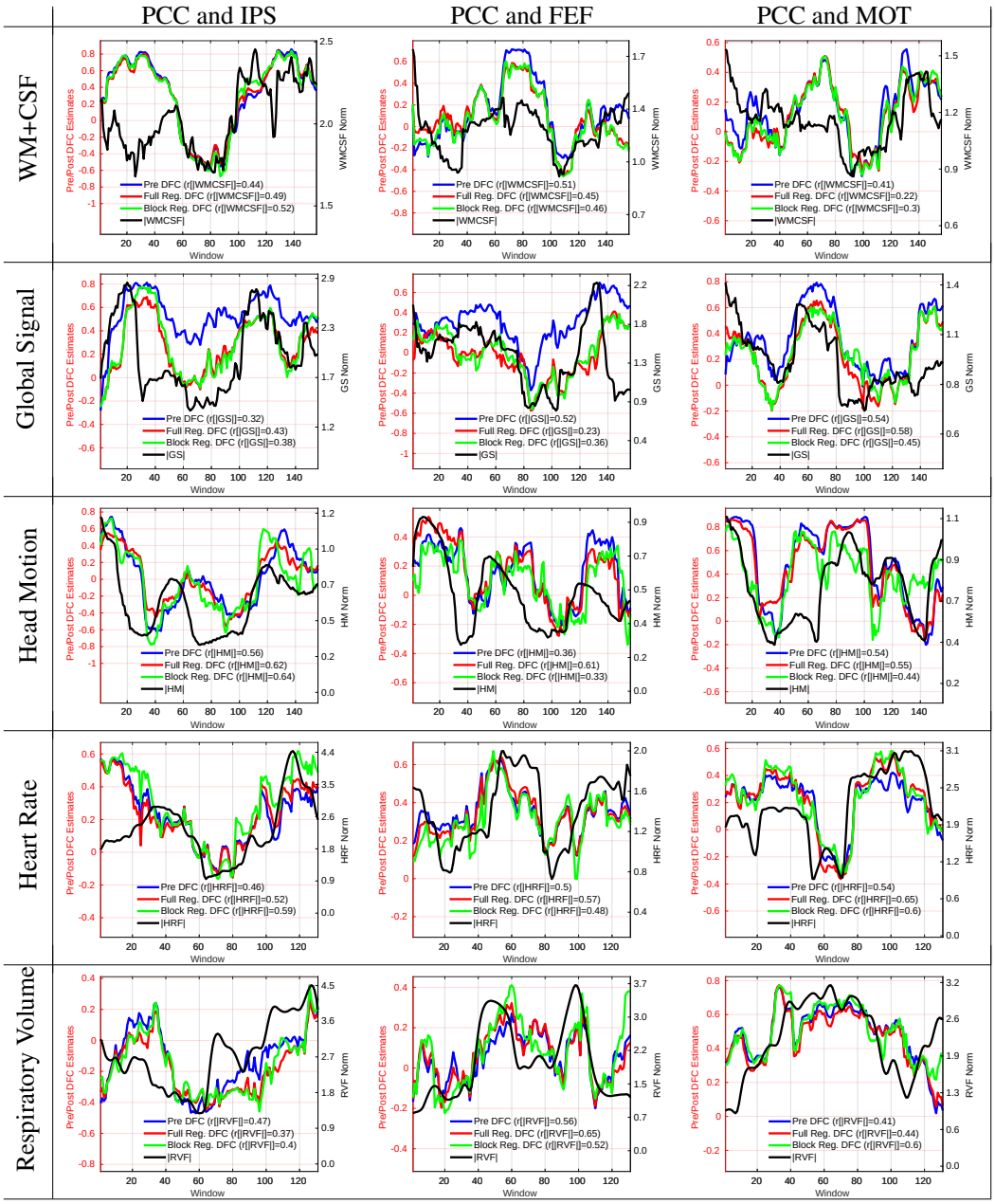
Supplementary Figure 7.6: The spatial patterns in the PCC correlation maps after GSR are similar to the patterns found in the average of the images retained by the GS censoring approach (with $g_C = 0.5$). To account for the occurrence of positive and negative values of the PCC seed signal, the set of retained images was divided into subsets with positive and negative PCC signal values and averages were computed for each subset. The average for the negative subset was multiplied by -1.0 to account for the sign change. The overall averages of the two subsets (after multiplying the negative subset by -1.0) are also shown. The PCC correlation map is scaled for display purposes.



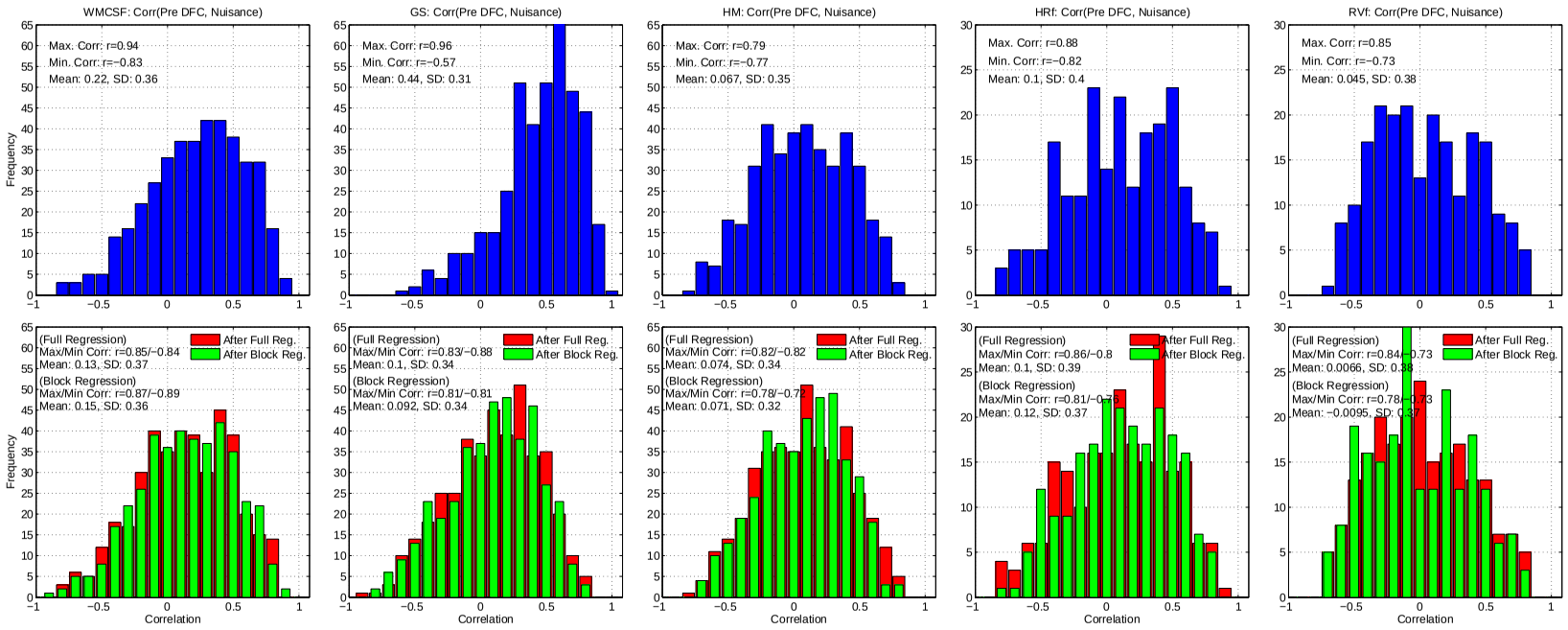
Supplementary Figure 7.7: PCC correlation maps obtained without regression of head motion covariates in the preprocessing stage. Maps are shown for the same scans used in Figure 3.12, with three low head motion runs (S2 R1, S8 R4, S8 R2) and three high motion runs (S16 R3, S12 R2, S6 R3). These maps were obtained before GSR and after the application of GSR, motion censoring, and GS censoring. The differences between motion and GS censoring are evident even when motion covariates are not regressed out.

Supplementary Table 7.1: Percentage of scans exhibiting significant correlations between the DFC estimates and nuisance norms across all the cases examined. We show the results for window sizes of ∼40 seconds, ∼58 seconds and ∼100 seconds as indicated in the top row. The significance thresholds are indicated in the leftmost column. For each significance level, the percentage of scans with significant correlations are shown for the Pre (prior to nuisance regression), Post (after full regression), and BlockReg (after block regression) conditions. The columns with the 'Overall' label show the significance results when considering all regressors. The columns with the 'GS' label display the percentage of scans with significant correlations when considering only the GS regressor. The columns with the label 'Non-GS' show the significance results when considering all regressors except for the GS.

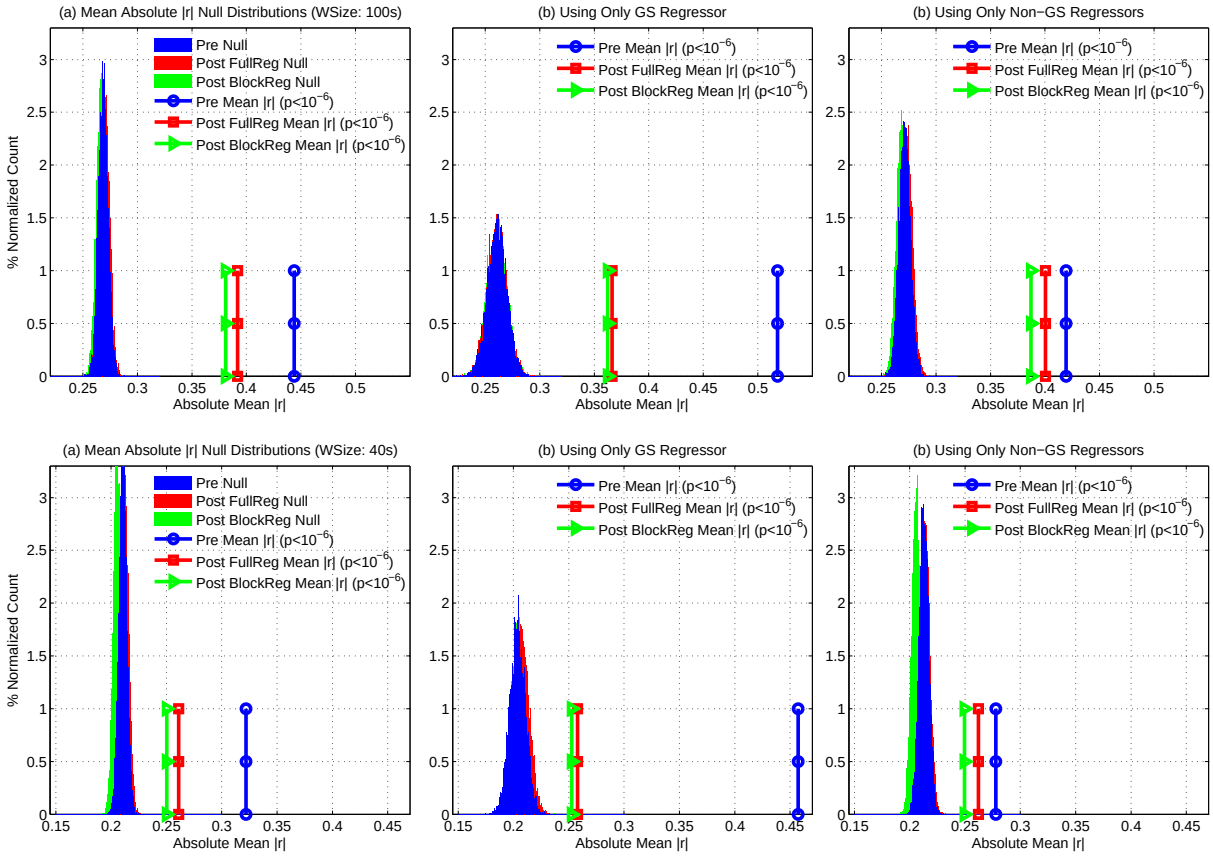
		% of Significant Scans (Window Size: 40 seconds)			**% of Significant Scans (Window Size: 58 seconds)**			**% of Significant Scans (Window Size: 100 seconds)**		
Significance Level	**State**	*GS*	*Non-GS*	***Overall***	*GS*	*Non-GS*	***Overall***	*GS*	*Non-GS*	***Overall***
$p \leq 0.05$	Pre	52%	15%	**24%**	44%	17%	**24%**	43%	24%	**29%**
	Post	12%	11%	**11%**	12%	15%	**14%**	20%	24%	**23%**
	BlockReg	11%	11%	**11%**	11%	15%	**14%**	18%	24%	**22%**
$p \leq 0.10$	Pre	59%	23%	**32%**	52%	23%	**30%**	53%	37%	**41%**
	Post	21%	20%	**20%**	19%	22%	**22%**	29%	35%	**33%**
	BlockReg	20%	19%	**20%**	19%	23%	**22%**	29%	33%	**32%**



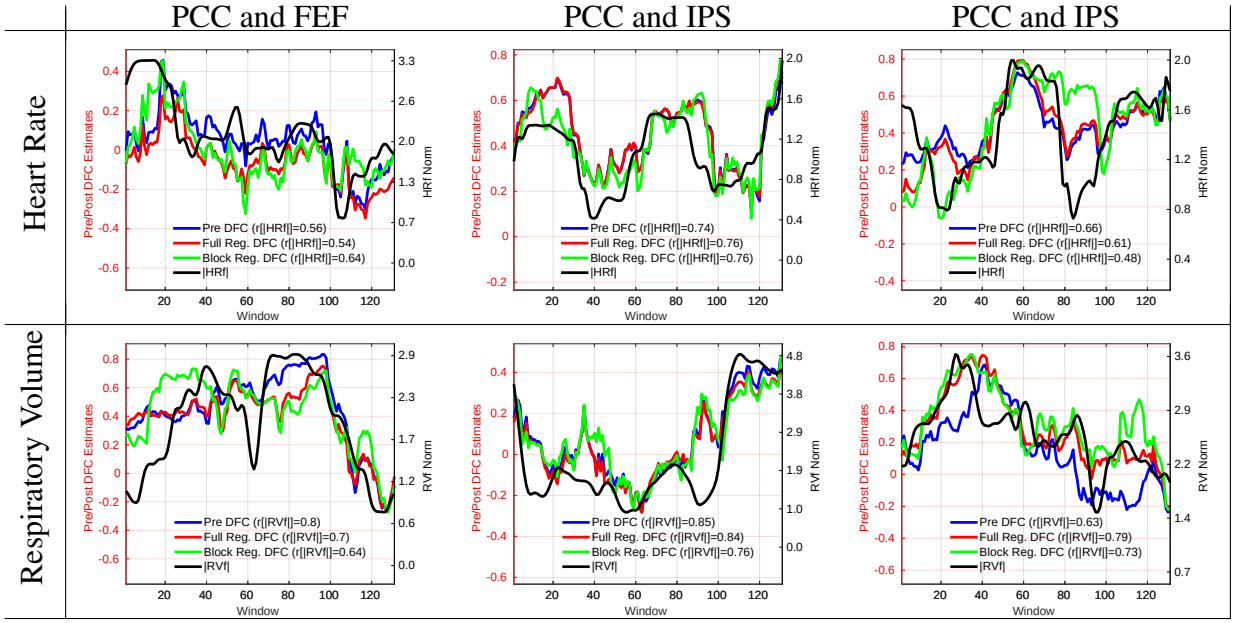
Supplementary Figure 7.8: 15 example scans that demonstrate a range of moderate to strong correlations ($0.30 < r < 0.65$) between the nuisance norms (black) and the DFC estimates before (blue) and after full (red) and block (green) regression. The type of nuisance regressor is indicated by the row label. The seed pair for the DFC estimate is indicated by the column label. There is some evidence of a delay between the nuisance norm and DFC estimates as in the first two columns for the HM regressors.



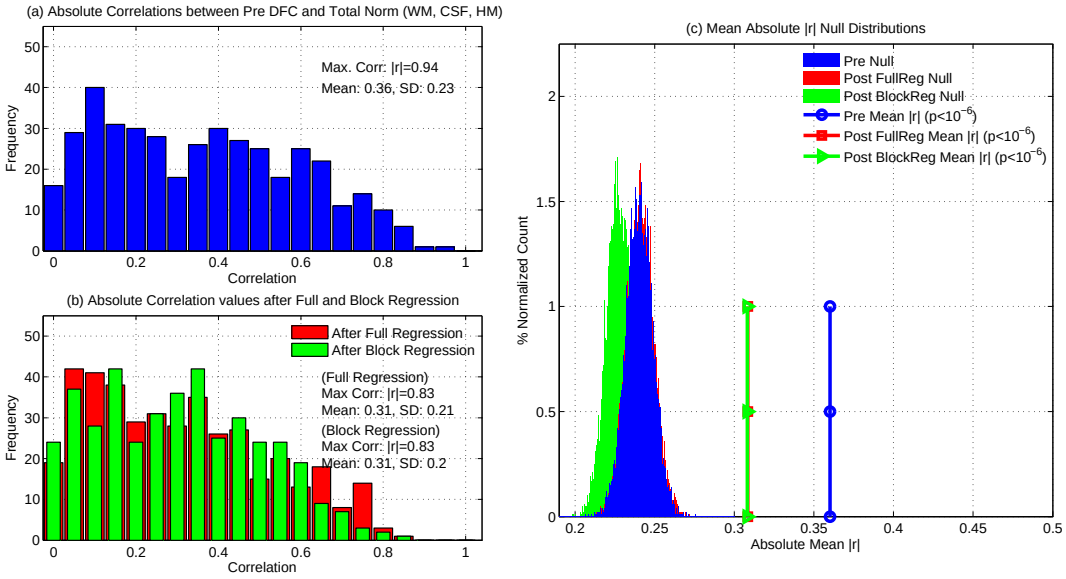
Supplementary Figure 7.9: This figure shows the correlations presented in Figure 4.4 grouped according to individual regressor norms. The first row in each column shows the histogram of correlation values obtained between the Pre DFC estimates and the nuisance norm across all scans, seed pairs for a specific nuisance regressor (columns from left to right correspond to WM+CSF, GS, HM, HRF, and RVf norms). The second row shows the correlation values after full/block regression for different regressors. The minimum and maximum correlations as well as the mean correlation value for each group is labeled in each subplot.



