Modelling Individual Variations in Brain Structure and Function Using Multimodal MRI

Ying-Qiu Zheng

St Edmund Hall

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2023

Acknowledgements

I would like to express my deepest gratitude to my incredible supervisors, Prof. Stephen Smith and Prof. Saad Jbabdi, for their constant guidance, support, and encouragement throughout my research journey. Their invaluable insights and expertise have been instrumental in shaping this thesis, and their patience and understanding during the most challenging moments have made all the difference. I am truly fortunate to have had the opportunity to learn from their wisdom and experience.

And to my loving parents – words cannot express my appreciation for the unconditional love and support they have provided me during this journey. Their constant presence and encouragement, especially during the dark times when I was struggling with depression, have been my source of strength and resilience. I am forever grateful for their sacrifices and unwavering faith in me, as well as their ability to lift me up when I needed it the most.

Last but not least, I would like to extend my heartfelt thanks to my friends and colleagues, who have been a constant source of inspiration, support, and camaraderie. Their friendship has made this challenging journey not only bearable but also enjoyable and memorable. I am grateful to have them in my life.

Abstract

Every brain is different. Understanding this variability is crucial for investigating the neural substrate underlying individuals' unique behaviour and developing personalised diagnosis and treatments. This thesis presents novel computational approaches to study individual variability in brain structure and function using magnetic resonance imaging (MRI) data. It comprises three main chapters, each addressing a specific challenge in the field.

In Chapter 3, the thesis proposes a novel Image Quality Transfer (IQT) technique, HQ-augmentation, to accurately localise a Deep Brain Stimulation (DBS) target in low-quality clinical-like data. Leveraging high-quality diffusion MRI datasets from the Human Connectome Project (HCP), the HQ-augmentation approach is robust to corruptions in data quality while preserving the individual anatomical variability of the DBS target. It outperforms existing alternatives and generalises to unseen low-quality diffusion MRI datasets with different acquisition protocols, such as the UK Biobank (UKB) dataset.

In Chapter 4, the thesis presents a framework for enhancing prediction accuracy of individual task-fMRI activation profiles using the variability of resting-state fMRI. Assuming resting-state functional modes underlie task-evoked activity, this chapter demonstrates that shape and intensity of individualised task activations can be separately modelled. This chapter introduced the concept of "residualisation" and showed that training on residuals leads to better individualised predictions. The framework's prediction accuracy, validated on HCP and UKB data, is on par with task-fMRI test-retest reliability, suggesting potential for supplementing traditional task localisers.

In Chapter 5, the thesis presents a novel framework for individualised retinotopic mapping using resting-state fMRI, from the primary visual cortex to visual cortex area 4. The proposed approach reproduces task-elicited retinotopy and captures individual differences in retinotopic organisation. The proposed framework delineates borders of early visual areas more accurately than group-average parcellation and is effective with both high-field 7T and more common 3T resting-state fMRI data, providing a valuable alternative to resource-intensive retinotopy task-fMRI experiments.

Overall, this thesis demonstrates the potential of advanced MRI analysis techniques to study individual variability in brain structure and function, paving the way for improved clinical applications tailored to individual patients and a better understanding of neural mechanisms underlying unique human behaviour.

Contents

List of Figures

List of Abbreviations

1-D, 2-D, 3-D	One-, two- or two-dimensional, referring in this thesis to spatial dimensions in an image.
BEDPOSTX	Bayesian estimation of diffusion parameters obtained using sampling techniques for crossing fibres
BOLD	Blood Oxygenation Level-Dependent
CRF	Conditional Random Field
CSF	Cerebrospinal fluid
DBS	Deep Brain Stimulation
dMRI	Diffusion Magnetic Resonance Imaging
DTI	Diffusion Tensor Imaging
FA	Fractional Anisotropy
FLIRT	FMRIB's Linear Image Registration Tool
fMRI	Functional Magnetic Resonance Imaging
FNIRT	FMRIB's Nonlinear Image Registration Tool
FSL	FMRIB Software Library
GLM	General linear model
GM	Grey matter
HCP	The Human Connectome Project
HRF	Hemodynamic Response Function
IQT	Image Quality Transfer
LE	Laplacian Eigenmap
M1	The primary motor cortex
MD	Mean diffusivity
MNI	Montreal Neurological Institute
MR	Magnetic Resonance
MRF	Markov Random Field

MRI . . . . . Magnetic Resonance Imaging

ROI . . . . . Region of Interest

SCP . . . . . Superior Cerebellar Peduncle

SCPCT . . . . Cerebellothalamic tract of Superior Cerebellar Peduncle

SNR . . . . . Signal-to-noise ratio

UKB . . . . . UK Biobank

V1 . . . . . The primary visual cortex

V2 . . . . . The secondary visual cortex

V3 . . . . . The third visual area

V4 . . . . . The forth visual area

Vim . . . . . Ventral Intermediate Nucleus of Thalamus

WM . . . . . White matter

1 Introduction

Contents

1.1 Background and Motivation

We all differ in how we perceive, think, and behave. The vast heterogeneity in human behaviour is often attributed to individual variability of the brain, the characteristics unique to individuals in terms of brain structure, function, and connectivity [1]. Individual variability of the brain manifests in various dimensions, including structural differences such as brain size and cortical thickness, and functional differences such as the organisation of neural pathways and macroscopic functional configurations.

