UC San Diego
UC San Diego Electronic Theses and Dissertations

Title
Optimization of Weights for Multi-echo fMRI Combination

Permalink
https://escholarship.org/uc/item/24b5k5md

Author
Li, Bochao

Publication Date
2019

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

Optimization of Weights for Multi-echo fMRI Combination

A Thesis submitted in partial satisfaction of the requirements for the degree Master of Science

in

Bioengineering

by

Bochao Li

Committee in charge:

Professor Thomas T. Liu, Chair
Professor Francisco Contijoch, Co-Chair
Professor Elliot R. McVeigh

2019

Copyright
Bochao Li, 2019
All rights reserved.

The Thesis of Bochao Li 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

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my special thanks of gratitude to my advisor, Professor Thomas Liu, who gave me the golden opportunity to do this wonderful project on the multi-echo fMRI. I was educated from him not only in doing a lot of research, designing plan but also other soft skills, such as presenting, writing and communicating with engineers. It was a great privilege and honor to work and study under his guidance.

I am extremely grateful to my parents for their love, caring and support for educating and preparing me for my future. Also I want to express my thanks to my friends and colleagues at CFMRI, Alican Nalci, Gokce Sarar, Maryam Falahpour, Conan Chen, Aaron Jacobson, who helped me a lot in teaching professional knowledge and collecting data. I am also very much thankful to Prof. Eric Wong who gave me the hand-on experience on MRI and Prof. Francisco Contijoch for his offer of first research experience at UCSD. And I give my special thanks to my friend, James Smith, who helped me improve my writing skills to have a better thesis.

Finally, my thanks go to all the people who have supported me to complete the research work directly or indirectly.

ABSTRACT OF THE THESIS

Optimization of Weights for Multi-echo fMRI Combination

by

Bochao Li

Master of Science in Bioengineering

University of California San Diego, 2019

Professor Thomas T. Liu, Chair
Professor Francisco Contijoch, Co-Chair

Due to the higher sensitivity to bold oxygenation level dependent (BOLD) signals and the ability to distinguish neurally related signals, multi-echo fMRI (ME-fMRI) has been a hot topic of research recently. Weighted combination in ME-fMRI is critical for the quality of the combined signal, which can be used for the analysis of brain functions. However, a convincing weighting method has never been put forward in previously published research. In this work, we optimized the weighting methods based on the Rayleigh quotient when using temporal signal-noise-ratio (tSNR) and multi-echo temporal signal-noise-ratio (metSNR) as metrics. These two metrics strongly represent the contrast-noise-ratio (CNR) which has been widely used to

estimate fMRI quality. In our case of steady-state fMRI, with the optimal weighting methods, both tSNR and metSNR of the combined signal were improved significantly as compared to other widely-used weighting methods. In the meantime, we obtained negative weights that were not used in the combination for previous research, which drove us to find the causes and re-define the normalization as 'L1 Normalization'. We also found the distribution characteristics of negative weights in terms of brain region and TE. The occurrence and features of negative weights were explained by the correlation between echoes' data and the change of the multipliers (e.g. the mean value). In addition, we found that the metSNR performs more robustly than tSNR when comparing non-optimal methods to optimal methods and explained their variations based on the ratio from a mathematical viewpoint.

Chapter 1

Introduction

For many years, functional MRI (fMRI) with echo planar imaging (EPI) [Mansfield, 1977] technology has been well-known for its sensitivity to blood oxygenation level dependent (BOLD) signals. However, the issue with signal echo fMRI is the difficulty to identify the sources of contrast signals, BOLD or other interference from physiological activities or hardware [Glover et al., 2000]. Since Multi-echo fMRI (ME-fMRI) was first proposed and verified that its high sensitivity to BOLD fluctuation could be increased by combining data [Posse et al., 1999], the characteristics of ME-fMRI drove more researchers to apply this emerging technique on steady-state fMRI processing and analysis. For example, a new method of multi-echo fMRI independent component analysis (ME-ICA) was introduced in [Kundu et al., 2012] to distinguish neuron activity signal fluctuations from nuisance signal by fitting T_2^* decay model. As one of the most important steps in the ME-fMRI pipeline, the combination of multi-echo data was realized by many weighting methods, among which T_2^* weighting method proposed in [Posse et al., 1999] was widely used. However, the comparison of different combination methods provided similar results[Kettinger et al., 2016] and a convincing weighting method was never put forward. What's more, when these methods were introduced, only specific and partial characteristics of signals were taken into account, such as noise being white and TE-dependent[Poser et al., 2006].

In this work, we describe a model for optimizing the weighting method based on Rayleigh quotient (RQ), which is distinct from most previous research because all the fluctuations' information of each echo data is taken into account in each echo. The optimized weighting methods are used for two metrics: temporal-signal-ratio (tSNR) and multi-echo temporal-signal-ratio (metSNR), which are strongly related to the widely-used contrast-noise-ratio (CNR). The quality of the combined signal is improved significantly compared to other alternate methods. In the meantime, we expanded the range of weight values which are only positive before and re-defined the normalization as 'L1 Normalization'.

In the present work, using metSNR as the evaluation metric shows more robustness than tSNR. Additionally, to examine the causes of uncommon negative weights from a mathematical viewpoint, we found the high correlation between two echoes' data is the critical factor. In terms of TE and brain region, the difference in the distribution of negative weights is also discussed based on the causes.

Chapter 2

Methods

2.1 Data Acquisition:

fMRI data used in the experiments were acquired from one healthy male participating in the resting-state experiment with the subject's eyes closed. This experiment was performed using a General Electric (GE) 3.0T whole-body MRI scanner with an 8-channel receive head coil. Functional images were obtained using a customized multi-echo version of 2D Gradient Echo EPI sequence with twofold ARC acceleration in the phase encoding direction and with the following parameters: TR = 1.3 s; flip angle = 52°; field of view (FOV) = 192 mm; matrix size = 64×64; pixel size = 3.0 mm; slice thickness = 3 mm; number of slices = 48. The data from 3 echoes were acquired with the following echo times: TE=12.2 ms, 30.1 ms and 48.0 ms. Partial Fourier was turned off during sampling k-space data. Pulse and respiratory signals were simultaneously acquired via the scanner integrated photoplethysmography and breathing belt, respectively. The duration of the whole scan session was 6 minutes.

2.2 Data Pre-processing:

Most steps of the fMRI data pre-processing pipeline were performed with the AFNI functions embedded in meica.py (version 2.5 (c) 2014 Prantik Kundu) [Cox, 1996, Kundu et al., 2012] with the following basic steps. Motion and obliquity parameters were calculated from the first echo's data and saved for motion correction. Each echo's dataset was aligned to the base dataset (*3dAllineate*) which was calculated from the first volume of the first echo data. Then, for each echo's data, the same processing steps were conducted: 1. Functional images were directly despiked with *3dDespike*; 2. The separate slices were aligned to the same temporal origin by using the 7th order Lagrange polynomial interpolation (*3dTshift*); 3. Data brick was oriented as axial slices (*3daxialize*); 4. Spatially smoothing was performed with a 5 mm FWHM Gaussian kernel(*3dBlurInMask*); 5. Linear least squares was used to remove the second polynomial trend of the time series data (*3dDtrend*).

To be more specific, the pre-processed data of each echo can be directly obtained in the section: 'Apply combined normalization/co-registration/motion correction parameter set to e***k***_ts+ori', in the generated '.sh' file (***k*** is for the ***k***th echo's command line). In this study, 'meica.py -e TE1,TE2,TE3-d Data_TE1.nii.gz,Data_TE2.nii.gz,Data_TE3.nii.gz –smooth 5 –detrend 2 –prefix OutputFile' was called to process data. Among these options, 'detrend 2' was used to realize removal of the second polynomial trend of data and 'smooth 5' was used for the 5mm FWHM Gaussian smooth. After these basic pre-processsing steps, the three new echo's datasets were obtained in 'e1_in.nii.gz', 'e2_in.nii.gz' and 'e3_in.nii.gz'. The whole process can be visualized in the flowchart in Appendix A.3.

2.3 Combination of multi-echo fMRI data:

For a given voxel, the pre-processed triple echo dataset was combined using a weighted summation to produce a single time series:

$$s_{comb} = \sum_{k=1}^{3} \frac{w_k}{\sum_{k=1}^{3} w_k} s_k(t),\ 0 \leq w_k < +\infty \tag{2.1}$$

where $s_k(t)$ is the signal amplitude of the kth echo at time t point and w_k is the set of weights corresponding to this echo, which were defined according to the different methods described below.

In [Kettinger et al., 2016, Kundu et al., 2012], positive weights were used, and they were normalized by their sum. In this work, however, both positive and negative weights were used, and they were normalized by the sum of absolute values so that the normalization always followed the L1-Norm rule. Thus, the weighted combination was generalized to:

$$s_{comb} = \sum_{k=1}^{3} \frac{w_k}{\sum_{k=1}^{3} |w_k|} s(t)_k,\ -\infty < w_k < +\infty \tag{2.2}$$

In addition, several proposed voxel-wise weighting methods in [Kettinger et al., 2016] will be compared with the result of the optimal weights in the Results Section. The detailed formula for each method is as follows:

<u>*Optimal Weighting* (*tOPT/mOPT*)</u>: For each TE time series, the optimal weight is calculated following Eq. 2.21 and Eq. 2.25:

$$w_k = w_{tOPT,\,k} = k\Sigma^{-1}\bar{S} \quad \text{for tSNR} \tag{2.3}$$

$$w_k = w_{mOPT,\,k} = k\Sigma^{-1}\Lambda\bar{S} \quad \text{for metSNR} \tag{2.4}$$

Suboptimal Weighting (tDIAG/mDIAG): For each TE time series, the suboptimal weight is calculated following Eq. 2.23 and Eq. 2.27:

$$w_k = w_{tDIAG,\,k} = kD^{-1}\bar{S} \quad \text{for tSNR} \tag{2.5}$$

$$w_k = w_{mDIAG,\,k} = kD^{-1}\Lambda\bar{S} \quad \text{for metSNR} \tag{2.6}$$

where D is the covariance matrix with only diagonal terms; Λ is the diagonal matrix with TE_1, TE_2, TE_3 as the diagonal terms.

Flat Weighting (Flat): For every TE time series, the weighting values are equal to 1:

$$w_k = 1 \tag{2.7}$$

tSNR Weighting (tSNRW): For a given voxel, the temporal signal-to-noise ratio (*tSNR*) of each echo time series represents the weight:

$$w_k = tSNR_k = \frac{\bar{s}_k}{\sigma_k} \tag{2.8}$$

where $\bar{s}_k$ is the mean signal intensity of the ith echo time-series and σ_k is its temporal standard deviation.

Signal Weighting (SIWGT): The mean signal intensity of each echo time series is used for each of the echo's weights:

$$w_k = \bar{s}_k \tag{2.9}$$

Mean Bold Sensitivity Weighting (mBSENS): The mean value of each echo time series is

multiplied by the corresponding echo time value [Deichmann et al., 2002, Posse et al., 1999]:

$$w_k = mBSENS_k = TE_k \bar{s}_k \tag{2.10}$$

<u>*Temporal Mean Bold Sensitivity weighting (tmBSENS)*</u>: The kth tBSENS is derived from the tSNR Weighting and Bold Sensitivity Weighting [Poser et al., 2006]:

$$w_k = tmBSENS_k = \frac{TE_k \bar{s}_k}{\sigma_k} \tag{2.11}$$

<u>*T_2^* Weighting ($T2W$)*</u>: As described in [Kundu et al., 2012], for each voxel, T_2^* was first found by fitting the mean of each of the three time courses to the MRI signal intensity formula: $\mathrm{S}(\mathrm{TE_n}) = \mathrm{S_0} \exp(-\mathrm{R_2^* TE_n})$, with log linear regression and the weight of T_2^* weighting scheme can be expressed as

$$w_k = TE_k \cdot e^{-TE_k / T^*_{2(fit)}} \tag{2.12}$$

For all the methods mentioned above, w_k represents the weight of the kth echo time series. The abbreviated names (e.g., *SIWGT*, *tSNRW*) represent different *type*s in the Results Section. These weighting methods are summarized in Table 2.1 for reference.

2.4 Combination Optimization:

2.4.1 Classic Rayleigh Quotient(RQ) Model

In mathematics, a generalized Rayleigh Quotient (RQ) is defined as a ratio of two quadratics as follows:

$$R(A,B,w) = \frac{W^T A W}{W^T B W} \tag{2.13}$$

where A and B are $n \times n$ Hermitian matrices, W is a nonzero column vector, and W^T is the conjugate transpose of W. RQ is used in the min-max theorem to get the exact values of all eigenvalues [Lischinski, 2007]. The maximization of RQ can eventually turn into a generalized eigenvalue problem with an established solution as:

$$B^{-1}AW = \lambda_{max}W \tag{2.14}$$

where λ_{max} is the largest eigenvalue to maximize tehe RQ equation. Therefore, W is the eigenvector corresponding to the largest eigenvalue of $B^{-1}A$, when the maximum is obtained.

Table 2.1: **List of weighting methods.**

Name	**Abbreviation**	**Definition of w_k**
Optimal Weighting	tOPT	$k\Sigma^{-1}\bar{S}$
	mOPT	$k\Sigma^{-1}\Lambda\bar{S}$
Suboptimal Weighting	tDIAG	$kD^{-1}\bar{S}$
	mDIAG	$kD^{-1}\Lambda\bar{S}$
Flat	Flat	1
Signal Weighting	SIGWT	$\frac{\bar{s}_k}{\sigma_k}$
tSNR Weighting	tSNRW	$\bar{s}_k$
Mean Bold Sensitivity Weighting	mBSENS	$TE_k\bar{s}_k$
Temporal Mean Bold Sensitivity Weighting	tmBSENS	$\frac{TE_k\bar{s}_k}{\sigma_k}$
T_2^* Weighting	T2W	$TE_k \cdot e^{-TE_k/T^*_{2(fit)}}$

2.4.2 Metric Optimization:

Contrast-to-noise-ratio (CNR) is a widely-used measure to reflect Blood Oxygen Level-Dependent (BOLD) signal change which is of significant interest to researchers. Two metrics for the evaluation of fMRI were used: temporal signal-to-noise ratio (tSNR) and multi-echo temporal

signal-to-noise ratio (metSNR) [Welvaert and Rosseel, 2013]. As can be seen in Appendix A.1 and A.2, both metrics have strong relationships with CNR and they are similar in that a signal measure is compared to its noise level. When considering that the fluctuations or noise in the combined signal are composed of the fluctuations of the original triple-echo signal in our study, the fluctuation of combined data can be expressed as:

$$\sigma_{comb}^2 = \sum_{k=1}^{3} \frac{w_k^2}{\sum_{k=1}^{3} |w_k|^2} Var\left(s_k\right) + 2 \sum_{1 \le k \le j \le 3} \frac{w_k w_j}{\sum_{k=1}^{3} |w_k|^2} Cov\left(s_k, s_j\right) \tag{2.15}$$

where $Var\left(s_k\right)$ is the variance of one time series and $Cov\left(s_k, s_j\right)$ is the covariance of two echoes' time series, s_k and s_j. The mean value of combined data can be expressed as:

$$\bar{s}_{comb} = \sum_{i=1}^{3} \frac{w_k}{\sum_{i=1}^{3} |w_k|} \bar{s}_k \tag{2.16}$$

Hence, in the current research, all the signal information of each echo was taken into account, and the optimal weight vector was found to make the two metrics reach the maximum value by fitting them with the RQ model.

tSNR

As an established quantity for the determination of SNR of fMRI data, tSNR in a given voxel was calculated from the mean signal intensity across the 272 time points divided by its temporal standard deviation, defined as

$$tSNR = \frac{\bar{s}}{\sigma} = \frac{\bar{s}}{\sqrt{\sigma^2}} \tag{2.17}$$

where σ^2 is the variance of each single echo time series, representing the noise level.