Supplementary Figure 7.10: Empirical null distributions and sample mean absolute correlation values for window sizes of 100s (top row) and 40s (bottom row). (a) Empirical null distributions for mean absolute correlation values when considering all regressors both before regression (blue) and after full regression (red) or block regression (green). The sample mean absolute correlation values for these conditions are indicated by the circles, squares, and triangles. These values were found to be significant $p < 10^{-6}$ both before and after full or block regression. (b) Empirical null distributions and sample mean absolute correlation values when considering only the GS regressor. There was a marked reduction in the sample mean absolute correlation values after regression. However, the sample mean absolute correlation values were still found to be significant $p < 10^{-6}$ both before and after nuisance regression. (c) Empirical null distributions and sample mean absolute correlation values for the Non-GS regressors. Sample mean absolute correlation values were found to be significant $p < 10^{-6}$ both before and after nuisance regression

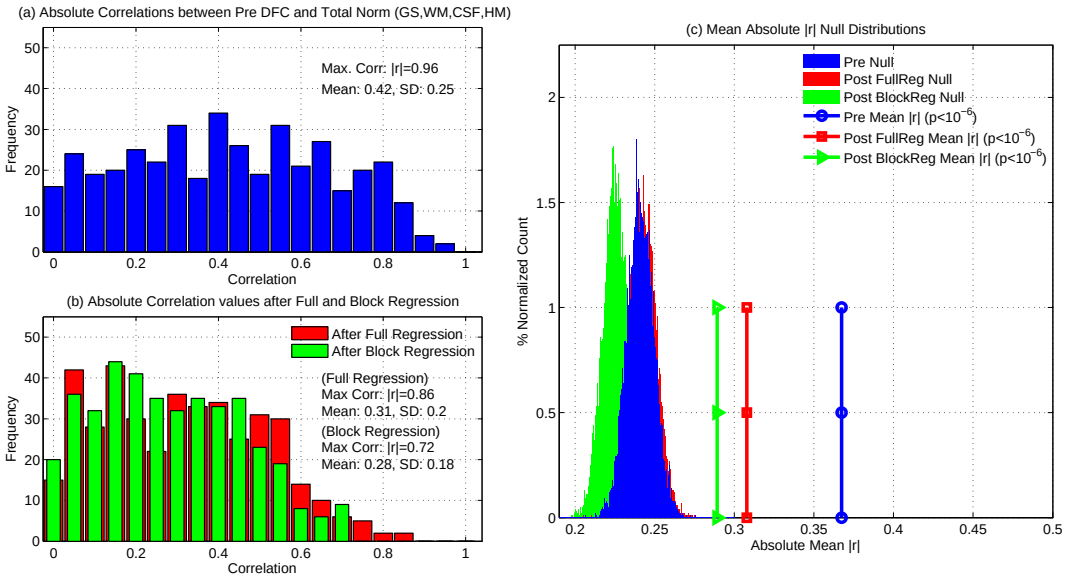


Supplementary Figure 7.11: 6 representative scans showing that both before and after nuisance regression the DFC estimates (e.g. between PCC and IPS, PCC and MOT and PCC and FEF) are strongly correlated with the respective nuisance norms.

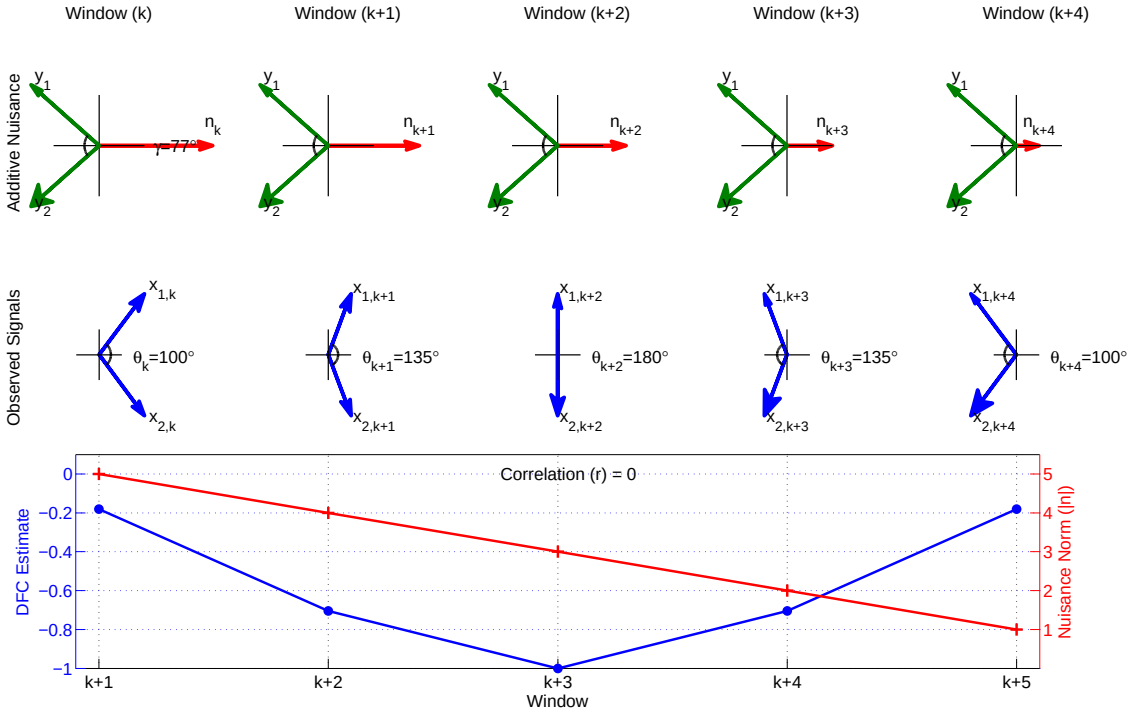


Supplementary Figure 7.12: Results obtained when using WM, CSF, and 6 HM regressors for nuisance regression. (a) Histogram of absolute correlations (before nuisance regression) between the total norm of the WM, CSF and 6 HM regressors and Pre DFC estimates obtained across all seed pairs. The mean of the absolute correlations across all cases was $|r| = 0.36$. The signed correlation values (not shown) ranged from $r = -0.83$ to $r = 0.94$ with a skewed distribution in which 72% of the correlations were positive and the remaining 28% were negative. (b) Histogram of absolute correlations between the total norm after full and block regression with sample means for both cases of $|r| = 0.31$. The signed correlations (not shown) ranged from $r = -0.80$ to $r = 0.83$ for full regression (65% of the correlations were positive) and from $r = -0.75$ to $r = 0.83$ for block regression (69% of the correlations were positive). The distribution of absolute correlation values before and after nuisance regression were similar with cosine similarity values of $S = 0.82$ and $S = 0.75$ for full and block regression, respectively. (c) Empirical null distributions and sample mean absolute correlation values before and after nuisance regression. All sample mean absolute correlation values were significant ($p < 10^{-6}$) as assessed using their respective null distribtuions.

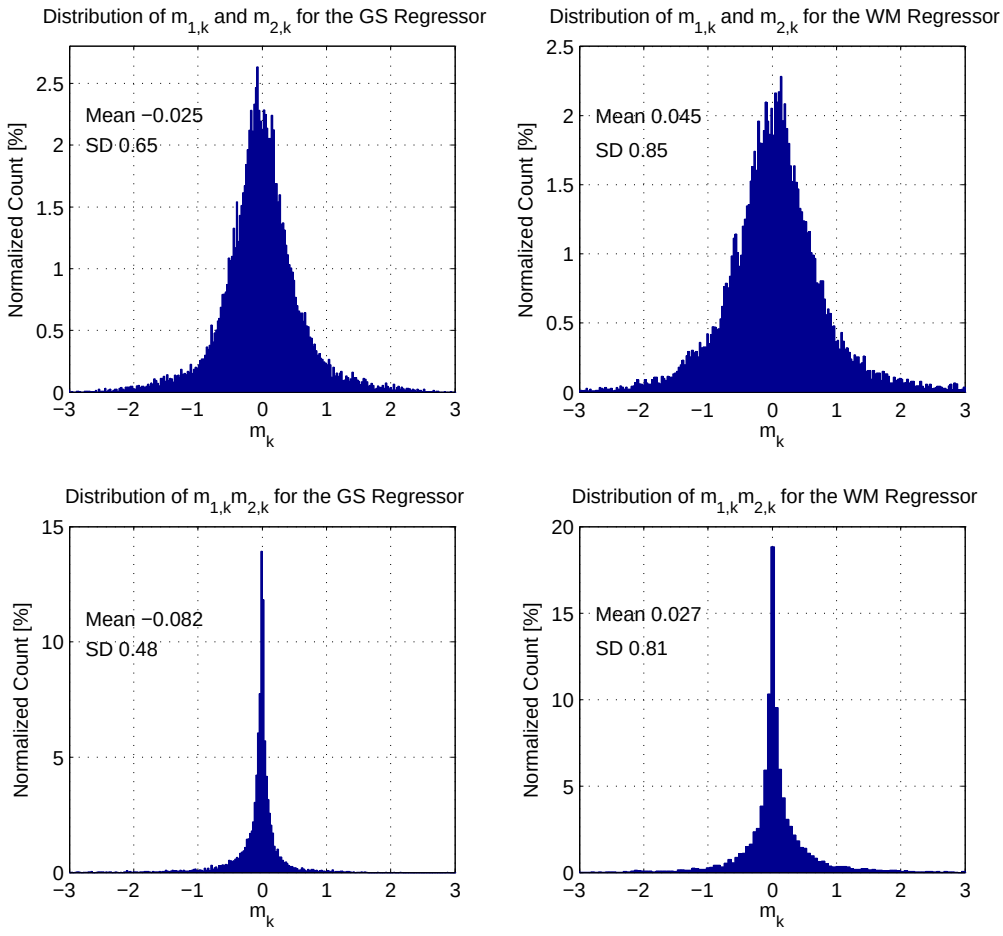
.



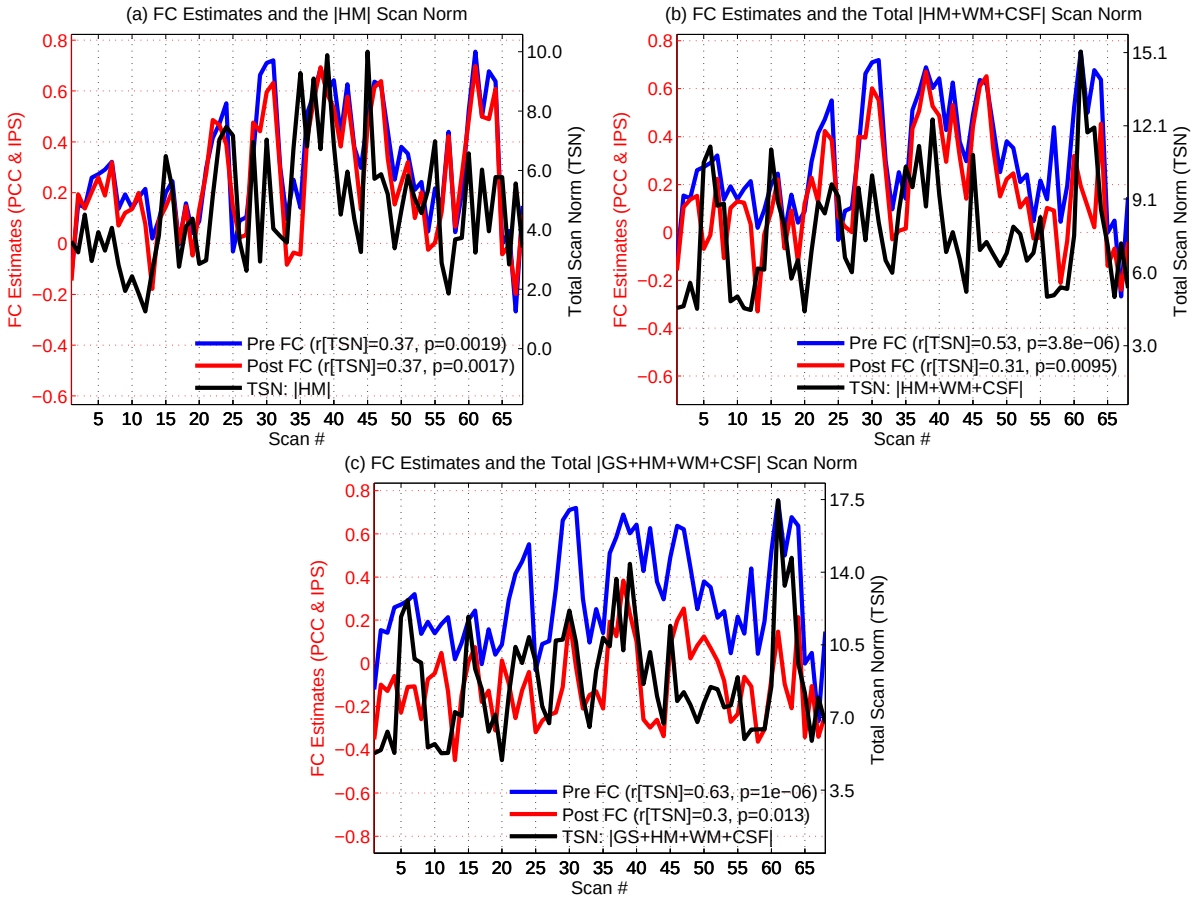
Supplementary Figure 7.13: Results obtained when using GS, WM, CSF, and 6 HM regressors for nuisance regression. (a) Histogram of absolute correlations (before nuisance regression) between the total norm of GS, WM, CSF and the 6 HM regressors and Pre DFC estimates obtained across all seed pairs. The mean of the absolute correlations across all cases was $|r| = 0.42$. The signed correlation values (not shown) ranged from $r = -0.74$ to $r = 0.95$ with a skewed distribution in which 96% of the correlations were positive and the remaining 4% were negative. (b) Histogram of absolute correlations after nuisance regression with sample means of $|r| = 0.31$ and $|r| = 0.29$ for full and block regression, respectively. The signed correlations (not shown) ranged from $r = -0.77$ to $r = 0.86$ for full regression (63% of the correlations were positive) and from $r = -0.72$ to $r = 0.71$ for block regression (67% of the correlations were positive). The distribution absolute correlation values before and after multiple regression were similar with a cosine similarity value of $S = 0.79$ for both full and block regression. (c) Empirical null distributions and sample mean absolute correlation values before and after nuisance regression. All sample mean absolute correlation values were significant ($p < 10^{-6}$) as assessed using their respective null distributions.



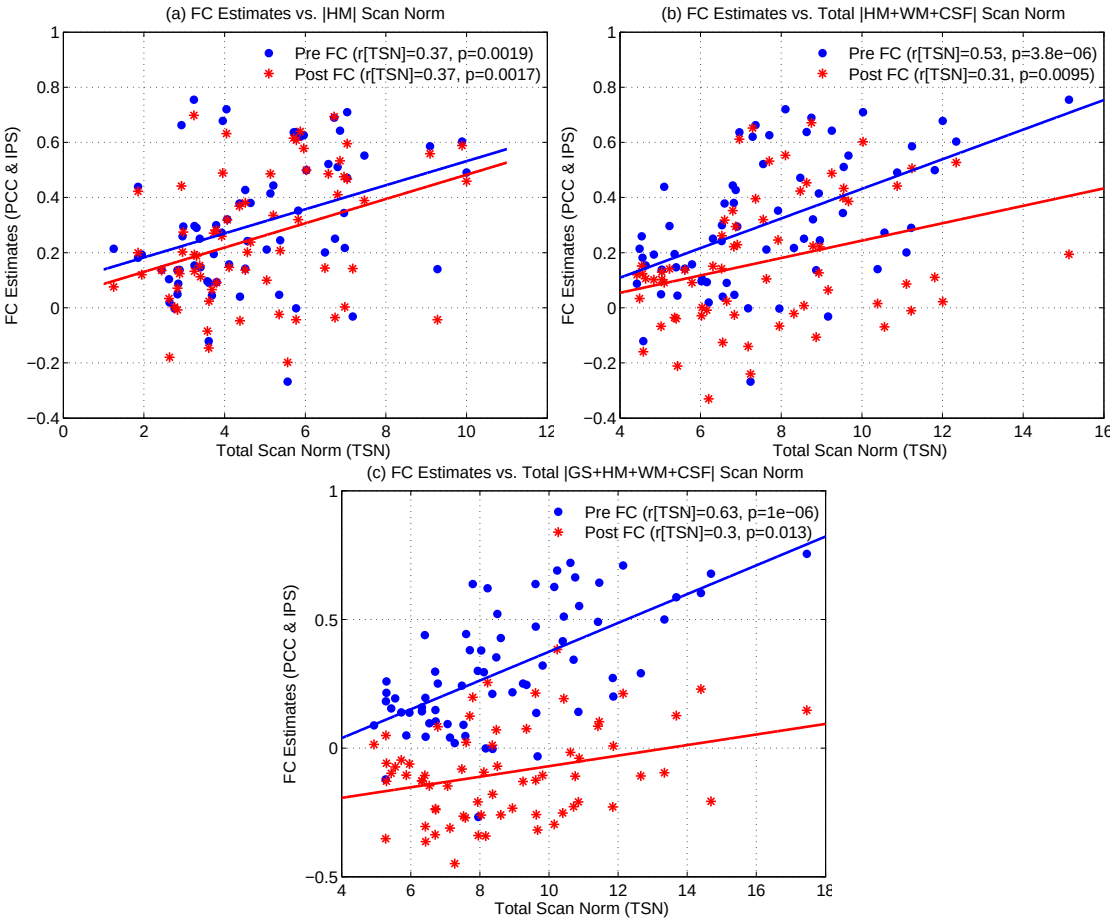
Supplementary Figure 7.14: In the first three windows (k, $k+1$, and $k+2$) a decrease in the nuisance norm corresponds to a decrease in the DFC estimates since the angle between the observed vectors increases as the norm decreases. In the last two windows ($k+3$ and $k+4$), the nuisance norm continues to decrease and a crossover occurs such that the angle between the observed vectors starts to decrease, leading to an increase in the DFC estimates. Overall, the nuisance term induces a decrease in DFC estimates in the first 3 windows and an increase in the last two windows. Thus the nuisance norm appears uncorrelated with the DFC estimates with $r = 0$, even though the nuisance term has a clear effect on the observed vectors.



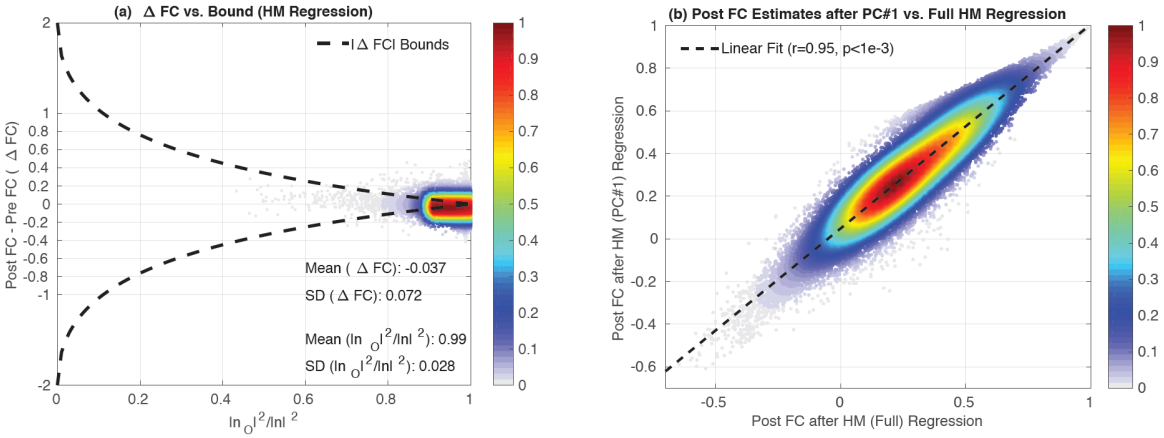
Supplementary Figure 7.15: Top row: Histograms of the difference terms $m_{1,k} = \beta_{1,F} - \beta_{1,k}$ and $m_{2,k} = \beta_{2,F} - \beta_{2,k}$ for GS and WM regressors. The plots combine the instances of the two terms. Bottom row: Histograms of the product term $m_{1,k}m_{2,k}$.