Integrating the definition of tSNR (Eq. 2.17), and the relationship between the combined signal and single echo signal (Eq. 2.15 and Eq. 2.16), the tSNR for the combined signal can be

expressed as :

$$tSNR_{comb} = \frac{\bar{s}_{comb}}{\sqrt{\sigma^2_{comb}}} \tag{2.18}$$

It can also be expressed by parameter matrices for a given voxel (A more detailed derivation can be found in Appendix A.1):

$$tSNR_{comb} = \frac{W^T \bar{S}}{\sqrt{W^T \Sigma W}} \tag{2.19}$$

where Σ is the 3×3 covariance matrix whose diagonal terms are variances of each echo's time series, off-diagonal terms are the covariances between two echoes' time series, and $\bar{S}$ is the 3×1 mean vector composed of the three echoes' mean values.

In order to find the optimal weight vector that maximizes tSNR, the optimization of the RQ model was applied on the square of tSNR to find the optimal weights:

$$tSNR^2_{comb} = \frac{W^T M W}{W^T \Sigma W} \tag{2.20}$$

where M is the 3×3 symmetric matrix such that $M = \bar{S}\bar{S}^T$.

In this case, however, the $tSNR^2_{comb}$ optimization can be calculated without solving the generalized eigenvalue problem. As proven in Appendix A.1, for the optimal weight vector of tSNR, there is a more direct and simple expression:

$$W_{tOPT} = k\Sigma^{-1}\bar{S} \tag{2.21}$$

where k is an arbitrary scalar. Then, W_{tOPT} was normalized by the L1-Norm of the weights for combination and the optimally combined data can be obtained from:

$$s_{comb} = \frac{\sum_{k=1}^{3} w_{tOPT,k} \cdot s_k(t)}{\sum_{k=1}^{3} |w_{tOPT,k}|} \tag{2.22}$$

where $w_{tOPT,k}$ is the kth term in W_{tOPT}. In addition, the suboptimal weights for tSNR can be

obtained by using only the diagonal terms of the covariance matrix:

$$W_{tDIAG} = kD^{-1}\bar{S} \tag{2.23}$$

where $D = diag(\Sigma)$.

metSNR

Multi-echo tSNR (metSNR) was derived from and represents contrast-to-noise ratio(CNR) for a multi-echo dataset. CNR has been used as a measure of image quality based on a contrast which is of most interest for fMRI research. The metSNR formula is quite similar to the tSNR expression, with the main difference being the mean value in metSNR is enlarged by the corresponding TE value:

$$metSNR = \frac{W^T \Lambda \bar{S}}{\sqrt{W^T \Sigma W}} \tag{2.24}$$

where Λ denotes $diag([TE_1, TE_2, TE_3])$. The derivation from multi-echo CNR to metSNR is covered in detail in Appendix A.2. As such, similarly to tSNR optimization, the optimal weights for metSNR were derived as:

$$W_{mOPT} = k\Sigma^{-1}\Lambda\bar{S} \tag{2.25}$$

Likewise, with W_{mOPT}, the combined data achieved the best signal quality when using metSNR as the metric.

$$s_{comb} = \frac{\sum_{k=1}^{3} w_{mOPT,k} \cdot s_k(t)}{\sum_{k=1}^{3} |w_{mOPT,k}|} \tag{2.26}$$

where $w_{mOPT,k}$ is the kth term in W_{mOPT}. Similarly, the suboptimal weights for metSNR can also be obtained by ignoring off-diagonal terms of the covariance matrix:

$$W_{mDIAG} = kD^{-1}\Lambda\bar{S} \tag{2.27}$$

where $D = diag(\Sigma)$.

Therefore, it is recommended to choose between the two metrics depending on the choice of evaluative metric, tSNR or metSNR which is an approximation of CNR.

Note: All detailed derivations for the expressions of the two metrics (Eq. 2.19 and Eq. 2.24) for evaluating the combined data and the derivation of optimal weights (Eq. 2.21 and Eq. 2.25) are found in Appendix A.1 and Appendix A.2, respectively. In the practical usage, k was assumed as 1 for ease of analysis of the calculation in the next sections, because k didn't affect the result of tSNR and metSNR.

Chapter 3

Results

3.1 Combined Signal Quality:

Fig. 3.2 and Fig. 3.3 are the scatter plots used to compare the performance of non-optimal and optimal methods, for tSNR and metSNR respectively. In each of the subplots. the tSNR or metSNR obtained from the non-optimal method and the optimal method were plotted against one another for each voxel. Because all points fall under the identity line in each subplot, all of the alternative combination methods were inferior to the optimal methods. The improved results of optimal methods can also be validated from the brain maps of performance ratio (PR) in Fig. 3.1 for two metrics from selected slices (S30∼ S34) as example. For each voxel, the two metric variation are defined as the ratios between the results of non-optimal and optimal weighting methods, Performance Ratio (PR):

$$PR_{tSNR} = \frac{tSNR_{type}}{tSNR_{tOPT}} = \frac{\frac{W_{type}^T \bar{S}}{\sqrt{W_{type}^T \Sigma W_{type}}}}{\frac{W_{tOPT}^T \bar{S}}{\sqrt{W_{tOPT}^T \Sigma W_{tOPT}}}} \quad \text{for tSNR} \tag{3.1a}$$

$$PR_{metSNR} = \frac{metSNR_{type}}{metSNR_{mOPT}} = \frac{\frac{W_{type}^T \Lambda \bar{S}}{\sqrt{W_{type}^T \Sigma W_{type}}}}{\frac{W_{mOPT}^T \Lambda \bar{S}}{\sqrt{W_{mOPT}^T \Sigma W_{mOPT}}}} \quad \text{for metSNR} \tag{3.1b}$$

where *type* is replaced by each of the other seven types of non-optimal weighting methods mentioned. The smaller PR value corresponds to larger variation from the metric results of non-optimal to optimal method.**Note:** When calculating Perforates Ratio for the two suboptimal methods, tDIAG and mDIAG are used for tSNR and metSNR, respectively. In the PR maps, for every non-optimal methods in the columns, all voxels are with PR values which are less than 1 indicated by the label in the upper-right corner. Thus, when using tSNR and metSNR as the evaluation metrics for the combined signal, combining multi-echo data with the optimal weights, W_{tOPT} and W_{mOPT}, can improve the signal quality of combined data significantly ($p - value < 0.0001$), when compared with other weighting methods.

However, the variance of the statistical results among all the non-optimal weighting methods indicated two more interesting points. First, the suboptimal methods, tDIAG for tSNR and mDIAG for metSNR, were superior to other weighting methods, which was indicated by the smallest effect size (0.88 for tDIAG in tSNR and 0.63 for mDIAG in metSNR) in scatter plots and the smaller mean and median value of percent decrease in Table 3.1. Secondly, mBSENS and tmBSENS performed better on metSNR than on tSNR, which was opposite of the performance of tSNRW and SIGWT, which performed better on tSNR. Furthermore, as observed from Fig. 3.1, compared with other regions of the brain, the improvement of tSNR was most obvious in the gray matter (GM) region with optimal methods tOPT, while the performance of mOPT is spatially homogeneous over the whole region of interest (ROI).

3.2 Robustness of metSNR:

Although both tSNR and metSNR showed an improvement with optimal methods, the variation of the two metrics differed. First, In Fig. 3.2 and Fig. 3.3, it was notable that when tSNR was used as the measurement (Fig. 3.2), the points were further distributed and dispersed under the identity line in each subplot, while when metSNR was used as the metric (Fig. 3.3),

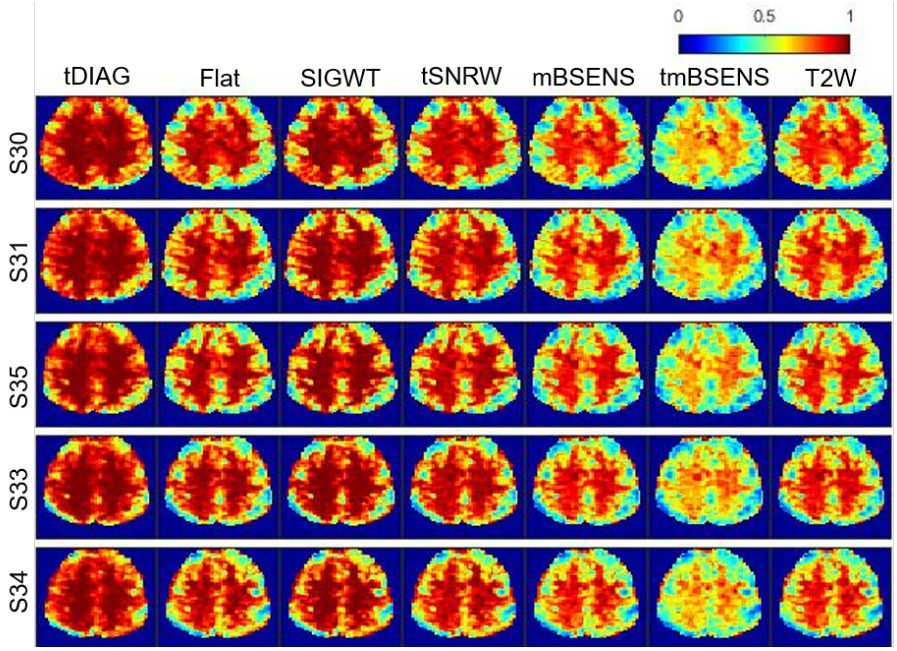
the points were closer to the identity line. Based on the statistical results labeled in each scatter plot, we can deduce for each specific non-optimal method, the t-value and effect size [Cohen, 1988] for each comparison (Fig. 3.2) are relatively larger in tSNR than in metSNR, which means the result of each non-optimal method varied more significantly from the result of the optimal solution when using tSNR as compared with metSNR. For example, the effect size between the tSNR results of 'mBSENS' and 'tOPT' is 1.45, while the effect size is 0.70 for the comparison of the metSNR result. The effect size rules can be referenced in [Sawilowsky, 2009].

In addition, the visualized results in the brain map (Fig. 3.1) indicate that the PR variations of every non-optimal method were smaller overall in metSNR than in tSNR because the metSNR PR map is closer to red overall. To be more specific, when comparing the GM region in two metric PR maps, there are higher percentage decrease in tSNR in such region which is indicated by the yellow voxels with larger negative values. Table 3.1 is also used for the comparison between tSNR PR and metSNR PR for each non-optimal method. In this table, as the mean and median become more negative, the difference between the result of the non-optimal method and that of the optimal method is greater, revealing that each non-optimal method presents less variation in metSNR than in tSNR in general. Hence, based on the visualized results and the statistical results, we can know that the metSNR performs more robustly than tSNR for the non-optimal methods.

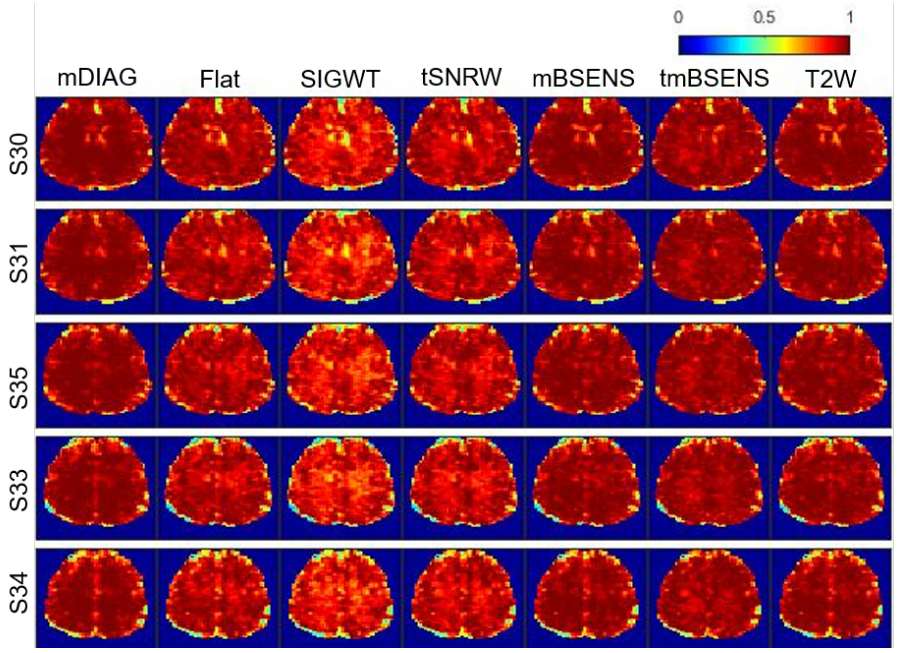
Table 3.1: **tSNR and metSNR PR for non-optimal methods**

	tSNR PR			metSNR PR		
Method	Mean	Median	Std	Mean	Median	Std
tDIAG	0.866	0.922	0.147	n/a	n/a	n/a
mDIAG	n/a	n/a	n/a	0.941	0.977	0.093
Flat	0.764	0.816	0.185	0.901	0.940	0.112
SIGWT	0.828	0.894	0.178	0.841	0.868	0.118
tSNRW	0.780	0.836	0.176	0.881	0.912	0.111
mBSENS	0.698	0.738	0.180	0.928	0.970	0.107
tmBSENS	0.608	0.636	0.157	0.932	0.960	0.089
T2W	0.697	0.737	0.180	0.928	0.970	0.107

Note: The metric'PRs were calculated according to Eq. 3.1. The mean values, standard deviations and median values of metrics' PRs were calculated on all voxels.



(a) tSNR performance ratio



(b) metSNR performance ratio

Figure 3.1: **Brain map of performance ratio (PR) for the two metrics.** For each method, the voxel-wise metric value is compared to the result of the optimal method from five selected slices (S30 $\sim$ S34). Each column represents the comparison result (PR) for each method. The closer the voxel was to red, the less the variation was.

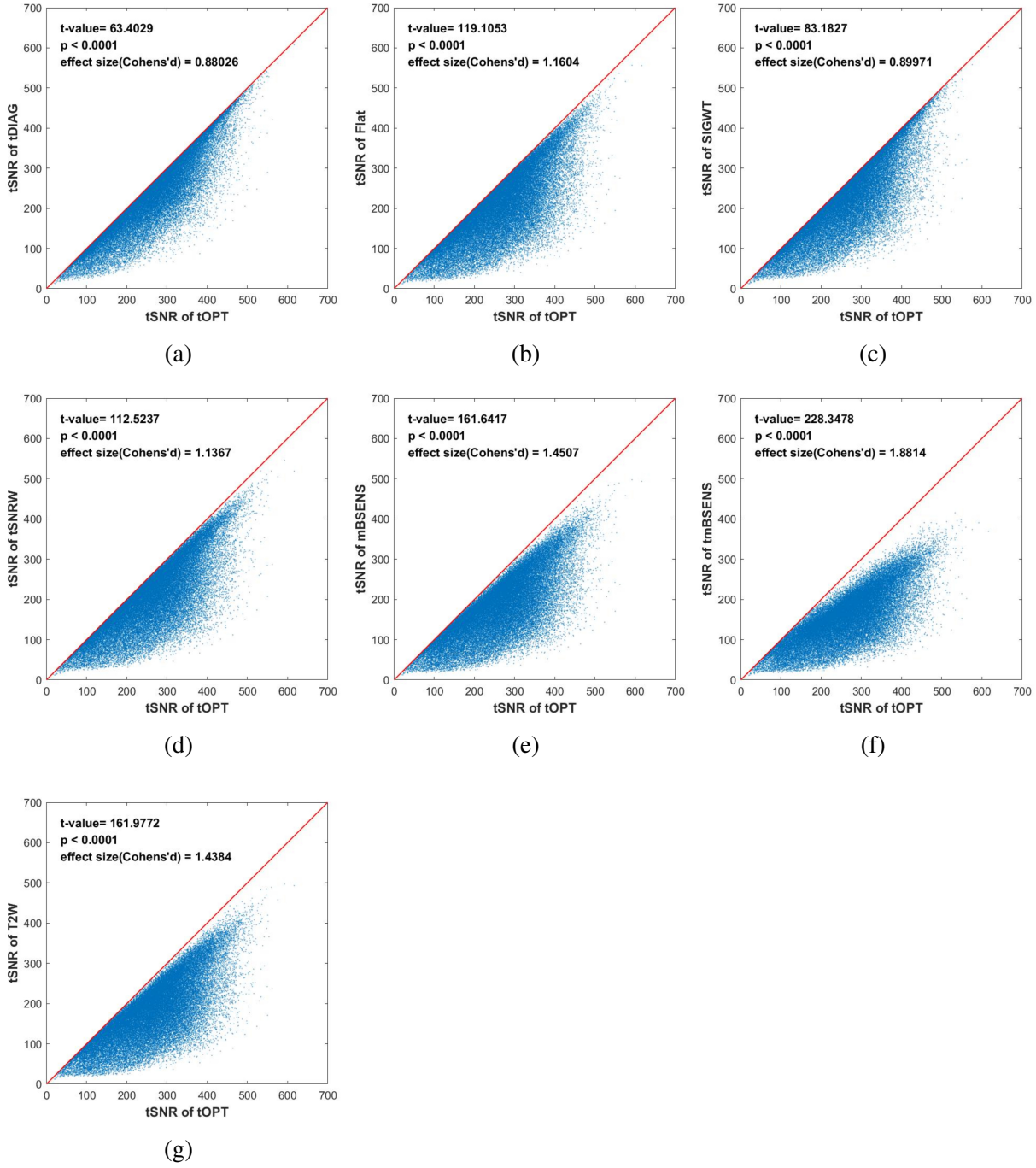


Figure 3.2: **Scatter Plots of the tSNR comparison.** Voxel-wise tSNR values resulting from each weighting method on the vertical axis are plotted against the results of optimal weights (tOPT) on the horizontal axis. The red line is the identity line ($y = x$). The vertical axis displays the tSNR resulting from (a) tDIAG, (b) Flat, (c) SIGWT, (d) tSNRW, (e) mBSENS, (f) tmBSENS and (g) T2W. The statistical analysis (t-value, p-value, effect size) is shown in the upper-left corner of each subplot.

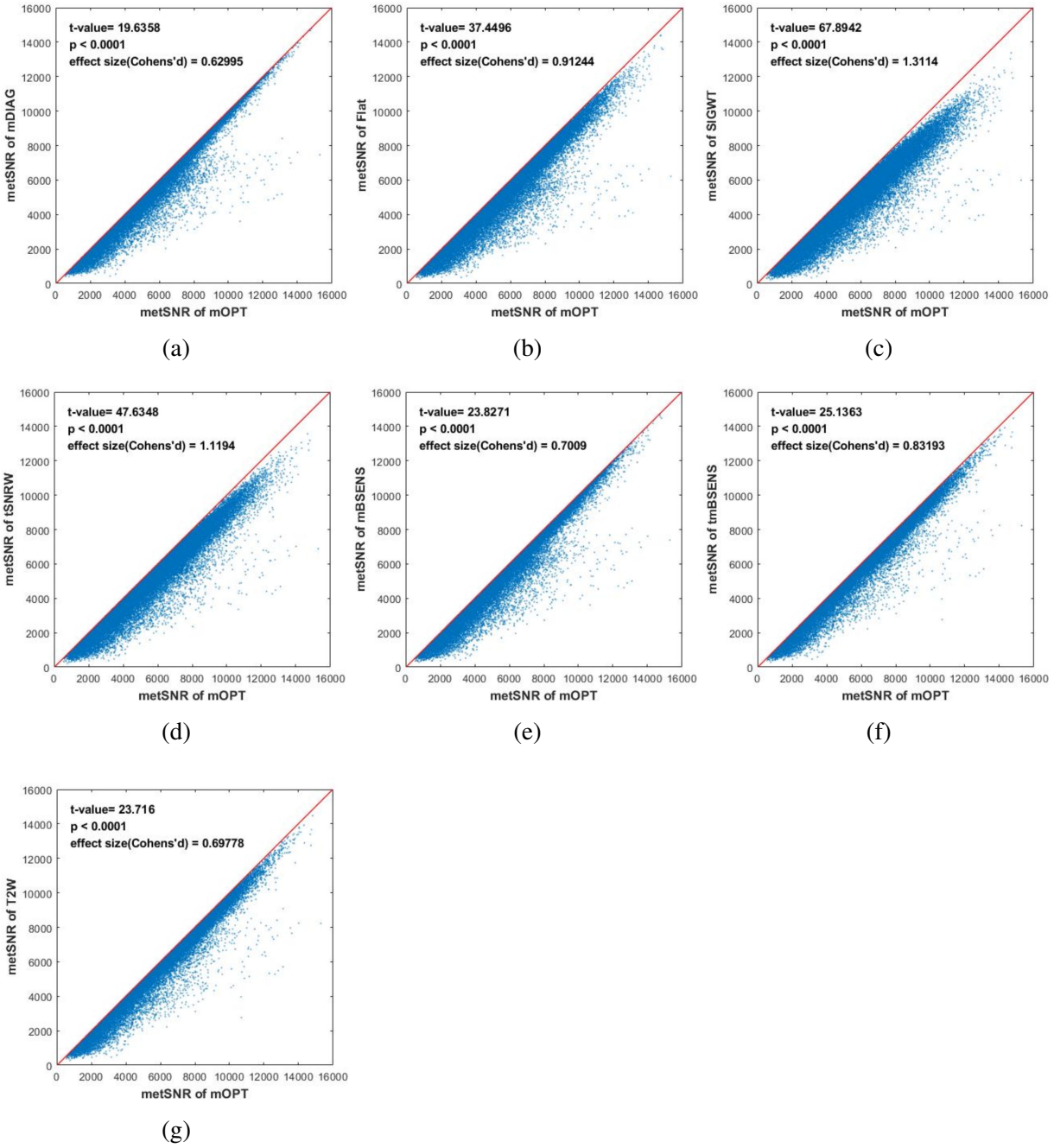


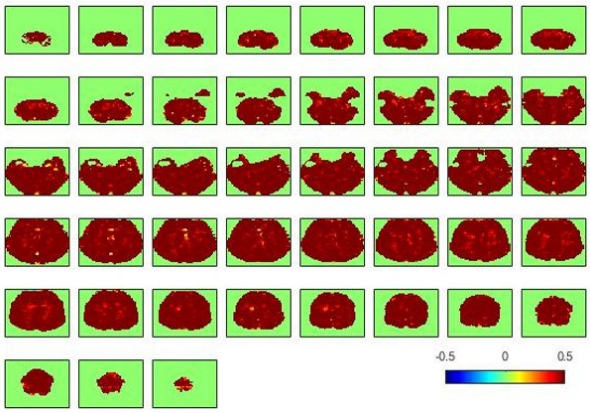
Figure 3.3: **Scatter plots of the metSNR comparison.** Differing from Fig 3.2, the vertical axis displays the metSNR resulting from (a) mDIAG, (b) Flat, (c) SIGWT,(d) tSNRW, (e) mBSENS, (f) tmBSENS and (f) T2W. The metSNR values obtained from mOPT are displayed on the horizontal axis in each subplot. The statistical analysis are shown in the upper-left corner of each subplot.

3.3 Negative Weights:

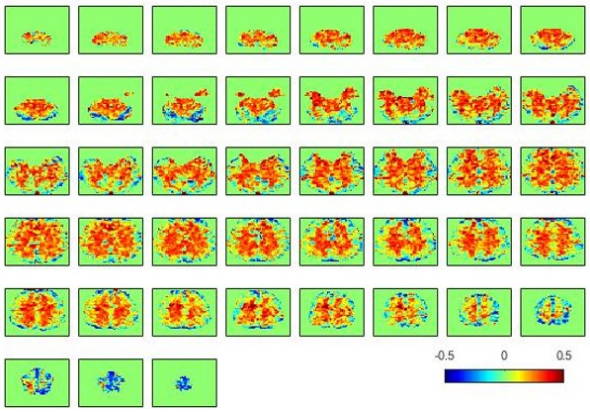
The maps of weights (Fig. 3.4 and Fig. 3.5) show that when calculating the optimal weights, negative weight values were obtained on many voxels for one or more weight terms. Negative weights have not been found in the various weighting methods that have been proposed before. This is also why the absolute values of weights were used with the L1 Normalization in our study, instead of the original values. There are 31363 (58.3%) voxels with negative W_tOPT and 21071 (39.2%) voxels with negative W_mOPT in total. It should be noted that in the case of W_{tOPT} for tSNR, a larger number of negative weights were present in the third weight term (27657).

Fig. 3.5 shows that the situation of W_{mOPT} is markedly different from the case of W_{tOPT}. The distribution of negative weights in terms of TE is different from the result of W_{tOPT} visually. There are more negative weights in the map of first weight term ($W_{mOPT,1}$) but fewer negative weights in the third map ($W_{mOPT,3}$), because there are blue voxels in Fig. 3.5a. than in Fig. 3.5c. As recorded in the captions of the weights maps, there are 44 and 27675 voxels with negative values in first and third term of W_{tOPT}, while there are 14850 and 417 voxels with negative values in first and third term of W_{mOPT}, which is a significant change of negative weights distribution.

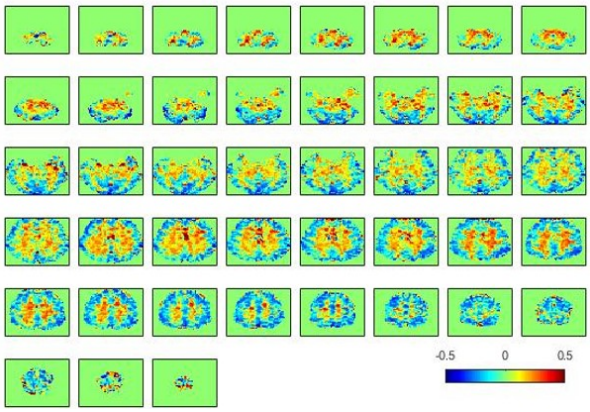
However, the similarity between these two kinds of optimal weights is that the regional distribution of negative weights is near the GM region as the W_{tOPT}, which is observed by the more blue voxels (negative values) near GM region in Fig 3.4 and Fig 3.5. The causes for the negative weights and the different distribution will be described in detail in Chapter. 4.