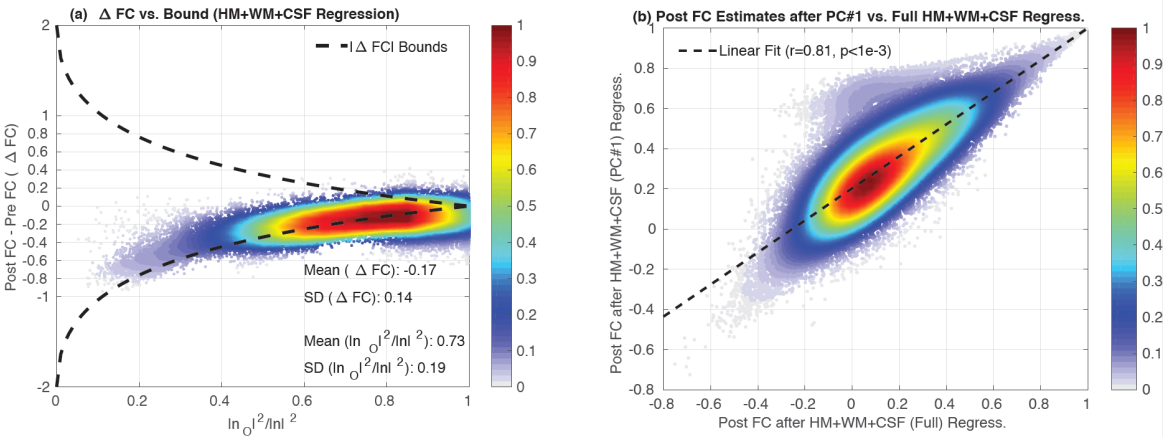
Supplementary Figure 7.16: Example of relation between static FC estimates and nuisance norms. In each of the panels, the blue line shows the per-scan estimates (prior to nuisance regression) of static FC between the PCC and IPS seeds versus the scan number. The black lines in each of the panels shows the per-scan nuisance norm as a function of scan number, where the nuisance terms used in computing the total nuisance norms were (a) 6 HM regressors, (b) WM, CSF, and 6 HM regressors, and (c) GS, WM, CSF, and 6 HM regressors. For each combination of regressors, the static FC estimates were significantly correlated ($r > 0.37; p < 2 \times 10^{-3}$) across scans with the total nuisance norms, with the correlation coefficients and p-values for each case indicated in each panel. The red lines in each of the panels show the static FC estimates after performing nuisance regression using the regressors previously described for each panel. After regression, the static FC estimates were still significantly correlated ($r > 0.30; p < 0.02$) across scans with the total nuisance norms, with the correlation coefficients and p-values for each case indicated in each panel. A scatter plot version of these results is provided in Supplementary Figure 7.17.



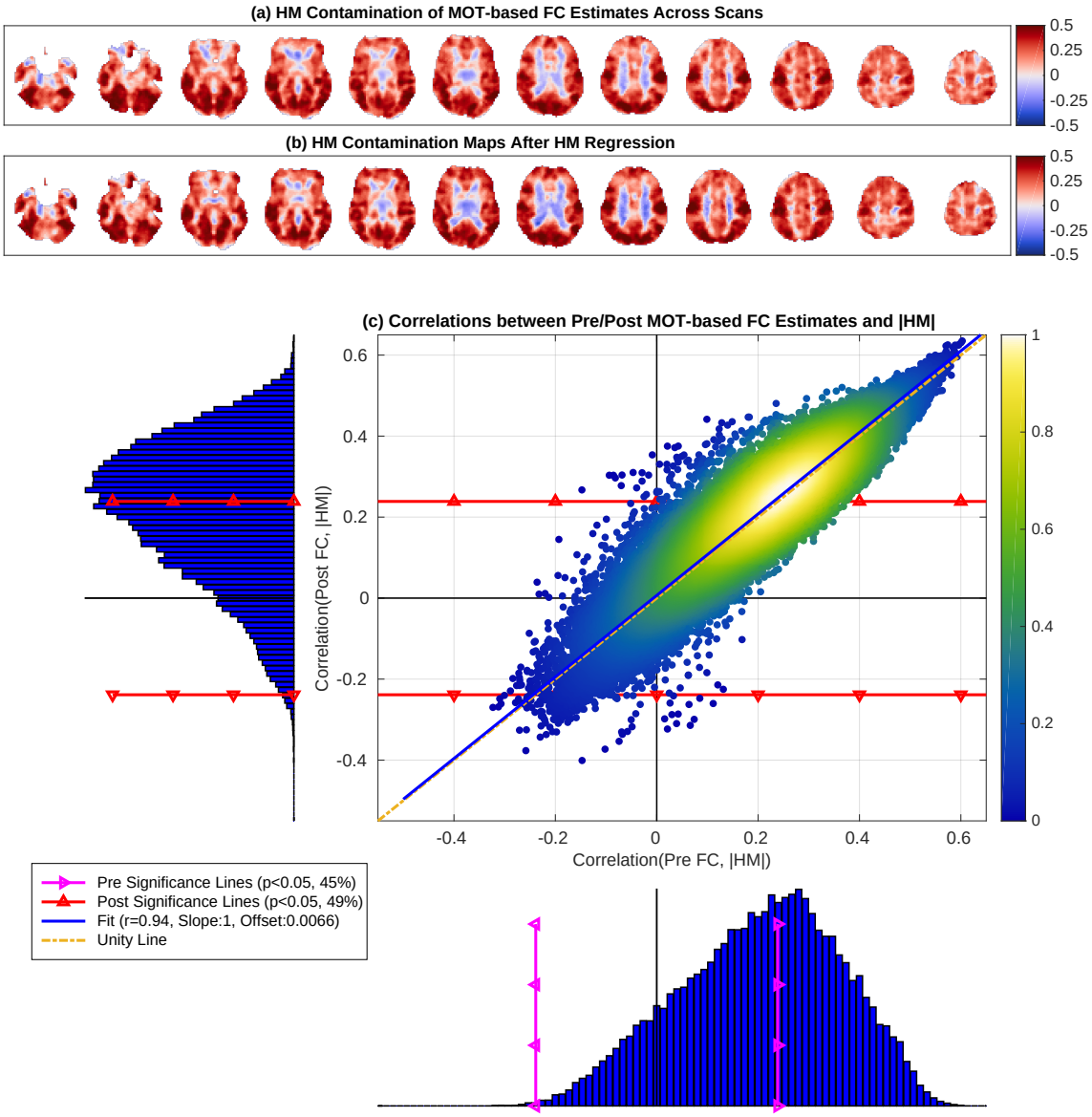
Supplementary Figure 7.17: Scatter plots detailing the relationship between the static FC estimates and nuisance norms. In each of the panels, the blue dots and red asterisks show the FC estimates before and after nuisance regression, respectively, with each point representing a single scan. Following the format of Supplementary Figure 7.16, the nuisance terms used in each panel were: (a) 6 HM regressors (b) WM, CSF, and 6 HM regressors, and (c) GS, WM, CSF, and 6 HM regressors. For each combination of regressors, we found a significant linear relationship between the static FC estimates and total nuisance norms ($r > 0.37; p < 2 \times 10^{-3}$ prior to nuisance regression and $r > 0.30; p < 0.02$ after nuisance regression).



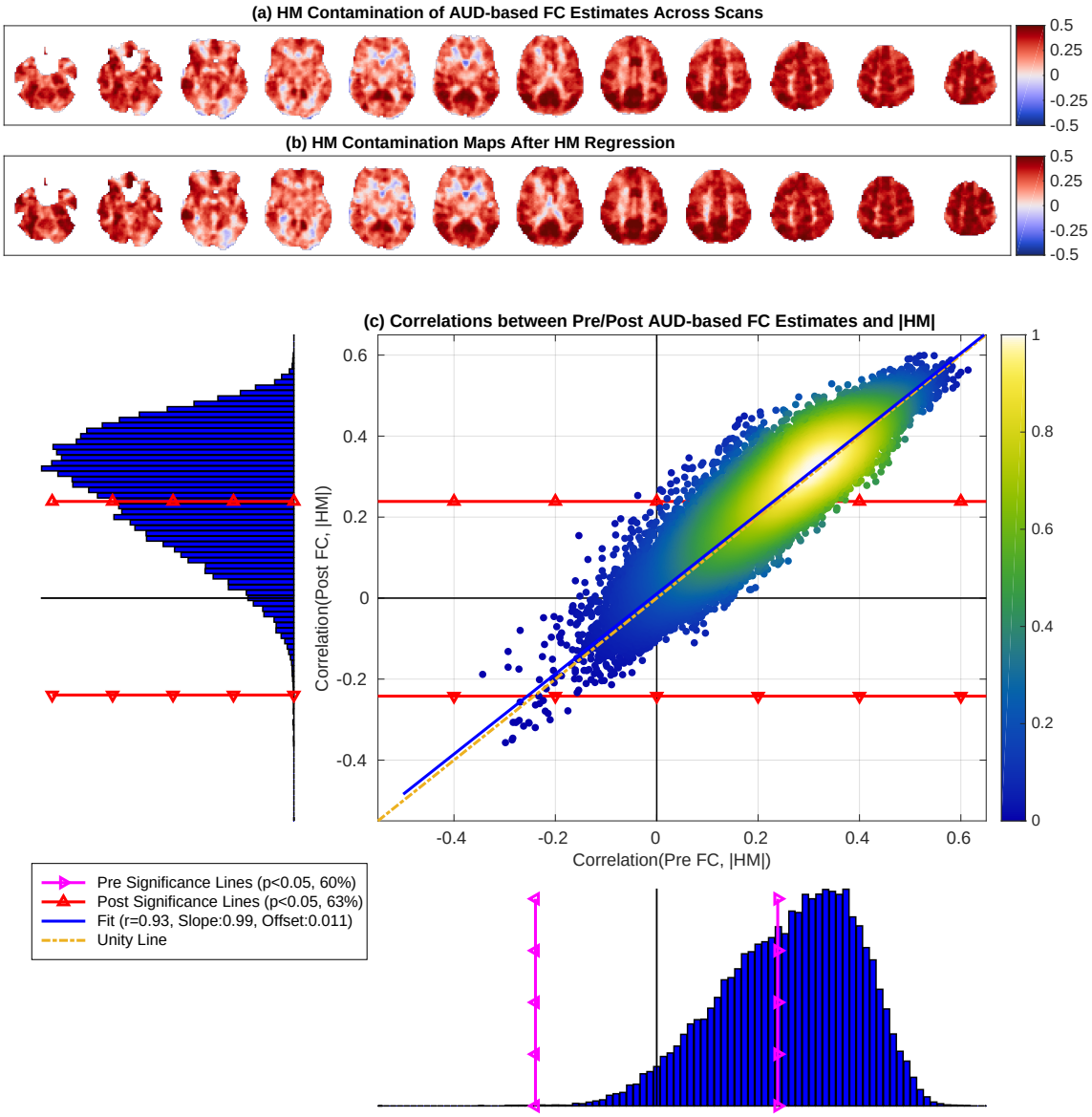
Supplementary Figure 7.18: (a)The empirical ΔFC versus the orthogonal nuisance fraction $|n_O|^2/|n|^2$ for the 6 HM regressors used in Figure 5.5a. This time full multiple HM regression was performed instead of using the 1st principal component (PC) of the 6 HM regressors to obtain the ΔFC values. The empirical values of ΔFC are clustered around 0 largely around the theoretical bounds similar to the ΔFC plot presented in Figure 5.5a. The ΔFC values slightly violate the bounds since the bounds for multiple regression is only approximate and correct bounds only apply for the analysis of a single regressor. (b) Post FC estimates obtained after regressing out the 1st PC of HM measurements and after regressing out all 6 HM regressors were strongly linear with respect to each other $r = 0.95$.



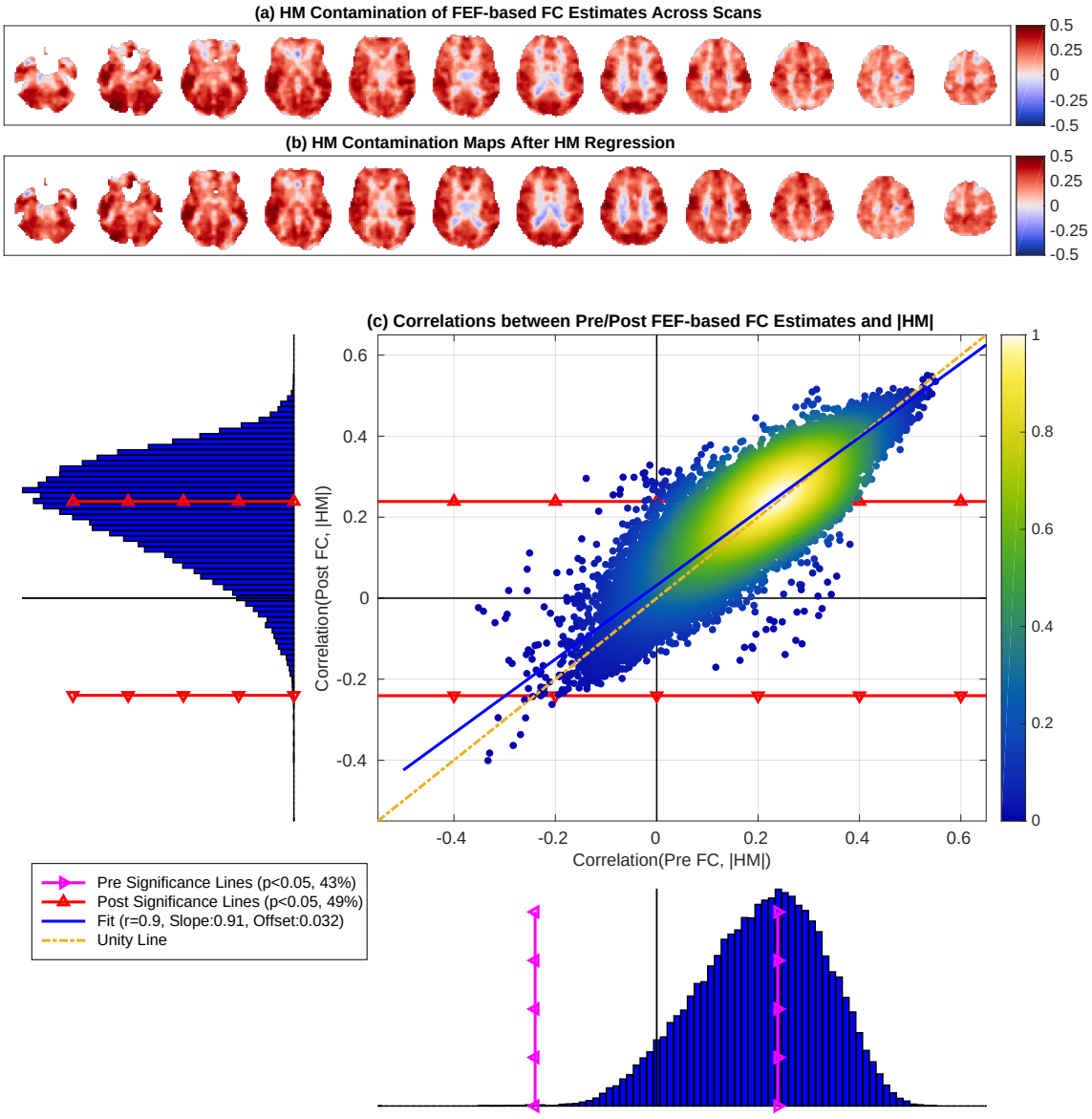
Supplementary Figure 7.19: (a)The empirical ΔFC versus the orthogonal nuisance fraction $|n_O|^2/|n|^2$ for the HM+WM+CSF regressors used in Figure 5.5d. This time multiple regression was performed instead of using the 1st principal component (PC) of all the regressors to obtain the ΔFC values. The empirical values of ΔFC are clustered around -0.17 and were mostly within the theoretical bounds similar to the ΔFC plot presented in Figure 5.5d. The ΔFC values slightly violate the bounds since the bounds for multiple regression is only approximate and correct bounds only apply for the analysis of a single regressor. (b) Post FC estimates obtained after regressing out the 1st PC of all HM+WM+CSF regressors and after performing multiple HM+WM+CSF regression were strongly linear with respect to each other $r = 0.81$.



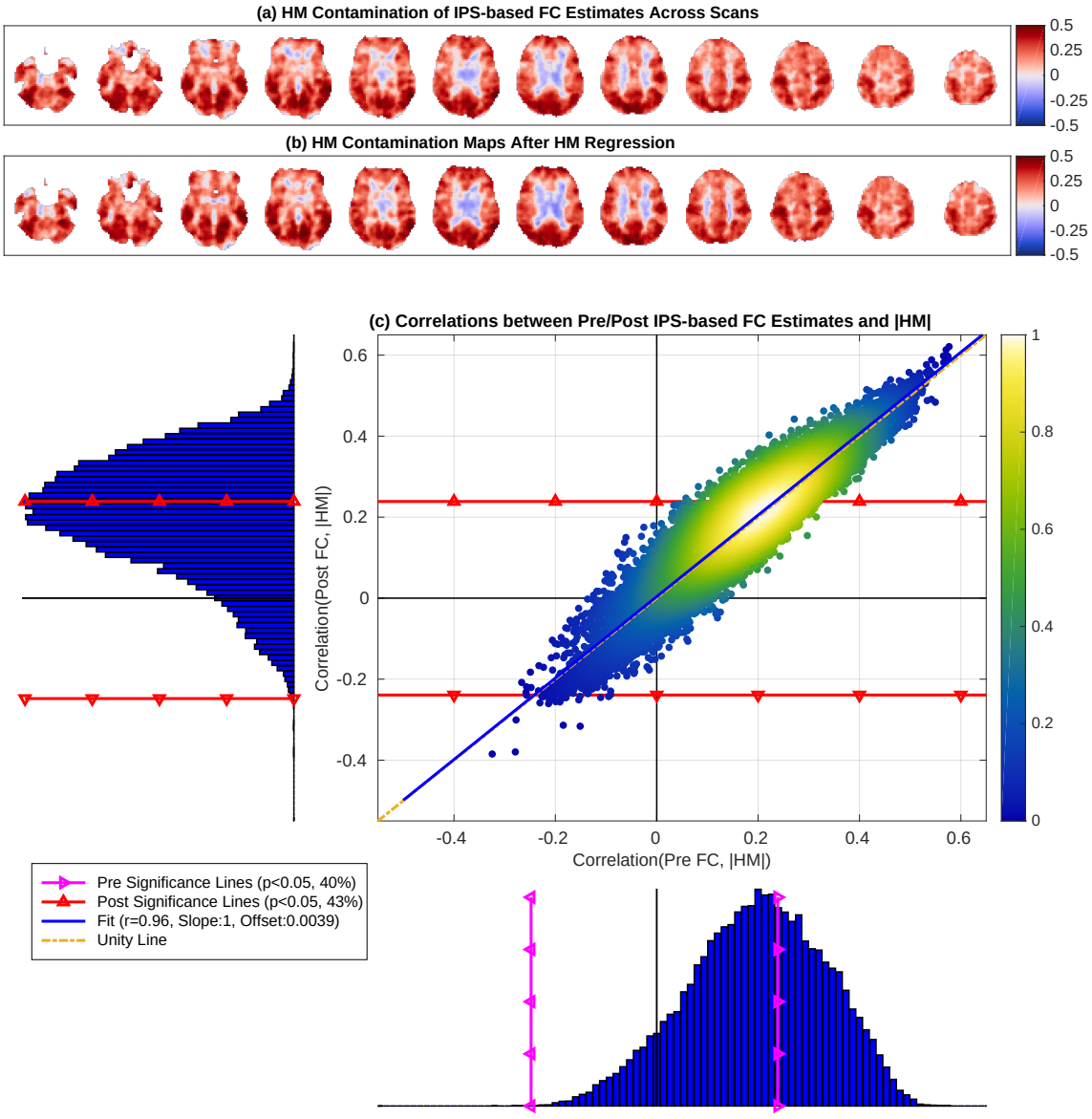
Supplementary Figure 7.20: MOT-based HM contamination maps before (a) and after (b) HM regression. These maps were fairly similar to each other (cosine similarity $S = 0.98$) and show widespread correlations HM norm and FC estimates across scans. The scatter plot in (c) shows the correlations between the Post FC estimates and HM norm after regression versus the correlations obtained between Pre FC estimates and HM norm. These correlation distributions were significantly related ($r = 0.94$, $p < 10^{-3}$) to each other. The linear fit (blue line, Slope$= 1$ and Offset$= 0.007$) between the two correlation distributions was nearly identical to the line of unity (dashed yellow line). At the bottom histogram, the correlations between the Pre FC estimates and HM norms ranged from $r = -0.32$ to $r = 0.60$ with mean 0.2 and 45% these correlations were significant ($p < 0.05$). In the sideways histogram on the left, the correlations between the Post FC estimates and HM norms ranged from $r = -0.40$ to $r = 0.64$ with mean 0.21. The post regression significance lines are shown with red lines with triangles. 49% of the Post FC estimates were significantly correlated with HM norms after regression.