(a) Weights for 1st TE: $W_{tOPT,\,1}$

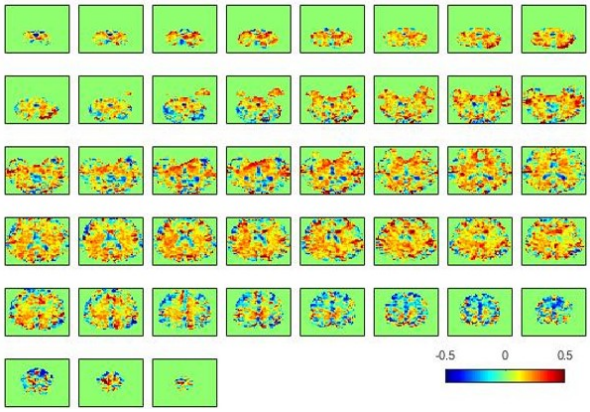


(b) Weights for 2nd TE: $W_{tOPT,\,2}$

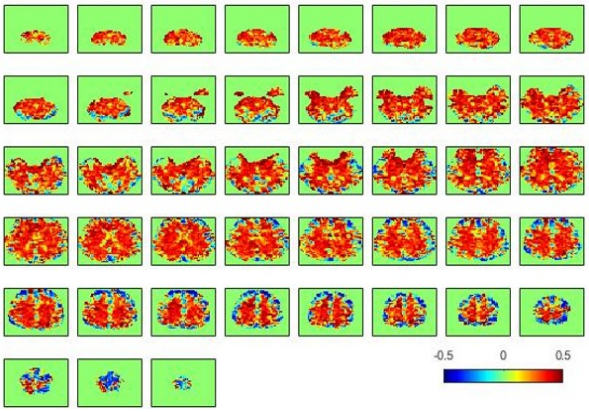


(c) Weights for 3rd TE: $W_{tOPT,\,3}$

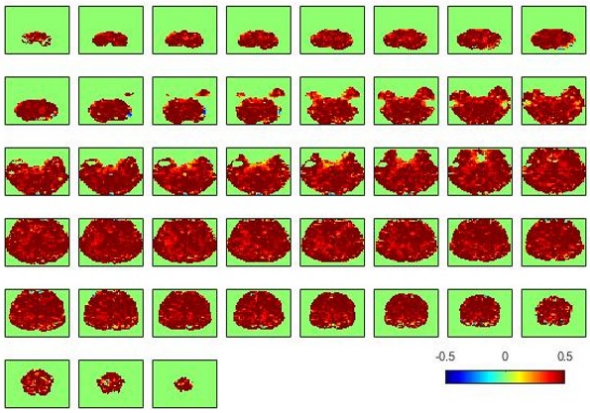
Figure 3.4: **Brain map of weights for W_{tOPT}.** Three optimal weight terms for tSNR are mapped. The colors represent the signs of each weight: blue is for negative and red is for positive. The shade of color represents the magnitude of the weight. (a), (b) and (c) show the maps of the first, second and third weight terms respectively. There are 44, 12213 and 27675 voxels with negative values in the first, second and third term of W_{tOPT}.



(a) Weights for 1st TE: $W_{mOPT,\,1}$



(b) Weights for 2nd TE: $W_{mOPT,\,2}$



(c) Weights for 3rd TE: $W_{mOPT,\,3}$

Figure 3.5: **Brain map of weights for W_{mOPT}.** Maps generated from the three optimal weights for metSNR. From (a) to (c) are the maps of weights for first to third TE, respectively. There are 14850, 9364 and 417 voxels with negative values in first, second and third term of W_{mOPT}.

Chapter 4

Discussion

The aim of this study was to find the optimal weighting method for the combination of multi-echo fMRI data, when using tSNR and metSNR as the metrics to evaluate the combined signal in steady-state fMRI. Based on the results, two models based on RQ for calculating optimal weights were proposed and verified by comparing with other alternative methods mentioned in [Posse et al., 1999] and [Kettinger et al., 2016]. However, there are two more unanticipated but also interesting findings. First, the variation of tSNR and the variation of metSNR have obvious differences in terms of intra-method comparison. Namely, the metSNR performed more robustly than tSNR for most non-optimal weighting methods as indicated by PR maps. The second notable finding is that negative weights were obtained from the proposed optimal weighting models for both tSNR and metSNR. The occurrence of negative weights in the combination process is not consistent with the non-negative weights found in previous research of multi-echo fMRI.

4.1 The difference in metric variation

Based on Eq. 3.1, to compare these two metrics' variation, we calculated the ratio between two PRs for a specific non-optimal weighting method, which was defined as Variation Ratio

(VarRatio):

$$
\begin{aligned}
VarRatio &= \frac{PR_{tSNR}}{PR_{metSNR}} \\
&= \frac{\frac{W_{type}^T \bar{S}}{W_{tOPT}^T \bar{S}} \sqrt{W_{tOPT}^T \Sigma W_{tOPT}}}{\frac{W_{type}^T \Lambda \bar{S}}{W_{mOPT}^T \Lambda \bar{S}} \sqrt{W_{mOPT}^T \Sigma W_{mOPT}}}
\end{aligned}
\tag{4.1}
$$

If we set $W_{tOPT} = k\bar{S}$ and $W_{mOPT} = k\Sigma^{-1}\Lambda\bar{S}$ in the above equation:

$$
\begin{aligned}
VarRatio &= \frac{\frac{W_{type}^T \bar{S}}{\left(\Sigma^{-1}\bar{S}\right)^T \bar{S}} \sqrt{\left(\Sigma^{-1}\bar{S}\right)^T \Sigma\Sigma^{-1}\bar{S}}}{\frac{W_{type}^T \bar{S}}{\left(\Sigma^{-1}\Lambda\bar{S}\right)^T \Lambda\bar{S}} \sqrt{\left(\Sigma^{-1}\Lambda\bar{S}\right)^T \Sigma\Sigma^{-1}\Lambda\bar{S}}} \\
&= \frac{\frac{W_{type}^T \bar{S}}{\sqrt{\bar{S}\Sigma^{-1}\bar{S}}}}{\frac{W_{type}^T \left(\Lambda\bar{S}\right)}{\sqrt{\left(\Lambda\bar{S}\right)\Sigma^{-1}\left(\Lambda\bar{S}\right)}}}
\end{aligned}
\tag{4.2}
$$

If the value of VarRatio is smaller than 1, then the value of PR_{tSNR} is smaller than the value of PR_{metSNR}, which means the variation between the tSNR results of one non-optimal method and $tOPT$ is greater than the variation between the metSNR results of that non-optimal method and $mOPT$. In our case, Table 4.1 shows the number of voxels which have larger PR in metSNR than the PR in tSNR (smaller variation in metSNR than in tSNR). Most of the non-optimal methods perform more robustly in metSNR than in tSNR because most voxels have smaller values of the ratio in Eq. 3.1b. In the map of VarRatio for a selected slice (Fig. 4.3), the values of VarRatio are less than 1 especially for Flat, mBSNES, tmBSENS and T2W, which is consistent with the results in Table. 4.1 and Fig. 3.1.

4.1.1 Theoretical Explanation

To examine the reasons for the different variation in tSNR and metSNR, in the simplified expression of the ratio in Eq. 4.3, we regard the two terms, $\frac{W_{type}^T \bar{S}}{\sqrt{\bar{S}\Sigma^{-1}\bar{S}}}$ and $\frac{W_{type}^T \left(\Lambda\bar{S}\right)}{\sqrt{\left(\Lambda\bar{S}\right)\Sigma^{-1}\left(\Lambda\bar{S}\right)}}$, as the outputs of the same function, $f(X) = \frac{W_{type}^T X}{\sqrt{X\Sigma^{-1}X}}$, when $X = \bar{S}$ and $X = \Lambda\bar{S}$ respectively. Eq. 4.3 can

be simplified as:

$$
\begin{aligned}
VarRatio &= \frac{\frac{W_{type}^T \bar{S}}{\sqrt{\bar{S}\Sigma^{-1}\bar{S}}}}{\frac{W_{type}^T(\Lambda\bar{S})}{\sqrt{(\Lambda\bar{S})\Sigma^{-1}(\Lambda\bar{S})}}} \\
&= \frac{f(\bar{S})}{f(\Lambda\bar{S})}
\end{aligned}
\tag{4.3}
$$

The function $f(X)$ is the same as the format of the square root of RQ and can be maximized to $f_{max}(X)$ by the RQ model when $X = X_{type}^* = \Sigma W_{type}$ which is the eigenvector corresponding to largest eigenvalue of $\Sigma W_{type} W_{type}^T$. The corresponding Rayleigh quotient $g(X) = \frac{X^T(W_{type}W_{type})^T X}{X\Sigma^{-1}X}$ is a continuous function on the unit sphere [Trefethen and Bau, 1997] as shown in Fig. 4.1, where the coordinate of each point is the normalized input vector of $g(X)$ and the value is the corresponding function value calculated from $g(X)$. And the normalized optimal input vector $X_{Norm,type}^*$ of X_{type}^* is the coordinate for the largest function value $g_{max}(X)$. Therefore, we can determine which function value, $f(\bar{S})$ or $f(\Lambda\bar{S})$, is closer to $f_{max}(X)$ using the following steps:

- Choose one non-optimal method ($type$) to compare the tSNR and metSNR variation which are indicated by PR.
- Normalize each of $\bar{S}$ and $\Lambda\bar{S}$ to $\bar{S}_{Norm}$ and $(\Lambda\bar{S})_{Norm}$ by L2 Norm so that they can used on the unit sphere.
- Use the angle between normalized vector and normalized eigenvector $X_{Norm,type}^*$ to compare which function value, $f(\bar{S})$ or $f(\Lambda\bar{S})$, is closer to $f(X_{type}^*)$ which is the largest function value on the unit sphere:
  - As shown in Fig. 4.2, which gives a cleaner view, we assume that θ_1 is the angle between $\bar{S}_{Norm}$ (green vector) and $X_{Norm,type}^*$ (yellow vector); θ_2 is the angle between $(\Lambda\bar{S})_{Norm}$ (orange vector) and $X_{Norm,type}^*$.
  - If $\theta_1 > \theta_2$, then the point at $(\Lambda\bar{S})_{Norm}$ is closer to the point at $X_{Norm,type}^*$ than the point at $\bar{S}_{Norm}$ on the unit sphere.

- Thus, $g((\Lambda\bar{S}) > g(\bar{S}$ which deduces $f(\Lambda\bar{S}) > f(\bar{S})$.
- In such case, $VarRatio = \frac{f(\bar{S})}{f(\Lambda\bar{S})} < 1$ when using the selected non-optimal method.

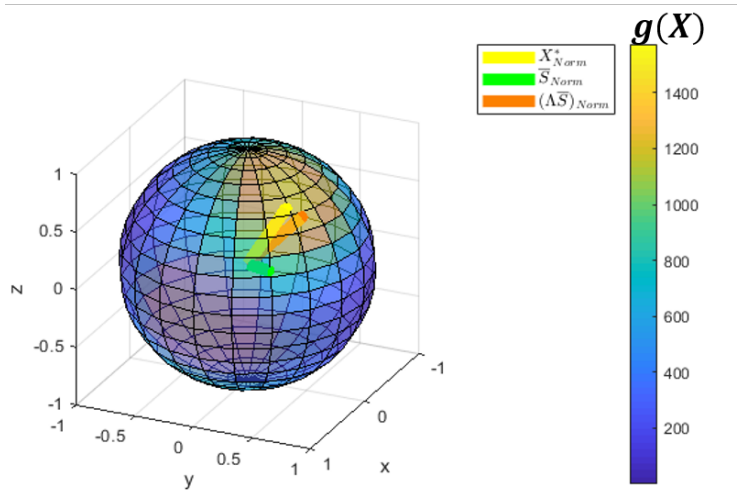


Figure 4.1: Unit sphere of Rayleigh quotient

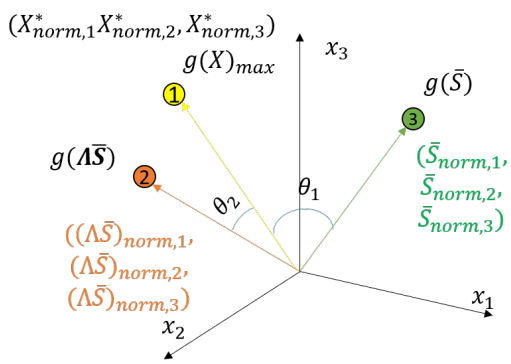


Figure 4.2: **Three input vector for** $f(x)$**.** θ_1 is the angle between $\bar{S}_{Norm}$ (green vector) and $X^*_{Norm,type}$ (yellow vector); θ_2 is the angle between $(\Lambda\bar{S})_{Norm}$ (orange vector) and $X^*_{Norm,type}$. The coordinate of each point is the corresponding normalized input vector by its L2 Norm.

Table 4.1: **The comparison between metSNR and tSNR**

Method	Number
Flat	41343 (76.9%)
SIGWT	24173 (45.0%)
tSNRW	35216 (65.5%)
mBSENS	49765 (92.6%)
tmBSENS	51910 (96.5%)
T2W	49799 (92.6%)

Note: There are 53772 voxels in total taken into account in the whole brain. The number of voxels performing less variation in metSNR than tSNR for each method is provided. The percent value is the proportion of these voxels in the total number.

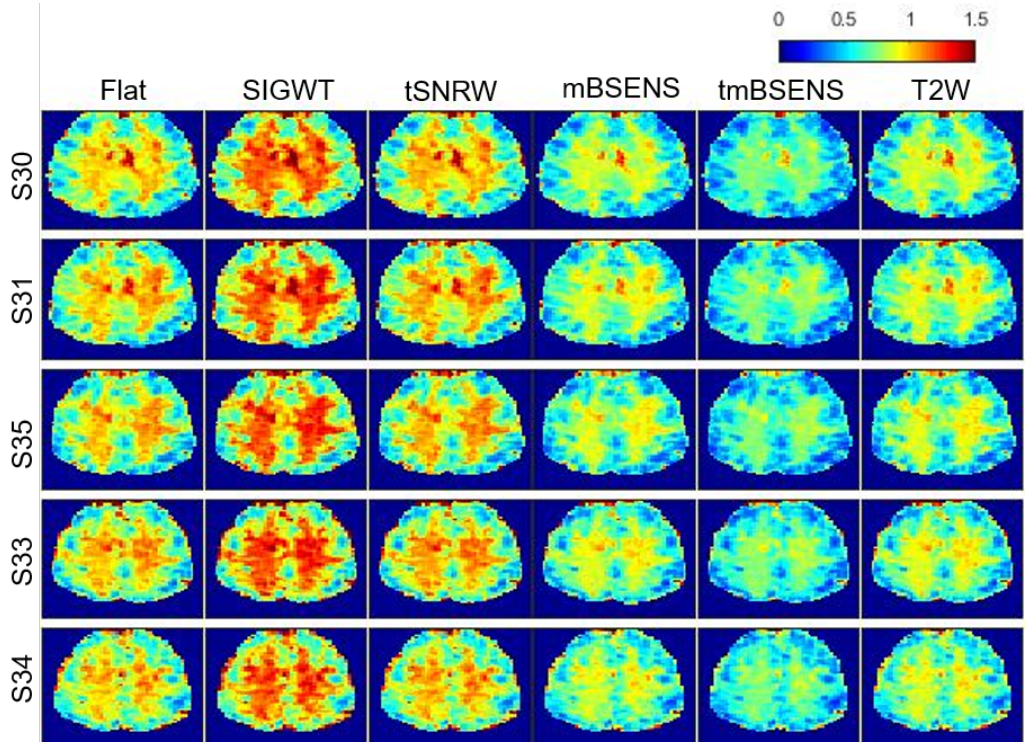


Figure 4.3: **Brain map of ratio between PRs of the two metrics.** For each method, the voxel-wise ratio between the PRs of two metrics was calculated from five selected slices (S30 $\sim$ S34) according to Eq. 4.1. Each column contains the maps of voxel-wise variation ratio for each non-optimal method.

4.1.2 Typical Voxels

To verify the argument above, five voxels with larger variation in metSNR (smaller PR_{metSNR}) and five voxels with larger variation in tSNR (smaller PR_{tSNR}) were selected. In Fig. 4.4a, θ_1 is smaller than θ_2 for these five selected voxels, which implies that the value of $f(\bar{S})$ is closer to $f(x)_{max}$ than $f(\Lambda\bar{S})$ so that $f(\bar{S})$ is larger than $f(\Lambda\bar{S})$ as indicated in the fourth and fifth column. So, the variation ratio is larger than 1. However, for the voxels in Fig. 4.4b with smaller PR_{tSNR} (larger tSNR variation), $f(\bar{S})$ is smaller which is consistent with the larger θ_1 so that the variation ratio is less than 1. Thus, based on these typical voxels' result, the reasoning of difference in variation ratio (PR) indicated by Eq. 4.3 can be explained from the mathematical viewpoint mentioned above.

(a) Voxels perform larger variation in metSNR when using SIGWT

Voxel	PR_{tSNR}	PR_{metSNR}	$\frac{W_{SIGWT}^T\bar{S}}{\sqrt{\bar{S}^T\Sigma^{-1}\bar{S}}}$	$\frac{W_{SIGWT}^T\Lambda\bar{S}}{\sqrt{(\Lambda\bar{S})^T\Sigma^{-1}\Lambda\bar{S}}}$	$VarRatio$	θ_1	θ_2
#85372	0.9882	0.7678	42.18	32.97	1.2792	7.24°	32.6°
#84581	0.9752	0.8560	27.00	23.69	1.1394	8.88°	21.3°
#85222	0.9032	0.6716	64.80	48.18	1.3450	10.2°	20.8°
#84666	0.9925	0.7999	24.26	19.52	1.2408	6.62°	37.1°
#85786	0.9984	0.8801	28.57	25.19	1.1344	3.21°	27.3°

S31

(b) Voxels perform larger variation in tSNR when using tmBSENS

Voxel	PR_{tSNR}	PR_{metSNR}	$\frac{W_{tmBSENS}^T\bar{S}}{\sqrt{\bar{S}^T\Sigma^{-1}\bar{S}}}$	$\frac{W_{tmBSENS}^T\Lambda\bar{S}}{\sqrt{(\Lambda\bar{S})^T\Sigma^{-1}\Lambda\bar{S}}}$	$VarRatio$	θ_1	θ_2
#85372	0.8103	0.9833	32.25	39.13	0.8241	30.6°	8.01°
#84588	0.5887	0.9916	20.75	34.96	0.5937	31.1°	3.18°
#84581	0.7078	0.9618	21.09	28.66	0.7359	30.2°	10.7°
#85060	0.4577	0.9662	39.54	83.49	0.4737	28.3°	3.58°
#86224	0.4012	0.9519	25.45	60.37	0.4215	36.5°	6.97°

S31

Figure 4.4: **Typical voxels for comparison.** The numbers on the right are the slices that voxels are in. The angles, θ_1 and θ_2 are the same as the angles proposed in Fig. 4.2. To get the metric PR, tmBSENS and SIGWT are used as the non-optimal methods respectively in (a) and (b).

4.2 Causes of Negative Weights

Negative weights in the optimal weighting methods have appeared in results in the normalized weighting method for the combination of multi-echo data. By decomposing and analyzing the calculation process of the optimal weighting method in detail, the appearance of negative weights can be reasonably explained, although the limited explanation of a more accurate relationship between cofactor and its multiplier remained. Thus, the range of weights can be extended to the range of real numbers and the new L1-Norm can be applied on the weighted summation.

4.2.1 Theoretical Explanation

The case of optimizing weights for tSNR was examined first. For each voxel, the optimal weights used in this study were calculated from

$$\mathrm{W}_{\mathrm{tOPT}} = k\Sigma^{-1}\bar{S} \tag{4.4}$$

Because the scalar k was canceled out in the numerator and the denominator in the calculation of tSNR or metSNR (Eq. 2.19 and Eq. 2.24), we assumed $k = 1$ for ease of analysis and use in this section. So, the optimal weight expression were re-written as

$$\mathrm{W}_{\mathrm{tOPT}} = \Sigma^{-1}\bar{S} = \frac{\mathrm{Adj}(\Sigma)}{\det(\Sigma)}\bar{S} \tag{4.5}$$

where $\mathrm{Adj}(\Sigma)$ is the cofactor matrix and $\det(\Sigma)$ is the determinant. The covariance matrix Σ was denoted as

$$\Sigma = \begin{vmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{vmatrix} \tag{4.6}$$

where σ_k^2 is the variance of the kth echo data and σ_{ij} is the covariance of the ith and jth data.

Since the determinants, $\det(\Sigma)$, are positive for all voxels, the optimal weight vector is positively related to the product of the cofactor matrix and the mean vector. Accordingly, each of the optimal weights is positively related to the product of the cofactor and the mean value

$$w_1 = K\left(\bar{s}_1 cof_{11} + \bar{s}_2 cof_{12} + \bar{s}_3 cof_{13}\right) \tag{4.7a}$$

$$w_2 = K\left(\bar{s}_1 cof_{12} + \bar{s}_2 cof_{22} + \bar{s}_3 cof_{23}\right) \tag{4.7b}$$

$$w_3 = K\left(\bar{s}_1 cof_{13} + \bar{s}_2 cof_{23} + \bar{s}_3 cof_{33}\right) \tag{4.7c}$$

where K is the reciprocal of the positive determinant for each relation and each of the cofactor, cof_{ij}, was defined in Eq 4.9.

The cofactor matrix of Σ was denoted as

$$\begin{aligned}
\mathrm{Adj}(\Sigma) &= \begin{vmatrix} \sigma_2^2\sigma_3^2 - \sigma_{23}^2 & \sigma_{12}\sigma_3^2 - \sigma_{13}\sigma_{23} & \sigma_{12}\sigma_{23} - \sigma_2^2\sigma_{13} \\ \sigma_{12}\sigma_3^2 - \sigma_{13}\sigma_{23} & \sigma_1^2\sigma_3^2 - \sigma_{13}^2 & \sigma_1^2\sigma_{23} - \sigma_{13}\sigma_{12} \\ \sigma_{12}\sigma_{23} - \sigma_2^2\sigma_{13} & \sigma_1^2\sigma_{23} - \sigma_{13}\sigma_{12} & \sigma_1^2\sigma_2^2 - \sigma_{12}^2 \end{vmatrix} \times \begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix} \\
&= \begin{vmatrix} cof_{11} & cof_{12} & cof_{13} \\ cof_{12} & cof_{22} & cof_{23} \\ cof_{13} & cof_{23} & cof_{33} \end{vmatrix}
\end{aligned} \tag{4.8}$$

Each entry in the cofactor matrix was expressed individually as

$$cof_{11} = \sigma_2^2\sigma_3^2 - \sigma_{23}^2 = \sigma_2^2\sigma_3^2\left(1 - r_{23}^2\right) = C_1\left(1 - r_{23}^2\right) \tag{4.9a}$$

$$cof_{22} = \sigma_1^2\sigma_3^2 - \sigma_{13}^2 = \sigma_1^2\sigma_3^2\left(1 - r_{13}^2\right) = C_2\left(1 - r_{13}^2\right) \tag{4.9b}$$

$$cof_{33} = \sigma_1^2\sigma_2^2 - \sigma_{12}^2 = \sigma_1^2\sigma_2^2\left(1 - r_{12}^2\right) = C_3\left(1 - r_{12}^2\right) \tag{4.9c}$$

$$cof_{12} = -\left(\sigma_3^2\sigma_{12} - \sigma_{13}\sigma_{23}\right) = -\sigma_1\sigma_2\sigma_3^2\sqrt{\left(1 - r_{23}^2\right)\left(1 - r_{13}^2\right)} \cdot r_{12,3} = -C_4 \cdot r_{12,3} \tag{4.9d}$$

$$cof_{13} = -\left(\sigma_2^2\sigma_{13} - \sigma_{12}\sigma_{23}\right) = -\sigma_1\sigma_3\sigma_2^2\sqrt{\left(1 - r_{12}^2\right)\left(1 - r_{23}^2\right)} \cdot r_{13,2} = -C_5 \cdot r_{13,2} \tag{4.9e}$$

$$cof_{23} = -\left(\sigma_1^2\sigma_{23} - \sigma_{13}\sigma_{12}\right) = -\sigma_2\sigma_3\sigma_1^2\sqrt{\left(1 - r_{12}^2\right)\left(1 - r_{13}^2\right)} \cdot r_{23,1} = -C_6 \cdot r_{23,1} \tag{4.9f}$$

where C_i represents a different positive constant for each expression. Among them, cof_{11}, cof_{22} and cof_{33} are defined as **diagonal cofactor terms**; cof_{11}, cof_{22} and cof_{33} are defined as **off-diagonal cofactor terms**; r_{ij} is the correlation coefficient between two echo time series; $r_{ij,k}$ is the partial correlation between two echo time series, while controlling the effect of the kth echo, e.g. $r_{12,3}$, which was calculated with the Matlab function 'partialcorr'. Integrating the relationships in Eq. 4.7 and Eq. 4.9, the optimal weights for tSNR was re-written as

$$w_1 = K\left(C_1\left(1 - r_{23}^2\right)\bar{s}_1 - C_4 \cdot \bar{s}_2 \cdot r_{12,3} - C_5 \cdot \bar{s}_3 \cdot r_{13,2}\right) \tag{4.10a}$$

$$w_2 = K\left(-C_4 \cdot \bar{s}_1 \cdot r_{12,3} + C_2\left(1 - r_{13}^2\right)\bar{s}_2 - C_6 \cdot \bar{s}_3 \cdot r_{23,1}\right) \tag{4.10b}$$

$$w_3 = K\left(-C_5 \cdot \bar{s}_1 \cdot r_{13,2} - C_6 \cdot \bar{s}_2 \cdot r_{23,1} + C_3\left(1 - r_{12}^2\right)\bar{s}_3\right) \tag{4.10c}$$

Given the expression of each cofactor term above, the diagonal terms are always positive as r_{ij} is always less than 1 and only the off-diagonal terms can bring about negative values. Actually, most of the voxels are in line with such a situation. There are 94.85% voxels with negative cof_{12}, 81.23% voxels with negative cof_{13} and 97.56% voxels with negative cof_{23} in our case. Thus, there are two necessary conditions for a negative weight:

1. Given that there exists a negative weight because the mean values and the diagonal cofactor

term are all positive in its expression, then at least one of two off-diagonal cofactors is negative.

2. As the multiplier of the cofactor term, the magnitude of the corresponding positive mean value shouldn't be too 'large' to make the positive value of the product more dominant in the sum expression (Eq. 4.7). How 'large' the multiplier is will be discussed in Necessary Condition 2.

Necessary Condition 1: The negative off-diagonal cofactor(s)

According to the relationships in Eq. 4.10, we have three initial conclusions about the relationship between the internal information of the three echoes' data and weight values:

1. The off-diagonal cofactor term value becomes more negative as the magnitude of the partial correlation is larger, e.g., $cof_{13} = -C_5 \cdot r_{13,2}$.

2. The diagonal cofactor term value becomes less positive as the magnitude of the correlation coefficient is larger, e.g. $cof_{33} = C_2 \left(1 - r_{13}^2\right)$.

3. The integration of conclusion 1 and 2 supports the notion that along with larger association between two echoes' data, the sum formula(s) in Eq. 4.7 tends to be negative, which results in a negative weight(s), e.g., $w_3 = K\left(-C_5 \cdot r_{13,2}\bar{s}_1 - C_6 \cdot r_{23,1}\bar{s}_2 + C_3\left(1 - r_{12}^2\right)\bar{s}_3\right)$

Necessary Condition 2: The mean value

In the case of w_{tOPT}, under Condition 1, if the multiplier of the positive diagonal cofactor is in a range so that the negative cofactors are able to be amplified by the corresponding multipliers, the weight can be negative. Because we only focused on the relative value of each term of the mean in the calculation of weight, normalized mean values were mapped on the brain in Fig. 4.7.

In our case, the mean values of all voxels across three TEs follow the relationship:

$$\bar{s}_1 > \bar{s}_2 > \bar{s}_3 > 0 \tag{4.11}$$

According to Eq. 4.7, given that one weight is negative, the multiplier ($\bar{s}_i$ in w_{tOPT}) of the positive diagonal cofactor would satisfy:

$$\bar{s}_1 < -\frac{\bar{s}_2 cof_{12} + \bar{s}_3 cof_{13}}{cof_{11}} = \frac{\bar{s}_2 \cdot C_4 \cdot r_{12,3} + \bar{s}_3 \cdot C_5 \cdot r_{13,2}}{C_1 \left(1 - r_{23}^2\right)} \tag{4.12}$$

when taking the expression of the first optimal weight (Eq.4.7a) as an example for the ease of explanation. This reasoning can be generalized for the other equations of weights.

4.2.2 Empirical evidence

There were two pieces of supporting evidence which agreed with the theoretical explanation hypothesized above. The first is the consistency between the regional characteristics of negative weights and correlation/partial correlation, and the second is the distribution of negative weights varying in terms of TE. Moreover, three typical voxels were selected as supplementary examples to support the above arguments.

Regional Characteristics

When examining the regional characteristic of negative weights, the maps of w_{tOPT} and w_{mOPT}, especially Fig. 3.4c and Fig. 3.5a, visually show that most voxels with negative weights (blue voxels) are located in the GM region, which is consistent with the localization of the voxels with high values in correlation and partial correlation shown in Fig. 4.6, where the red voxels with high values are mostly located in the GM region. This consistency in regional distribution agrees with the inferences mentioned in Condition 1 that the high correlation and partial correlation

are responsible for negative weights. However, it is noticeable that regional variation of the normalized mean value across the brain is not visually obvious as indicated in Fig. 4.7 when compared to the maps of correlation and partial correlation. The normalized mean value doesn't appear to be relatively higher value in the GM region. Hence, the regional characteristics of correlation and partial correlation are the main factors of such characteristic of negative weights.