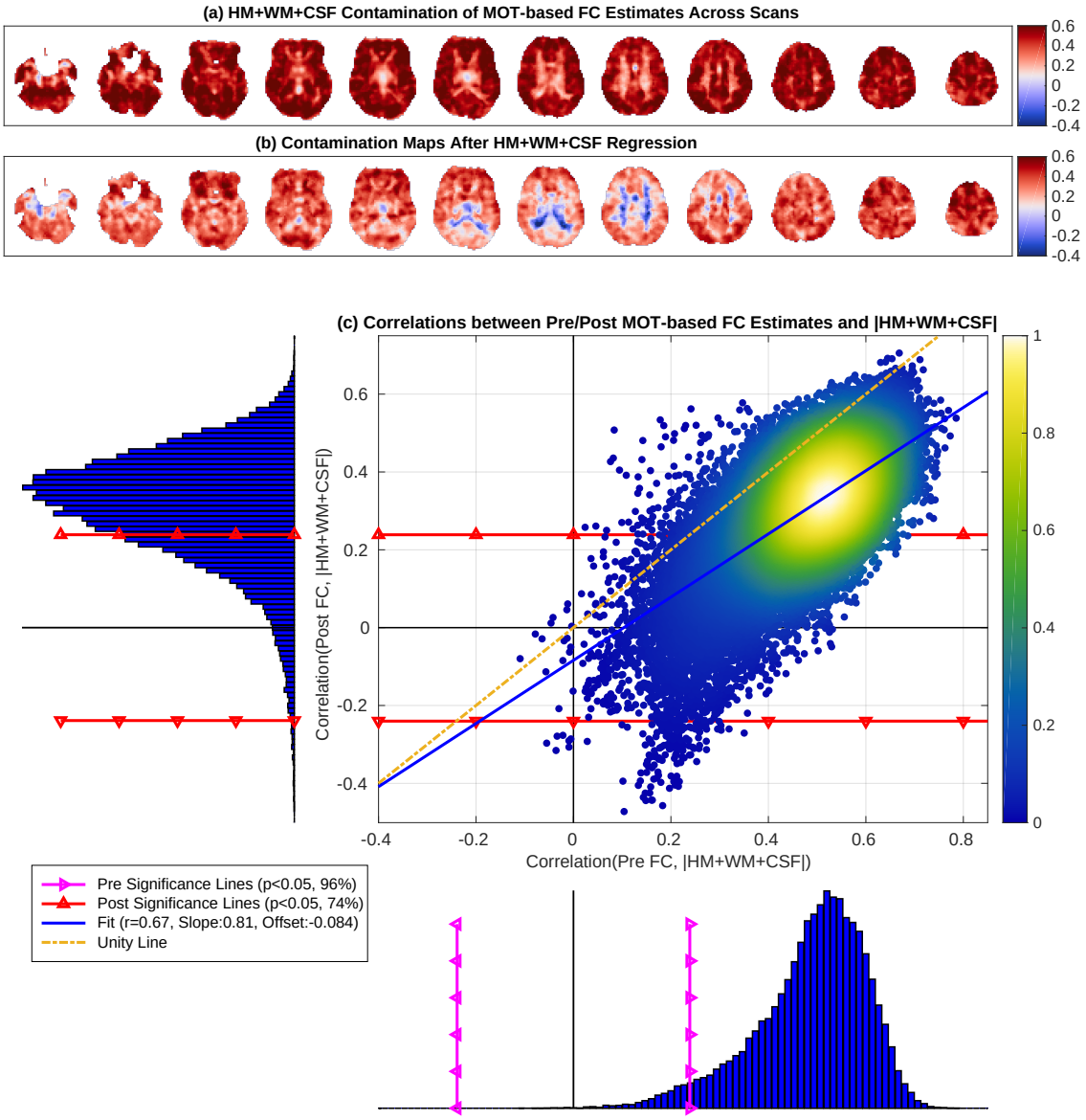
Supplementary Figure 7.21: AUD-based HM contamination maps before (a) and after (b) HM regression. These maps were fairly similar to each other (cosine similarity $S = 0.99$) and show widespread correlations HM norm and FC estimates across scans. The scatter plot in (c) shows the correlations between the Post FC estimates and HM norm after regression versus the correlations obtained between Pre FC estimates and HM norm. These correlation distributions were significantly related ($r = 0.93$, $p < 10^{-3}$) to each other. The linear fit (blue line, Slope= 0.99 and Offset= 0.01) between the two correlation distributions was nearly identical to the line of unity (dashed yellow line). At the bottom histogram, the correlations between the Pre FC estimates and HM norms ranged from $r = -0.34$ to $r = 0.61$ with mean 0.26 and 60% these correlations were significant ($p < 0.05$). In the sideways histogram on the left, the correlations between the Post FC estimates and HM norms ranged from $r = -0.36$ to $r = 0.60$ with mean 0.27. The post regression significance lines are shown with red lines with triangles. 63% of the Post FC estimates were significantly correlated with HM norms after regression.



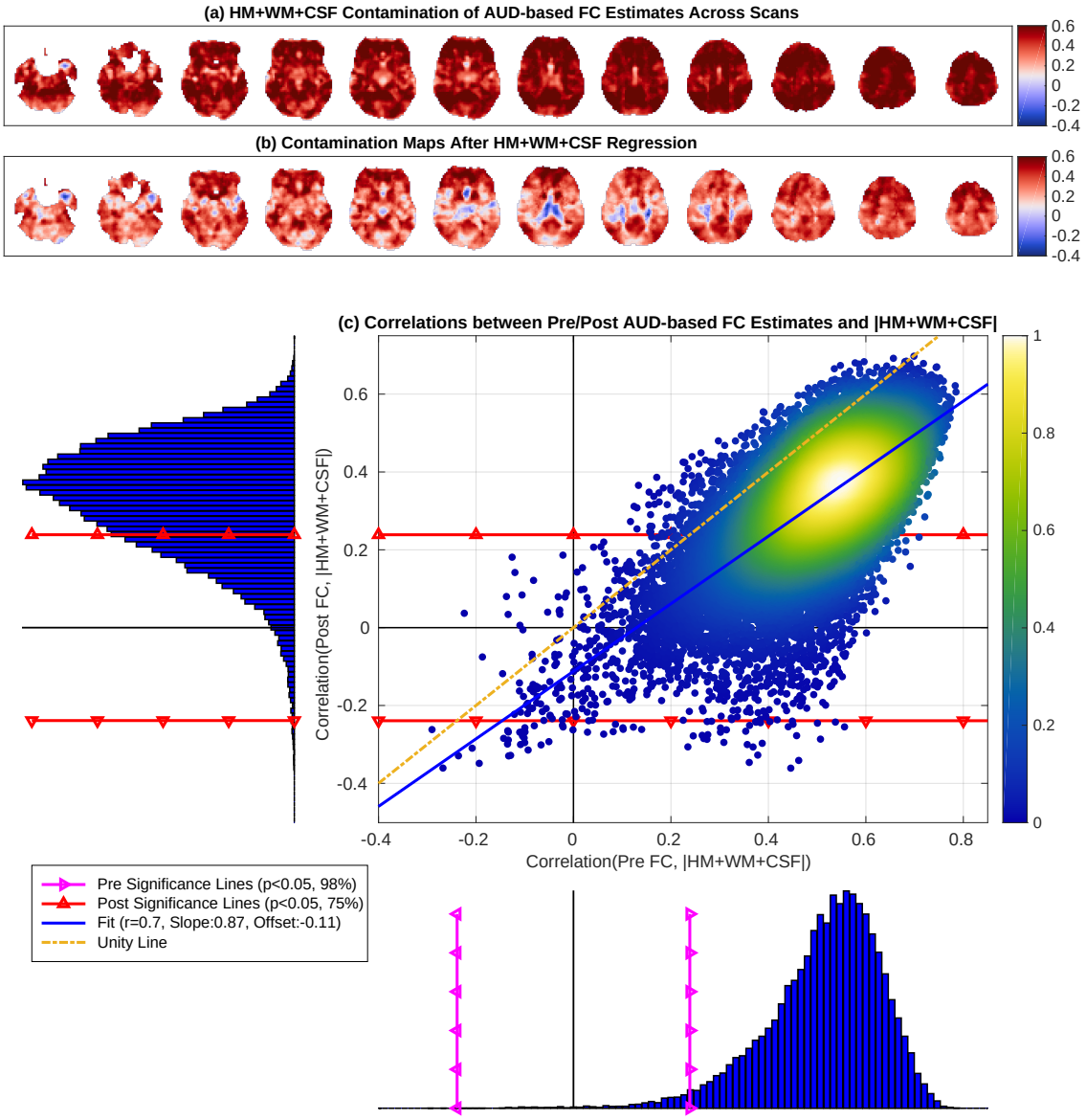
Supplementary Figure 7.22: FEF-based HM contamination maps before (a) and after (b) HM regression. These maps were fairly similar to each other (cosine similarity $S = 0.97$) and show widespread correlations HM norm and FC estimates across scans. The scatter plot in (c) shows the correlations between the Post FC estimates and HM norm after regression versus the correlations obtained between Pre FC estimates and HM norm. These correlation distributions were significantly related ($r = 0.9$, $p < 10^{-3}$) to each other. The linear fit (blue line, Slope$= 0.91$ and Offset$= 0.032$) between the two correlation distributions was nearly identical to the line of unity (dashed yellow line). At the bottom histogram, the correlations between the Pre FC estimates and HM norms ranged from $r = -0.34$ to $r = 0.55$ with mean 0.21 and 43% these correlations were significant ($p < 0.05$). In the sideways histogram on the left, the correlations between the Post FC estimates and HM norms ranged from $r = -0.40$ to $r = 0.55$ with mean 0.22. The post regression significance lines are shown with red lines with triangles. 49% of the Post FC estimates were significantly correlated with HM norms after regression.



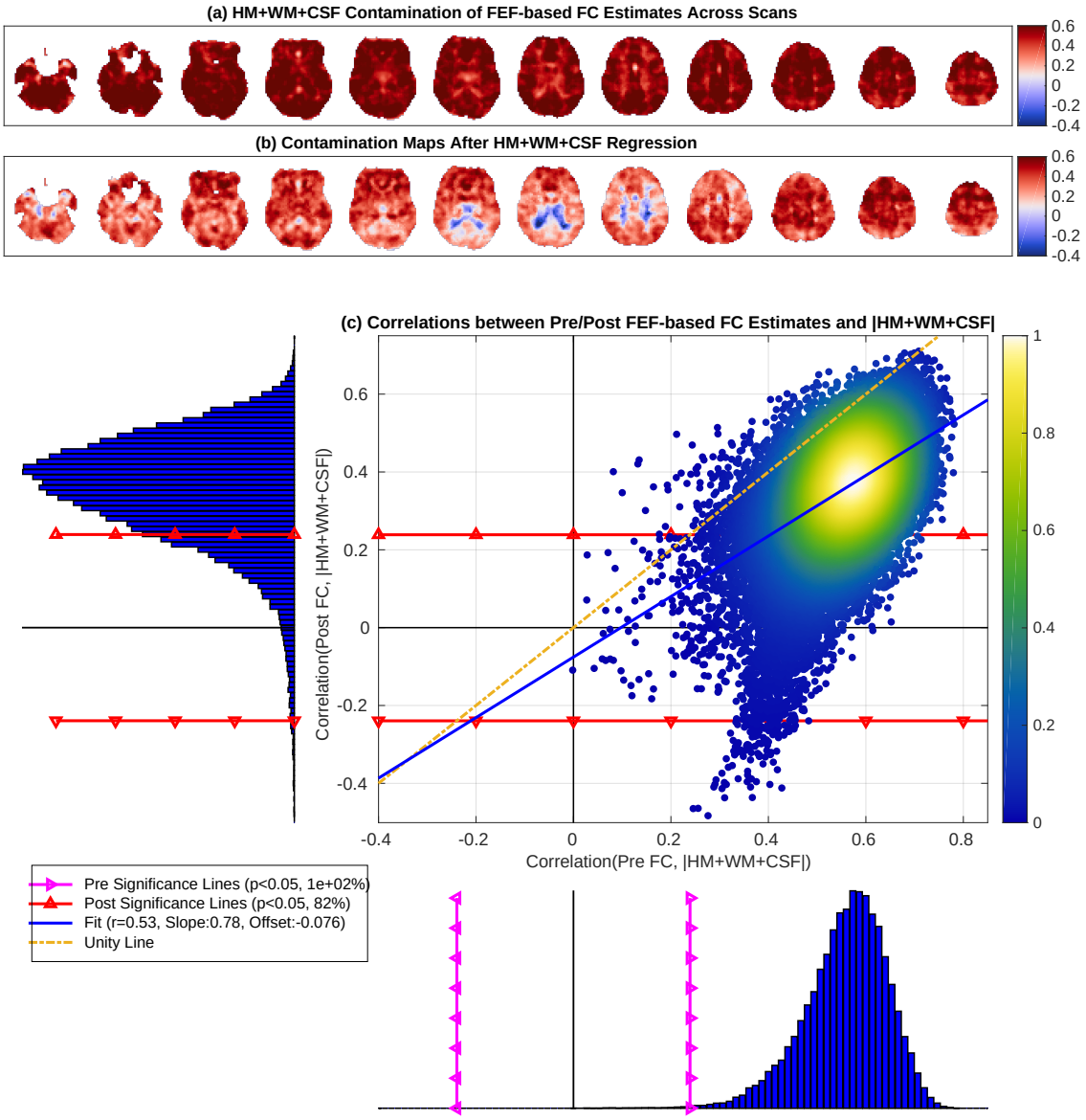
Supplementary Figure 7.23: IPS-based HM contamination maps before (a) and after (b) HM regression. These maps were fairly similar to each other (cosine similarity $S = 0.99$) and show widespread correlations HM norm and FC estimates across scans. The scatter plot in (c) shows the correlations between the Post FC estimates and HM norm after regression versus the correlations obtained between Pre FC estimates and HM norm. These correlation distributions were significantly related ($r = 0.96$, $p < 10^{-3}$) to each other. The linear fit (blue line, Slope= 1.0 and Offset= 0.0039) between the two correlation distributions was nearly identical to the line of unity (dashed yellow line). At the bottom histogram, the correlations between the Pre FC estimates and HM norms ranged from $r = -0.33$ to $r = 0.58$ with mean 0.20 and 40% these correlations were significant ($p < 0.05$). In the sideways histogram on the left, the correlations between the Post FC estimates and HM norms ranged from $r = -0.36$ to $r = 0.62$ with mean 0.21. The post regression significance lines are shown with red lines with triangles. 43% of the Post FC estimates were significantly correlated with HM norms after regression.



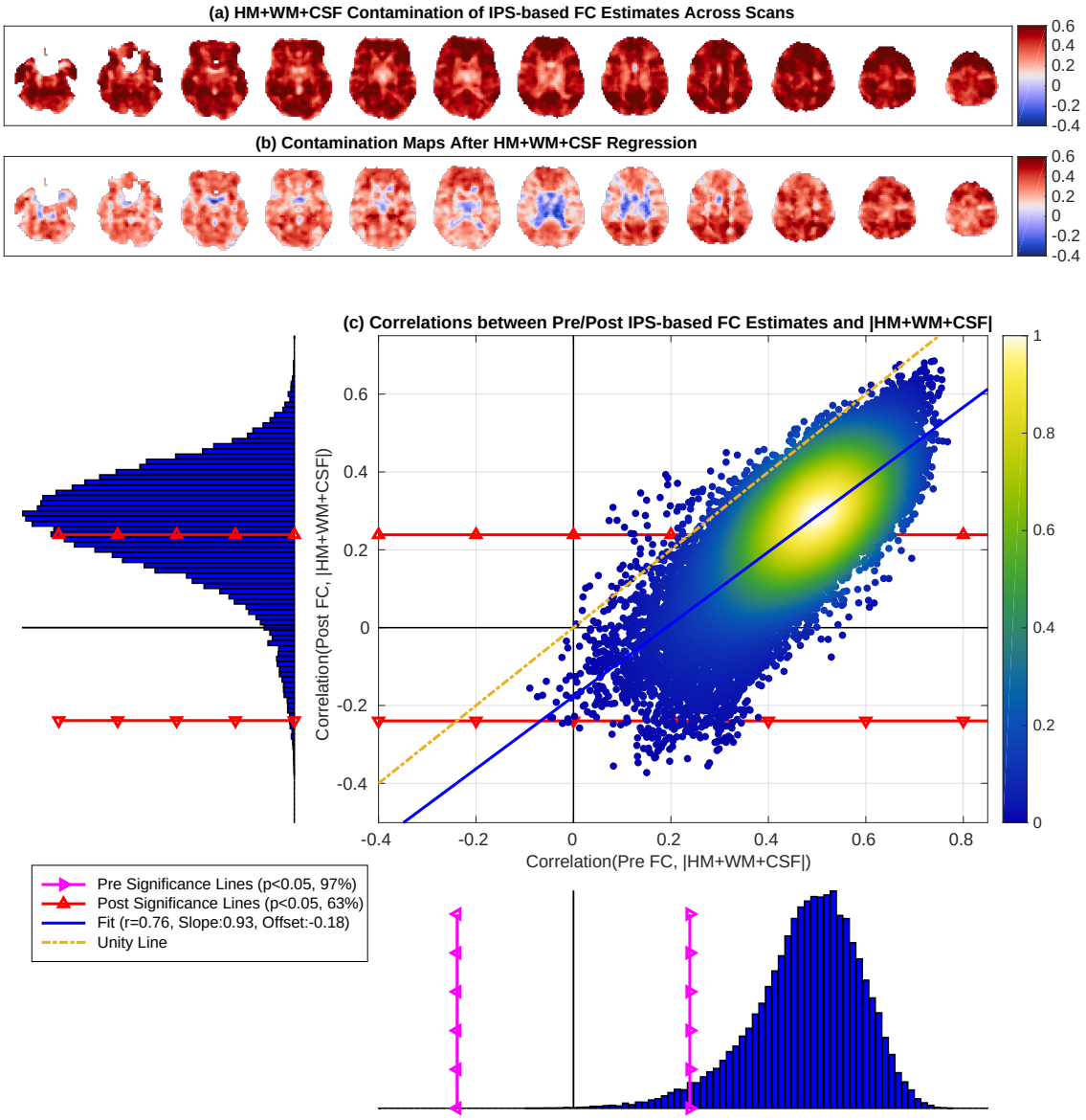
Supplementary Figure 7.24: MOT-based nuisance contamination maps before (a) and after (b) HM+WM+CSF regression. There was a visible reduction in the correlation values after regression with a slight increase in anti-correlations with blue regions appearing in (b). In (c) the correlations between the Pre FC estimates and the nuisance norm (x-axis) ranged from $= -0.11$ to $r = 0.76$ with mean 0.49. The correlations between the Post FC estimates and nuisance norm (y-axis) ranged from $= -0.55$ to $r = 0.71$ with mean 0.31. There was a strong linear relation between the two correlation distributions (linear fit $r = 0.67$, $p < 10^{-3}$). The linear fit between the two correlation distributions was close to the line of unity with a slight reduction in the slope (Slope$= 0.81$).



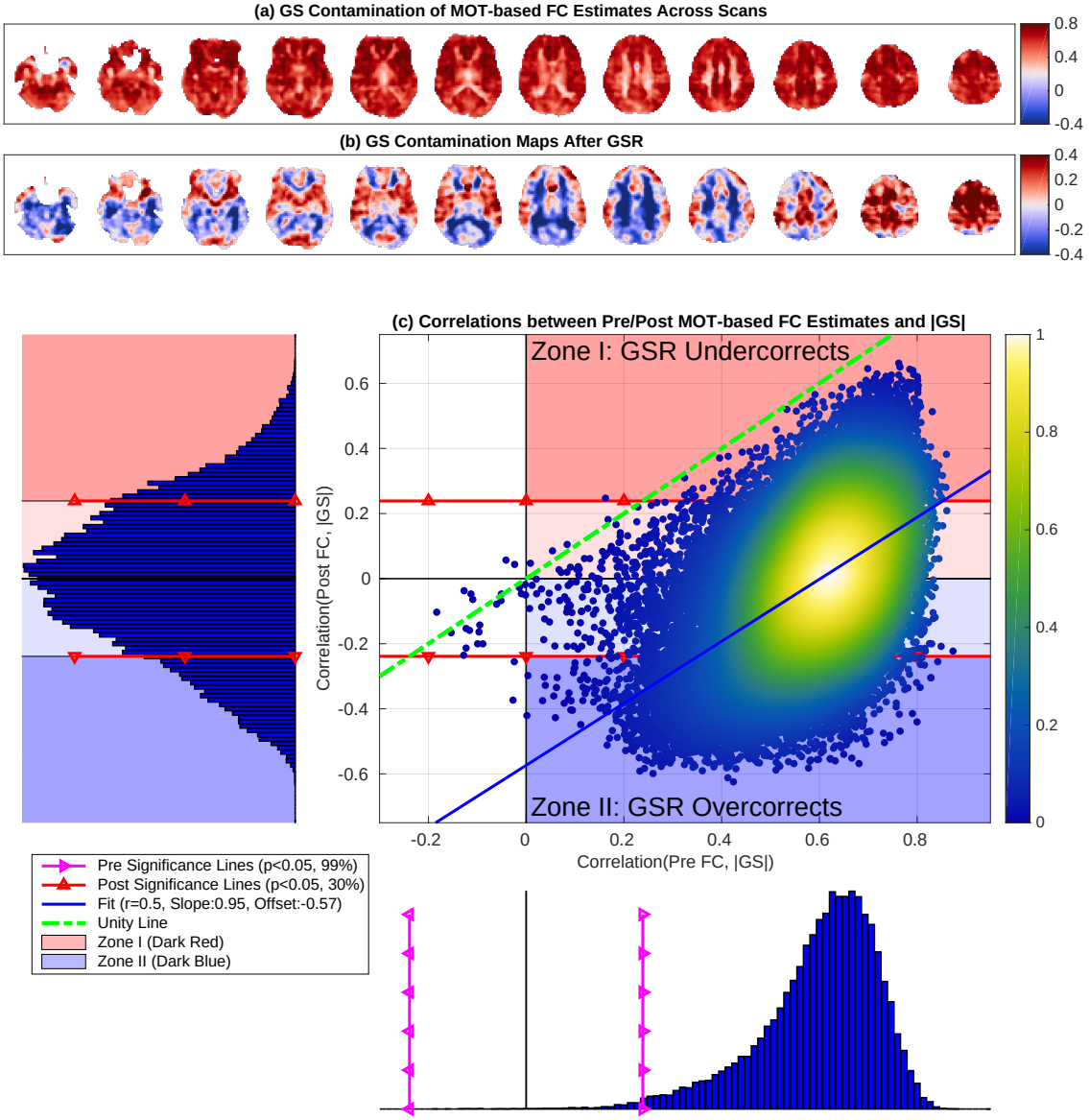
Supplementary Figure 7.25: AUD-based nuisance contamination maps before (a) and after (b) HM+WM+CSF regression. In (c) the correlations between the Pre FC estimates and the nuisance norm (x-axis) ranged from $= -0.29$ to $r = 0.79$ with mean 0.51 and 98% of these correlations were significant ($p < 0.05$). The correlations between the Post FC estimates and nuisance norm (y-axis) ranged from $= -0.36$ to $r = 0.70$ with mean 0.33 and 74% of these correlations were significant ($p < 0.05$). There was a strong linear relation between the two correlation distributions (linear fit $r = 0.70$, $p < 10^{-3}$). The linear fit between the two correlation distributions was close to the line of unity with a slight reduction in the slope (Slope$= 0.87$) and a small negative offset -0.11.



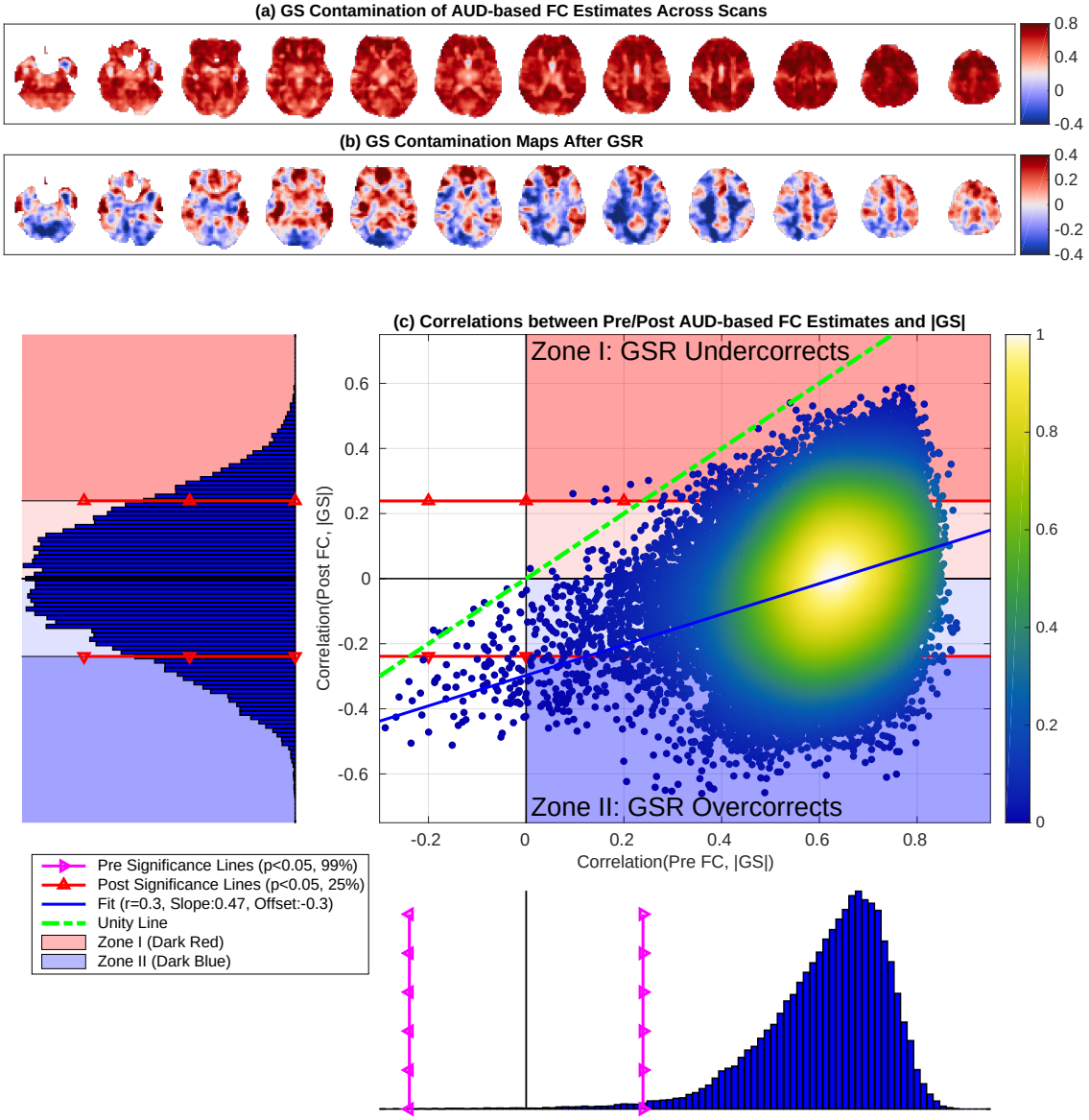
Supplementary Figure 7.26: FEF-based nuisance contamination maps before (a) and after (b) HM+WM+CSF regression. In (c) the correlations between the Pre FC estimates and the nuisance norm (x-axis) ranged from 0 to $r = 0.78$ with mean 0.56 and 98% of these correlations were significant ($p < 0.05$). The correlations between the Post FC estimates and nuisance norm (y-axis) ranged from $= -0.51$ to $r = 0.71$ with mean 0.35 and 75% of these correlations were significant ($p < 0.05$). There was a strong linear relation between the two correlation distributions (linear fit $r = 0.53$, $p < 10^{-3}$). The linear fit between the two correlation distributions was close to the line of unity with a slight reduction in the slope (Slope$= 0.78$) and a small negative offset -0.08.



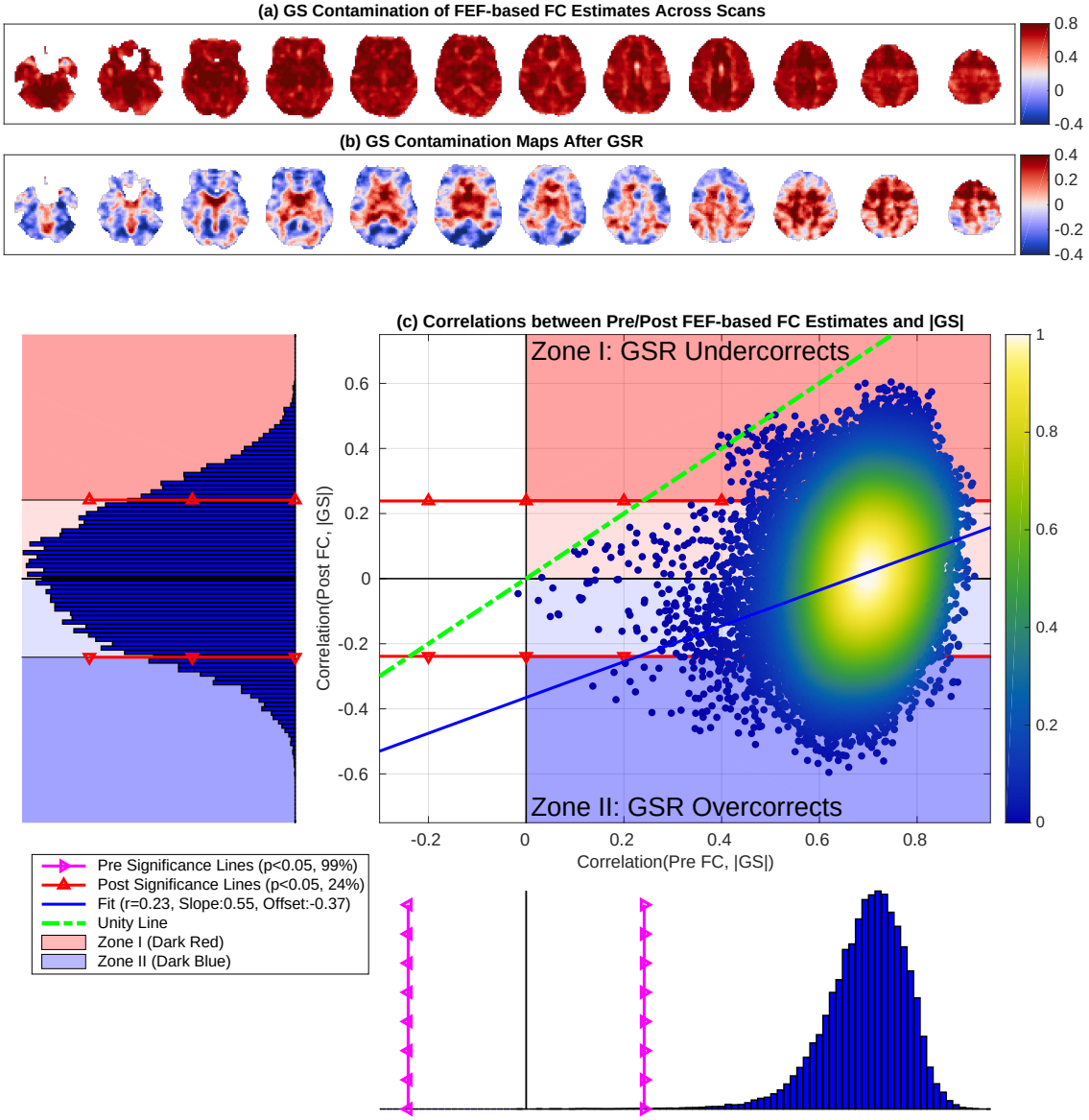
Supplementary Figure 7.27: IPS-based nuisance contamination maps before (a) and after (b) HM+WM+CSF regression. In (c) the correlations between the Pre FC estimates and the nuisance norm (x-axis) ranged from -0.1 to $r = 0.77$ with mean 0.48 and 97% of these correlations were significant ($p < 0.05$). The correlations between the Post FC estimates and nuisance norm (y-axis) ranged from $= -0.37$ to $r = 0.68$ with mean 0.27 and 63% of these correlations were significant ($p < 0.05$). There was a strong linear relation between the two correlation distributions (linear fit $r = 0.76$, $p < 10^{-3}$). The linear fit between the two correlation distributions was close to the line of unity with a slight reduction in the slope (Slope$= 0.93$) and a small negative offset -0.18.



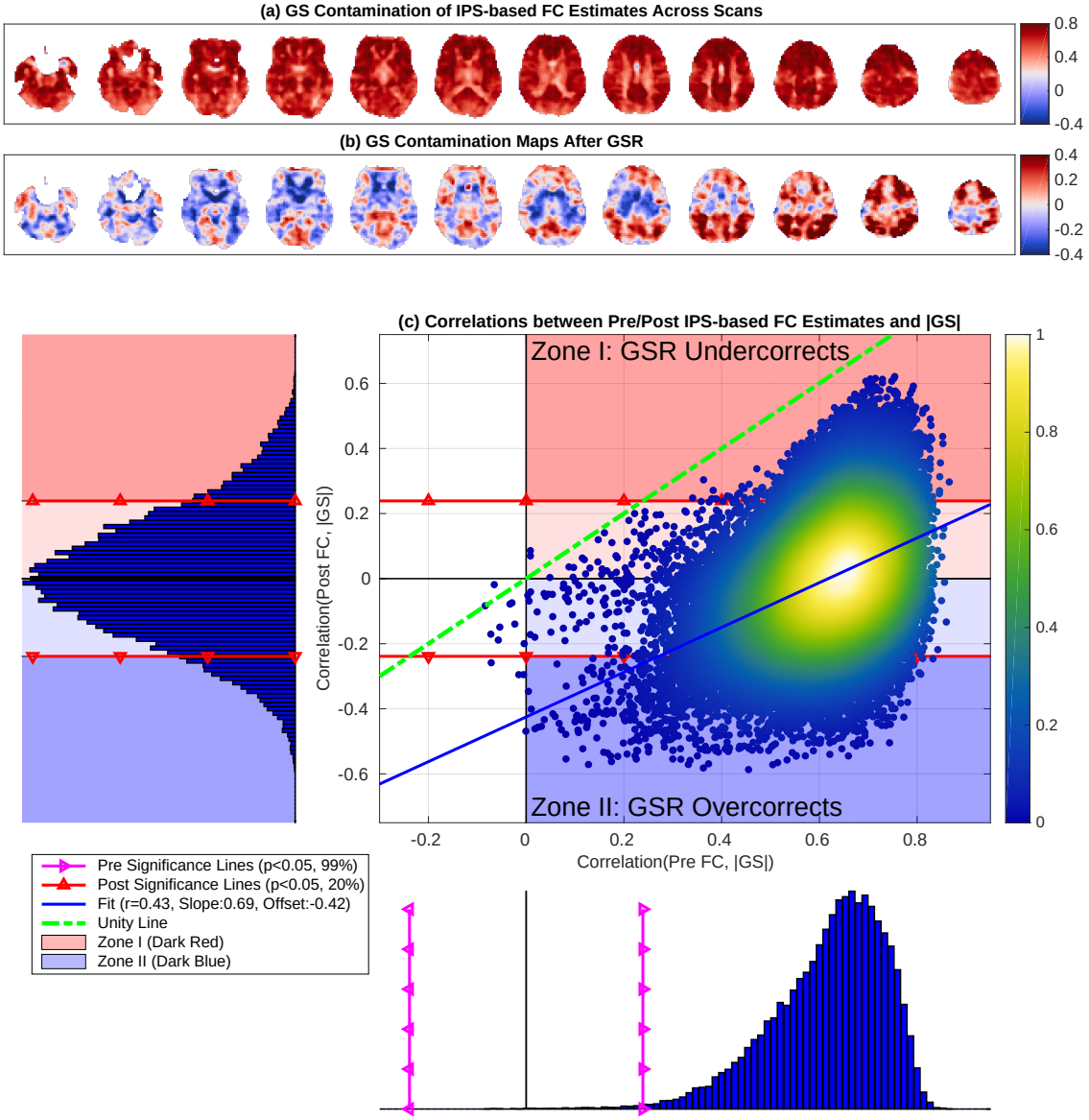
Supplementary Figure 7.28: MOT-based GS contamination maps obtained both before (a) and after (b) GSR. In panel (a) the contamination map involved positive correlations (99%) between Pre FC estimates and GS norm. In (b) the contamination map after GSR involves both positive (50%) and negative (50%) correlations between the GS norms and Post FC estimates. In (c) the correlations between the Pre FC estimates and the nuisance norm (x-axis) ranged from -0.18 to $r = 0.87$ with mean 0.60 and 99% of these correlations were significant ($p < 0.05$). The correlations between the Post FC estimates and nuisance norm (y-axis) ranged from $= -0.62$ to $r = 0.66$ with mean 0 and 30% of these correlations were significant ($p < 0.05$). There was a strong linear relation between the two correlation distributions (linear fit $r = 0.50$, $p < 10^{-3}$) and the linear fit (blue line) was fairly parallel to the line of unity with slope 0.95 and a very large negative offset -0.57.