Distribution Characteristics of TE

In Fig. 3.4, there is an apparent increase in the number of negative $w_{tOPT,3}$ (27675) compared to $w_{tOPT,1}$ (44). The reason for such a distribution of negative weights in terms of TE is mainly the magnitude of the multiplier ($\bar{s}_1$, $\bar{s}_2$ and $\bar{s}_3$) for each cofactor term in the three relations in Eq. 4.7. In Eq. 4.7a, the positive diagonal cofactor term cof_{11} is multiplied by the largest mean value $\bar{s}_1$ relative to $\bar{s}_2$ and $\bar{s}_3$. In our case, there are only 44 voxels with $\bar{s}_1$ following the relation as Eq.4.12. But the positive cofactor (cof_{33}) is amplified by the smallest mean values $\bar{s}_3$ in Eq. 4.7c and there are 27675 voxels with $\bar{s}_3$ following the restriction like Eq. 4.7c:

$$\bar{s}_3 < -\frac{\bar{s}_1 cof_{13} + \bar{s}_2 cof_{23}}{cof_{33}} = \frac{\bar{s}_1 \cdot C_5 \cdot r_{13,2} + \bar{s}_2 \cdot C_6 \cdot r_{23,1}}{C_3\left(1 - r_{12}^2\right)} \tag{4.13}$$

Hence, more voxels follow such a relation for the multiplier of positive cofactor in the calculation of $w_{tOPT,3}$ compared to number of voxels in the calculation of $w_{tOPT,1}$, which led to more voxels with negative $w_{tOPT,3}$.

Difference between negative w_{tOPT} and w_{mOPT}

The difference between the appearance of negative w_{tOPT} and w_{mOPT} in terms of TE, which was described in the results section, is mainly due to the change of the multipliers for

cofactor terms in Eq. 4.7, from $\bar{s}_i$ to $TE_i\bar{s}_i$. Then, Eq. 4.7 was re-written as:

$$w_1 = K\left(TE_1\bar{s}_1 cof_{11} + TE_2\bar{s}_2 cof_{12} + TE_3\bar{s}_3 cof_{13}\right) \tag{4.14a}$$

$$w_2 = K\left(TE_1\bar{s}_1 cof_{12} + TE_2\bar{s}_2 cof_{22} + TE_3\bar{s}_3 cof_{23}\right) \tag{4.14b}$$

$$w_3 = K\left(TE_1\bar{s}_1 cof_{13} + TE_2\bar{s}_2 cof_{23} + TE_3\bar{s}_3 cof_{33}\right) \tag{4.14c}$$

They can also be written as the expressions of the correlations and partial correlation:

$$w_1 = K\left(C_1\left(1-r_{23}^2\right)\ TE_1\bar{s}_1 - C_4\cdot\ TE_2\bar{s}_2\cdot r_{12,3} - C_5\cdot\ TE_3\bar{s}_3\cdot r_{13,2}\right) \tag{4.15a}$$

$$w_2 = K\left(-C_4\cdot TE_1\bar{s}_1\cdot r_{12,3} + C_2\left(1-r_{13}^2\right)\ TE_2\bar{s}_2 - C_6\cdot TE_3\bar{s}_3\cdot r_{23,1}\right) \tag{4.15b}$$

$$w_3 = K\left(-C_5\cdot TE_1\bar{s}_1\cdot r_{13,2} - C_6\cdot TE_2\bar{s}_2\cdot r_{23,1} + C_3\left(1-r_{12}^2\right)TE_3\bar{s}_3\right) \tag{4.15c}$$

As indicated in Fig. 4.8, for most voxels (80.35%), the relationship of the new multiplier $TE_i\bar{s}_i$ is opposite to the relationship of the mean value used for w_{tOPT}

$$TE_3\bar{s}_3 > TE_2\bar{s}_2 > TE_1\bar{s}_1 > 0 \tag{4.16}$$

which results in more negative weights in $w_{mOPT,1}$ instead of in $w_{mOPT,3}$ as demonstrated in the results. The relatively larger multiplier, $TE_3\bar{s}_3$ was used for positive cof_{33} in the calculation of the first term of Eq. 4.14 compared to $\bar{s}_3$ in Eq. 4.7, but smaller multiplier ($TE_1\bar{s}_1$) was used for positive cof_{11} compared to $\bar{s}_1$. More voxels' $TE_1\bar{s}_1$ (14850) satisfy the relation to result in negative $w_{mOPT,1}$

$$TE_1\bar{s}_1 < -\frac{TE_2\bar{s}_2 cof_{12} + TE_3\bar{s}_3 cof_{13}}{cof_{11}} = \frac{TE_2\bar{s}_2\cdot C_4\cdot r_{12,3} + TE_3\bar{s}_3\cdot C_5\cdot r_{13,2}}{C_1\left(1-r_{23}^2\right)} \tag{4.17}$$

But fewer voxels' $TE_3\bar{s}_3$ (471) satisfy the relation to result in negative $w_{mOPT,3}$

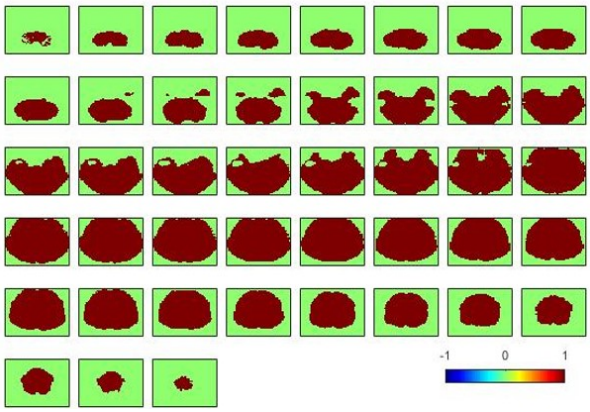
$$TE_3\bar{s}_3 < -\frac{TE_1\bar{s}_1 cof_{13} + \bar{s}_2 cof_{23}}{cof_{33}} = \frac{TE_1 TE_2 \bar{s}_1 \cdot C_5 \cdot r_{13,2} + TE_2\bar{s}_2 \cdot C_6 \cdot r_{23,1}}{C_3\left(1 - r_{12}^2\right)} \tag{4.18}$$

Thereby, the change of relatively larger multiplier caused more voxels to satisfy the condition to make the first term negative (14850), but less voxels to satisfy for the third term (417), which is opposite to the case of w_{tOPT}.

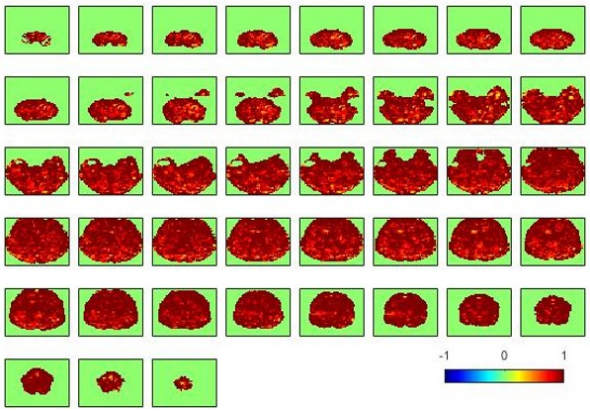
However, the cause of the similarity in the regional distribution of negative weights for w_{tOPT} and w_{mOPT} is that voxels with high partial correlation and correlation between two TEs' data are similarly near the GM region.

Typical Voxels

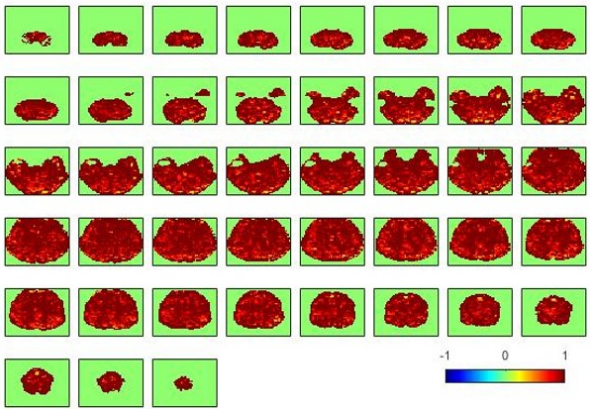
In Fig. 4.9, three typical voxels were selected to be examined and compared with respect to their internal information, such as the mean, the partial correlation and the cofactor which give us a more straightforward form of evidence to support the reasoning. The comparison between Voxel #33375 and Voxel #71610 indicates that due to the relatively larger partial correlation and Pearson correlation coefficients between two echoes' data in Voxel #3337, larger negative off-diagonal cofactors (cof_{12},cof_{13},cof_{23}) are produced which lead to more negative weights. For Voxel #66194, in terms of TE, the occurrence of negative weights in W_{tOPT} and W_{mOPT} have the opposite behavior, which shows that the second and third W_{tOPT} are negative while the first and second W_{mOPT} were negative. This difference can be due to the different relationships that $\bar{s}_i$ and $TE_i\bar{s}_i$ follow, which have been proposed in Eq. 4.11 and Eq. 4.16. Relatively smaller $\bar{s}_3$ and $TE_1\bar{s}_1$ as the multipliers make the results of Eq. 4.10c and Eq. 4.15a negative respectively, because they follow the restrictions in Eq. 4.13 and Eq. 4.17.



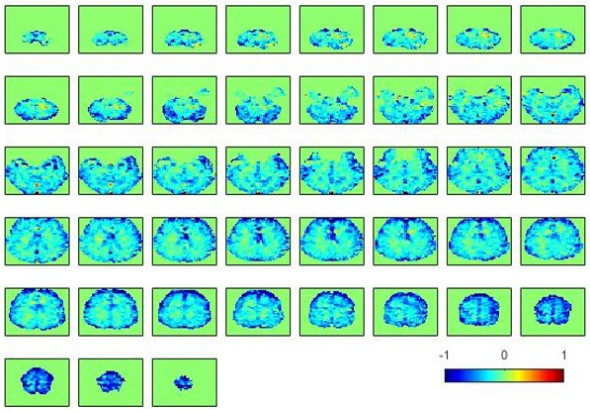
(a) cof_{11}



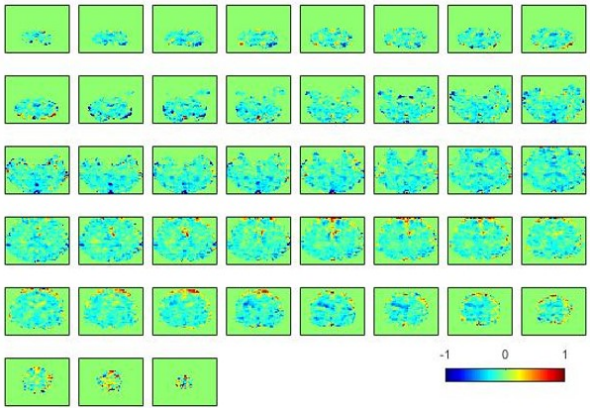
(b) cof_{22}



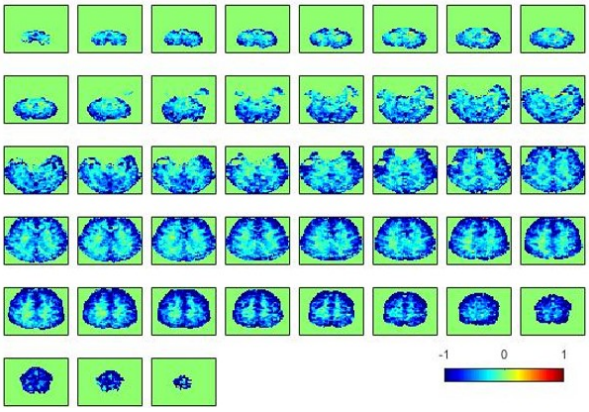
(c) cof_{33}



(d) cof_{12}

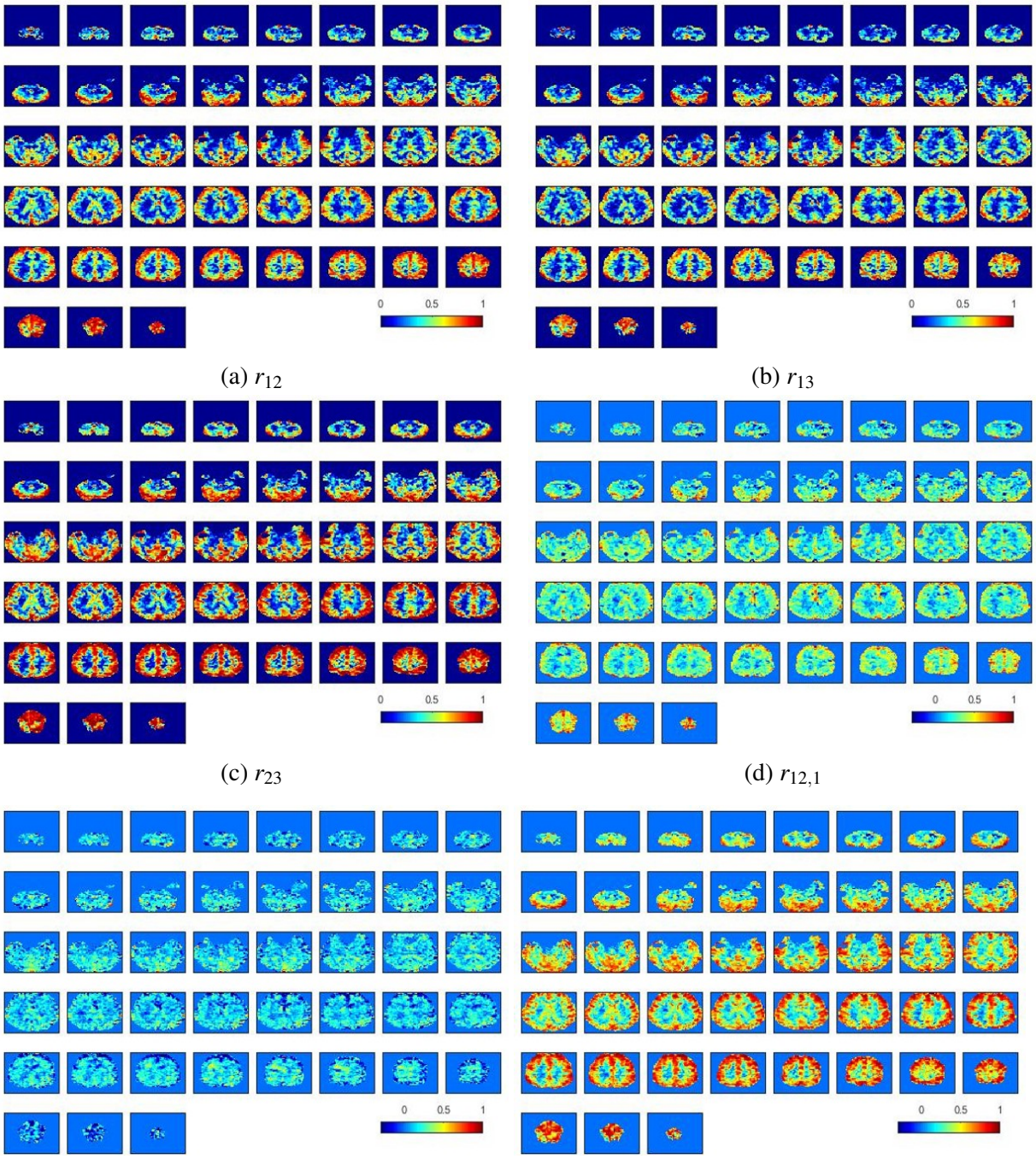


(e) cof_{13}



(f) cof_{23}

Figure 4.5: **Brain maps of cofactors.** (a) to (c) are the maps for diagonal cofactor terms. (d) to (f) for the off-diagonal cofactor terms. Each cofactor term of each voxel was normalized by its cof_{11}.