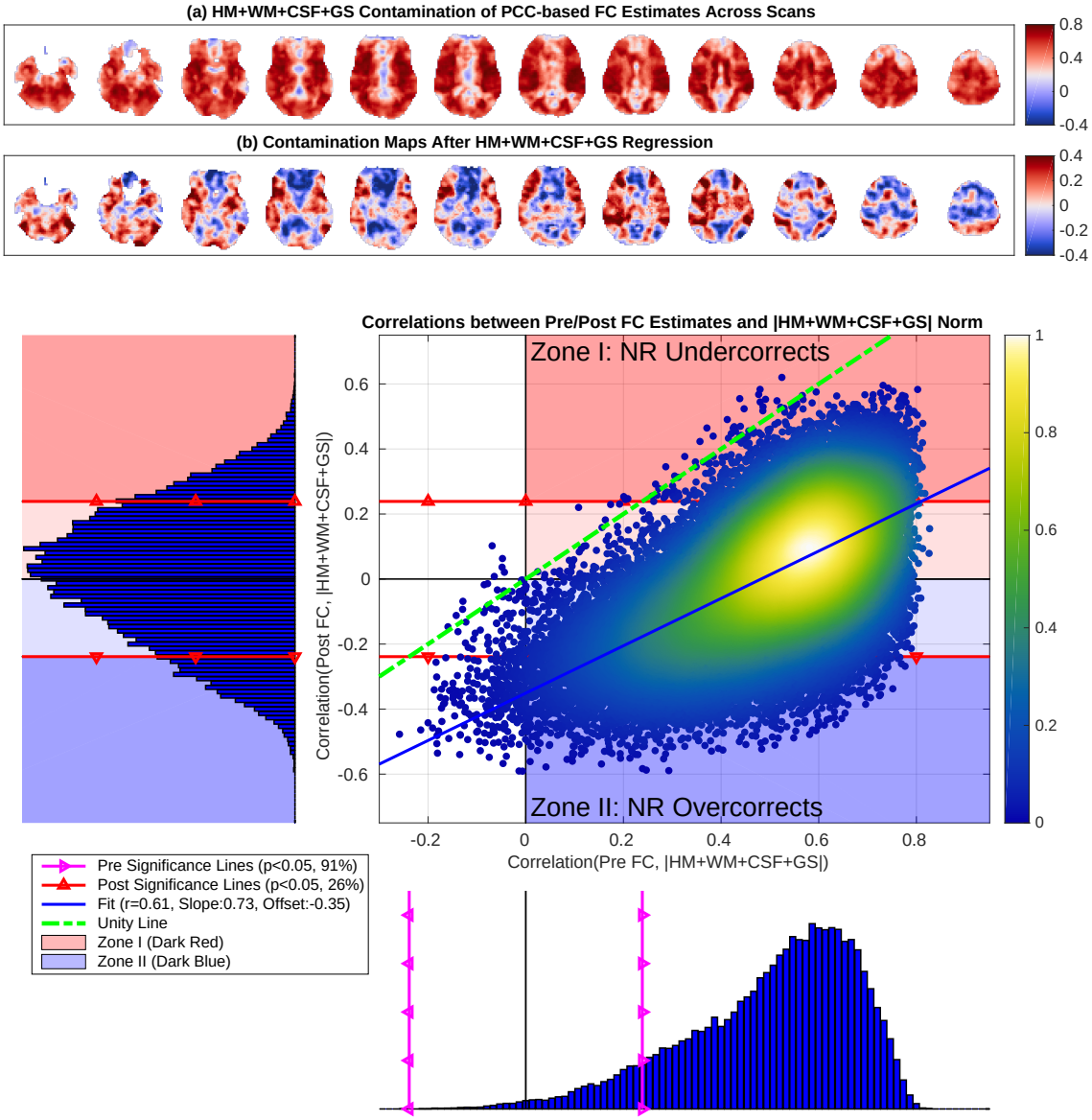
Supplementary Figure 7.29: AUD-based GS contamination maps obtained both before (a) and after (b) GSR. In panel (a) the contamination map involved positive correlations (99%) between Pre FC estimates and GS norm. In (b) the contamination map after GSR involves both positive (48%) and negative (52%) correlations between the GS norms and Post FC estimates. In (c) the correlations between the Pre FC estimates and the nuisance norm (x-axis) ranged from -0.35 to $r = 0.88$ with mean 0.61 and 99% of these correlations were significant ($p < 0.05$). The correlations between the Post FC estimates and nuisance norm (y-axis) ranged from $= -0.66$ to $r = 0.59$ with mean 0.01 and 25% of these correlations were significant ($p < 0.05$). There was a linear relation between the two correlation distributions (linear fit $r = 0.30$, $p < 10^{-3}$) and the linear fit (blue line) was fairly parallel to the line of unity with slope 0.47 and large negative offset -0.30.



Supplementary Figure 7.30: FEF-based GS contamination maps obtained both before (a) and after (b) GSR. In panel (a) the contamination map involved positive correlations (100%) between Pre FC estimates and GS norm. In (b) the contamination map after GSR involves both positive (53%) and negative (48%) correlations between the GS norms and Post FC estimates. In (c) the correlations between the Pre FC estimates and the nuisance norm (x-axis) ranged from -0.02 to $r = 0.92$ with mean 0.70 and 100% of these correlations were significant ($p < 0.05$). The correlations between the Post FC estimates and nuisance norm (y-axis) ranged from $= -0.60$ to $r = 0.60$ with mean 0 and 24% of these correlations were significant ($p < 0.05$). There was a weak linear relation between the two correlation distributions (linear fit $r = 0.23$, $p < 10^{-3}$) and the linear fit (blue line) was moderately parallel to the line of unity with a slope 0.55 and a large negative offset -0.37.



Supplementary Figure 7.31: IPS-based GS contamination maps obtained both before (a) and after (b) GSR. In panel (a) the contamination map involved positive correlations (100%) between Pre FC estimates and GS norm. In (b) the contamination map after GSR involves both positive (49%) and negative (51%) correlations between the GS norms and Post FC estimates. In (c) the correlations between the Pre FC estimates and the nuisance norm (x-axis) ranged from -0.1 to $r = 0.87$ with mean 0.62 and 100% of these correlations were significant ($p < 0.05$). The correlations between the Post FC estimates and nuisance norm (y-axis) ranged from $= -0.59$ to $r = 0.62$ with mean 0 and 20% of these correlations were significant ($p < 0.05$). There was a linear relation between the two correlation distributions (linear fit $r = 0.43$, $p < 10^{-3}$) and the linear fit (blue line) was fairly parallel to the line of unity with a slope 0.69 and a large negative offset -0.42.



Supplementary Figure 7.32: PCC-based HM+WM+CSF+GS contamination maps obtained both before (a) and after (b) multiple regression. In panel (a) the contamination map mostly involved positive correlations (99%) between Pre FC estimates and nuisance norm. In (b) the contamination map after nuisance regression involves both positive (54%) and negative (46%) correlations between the nuisance norm and Post FC estimates. In (c) the correlations between the Pre FC estimates and the nuisance norm (x-axis) ranged from -0.26 to $r = 0.83$ with mean 0.50 and 99% of these correlations were significant ($p < 0.05$). The correlations between the Post FC estimates and nuisance norm (y-axis) ranged from $= -0.59$ to $r = 0.62$ with mean 0.01 and 26% of these correlations were significant ($p < 0.05$). There was a linear relation between the two correlation distributions (linear fit $r = 0.61$, $p < 10^{-3}$) and the linear fit (blue line) was fairly parallel to the line of unity with a slope 0.73 and a large negative offset -0.35. These results were very similar to performing GSR alone.

Bibliography

[1] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *The Journal of Machine Learning Research*, vol. 1, pp. 211–244, 2001.

[2] R. Rubinstein, M. Zibulevsky, and M. Elad, "Double sparsity: Learning sparse dictionaries for sparse signal approximation," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1553–1564, 2010.

[3] R. Jenatton, J. Mairal, G. Obozinski, and F. R. Bach, "Proximal methods for sparse hierarchical dictionary learning." in *ICML*, no. 2010, 2010, pp. 487–494.

[4] J. Huang, T. Zhang, and D. Metaxas, "Learning with structured sparsity," *Journal of Machine Learning Research*, vol. 12, no. Nov, pp. 3371–3412, 2011.

[5] R. Peharz and F. Pernkopf, "Sparse nonnegative matrix factorization with $\ell 0$-constraints," *Neurocomputing*, vol. 80, pp. 38–46, 2012.

[6] W. Wen, C. Wu, Y. Wang, Y. Chen, and H. Li, "Learning structured sparsity in deep neural networks," in *Advances in Neural Information Processing Systems*, 2016, pp. 2074–2082.

[7] D. P. Wipf and B. D. Rao, "An empirical Bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3704–3716, 2007.

[8] I. F. Gorodnitsky, J. S. George, and B. D. Rao, "Neuromagnetic source imaging with focuss: a recursive weighted minimum norm algorithm," *Electroencephalography and Clinical Neurophysiology*, vol. 95, no. 4, pp. 231–251, 1995.

[9] S. F. Cotter and B. D. Rao, "Sparse channel estimation via matching pursuit with application to equalization," *IEEE Transactions on Communications*, vol. 50, no. 3, pp. 374–377, 2002.

[10] M. Aharon, M. Elad, and A. M. Bruckstein, "K-SVD and its non-negative variant for dictionary design," in *Optics & Photonics 2005*, 2005, pp. 591 411–591 411.

[11] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[12] Y. Ding and B. D. Rao, "Dictionary learning-based sparse channel representation and estimation for fdd massive mimo systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5437–5451, 2018.

[13] R. G. Baraniuk, V. Cevher, and M. B. Wakin, "Low-dimensional models for dimensionality reduction and signal recovery: A geometric perspective," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 959–971, 2010.

[14] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.

[15] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via $\ell 1$ minimization," *Proceedings of the National Academy of Sciences*, vol. 100, no. 5, pp. 2197–2202, 2003.

[16] D. Ge, X. Jiang, and Y. Ye, "A note on the complexity of ℓ_p minimization," *Mathematical Programming*, vol. 129, no. 2, pp. 285–299, 2011.

[17] M. Elad, *Sparse and Redundant Representations*. Springer New York, 2010.

[18] Y. C. Pati, R. Rezaiifar, and P. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Asilomar Conference on Signals, Systems and Computers*. IEEE, 1993, pp. 40–44.

[19] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on information theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[20] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.

[21] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.

[22] D. Needell and R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," *Foundations of Computational Mathematics*, vol. 9, no. 3, pp. 317–334, 2009.

[23] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, no. 1, pp. 34–81, 2009.

[24] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[25] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via 1 minimization," *Proceedings of the National Academy of Sciences*, vol. 100, no. 5, pp. 2197–2202, 2003.

[26] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[27] R. Giri and B. D. Rao, "Type I and Type II Bayesian Methods for Sparse Signal Recovery using Scale Mixtures," *IEEE Transactions on Signal Processing*, vol. 64, 2016.

[28] T. Park and G. Casella, "The bayesian lasso," *Journal of the American Statistical Association*, vol. 103, no. 482, pp. 681–686, 2008.

[29] J. Palmer, K. Kreutz-Delgado, B. D. Rao, and D. P. Wipf, "Variational em algorithms for non-gaussian latent variable models," in *Advances in Neural Information Processing Systems*, 2006, pp. 1059–1066.

[30] P. Garrigues and B. A. Olshausen, "Group sparse coding with a laplacian scale mixture prior," in *Advances in neural information processing systems*, 2010, pp. 676–684.

[31] R. Giri and B. D. Rao, "Type I and Type II Bayesian Methods for Sparse Signal Recovery Using Scale Mixtures." *IEEE Transactions on Signal Processing*, vol. 64, no. 13, pp. 3418–3428, 2016.

[32] C. L. Lawson and R. J. Hanson, *Solving least squares problems*. SIAM, 1974, vol. 161.

[33] B. M. Jedynak and S. Khudanpur, "Maximum likelihood set for estimating a probability mass function," *Neural Computation*, vol. 17, no. 7, pp. 1508–1530, 2005.

[34] H. Kim and H. Park, "Nonnegative matrix factorization based on alternating nonnegativity constrained least squares and active set method," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 2, pp. 713–730, 2008.

[35] V. P. Pauca, F. Shahnaz, M. W. Berry, and R. J. Plemmons, "Text mining rsing non-negative matrix factorizations," in *Proceedings of the 2004 SIAM International Conference on Data Mining*, vol. 4, 2004, pp. 452–456.

[36] V. Monga and M. K. Mihçak, "Robust and secure image hashing via non-negative matrix factorizations," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 3, pp. 376–390, 2007.

[37] P. C. Loizou, "Speech enhancement based on perceptually motivated Bayesian estimators of the magnitude spectrum," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 5, pp. 857–869, 2005.

[38] C. Févotte, N. Bertin, and J.-L. Durrieu, "Nonnegative matrix factorization with the Itakura-Saito divergence: With application to music analysis," *Neural Computation*, vol. 21, no. 3, pp. 793–830, 2009.

[39] Y. Lin and D. D. Lee, "Bayesian regularization and nonnegative deconvolution for room impulse response estimation," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 839–847, 2006.

[40] A. Ghosh, T. Megherbi, F. O. Boumghar, and R. Deriche, "Fiber orientation distribution from non-negative sparse recovery," in *10th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2013, pp. 254–257.

[41] H. Kim and H. Park, "Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis," *Bioinformatics*, vol. 23, no. 12, pp. 1495–1502, 2007.

[42] I. Fedorov, A. Nalci, R. Giri, B. D. Rao, T. Q. Nguyen, and H. Garudadri, "A unified framework for sparse non-negative least squares using multiplicative updates and the non-negative matrix factorization problem," *Signal Processing*, 2018.

[43] S. Ogawa, T.-M. Lee, A. R. Kay, and D. W. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation," *Proceedings of the National Academy of Sciences*, vol. 87, no. 24, pp. 9868–9872, 1990.

[44] R. B. Buxton, E. C. Wong, and L. R. Frank, "Dynamics of blood flow and oxygenation changes during brain activation: the balloon model," *Magnetic Resonance in Medicine*, vol. 39, no. 6, pp. 855–864, 1998.

[45] P. T. Fox and M. E. Raichle, "Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects," *Proceedings of the National Academy of Sciences*, vol. 83, no. 4, pp. 1140–1144, 1986.

[46] P. T. Fox, M. E. Raichle, M. A. Mintun, and C. Dence, "Nonoxidative glucose consumption during focal physiologic neural activity," *Science*, vol. 241, no. 4864, pp. 462–464, 1988.

[47] L. Pauling and C. D. Coryell, "The magnetic properties and structure of hemoglobin, oxyhemoglobin and carbonmonoxyhemoglobin," *Proceedings of the National Academy of Sciences*, vol. 22, no. 4, pp. 210–216, 1936.

[48] D. G. Nishimura, *Principles of Magnetic Resonance Imaging*. Stanford University, 2010.

[49] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman, "A default mode of brain function," *Proceedings of the National Academy of Sciences*, vol. 98, no. 2, pp. 676–682, 2001.

[50] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle, "The human brain is intrinsically organized into dynamic, anticorrelated functional networks," *Proceedings of the National Academy of Sciences*, vol. 102, no. 27, pp. 9673–9678, 2005.

[51] M. D. Fox, D. Zhang, A. Z. Snyder, and M. E. Raichle, "The global signal and observed anticorrelated resting state brain networks." *Journal of Neurophysiology*, vol. 101, no. 6, pp. 3270–3283, Jun. 2009.

[52] K. Murphy, R. M. Birn, D. A. Handwerker, T. B. Jones, and P. A. Bandettini, "The impact of global signal regression on resting state correlations: are anti-correlated networks introduced?" *NeuroImage*, vol. 44, no. 3, pp. 893–905, Feb. 2009.

[53] K. Murphy and M. D. Fox, "Towards a consensus regarding global signal regression for resting state functional connectivity mri," *NeuroImage*, vol. 154, pp. 169–173, 2017.

[54] A. Nalci, B. D. Rao, and T. T. Liu, "Global signal regression acts as a temporal down-weighting process in resting-state fMRI," *NeuroImage*, vol. 152, pp. 602–618, 2017.

[55] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun, "Tracking whole-brain connectivity dynamics in the resting state," *Cerebral Cortex*, vol. 24, no. 3, pp. 663–676, 2014.

[56] R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. Della Penna, J. H. Duyn, G. H. Glover, J. Gonzalez-Castillo, D. A. Handwerker, S. Keilholz, V. Kiviniemi, D. A. Leopold, F. de Pasquale, O. Sporns, M. Walter, and C. Chang, "Dynamic functional connectivity: Promise, issues, and interpretations." *NeuroImage*, vol. 80, pp. 360–378, Oct. 2013.

[57] M. G. Preti, T. A. Bolton, and D. Van De Ville, "The dynamic functional connectome: State-of-the-art and perspectives." *NeuroImage*, vol. 160, pp. 41–54, Oct. 2017.

[58] R. B. Buxton, K. Uludağ, D. J. Dubowitz, and T. T. Liu, "Modeling the hemodynamic response to brain activation," *NeuroImage*, vol. 23, pp. S220–S233, 2004.

[59] T. T. Liu, "Noise contributions to the fMRI signal: An overview," *NeuroImage*, vol. 143, pp. 141–151, 2016.

[60] M. G. Bright and K. Murphy, "Is fMRI "noise" really noise? Resting state nuisance regressors remove variance with network structure." *NeuroImage*, vol. 114, pp. 158–169, Jul. 2015.

[61] M. G. Bright, C. R. Tench, and K. Murphy, "Potential pitfalls when denoising resting state fmri data using nuisance regression," *NeuroImage*, vol. 154, pp. 159–168, 2017.

[62] M. N. Hallquist, K. Hwang, and B. Luna, "The nuisance of nuisance regression: spectral misspecification in a common approach to resting-state fMRI preprocessing reintroduces noise and obscures functional connectivity." *NeuroImage*, vol. 82, pp. 208–225, Nov. 2013.

[63] R. M. Birn, M. A. Smith, T. B. Jones, and P. A. Bandettini, "The respiration response function: the temporal dynamics of fMRI signal fluctuations related to changes in respiration." *NeuroImage*, vol. 40, no. 2, pp. 644–654, Apr. 2008.

[64] C. Chang, J. P. Cunningham, and G. H. Glover, "Influence of heart rate on the BOLD signal: the cardiac response function." *NeuroImage*, vol. 44, no. 3, pp. 857–869, Feb. 2009.

[65] M. F. Glasser, T. S. Coalson, J. D. Bijsterbosch, S. J. Harrison, M. P. Harms, A. Anticevic, D. C. Van Essen, and S. M. Smith, "Using temporal ICA to selectively remove global noise while preserving global signal in functional MRI data," *bioRxiv*, p. 193862, 2018.

[66] 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," *Proceedings of the National Academy of Sciences*, vol. 101, no. 13, pp. 4637–4642, 2004.

[67] G. J. Yang, J. D. Murray, G. Repovs, M. W. Cole, A. Savic, M. F. Glasser, C. Pittenger, J. H. Krystal, X.-J. Wang, G. D. Pearlson, D. C. Glahn, and A. Anticevic, "Altered global brain signal in schizophrenia." *Proceedings of the National Academy of Sciences*, vol. 111, no. 20, pp. 7438–7443, May 2014.

[68] K. Wang, M. Liang, L. Wang, L. Tian, X. Zhang, K. Li, and T. Jiang, "Altered functional connectivity in early Alzheimer's disease: A resting-state fMRI study," *Human Brain Mapping*, vol. 28, no. 10, pp. 967–978, 2007.

[69] S. Baudrexel, T. Witte, C. Seifried, F. von Wegner, F. Beissner, J. C. Klein, H. Steinmetz, R. Deichmann, J. Roeper, and R. Hilker, "Resting state fMRI reveals increased subthalamic nucleus-motor cortex connectivity in Parkinson's disease," *NeuroImage*, vol. 55, no. 4, pp. 1728–1738, 2011.

[70] M. D. Greicius, B. H. Flores, V. Menon, G. H. Glover, H. B. Solvason, H. Kenna, A. L. Reiss, and A. F. Schatzberg, "Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus," *Biological Psychiatry*, vol. 62, no. 5, pp. 429–437, 2007.

[71] Y. Liu, M. Liang, Y. Zhou, Y. He, Y. Hao, M. Song, C. Yu, H. Liu, Z. Liu, and T. Jiang, "Disrupted small-world networks in schizophrenia," *Brain*, vol. 131, no. 4, pp. 945–961, 2008.

[72] P. Sajda, S. Du, T. R. Brown, R. Stoyanova, D. C. Shungu, X. Mao, and L. C. Parra, "Nonnegative matrix factorization for rapid recovery of constituent spectra in magnetic resonance chemical shift imaging of the brain," *IEEE Transactions on Medical Imaging*, vol. 23, no. 12, pp. 1453–1465, 2004.

[73] L. C. Potter, E. Ertin, J. T. Parker, and M. Cetin, "Sparsity and compressed sensing in radar imaging," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1006–1020, 2010.

[74] A. Hurmalainen, R. Saeidi, and T. Virtanen, "Group sparsity for speaker identity discrimination in factorisation-based speech recognition," in *Interspeech*, 2012.

[75] M. Lustig, J. M. Santos, D. L. Donoho, and J. M. Pauly, "kt SPARSE: High frame rate dynamic MRI exploiting spatio-temporal sparsity," in *Proceedings of the 13th Annual Meeting of ISMRM*, vol. 2420, 2006.

[76] J. Meng, J. M. Zhang, Y. Chen, and Y. Huang, "Bayesian non-negative factor analysis for reconstructing transcription factor mediated regulatory networks," *Proteome Science*, vol. 9, no. 1, p. S9, 2011.

[77] A. Nalci, B. Rao, and T. T. Liu, "Sparse Estimation of Quasi-periodic Spatiotemporal Components in Resting-State fMRI," in *Proceedings of the 24th Annual Meeting of the ISMRM*, 2016, p. 3824.

[78] T. T. Liu, A. Nalci, and M. Falahpour, "The global signal in fMRI: Nuisance or Information?" *NeuroImage*, vol. 150, pp. 213–229, 2017.

[79] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani *et al.*, "Least angle regression," *The Annals of Statistics*, vol. 32, no. 2, pp. 407–499, 2004.

[80] M. A. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586–597, 2007.

[81] S. J. Wright, R. D. Nowak, and M. A. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2479–2493, 2009.

[82] D. L. Donoho and J. Tanner, "Sparse nonnegative solution of underdetermined linear equations by linear programming," *Proceedings of the National Academy of Sciences*, vol. 102, no. 27, pp. 9446–9451, 2005.

[83] J. Nocedal and S. Wright, *Numerical Optimization*. Springer Science & Business Media, 2006.

[84] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[85] M. A. Khajehnejad, A. G. Dimakis, W. Xu, and B. Hassibi, "Sparse recovery of nonnegative signals with minimal expansion," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 196–208, 2011.

[86] C.-b. Lin, "Projected gradient methods for nonnegative matrix factorization," *Neural Computation*, vol. 19, no. 10, pp. 2756–2779, 2007.

[87] P. D. Grady and S. T. Rickard, "Compressive sampling of non-negative signals," in *IEEE Workshop on Machine Learning for Signal Processing*. IEEE, 2008, pp. 133–138.

[88] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2008, pp. 3869–3872.

[89] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian compressive sensing using Laplace priors," *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 53–63, 2010.

[90] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2346–2356, 2008.

[91] D. F. Andrews and C. L. Mallows, "Scale mixtures of normal distributions," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 99–102, 1974.

[92] J. Palmer, K. Kreutz-Delgado, B. D. Rao, and D. P. Wipf, "Variational EM algorithms for non-Gaussian latent variable models," in *Advances in Neural Information Processing Systems*, 2005, pp. 1059–1066.

[93] K. Lange and J. S. Sinsheimer, "Normal/independent distributions and their applications in robust regression," *Journal of Computational and Graphical Statistics*, vol. 2, no. 2, pp. 175–198, 1993.

[94] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Iteratively reweighted least squares for linear regression when errors are normal/independent distributed," *Multivariate Analysis V*, pp. 35–57, 1980.

[95] ——, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society. Series B*, pp. 1–38, 1977.

[96] D. P. Wipf, B. D. Rao, and S. Nagarajan, "Latent variable Bayesian models for promoting sparsity," *IEEE Transactions on Information Theory*, vol. 57, no. 9, pp. 6236–6255, 2011.

[97] M. Al-Shoukairi, P. Schniter, and B. D. Rao, "A GAMP-based low complexity sparse Bayesian learning algorithm," *IEEE Transactions on Signal Processing*, vol. 66, no. 2, pp. 294–308, 2018.

[98] D. P. Wipf and B. D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Transactions on Signal Processing*, vol. 52, no. 8, pp. 2153–2164, 2004.

[99] L. Zhang, M. Yang, and X. Feng, "Sparse representation or collaborative representation: Which helps face recognition?" in *IEEE International Conference on Computer Vision (ICCV)*. IEEE, 2011, pp. 471–478.

[100] I. Fedorov, B. D. Rao, and T. Q. Nguyen, "Multimodal sparse Bayesian dictionary learning applied to multimodal data classification," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 2237–2241.

[101] M. Harva and A. Kabán, "Variational learning for rectified factor analysis," *Signal Processing*, vol. 87, no. 3, pp. 509–527, 2007.

[102] J. W. Miskin, "Ensemble learning for independent component analysis," in *Advances in Independent Component Analysis*, 2000.

[103] M. A. Figueiredo, "Adaptive sparseness for supervised learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1150–1159, 2003.

[104] J. P. Vila and P. Schniter, "An empirical-Bayes approach to recovering linearly constrained non-negative sparse signals," *IEEE Transactions on Signal Processing*, vol. 62, no. 18, pp. 4689–4703, 2014.

[105] J.-L. Gauvain and C.-H. Lee, "Maximum a posteriori estimation for multivariate gaussian mixture observations of markov chains," *IEEE Transactions on Speech and Audio Processing*, vol. 2, no. 2, pp. 291–298, 1994.

[106] C. M. Bishop, "Pattern Recognition," *Machine Learning*, 2006.

[107] Z. Zhang, T.-P. Jung, S. Makeig, Z. Pi, and B. Rao, "Spatiotemporal sparse Bayesian learning with applications to compressed sensing of multichannel physiological signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 6, pp. 1186–1197, 2014.

[108] W. C. Horrace, "Some results on the multivariate truncated normal distribution," *Journal of Multivariate Analysis*, vol. 94, no. 1, pp. 209–221, 2005.

[109] C. Robert and G. Casella, *Monte Carlo statistical methods*. Springer Science & Business Media, 2013.

[110] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, "Hybrid Monte Carlo," *Physics Letters B*, vol. 195, no. 2, pp. 216–222, 1987.

[111] A. Pakman and L. Paninski, "Exact Hamiltonian Monte Carlo for truncated multivariate Gaussians," *Journal of Computational and Graphical Statistics*, vol. 23, no. 2, pp. 518–542, 2014.

[112] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, "An introduction to MCMC for machine learning," *Machine Learning*, vol. 50, no. 1-2, pp. 5–43, 2003.

[113] R. P. Sherman, Y.-Y. K. Ho, and S. R. Dalal, "Conditions for convergence of Monte Carlo EM sequences with an application to product diffusion modeling," *The Econometrics Journal*, vol. 2, no. 2, pp. 248–267, 1999.

[114] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 6, pp. 721–741, 1984.

[115] Y. Li and S. K. Ghosh, "Efficient sampling methods for truncated multivariate normal and Student-t distributions subject to linear inequality constraints," *Journal of Statistical Theory and Practice*, vol. 9, no. 4, pp. 712–732, 2015.

[116] R. C. Neath *et al.*, "On convergence properties of the Monte Carlo EM algorithm," in *Advances in Modern Statistical Theory and Applications*. Institute of Mathematical Statistics, 2013, pp. 43–62.

[117] P. J. Bickel and E. Levina, "Covariance regularization by thresholding," *The Annals of Statistics*, pp. 2577–2604, 2008.

[118] T. Tong, C. Wang, and Y. Wang, "Estimation of variances and covariances for high-dimensional data: a selective review," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 6, no. 4, pp. 255–264, 2014.

[119] G. Celeux and J. Diebolt, "A stochastic approximation type EM algorithm for the mixture problem," *Stochastics: An International Journal of Probability and Stochastic Processes*, vol. 41, no. 1-2, pp. 119–134, 1992.

[120] T. J. Fisher and X. Sun, "Improved Stein-type shrinkage estimators for the high-dimensional multivariate normal covariance matrix," *Computational Statistics & Data Analysis*, vol. 55, no. 5, pp. 1909–1918, 2011.

[121] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of Multivariate Analysis*, vol. 88, no. 2, pp. 365–411, 2004.

[122] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Prentice Hall Upper Saddle River, NJ, 2000, vol. 1.

[123] J. W. Miskin, "Ensemble learning for independent component analysis," in *Advances in Independent Component Analysis*, 2000.

[124] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, 2011, pp. 2168–2172.

[125] S. Rangan, P. Schniter, and A. Fletcher, "On the convergence of approximate message passing with arbitrary matrices," in *IEEE International Symposium on Information Theory (ISIT)*, 2014, pp. 236–240.

[126] F. Caltagirone, L. Zdeborová, and F. Krzakala, "On convergence of approximate message passing," in *IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2014, pp. 1812–1816.

[127] S. Rangan, P. Schniter, and A. Fletcher, "On the convergence of approximate message passing with arbitrary matrices," in *IEEE International Symposium on Information Theory (ISIT)*, June 2014, pp. 236–240.

[128] A. Javanmard and A. Montanari, "State evolution for general approximate message passing algorithms with applications to spatial coupling," *Information and Inference*, 2013.

[129] J. P. Vila and P. Schniter, "Expectation-maximization Gaussian-mixture approximate message passing," *IEEE Transactions on Signal Processing*, vol. 61, no. 19, pp. 4658–4672, 2013.

[130] W. C. Horrace, "On ranking and selection from independent truncated normal distributions," *Journal of Econometrics*, vol. 126, no. 2, pp. 335–354, 2005.

[131] M. Magdon-Ismail and J. T. Purnell, "Approximating the covariance matrix of GMMs with low-rank perturbations," in *Intelligent Data Engineering and Automated Learning*. Springer, 2010, pp. 300–307.

[132] J. Liu, S. Ji, J. Ye *et al.*, "SLEP: Sparse learning with efficient projections," *Arizona State University*, vol. 6, p. 491, 2009.

[133] A. M. Bruckstein, M. Elad, and M. Zibulevsky, "On the uniqueness of nonnegative sparse solutions to underdetermined systems of equations," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 4813–4820, 2008.

[134] E. J. Candes, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[135] J. Vila, P. Schniter, S. Rangan, F. Krzakala, and L. Zdeborová, "Adaptive damping and mean removal for the generalized approximate message passing algorithm," in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*, 2015, pp. 2021–2025.

[136] T. Li and Z. Zhang, "Robust face recognition via block sparse bayesian learning," *Mathematical Problems in Engineering*, vol. 2013, 2013.

[137] R. He, W.-S. Zheng, B.-G. Hu, and X.-W. Kong, "Two-stage nonnegative sparse representation for large-scale face recognition," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 1, pp. 35–46, 2013.

[138] R. He, W.-S. Zheng, and B.-G. Hu, "Maximum correntropy criterion for robust face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1561–1576, 2011.

[139] N. Vo, B. Moran, and S. Challa, "Nonnegative-least-square classifier for face recognition," *Advances in Neural Networks–ISNN 2009*, pp. 449–456, 2009.

[140] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2009.

[141] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[142] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using laplacianfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 328–340, 2005.

[143] A. M. Martinez, "The AR face database," *CVC technical report*, 1998.

[144] C. Chang and G. H. Glover, "Time-frequency dynamics of resting-state brain connectivity measured with fMRI," *NeuroImage*, vol. 50, no. 1, pp. 81–98, 2010.

[145] D. A. Handwerker, V. Roopchansingh, J. Gonzalez-Castillo, and P. A. Bandettini, "Periodic changes in fMRI connectivity," *NeuroImage*, vol. 63, no. 3, pp. 1712–1719, 2012.

[146] W. Majeed, M. Magnuson, W. Hasenkamp, H. Schwarb, E. H. Schumacher, L. Barsalou, and S. D. Keilholz, "Spatiotemporal dynamics of low frequency BOLD fluctuations in rats and humans." *NeuroImage*, vol. 54, no. 2, pp. 1140–1150, Jan. 2011.

[147] W. Majeed, M. Magnuson, and S. D. Keilholz, "Spatiotemporal dynamics of low frequency fluctuations in BOLD fMRI of the rat." *Journal of Magnetic Resonance Imaging*, vol. 30, no. 2, pp. 384–393, Aug. 2009.

[148] G. J. Thompson, W.-J. Pan, M. E. Magnuson, D. Jaeger, and S. D. Keilholz, "Quasi-periodic patterns (QPP): Large-scale dynamics in resting state fMRI that correlate with local infraslow electrical activity," *NeuroImage*, vol. 84, no. C, pp. 1018–1031, Jan. 2014.

[149] A. Nalci and T. T. Liu, "Deterministic Estimation of Spatiotemporal Motifs in Resting-State fMRI," in *Proceedings of the 24th Annual Meeting of the ISMRM*, 2016, p. 4586.

[150] M. D. Fox and M. E. Raichle, "Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging," *Nature Reviews Neuroscience*, vol. 8, no. 9, pp. 700–711, Sep. 2007.

[151] B. Biswal, F. Z. Yetkin, V. M. Haughton, and J. S. Hyde, "Functional connectivity in the motor cortex of resting human brain using echo-planar MRI." *Magnetic Resonance in Medicine*, vol. 34, no. 4, pp. 537–541, Oct. 1995.

[152] D. Cordes, V. M. Haughton, K. Arfanakis, G. J. Wendt, P. A. Turski, C. H. Moritz, M. A. Quigley, and M. E. Meyerand, "Mapping functionally related regions of brain with functional connectivity MR imaging," *American Journal of Neuroradiology*, vol. 21, no. 9, pp. 1636–1644, 2000.

[153] D. Cordes, V. M. Haughton, K. Arfanakis, J. D. Carew, P. A. Turski, C. H. Moritz, M. A. Quigley, and M. E. Meyerand, "Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data." *American Journal of Neuroradiology*, vol. 22, no. 7, pp. 1326–1333, Aug. 2001.

[154] M. Hampson, B. S. Peterson, P. Skudlarski, J. C. Gatenby, and J. C. Gore, "Detection of functional connectivity using temporal correlations in MR images," *Human Brain Mapping*, vol. 15, no. 4, pp. 247–262, 2002.

[155] A. Weissenbacher, C. Kasess, F. Gerstl, R. Lanzenberger, E. Moser, and C. Windischberger, "Correlations and anticorrelations in resting-state functional connectivity MRI: a quantitative comparison of preprocessing strategies," *NeuroImage*, vol. 47, no. 4, pp. 1408–1416, Oct. 2009.

[156] J. S. Anderson, T. J. Druzgal, M. Lopez-Larson, E.-K. Jeong, K. Desai, and D. Yurgelun-Todd, "Network anticorrelations, global regression, and phase-shifted soft tissue correction." *Human Brain Mapping*, vol. 32, no. 6, pp. 919–934, Jun. 2011.

[157] Z. S. Saad, S. J. Gotts, K. Murphy, G. Chen, H. J. Jo, A. Martin, and R. W. Cox, "Trouble at rest: how correlation patterns and group differences become distorted after global signal regression." *Brain Connectivity*, vol. 2, no. 1, pp. 25–32, 2012.

[158] X. J. Chai, A. N. Castañón, D. Ongür, and S. Whitfield-Gabrieli, "Anticorrelations in resting state networks without global signal regression," *NeuroImage*, vol. 59, no. 2, pp. 1420–1428, Jan. 2012.

[159] C. Chang and G. H. Glover, "Effects of model-based physiological noise correction on default mode network anti-correlations and correlations." *NeuroImage*, vol. 47, no. 4, pp. 1448–1459, Oct. 2009.

[160] C. W. Wong, V. Olafsson, O. Tal, and T. T. Liu, "Anti-correlated networks, global signal regression, and the effects of caffeine in resting-state functional MRI." *NeuroImage*, vol. 63, no. 1, pp. 356–364, Oct. 2012.

[161] Z. Liang, J. King, and N. Zhang, "Anticorrelated resting-state functional connectivity in awake rat brain." *NeuroImage*, vol. 59, no. 2, pp. 1190–1199, Jan. 2012.

[162] H. He and T. T. Liu, "A geometric view of global signal confounds in resting-state functional MRI," *NeuroImage*, vol. 59, no. 3, pp. 2339–48, Feb 2012.

[163] F. Carbonell, P. Bellec, and A. Shmuel, "Quantification of the impact of a confounding variable on functional connectivity confirms anti-correlated networks in the resting-state." *NeuroImage*, vol. 86, pp. 343–353, Feb. 2014.

[164] M. L. Schölvinck, A. Maier, F. Q. Ye, J. H. Duyn, and D. A. Leopold, "Neural basis of global resting-state fMRI activity." *Proceedings of the National Academy of Sciences*, vol. 107, no. 22, pp. 10 238–10 243, Jun. 2010.

[165] C. W. Wong, V. Olafsson, O. Tal, and T. T. Liu, "The amplitude of the resting-state fMRI global signal is related to EEG vigilance measures." *NeuroImage*, vol. 83, pp. 983–990, Dec. 2013.

[166] H. Wen and Z. Liu, "Broadband Electrophysiological Dynamics Contribute to Global Resting-State fMRI Signal." *Journal of Neuroscience*, vol. 36, no. 22, pp. 6030–6040, Jun. 2016.

[167] J. D. Power, B. L. Schlaggar, and S. E. Petersen, "Recent progress and outstanding issues in motion correction in resting state fMRI." *NeuroImage*, vol. 105, pp. 536–551, Jan. 2015.

[168] J. M. Li, W. J. Bentley, A. Z. Snyder, M. E. Raichle, and L. H. Snyder, "Functional connectivity arises from a slow rhythmic mechanism." *Proceedings of the National Academy of Sciences*, vol. 112, no. 19, pp. E2527–35, May 2015.

[169] A. Hahamy, V. Calhoun, G. Pearlson, M. Harel, N. Stern, F. Attar, R. Malach, and R. Salomon, "Save the global: global signal connectivity as a tool for studying clinical populations with functional magnetic resonance imaging." *Brain Connectivity*, vol. 4, no. 6, pp. 395–403, Aug. 2014.

[170] S. J. Gotts, Z. S. Saad, H. J. Jo, G. L. Wallace, R. W. Cox, and A. Martin, "The perils of global signal regression for group comparisons: a case study of Autism Spectrum Disorders." *Frontiers in Human Neuroscience*, vol. 7, p. 356, 2013.

[171] D. C. Hoaglin and R. E. Welsch, "The Hat Matrix in Regression and ANOVA," *The American Statistician*, vol. 32, no. 1, pp. 17–22, Feb. 1978.