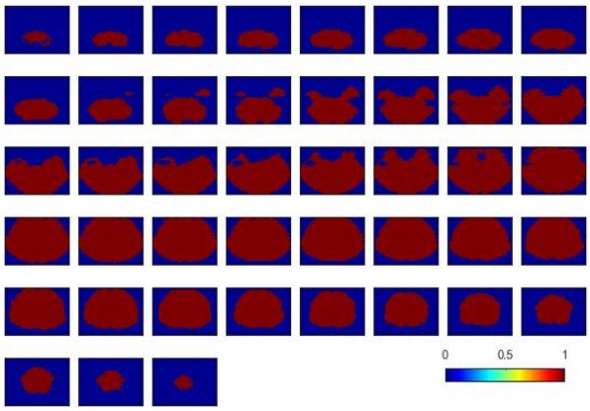


(a) r_{12} (b) r_{13}

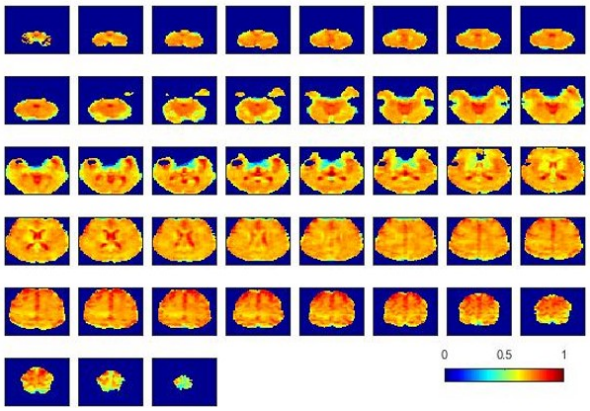
(c) r_{23} (d) $r_{12,1}$

(e) $r_{13,2}$ (f) $r_{23,1}$

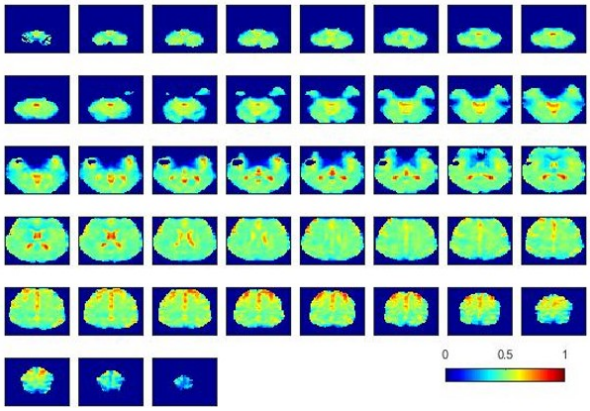
Figure 4.6: **Brain maps of correlation coefficient and partial correlation.** (a) to (c) are the maps for correlation coefficients. (d) to (f) are the maps for partial correlation between two echoes' data, controlling the third echo's data.



(a) $\bar{s}_1$

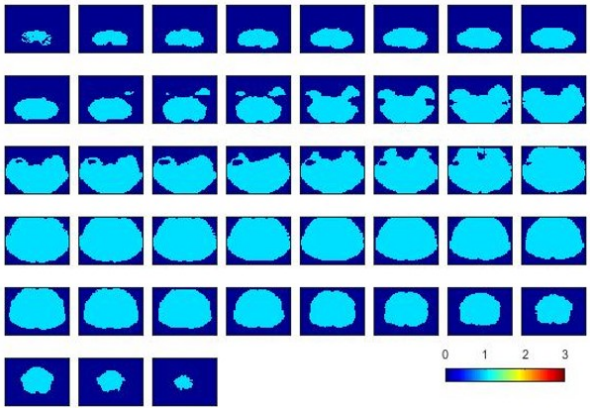


(b) $\bar{s}_2$

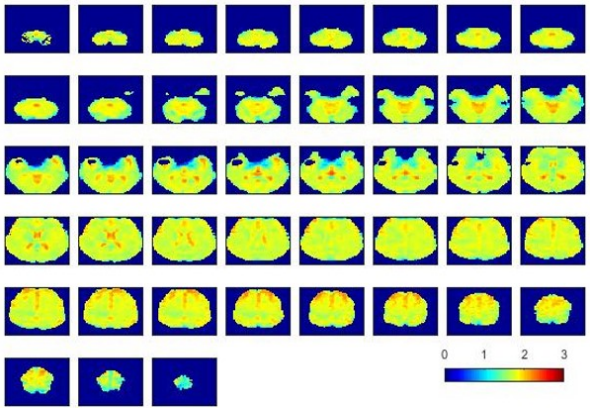


(c) $\bar{s}_3$

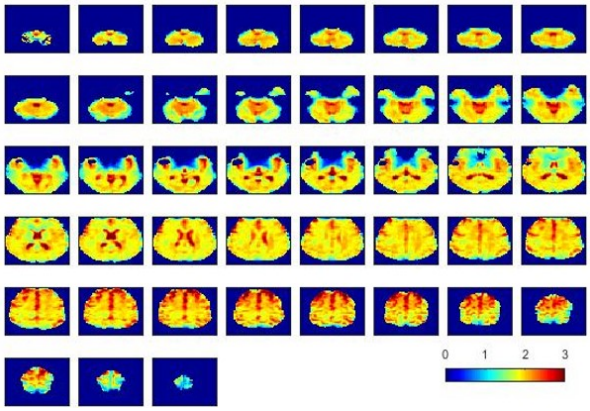
Figure 4.7: **Brain map of mean value.** For each voxel, each term of mean value was normalized by the first term of the mean value. For all voxels, the mean value follows $\bar{s}_1 > \bar{s}_2 > \bar{s}_3 > 0$.



(a) $TE_1\bar{s}_1$



(b) $TE_2\bar{s}_2$

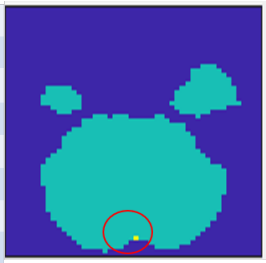


(c) $TE_3\bar{s}_3$

Figure 4.8: **Brain map of TE $\times$ Mean.** For each voxel, each $TE_i\bar{s}_i$ was normalized by $TE_1\bar{s}_1$ and the mean of each normalized value is 1.00, 1.71, 1.85. For the majority of voxels, $TE_k\bar{s}_k$ follows the relations: for 99.98% of voxels, $TE_2\bar{s}_2 > TE_1\bar{s}_1 > 0$; for 95.73% of voxels, $TE_3\bar{s}_3 > TE_1\bar{s}_1 > 0$; for 80.39% of voxels, $TE_3\bar{s}_3 > TE_2\bar{s}_2 > 0$.

(a) Voxel #33375

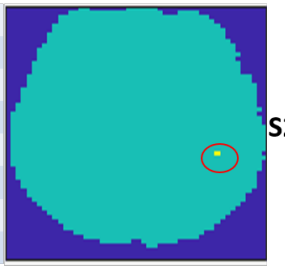
W_{tOPT1}		W_{tOPT2}		W_{tOPT3}	
0.4526		-0.1951		-0.3523	
$\mathbf{Mean_{TE_1}}$		$\mathbf{Mean_{TE_2}}$		$\mathbf{Mean_{TE_3}}$	
1		0.4663		0.1851	
$corr_{12}$	$corr_{13}$	$corr_{23}$	$r_{12,3}$	$r_{13,2}$	$r_{23,1}$
0.8652	0.8263	0.7612	0.6458	0.5142	0.1670
$Cofactor_{11}$	$Cofactor_{22}$	$Cofactor_{33}$	$Cofactor_{12}$	$Cofactor_{13}$	$Cofactor_{23}$
1	0.7471	2.5940	-0.5582	-0.8282	-0.2325



S12

(b) Voxel #71610

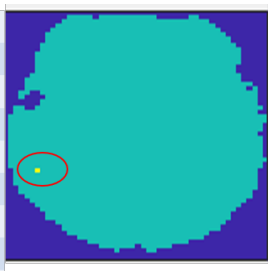
$W_{tOPT,1}$		$W_{tOPT,2}$		$W_{tOPT,3}$	
0.3950		0.3275		0.2774	
$\mathbf{Mean_1}$		$\mathbf{Mean_2}$		$\mathbf{Mean_3}$	
1		0.6768		0.4611	
r_{12}	r_{13}	r_{23}	$r_{12,3}$	$r_{13,2}$	$r_{23,1}$
0.3585	0.1200	0.0453	0.3500	0.1113	0.0025
$Cofactor_{11}$	$Cofactor_{22}$	$Cofactor_{33}$	$Cofactor_{12}$	$Cofactor_{13}$	$Cofactor_{23}$
1	1.4328	1.2728	-0.4261	-0.1255	-0.0033



S26

(c) Voxel #66194

$W_{tOPT,1}$		$W_{tOPT,2}$		$W_{tOPT,3}$	
0.6609		-0.1309		-0.1992	
$W_{mOPT,1}$		$W_{mOPT,2}$		$W_{mOPT,3}$	
-0.0197		-0.2776		0.7028	
$\mathbf{Mean_1}$		$\mathbf{Mean_2}$		$\mathbf{Mean_3}$	
1		0.8618		0.7370	
$\mathbf{TE_1 \cdot Mean_1}$		$\mathbf{TE_2 \cdot Mean_2}$		$\mathbf{TE_3 \cdot Mean_3}$	
1		1.6502		1.6582	
r_{12}	r_{13}	r_{23}	$r_{12,3}$	$r_{13,2}$	$r_{23,1}$
0.8360	0.8077	0.9168	0.4068	0.1873	0.7465
$Cofacto_{11}$	$Cofactor_{22}$	$Cofactor_{33}$	$Cofactor_{12}$	$Cofactor_{13}$	$Cofactor_{23}$
1	0.9286	1.4506	-0.3902	-0.2256	-0.8664



S24

Figure 4.9: **Typical voxels for comparison.** The numbers on the right are the slices that voxels are in. (a) Voxel #33375 has two negative W_{tOPT}; (b) Voxel #71610 only has positive W_{tOPT}; (c) Voxel #66194 has two negative W_{tOPT} in W2 and W3 and also has two negative W_{mOPT} in W1 and W3. **Note:** To ease the display and analysis, the mean and the cofactor are normalized by the first term of each variable. Weights shown here were normalized by the L1-Norm. The calculation of the weights for the selected voxels can be referenced in Appendix A.4.

Chapter 5

Conclusion

The quality of the combined signal was quantified by tSNR and metSNR, because both metrics can represent the CNR firmly. Taking all the variance information of multiple echoes' into account, the optimized weighting method in the combination step for each voxel in the brain region showed great improvement in these two evaluation metrics. Although only one dataset was used for testing, we believe that the optimal weighting methods can be suitable for any multi-echo fMRI dataset due to the rigorous mathematical derivation process provided.

The finding of negative weights led us to re-consider the range of weights in previous research, expand the non-negative values to real values and propose the new definition of normalization in weighting summation as 'L1-Norm Normalization'. For one voxel with negative weight(s), the causes of negative weight(s) are mainly two properties of its echo's time series, which are responsible for the calculation of the optimal weight: the high correlation between two echo's data resulting in the negative cofactors and the relatively small multiplier ($\bar{s}$ or $TE\bar{s}$) for the positive cofactors. These causes can also explain the interesting distribution of negative weights in terms of brain region and TE.

In addition, for most voxels, the metSNR varied less than tSNR when comparing the results of non-optimal to the optimal method, especially when using Bold Sensitivity weighting

method, such as mBSENS, tmBSENS and T2W. The comparison of two metrics was represented by their variation ratio which was less than 1. According to the geometric property of Rayleigh quotient, we found that the $\bar{S}$ and $\Lambda\bar{S}$ used for tSNR and metSNR were suitable for different non-optimal methods which can be determined by their similarity to the optimal input vector X of the corresponding Rayleigh quotient.

Appendix A

Note: In all expressions, matrices will be represented with uppercase letters, while scalars will be lowercase.