[172] N. R. Draper and H. Smith, *Applied Regression Analysis*, 3rd ed. John Wiley & Sons, Aug. 2014.

[173] J. D. Power, M. Plitt, T. O. Laumann, and A. Martin, "Sources and implications of whole-brain fMRI signals in humans," *NeuroImage*, vol. 146, pp. 609–625, 2017.

[174] M. D. Fox, A. Z. Snyder, J. L. Vincent, and M. E. Raichle, "Intrinsic fluctuations within cortical systems account for intertrial variability in human behavior," *Neuron*, vol. 56, no. 1, pp. 171–184, 2007.

[175] R. W. Cox, "AFNI: software for analysis and visualization of functional magnetic resonance neuroimages." *Computers and Biomedical Research*, vol. 29, no. 3, pp. 162–173, Jun. 1996.

[176] L. Yan, Y. Zhuo, Y. Ye, S. X. Xie, J. An, G. K. Aguirre, and J. Wang, "Physiological origin of low-frequency drift in blood oxygen level dependent (BOLD) functional magnetic resonance imaging (fMRI)." *Magnetic Resonance in Medicine*, vol. 61, no. 4, pp. 819–827, Apr. 2009.

[177] P. M. Macey, K. E. Macey, R. Kumar, and R. M. Harper, "A method for removal of global effects from fMRI time series," *NeuroImage*, vol. 22, no. 1, pp. 360–366, May 2004.

[178] K. R. Van Dijk, T. Hedden, A. Venkataraman, K. C. Evans, S. W. Lazar, and R. L. Buckner, "Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization," *Journal of Neurophysiology*, vol. 103, no. 1, pp. 297–321, 2010.

[179] C. M. Lacadie, R. K. Fulbright, N. Rajeevan, R. T. Constable, and X. Papademetris, "More accurate Talairach coordinates for neuroimaging using non-linear registration," *NeuroImage*, vol. 42, no. 2, pp. 717–725, 2008.

[180] M. D. Fox, M. Corbetta, A. Z. Snyder, J. L. Vincent, and M. E. Raichle, "Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems," *Proceedings of the National Academy of Sciences*, vol. 103, no. 26, pp. 10 046–10 051, 2006.

[181] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median," *Journal of Experimental Social Psychology*, vol. 49, no. 4, pp. 764–766, 2013.

[182] C. Reimann, P. Filzmoser, and R. G. Garrett, "Background and threshold: critical comparison of methods of determination," *Science of the Total Environment*, vol. 346, no. 1, pp. 1–16, 2005.

[183] W. L. Martinez, A. R. Martinez, A. Martinez, and J. Solka, *Exploratory data analysis with MATLAB*. CRC Press, 2010.

[184] T. D. Wager, M. C. Keller, S. C. Lacey, and J. Jonides, "Increased sensitivity in neuroimaging analyses using robust regression," *NeuroImage*, vol. 26, no. 1, pp. 99–113, 2005.

[185] J. D. Power, K. A. Barnes, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen, "Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion," *NeuroImage*, vol. 59, no. 3, pp. 2142–2154, Feb. 2012.

[186] F. I. Karahanoğlu and D. Van De Ville, "Transient brain activity disentangles fMRI resting-state dynamics in terms of spatially and temporally overlapping networks," *Nature Communications*, vol. 6, no. 7751, July 2015.

[187] A. Walther, H. Nili, N. Ejaz, A. Alink, N. Kriegeskorte, and J. Diedrichsen, "Reliability of dissimilarity measures for multi-voxel pattern analysis," *NeuroImage*, vol. 137, pp. 188–200, December 2015.

[188] J. S. Siegel, J. D. Power, J. W. Dubis, A. C. Vogel, J. A. Church, B. L. Schlaggar, and S. E. Petersen, "Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points." *Human Brain Mapping*, vol. 35, no. 5, pp. 1981–1996, May 2014.

[189] M. Falahpour, C. W. Wong, and T. T. Liu, "The resting state fMRI global signal is negatively correlated with time-varying EEG vigilance," in *Proceedings of the 24th Annual Meeting of the ISMRM*, 2016, p. 641.

[190] C. Chang, D. A. Leopold, M. L. Schölvinck, H. Mandelkow, D. Picchioni, X. Liu, F. Q. Ye, J. N. Turchi, and J. H. Duyn, "Tracking brain arousal fluctuations with fMRI." *Proceedings of the National Academy of Sciences*, vol. 113, no. 16, pp. 4518–4523, Apr. 2016.

[191] X. Liu, T. Yanagawa, D. Leopold, C. Chang, H. Ishida, i. N. Fuji, and J. Duyn, "Arousal transitions in sleep, wakefulness, and anesthesia are characterized by an orderly sequence of cortical events," *NeuroImage*, vol. 116, pp. 222–231, Aug. 2015.

[192] M. A. Pisauro, A. Benucci, and M. Carandini, "Local and global contributions to hemodynamic activity in mouse cortex." *Journal of Neurophysiology*, vol. 115, no. 6, pp. 2931–2936, Jun. 2016.

[193] B. E. Jones, "From waking to sleeping: neuronal and chemical substrates." *Trends in Pharmacological Sciences*, vol. 26, no. 11, pp. 578–586, Nov. 2005.

[194] D. Picchioni, J. H. Duyn, and S. G. Horovitz, "Sleep and the functional connectome." *NeuroImage*, vol. 80, pp. 387–396, Oct. 2013.

[195] E. Tagliazucchi, P. Balenzuela, D. Fraiman, and D. R. Chialvo, "Criticality in large-scale brain fMRI dynamics unveiled by a novel point process analysis." *Frontiers in Physiology*, vol. 3, no. 15, 2012.

[196] X. Liu and J. H. Duyn, "Time-varying functional network information extracted from brief instances of spontaneous brain activity," *Proceedings of the National Academy of Sciences*, vol. 110, no. 11, pp. 4392–4397, 2013.

[197] J. J. Remes, T. Starck, J. Nikkinen, E. Ollila, C. F. Beckmann, O. Tervonen, V. Kiviniemi, and O. Silven, "Effects of repeatability measures on results of fMRI sICA: a study on simulated and real resting-state effects," *NeuroImage*, vol. 56, no. 2, pp. 554–569, May 2011.

[198] F. Carbonell, P. Bellec, and A. Shmuel, "Global and system-specific resting-state fMRI fluctuations are uncorrelated: principal component analysis reveals anti-correlated networks." *Brain Connectivity*, vol. 1, no. 6, pp. 496–510, 2011.

[199] V. D. Calhoun, R. Miller, G. Pearlson, and T. Adalı, "The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery," *Neuron*, vol. 84, no. 2, pp. 262–274, 2014.

[200] M. Yaesoubi, E. A. Allen, R. L. Miller, and V. D. Calhoun, "Dynamic coherence analysis of resting fMRI data to jointly capture state-based phase, frequency, and time-domain information," *NeuroImage*, vol. 120, pp. 133–142, 2015.

[201] D. Vidaurre, S. M. Smith, and M. W. Woolrich, "Brain network dynamics are hierarchically organized in time." *Proceedings of the National Academy of Sciences*, vol. 114, no. 48, pp. 12 827–12 832, Nov. 2017.

[202] K. Murphy, R. M. Birn, and P. A. Bandettini, "Resting-state fMRI confounds and cleanup," *NeuroImage*, vol. 80, pp. 349–359, 2013.

[203] F. Nikolaou, C. Orphanidou, P. Papakyriakou, K. Murphy, R. G. Wise, and G. D. Mitsis, "Spontaneous physiological variability modulates dynamic functional connectivity in resting-state functional magnetic resonance imaging." *Philosophical Transactions Series A, Mathematical, Physical, and Engineering Sciences*, vol. 374, no. 2067, p. 20150183, May 2016.

[204] K. Glomb, A. Ponce-Alvarez, M. Gilson, P. Ritter, and G. Deco, "Stereotypical modulations in dynamic functional connectivity explained by changes in BOLD variance." *NeuroImage*, vol. 171, pp. 40–54, Dec. 2017.

[205] T. T. Liu, "Noise contributions to the fMRI signal: An overview." *NeuroImage*, vol. 143, pp. 141–151, Sep. 2016.

[206] R. Ciric, D. H. Wolf, J. D. Power, D. R. Roalf, G. L. Baum, K. Ruparel, R. T. Shinohara, M. A. Elliott, S. B. Eickhoff, C. Davatzikos, R. C. Gur, R. E. Gur, D. S. Bassett, and T. D. Satterthwaite, "Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity." *NeuroImage*, vol. 154, pp. 174–187, Jul. 2017.

[207] H. Xu, J. Su, J. Qin, M. Li, L.-L. Zeng, D. Hu, and H. Shen, "Impact of global signal regression on characterizing dynamic functional connectivity and brain states," *NeuroImage*, p. In Press, 2018.

[208] A. Nalci, M. Falahpour, and T. T. Liu, "On dynamic functional connectivity and global signal regression," in *23rd Annual Meeting of the Organization for Human Brain Mapping*, 2017, p. 1919.

[209] A. Nalci and T. T. Liu, "Regression does not eliminate the effects of nuisance terms in dynamic functional connectivity estimates," in *Proceedings of the 26th Annual Meeting of the ISMRM*, 2018, p. 3579.

[210] M. Woolrich, S. Jbabdi, B. Patenaude, M. Chappell, S. Makni, T. Behrens, C. Beckmann, M. Jenkinson, and S. Smith, " Bayesian analysis of neuroimaging data in FSL," *NeuroImage*, vol. 45, pp. S173–S186, Mar. 2009.

[211] S. Smith, M. Jenkinson, M. Woolrich, C. Beckmann, T. Behrens, P. Johansen-Berg, H.and Bannister, I. De Luca, M.and Drobnjak, D. Flitney, R. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, J. Brady, and P. Matthews, "Advances in functional and structural MR image analysis and implementation as FSL," *NeuroImage*, vol. 23, pp. S208–S219, Sep. 2004.

[212] A. Rack-Gomer and T. Liu, "Caffeine increases the temporal variability of resting state BOLD connectivity in the motor cortex," *NeuroImage*, vol. 59, pp. 2994–3002, Oct. 2012.

[213] B. Efron and R. Tibshirani, "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy," *Statistical science*, pp. 54–75, 1986.

[214] A. Zalesky, A. Fornito, L. Cocchi, L. L. Gollo, and M. Breakspear, "Time-resolved resting-state brain networks," *Proceedings of the National Academy of Sciences*, vol. 111, no. 28, pp. 10 341–10 346, 2014.

[215] T. O. Laumann, A. Z. Snyder, A. Mitra, E. M. Gordon, C. Gratton, B. Adeyemo, A. W. Gilmore, S. M. Nelson, J. J. Berg, D. J. Greene, J. E. McCarthy, E. Tagliazucchi, H. Laufs, B. L. Schlaggar, N. U. F. Dosenbach, and S. E. Petersen, "On the Stability of BOLD fMRI Correlations." *Cerebral Cortex*, vol. 27, pp. 4719–4732, Sep. 2017.

[216] G. Sugihara, R. May, H. Ye, C.-h. Hsieh, E. Deyle, M. Fogarty, and S. Munch, "Detecting causality in complex ecosystems," *Science*, vol. 338, no. 6106, pp. 496–500, 2012.

[217] J. D. Power, A. Mitra, T. O. Laumann, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen, "Methods to detect, characterize, and remove motion artifact in resting state fMRI." *NeuroImage*, vol. 84, pp. 320–341, Jan. 2014.

[218] G. C. Burgess, S. Kandala, D. Nolan, T. O. Laumann, J. Power, B. Adeyemo, M. P. Harms, S. E. Petersen, and D. M. Barch, "Evaluation of Denoising Strategies To Address Motion-Correlated Artifact in Resting State fMRI Data from the Human Connectome Project." *Brain Connectivity*, Aug. 2016.

[219] J. D. Power, M. Plitt, S. J. Gotts, P. Kundu, V. Voon, P. Bandettini, and A. Martin, "Ridding fMRI data of motion-related influences: Removal of signals with distinct spatial and physical bases in multiecho data," *Proceedings of the National Academy of Sciences; DOI: 10.1073/pnas.1720985115*, 2018.

[220] M. Falahpour, C. Chang, C. W. Wong, and T. T. Liu, "Template-based prediction of vigilance fluctuations in resting-state fMRI," *NeuroImage*, vol. 174, pp. 317–327, 2018.

[221] A. Haimovici, E. Tagliazucchi, P. Balenzuela, and H. Laufs, "On wakefulness fluctuations as a source of BOLD functional connectivity dynamics," *Sci Rep*, vol. 7, p. 5908, 2017.

[222] C. Chang, P. Ozbay, J. Zwart de, D. Picchioni, M. Chappell, H. Mandelkow, and J. Duyn, "Covariation of Pulse Oximetry Amplitude and BOLD fMRI Across Vigilance States," in *Proceedings of the 26th Annual Meeting of the ISMRM*, 2018, p. 46.

[223] J. S. Siegel, A. Mitra, T. O. Laumann, B. A. Seitzman, M. Raichle, M. Corbetta, and A. Z. Snyder, "Data Quality Influences Observed Links Between Functional Connectivity and Behavior." *Cerebral Cortex*, vol. 27, no. 9, pp. 4492–4502, Sep. 2017.

[224] M. Pannunzi, R. Hindriks, R. G. Bettinardi, E. Wenger, N. Lisofsky, J. Martensson, O. Butler, E. Filevich, M. Becker, M. Lochstet, S. Kühn, and G. Deco, "Resting-state fMRI correlations: From link-wise unreliability to whole brain stability." *NeuroImage*, vol. 157, pp. 250–262, Aug. 2017.

[225] E. Tagliazucchi and H. Laufs, "Decoding Wakefulness Levels from Typical fMRI Resting-State Data Reveals Reliable Drifts between Wakefulness and Sleep," *Neuron*, vol. 82, pp. 695–708, 2014.

[226] R. Liégeois, T. O. Laumann, A. Z. Snyder, J. Zhou, and B. T. T. Yeo, "Interpreting temporal fluctuations in resting-state functional connectivity MRI." *NeuroImage*, vol. 163, pp. 437–455, Dec. 2017.

[227] V. Iacovella and U. Hasson, "The relationship between BOLD signal and autonomic nervous system functions: implications for processing of "physiological noise"." *Magnetic Resonance Imaging*, vol. 29, no. 10, pp. 1338–1345, Dec. 2011.

[228] G. U. Yule, "On the Theory of Correlation," *Journal of the Royal Statistical Society*, vol. 60, no. 4, pp. 812–854, Dec. 1897.

[229] S. M. Smith, K. L. Miller, S. Moeller, J. Xu, E. J. Auerbach, M. W. Woolrich, C. F. Beckmann, M. Jenkinson, J. Andersson, M. F. Glasser, D. C. Van Essen, D. A. Feinberg, E. S. Yacoub, and K. Ugurbil, "Temporally-independent functional modes of spontaneous brain activity." *Proceedings of the National Academy of Sciences*, vol. 109, no. 8, pp. 3131–3136, Feb. 2012.

[230] T. T. Liu, G. H. Glover, B. A. Mueller, D. N. Greve, J. Rasmussen, J. T. Voyvodic, J. A. Turner, T. G. M. van Erp, D. H. Mathalon, K. Andersen, L. Kun, G. G. Brown, D. B. Keator, V. D. Calhoun, H. J. Lee, J. M. Ford, M. Diaz, D. S. O'Leary, S. Gadde, A. Preda, K. O. Lim, C. G. Wible, H. S. Stern, A. Belger, G. McCarthy, B. Ozyurt, and S. G. Potkin, "Quality Assurance in Functional MRI," in *fMRI: From Nuclear Spins to Brain Function*, K. Ugurbil, L. Berliner, and K. Uludag, Eds. Springer, 2015, ch. 10, pp. 245–270.

[231] A. Nalci, B. D. Rao, and T. T. Liu, "Nuisance effects and the limitations of nuisance regression in dynamic functional connectivity fMRI," *NeuroImage*, vol. 184, pp. 1005–1031, 2019.

[232] T. D. Satterthwaite, D. H. Wolf, J. Loughead, K. Ruparel, M. A. Elliott, H. Hakonarson, R. C. Gur, and R. E. Gur, "Impact of in-scanner head motion on multiple measures of functional connectivity: relevance for studies of neurodevelopment in youth." *NeuroImage*, vol. 60, no. 1, pp. 623–632, Mar. 2012.

[233] K. R. A. Van Dijk, M. R. Sabuncu, and R. L. Buckner, "The influence of head motion on intrinsic functional connectivity MRI," *NeuroImage*, vol. 59, no. 1, pp. 431–438, Jan. 2012.

[234] M. P. Van Den Heuvel and H. E. H. Pol, "Exploring the brain network: a review on resting-state fMRI functional connectivity," *European Neuropsychopharmacology*, vol. 20, no. 8, pp. 519–534, 2010.

[235] S. A. Rombouts, J. S. Damoiseaux, R. Goekoop, F. Barkhof, P. Scheltens, S. M. Smith, and C. F. Beckmann, "Model-free group analysis shows altered BOLD FMRI networks in dementia." *Human Brain Mapping*, vol. 30, no. 1, pp. 256–266, 2009.

[236] B. Mohammadi, K. Kollewe, A. Samii, K. Krampfl, R. Dengler, and T. F. Münte, "Changes of resting state brain networks in amyotrophic lateral sclerosis," *Experimental Neurology*, vol. 217, no. 1, pp. 147–153, 2009.

[237] F. Agosta, E. Canu, P. Valsasina, N. Riva, A. Prelle, G. Comi, and M. Filippi, "Divergent brain network connectivity in amyotrophic lateral sclerosis," *Neurobiology of Aging*, vol. 34, no. 2, pp. 419–427, 2013.

[238] J. Muschelli, M. B. Nebel, B. S. Caffo, A. D. Barber, J. J. Pekar, and S. H. Mostofsky, "Reduction of motion-related artifacts in resting state fMRI using aCompCor." *NeuroImage*, vol. 96, pp. 22–35, Aug. 2014.

[239] Y. Behzadi, K. Restom, J. Liau, and T. T. Liu, "A component based noise correction method (CompCor) for BOLD and perfusion based fMRI," *NeuroImage*, vol. 37, no. 1, pp. 90–101, Aug 2007.

[240] G. H. Glover, T. Q. Li, and D. Ress, "Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR," *Magnetic Resonance in Medicine*, vol. 44, no. 1, pp. 162–167, Jul. 2000.

[241] R. H. R. Pruim, M. Mennes, D. van Rooij, A. Llera, J. K. Buitelaar, and C. F. Beckmann, "ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data." *NeuroImage*, vol. 112, pp. 267–277, May 2015.

[242] C. Caballero-Gaudes and R. Reynolds, "Methods for cleaning the BOLD fMRI signal," *NeuroImage*, vol. 154, pp. 128–149, 2017.

[243] C. W. Wong, P. N. DeYoung, and T. T. Liu, "Differences in the resting-state fMRI global signal amplitude between the eyes open and eyes closed states are related to changes in EEG vigilance." *NeuroImage*, vol. 124, no. Pt A, pp. 24–31, Aug. 2015.

[244] M. Falahpour, A. Nalci, and T. T. Liu, "The effects of global signal regression on estimates of resting-state blood oxygen-level-dependent functional magnetic resonance imaging and electroencephalogram vigilance correlations," *Brain Connectivity*, vol. 8, no. 10, pp. 618–627, 2018.