A.1 Derivation of Optimal Weights for tSNR of Combined Data

For a single echo acquisition, the measure for contrast-to-noise ratio (CNR) [Welvaert and Rosseel, 2013] is:

$$\mathrm{CNR} = \frac{\Delta s_1}{\sigma_1} \tag{A.1}$$

where Δs_1 is the signal change and σ_1 is the standard deviation. When dealing with fMRI data, we typically think in terms of percent BOLD change $\%\Delta s_1$, and so it is useful to rewrite the CNR as follows:

$$\frac{\Delta s_1}{\sigma_1} = \frac{(\%\Delta s_1)(\bar{s}_1)}{\sigma_1} = (\%\Delta s_1)\frac{\bar{s}_1}{\sigma_1} = (\%\Delta s_1)\,\mathrm{tSNR}_1 \tag{A.2}$$

where $\bar{s}_1$ denotes the temporal mean of the signal. Eq. A.2 shows *tSNR* is a useful measure. Recall from Equation (2.9) and (2.10), *tSNR* can be expressed as:

$$tSNR = \frac{\bar{s}}{\sqrt{\sigma^2}} \tag{A.3}$$

Thus, the detailed expression of $tSNR$ for the combined signal is

$$\begin{aligned} tSNR_{comb} &= \frac{\bar{s}_{comb}}{\sqrt{\sigma^2_{comb}}} \\ &= \frac{\sum_{i=1}^{3} \frac{w_k}{\sum_{i=1}^{3}|w_k|}\bar{s}_k}{\sqrt{\sum_{k=1}^{3} \frac{w_k^2}{\sum_{k=1}^{3}|w_k|^2} Var\left(s_k\right) + 2\sum_{1\le k\le j\le 3} \frac{w_k w_j}{\sum_{k=1}^{3}|w_k|^2} Cov\left(s_k, s_j\right)}} \\ &= \frac{\sum_{i=1}^{3} w_k \bar{s}_k}{\sqrt{\sum_{k=1}^{3} w_k^2 Var\left(s_k\right) + 2\sum_{1\le k\le j\le 3} w_k w_j Cov\left(s_k, s_j\right)}} \end{aligned} \tag{A.4}$$

The long fraction can be rewritten as an expression of parameter matrices. $\sum_{i=1}^{3} w_k \bar{s}_k$ can be written as $W^T \bar{S}$ and the sum form in the denominator can be written as $W^T \Sigma W$. So, the $tSNR$ for the combined time series can be expressed with parameter matrices as:

$$tSNR_{comb} = \frac{W^T \bar{S}}{\sqrt{W^T \Sigma W}} \tag{A.5}$$

where W is the 3×1 un-normalized weights vector, $\bar{S}$ is the 3×1 mean-value vector and Σ is the covariance matrix which is introduced in more detail in 'Methods' section and defined as:

$$\Sigma = \begin{vmatrix} Var\left(s_1\right) & Cov\left(s_1, s_2\right) & Cov\left(s_1, s_3\right) \\ Cov\left(s_1, s_2\right) & Var\left(s_2\right) & Cov\left(s_2, s_3\right) \\ Cov\left(s_1, s_3\right) & Cov\left(s_2, s_3\right) & Var\left(s_3\right) \end{vmatrix} \tag{A.6}$$

Then, the square of tSNR:

$$tSNR^2_{comb} = \frac{W^T \bar{S}\bar{S}^T W}{W^T \Sigma W} = \frac{W^T M W}{W^T \Sigma W} \tag{A.7}$$

is considered first because its expression form is similar to the general RQ model. According to

the RQ optimization [Lischinski, 2007], maximization of $tSNR^2$ can be converted to

$$\begin{aligned} &\max W^T M W \\ &\text{s.t.} \quad W^T \Sigma W = c \end{aligned} \tag{A.8}$$

In other words, we wish to find the critical points of the Lagrangian expression:

$$L = W^T M W - \lambda(W^T \Sigma W - c) \tag{A.9}$$

by taking the derivative with respect to W:

$$\frac{dL}{dW} = 2(M - \lambda\Sigma)W = 0 \tag{A.10}$$

which results in

$$MW = \lambda_{max}\Sigma W \tag{A.11}$$

to maximize $tSNR^2$ with largest eigenvalue λ_{max}. In our case, where $M = \bar{S}\bar{S}^T$, the optimal weights can be expressed in an alternative way without calculating eigenvectors and eigenvalues. Substituting $M = \bar{S}\bar{S}^T$ into Eq. A.11 gives:

$$\Sigma^{-1}\bar{S}\bar{S}^T W = \lambda_{max} W \tag{A.12}$$

And $\bar{S}^T W$ can be seen as a scalar, so the optimal W can be written as:

$$W_{tOPT} = \frac{\bar{S}^T W}{\lambda_{max}}\Sigma^{-1}\bar{S} = k\Sigma^{-1}\bar{S} \tag{A.13}$$

where k is an arbitrary scalar. Because $tSNR^2$ is always positive and maximized with the optimal weight vector W_{tOPT}, $tSNR$, the positive square root of $tSNR^2$, can reach the maximum when using the calculated optimal weights for combination according to the feature of square root of

functions [Roberts, 2012].

A.2 Derivation of metSNR for Combined Data and its Optimal Weights

Following Equation A.2, for the weighted sum of signals $s_{comb} = \sum_{i=1}^{3} w_{mOPT,\,k} \cdot s_k(t)$, the desired expression of *CNR* is

$$CNR = \frac{\Delta s_{comb}}{std\left(s_{comb}\right)} \tag{A.14}$$

Then, the numerator term can be expanded to:

$$\begin{aligned} \Delta s_{comb} &= \sum_{k=1}^{N_E} w_k \Delta s_k \\ &= \sum_{k=1}^{N_E} w_k \frac{\Delta s_k}{\bar{s}_k} \bar{s}_k \end{aligned} \tag{A.15}$$

For multiple echo acquisitions, we must consider the fact that for the kth echo, the percent signal change $\%\Delta s_k = \frac{\Delta s_k}{\bar{s}_k} \approx -\Delta R_2^* \cdot TE_k$ [Kundu et al., 2012]. Thus, the combined signal change can be approximated as:

$$\begin{aligned} \Delta s_{comb} &\approx \left(-\Delta R_2^*\right) \sum_{k=1}^{N_E} w_k \mathrm{TE}_k \bar{s}_k \\ &= \left(-\Delta R_2^*\right) W^T \Lambda \bar{S} \end{aligned} \tag{A.16}$$

where $\Lambda = \mathrm{diag}\left(\begin{bmatrix} TE_1 & TE_2 & \cdots & TE_{N_E} \end{bmatrix}\right)$. Putting everything together, we may approximate the expression for *CNR* as:

$$\begin{aligned} CNR &\approx \left(-\Delta R_2^*\right) \frac{W^T \Lambda \bar{S}}{\mathrm{std}\left(W^T S\right)} \\ &= \left(-\Delta R_2^*\right) \cdot metSNR \end{aligned} \tag{A.17}$$

where *metSNR* denotes multi-echo tSNR.

The detailed expression of *metSNR* is also provided for reference:

$$
\begin{aligned}
metSNR_{comb} &= \frac{\sum_{i=1}^{3} \frac{w_k}{\sum_{i=1}^{3}|w_k|} TE_k \bar{s}_k}{\sqrt{\sum_{k=1}^{3} \frac{w_k^2}{\sum_{k=1}^{3}|w_k|^2} Var\left(s_k\right) + 2\sum_{1\le k\le j\le 3} \frac{w_k w_j}{\sum_{k=1}^{3}|w_k|^2} Cov\left(s_k, s_j\right)}} \\
&= \frac{\sum_{i=1}^{3} w_k TE_k \bar{s}_k}{\sqrt{\sum_{k=1}^{3} w_k^2 Var\left(s_k\right) + 2\sum_{1\le k\le j\le 3} w_k w_j Cov\left(s_k, s_j\right)}}
\end{aligned}
\tag{A.18}
$$

Since *metSNR* equals $\frac{W^T \Lambda \bar{S}}{\mathrm{std}(W^T S)}$, whose denominator terms are exactly the same as *tSNR*'s and the numerator terms have a similar layout, the process of derivation is similar to the derivation of W_{tOPT} and the optimal weights for *metSNR*, W_{mOPT}, can be proven to be:

$$
W_{mOPT} = \frac{\bar{S}^T W}{\lambda_{max}} \Sigma^{-1} \Lambda \bar{S} = k \Sigma^{-1} \Lambda \bar{S} \tag{A.19}
$$

A.3 Flowchart of Data Pre-processing

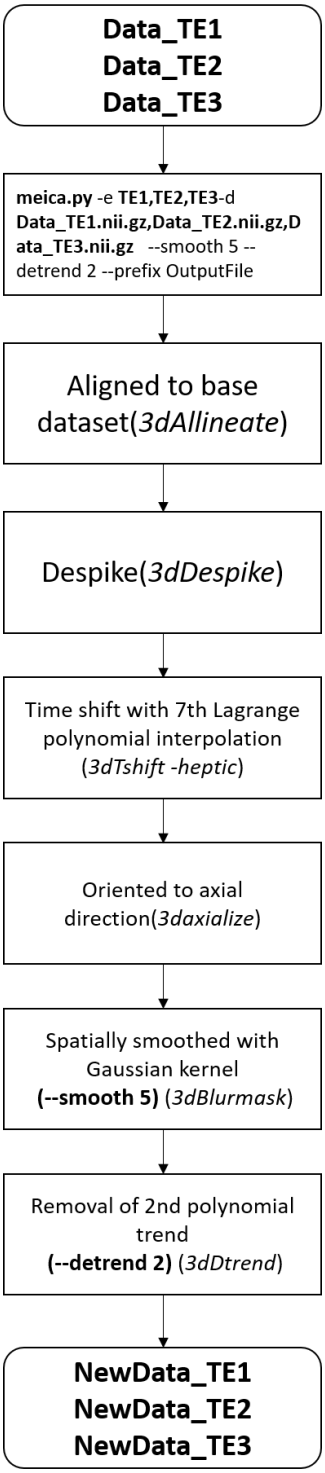


Figure A.1: Flowchart of Data Pre-processing

A.4 Optimal weights for the selected voxels

This part is used for the calculation of weights for the selected voxels for reference.

Voxel #33375:

Table A.1: Parameters for the calculation of w_{tOPT} of Voxel #33375

$K = 5.861 \times 10^{-13}$	$C_1 = 4.841 \times 10^8$	$C_2 = 4.780 \times 10^8$	$C_3 = 2.094 \times 10^9$
$C_4 = 1.754 \times 10^8$	$C_5 = 3.268 \times 10^8$	$C_6 = 2.825 \times 10^8$	$r_{12} = 0.865$
$r_{13} = 0.826$	$r_{23} = 0.761$	$r_{12,3} = 0.646$	$r_{13,2} = 0.514$
$r_{23,1} = 0.167$	$\bar{s}_1 = 9.545 \times 10^3$	$\bar{s}_2 = 4.450 \times 10^3$	$\bar{s}_3 = 1.767 \times 10^3$

Voxel #71610:

Table A.2: Parameters for the calculation of w_{tOPT} of Voxel #71610

$K = 3.227 \times 10^{-10}$	$C_1 = 1.833 \times 10^6$	$C_2 = 2.659 \times 10^6$	$C_3 = 2.671 \times 10^6$
$C_4 = 2.189 \times 10^6$	$C_5 = 2.063 \times 10^6$	$C_6 = 2.470 \times 10^6$	$r_{12} = 0.359$
$r_{13} = 0.120$	$r_{23} = 0.045$	$r_{12,3} = 0.356$	$r_{13,2} = 0.111$
$r_{23,1} = 0.003$	$\bar{s}_1 = 1.289 \times 10^4$	$\bar{s}_2 = 8.724 \times 10^3$	$\bar{s}_3 = 5.943 \times 10^3$

Voxel #66194:

Table A.3: Parameters for the calculation of w_{tOPT} of Voxel #66194

$K = 1.757 \times 10^{-11}$	$TE_1\bar{s}_1 = 1.882 \times 10^5$	$TE_2\bar{s}_2 = 3.105 \times 10^5$	$TE_2\bar{s}_2 = 3.120 \times 10^5$
$C_1 = 1.665 \times 10^8$	$C_2 = 7.094 \times 10^7$	$C_3 = 1.281 \times 10^8$	$C_4 = 2.559 \times 10^7$
$C_5 = 3.198 \times 10^7$	$C_6 = 3.082 \times 10^7$	$r_{12} = 0.836$	$r_{13} = 0.808$
$r_{23} = 0.917$	$r_{12,3} = 0.407$	$r_{13,2} = 0.187$	$r_{23,1} = 0.747$
$\bar{s}_1 = 9.545 \times 10^3$	$\bar{s}_2 = 4.450 \times 10^3$	$\bar{s}_3 = 1.767 \times 10^3$	

These parameters can be easily used for the calculation of w_{tOPT} and w_{mOPT} according to Eq. 4.10 and Eq. 4.15. And these calculated weights were normalized by L1-Norm which

are shown in Fig. A.4. The constants K and C_i were directly calculated from determinant of covariance matrix and Eq. 4.9.

Bibliography

J. Cohen. Statistical Power Analysis for the Behavioral Sciences. L. Erlbaum Associates, 1988.

R. W. Cox. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. Computers and Biomedical Research, 29(3):162 – 173, 1996. ISSN 0010-4809. doi: https://doi.org/10.1006/cbmr.1996.0014.

R. Deichmann, O. Josephs, C. Hutton, D. Corfield, and R. Turner. Compensation of susceptibility-induced BOLD sensitivity losses in echo-planar fMRI imaging. NeuroImage, 15(1):120 – 135, 2002. ISSN 1053-8119. doi: https://doi.org/10.1006/nimg.2001.0985.

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, 44(1): 162–167, 2000. doi: 10.1002/1522-2594(200007)44:1⟨162::AID-MRM23⟩3.0.CO;2-E.

A. Kettinger, C. Hill, Z. Vidnynszky, C. Windischberger, and Z. Nagy. Investigating the group-level impact of advanced dual-echo fMRI combinations. Frontiers in Neuroscience, 10:571, 2016. ISSN 1662-453X. doi: 10.3389/fnins.2016.00571.

P. Kundu, S. J. Inati, J. W. Evans, W.-M. Luh, and P. A. Bandettini. Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. NeuroImage, 60(3):1759 – 1770, 2012. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2011.12.028.

D. Lischinski. Lagrange multipliers and the Rayleigh quotient, 2007. URL `http://www.cs.huji.ac.il/~csip/lagrange.pdf`.

P. Mansfield. Multi-planar image formation using NMR spin echoes. Journal of Physics C: Solid State Physics, 10(3):L55–L58, feb 1977. doi: 10.1088/0022-3719/10/3/004. URL `https://doi.org/10.1088%2F0022-3719%2F10%2F3%2F004`.

B. A. Poser, M. J. Versluis, J. M. Hoogduin, and D. G. Norris. BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: Parallel-acquired inhomogeneity-desensitized fMRI. Magnetic Resonance in Medicine, 55(6):1227–1235, 2006. doi: 10.1002/mrm.20900.

S. Posse, S. Wiese, D. Gembris, K. Mathiak, C. Kessler, M.-L. Grosse-Ruyken, B. Elghahwagi, T. Richards, S. R. Dager, and V. G. Kiselev. Enhancement of BOLD-contrast sensitivity

by single-shot multi-echo functional MR imaging. Magnetic Resonance in Medicine, 42(1): 87–97, 1999. doi: 10.1002/(SICI)1522-2594(199907)42:1⟨87::AID-MRM13⟩3.0.CO;2-O.

D. M. Roberts. Square root & cube root functions. 2012. URL https://mathbitsnotebook.com/Algebra1/FunctionGraphs/FNGTypeSquareRoot.html.

S. Sawilowsky. New effect size rules of thumb. Journal of Modern Applied Statistical Methods, 8:597–599, 11 2009. doi: 10.22237/jmasm/1257035100.

L. N. Trefethen and D. Bau. Numerical Linear Algebra. SIAM, 1997. ISBN 0898713617.

M. Welvaert and Y. Rosseel. On the definition of signal-to-noise ratio and contrast-to-noise ratio for fMRI data. PLOS ONE, 8:1–10, 11 2013.