Understanding this variability is important in several ways. First, it can help unravel the complexities of the human mind and identify the factors that contribute to a wide range of cognitive abilities, behaviours, and psychological traits. For example, variations in the information exchange efficiency between specific brain regions may underlie heterogeneity in cognitive abilities, such as memory, attention,

or problem-solving skills [2, 3]; additionally, differences in brain structure and function may influence an individual's susceptibility to psychological disorders [4, 5]. By investigating the factors contributing to these unique characteristics and abilities, researchers can gain valuable insights into the influence of neural architecture on various aspects of human experience. Second, this understanding can facilitate the development of personalised medicine and targeted interventions for neurological and psychological disorders, leading to more effective and precise treatments tailored to the unique needs of each patient [6]. An interesting example is the use of personalised transcranial magnetic stimulation (TMS) for the treatment of depression. Recent research has shown that identifying the optimal stimulation site for each patient, based on their individual brain connectivity patterns, can result in more effective treatment outcomes [7]. Such developments in personalised treatments will ultimately improve patient outcomes and overall well-being.

Historically, individual variability in the brain has been treated as noise or an inconvenient confound in basic and clinical neuroscience. This was primarily due to the focus on discovering general principles of brain organisation and function, leading researchers to aggregate data across individuals to identify common patterns. This type of group-level or group-average analysis has some advantages that have helped neuroscience research prosper when data quality and size has often been limited. Specifically, averaging data across subjects reduces sources of unwanted variability, such as head motion and physiological noise [8], and thus improves signal-to-noise ratio (SNR), making it easier to detect common patterns across individuals [9]. Furthermore, pooling data across individuals usually increases the statistical power of a study, making it more likely to detect true effects and improve the reliability of findings [10].

However, focusing on group-average statistics and ignoring individual differences has led to several important disadvantages. For example, neglecting individual variability in the brain and treating it as noise may result in conclusions that are not applicable to all individuals, ultimately limiting the generalisability of the findings [1, 11]; moreover, the differences between individuals may be erroneously attributed

to experimental manipulations or other factors in group-average inference, leading to inaccurate conclusions. More importantly, ignoring individual variability in the brain hinders understanding of the factors underlying disparities in cognitive abilities, personality traits, and vulnerability to psychological disorders [1]. A deeper understanding of these disparities is crucial to the development of personalised medicine and interventions tailored to each individual's unique biological and genetic fingerprints.

Thanks to the advancement of multimodal magnetic resonance imaging (MRI) techniques and emergence of large-scale neuroimaging consortium such as the Human Connectome Projects (HCP) and UK Biobank (UKB), researchers have been enabled to noninvasively examine the individual brain in closer detail with enhanced statistical power. In the last decade, numerous lines of research have emerged, utilising MRI of different modalities to study various aspects of individual brain variability. For example, high-resolution structural MRI has been employed to assess individual differences in grey matter volume, cortical thickness, and surface area, etc., providing insights into their relationships with a range of cognitive abilities, behaviour, and susceptibility to disorders [1, 12–16]. Functional MRI (fMRI) has been employed to examine individual variability in brain functional configurations during cognitive tasks and resting-state, elucidating the functional underpinnings of cognition and behaviour as well as their implications in brain disorders [3, 17–20]. Diffusion MRI has been employed to study matter microstructure and anatomical connectivity between brain regions, revealing their influence on cognitive performance and predisposition to neurological and psychiatric disorders [21–26]. Furthermore, multimodal MRI data has been integrated by more complicated modelling to reveal complex interactions between different aspects of brain structure/function and behavioural/genetic measures, shedding light on the mechanisms underlying individual variability in the brain and certain neurological disorders [27–30]. Longitudinal MRI studies at individual-level have also emerged to track changes in brain structure and function over time, offering valuable insights

into individual differences in brain development and ageing [16], as well as their associations with cognitive, disease, and emotional outcomes [31–33].

However, current research on individual brain variability faces several limitations, such as lack of methodology, strict requirements on data quality, heterogeneity in study designs and acquisition/preprocessing protocols, particularly in terms of specific "fine-grained" aspects of brain anatomy and function. One example is localising thalamic nuclei at an individual basis for personalised surgical targeting. While many studies have used diffusion MRI and tractography methods to localise thalamic nuclei at an individual basis [34–36], a well established and clinically reliable pipeline is still lacking, and significant methodological heterogeneity is present across these studies; in addition, localising deep brain nuclei often requires high-quality data not readily available. Another example is functional localisers that use resting-state fMRI to identify brain regions activated in response to an external stimuli or when the individual is executing a cognitive task. Although numerous studies have successfully predicted individual differences in task-evoked brain activity using resting-state fMRI [19, 37–40], they usually involved high-quality datasets such as HCP and did not investigate routine-quality data, which are more commonly available. Furthermore, these studies did not investigate more fine-grained functional profiles, e.g., the visuotopic configurations in the visual cortex.

Using MRI of various modalities, this thesis examines three aspects of individual variability in the brain that remain unaddressed or inadequately explored in previous studies, in response to the aforementioned limitations. The three aspects share a common theme of "localisation", emphasising methods development for localising specific brain structures or regions at the individual-level for personalised surgical targeting. First, we demonstrate a novel approach that reliably localises a Deep Brain Stimulation (DBS) target, which is not readily discernible on conventional MRI. Notably, the approach yields consistent results despite compromised data quality, demonstrating its potential for translating into a clinically reliable routine. Second, we present a framework that uses resting-state brain configurations to reconstruct individual task-evoked brain activation profiles. This framework is not

only effective on the high-quality HCP data but also applicable to routine-quality UKB subjects. The framework may serve as a functional localiser to supplement task localisers when task-fMRI is unavailable. Lastly, we investigate individual differences in the retinotopic organisation of visual areas using resting-state fMRI. While this variability holds important clinical implications, obtaining it typically requires dedicated retinotopy task paradigms and high-quality 7T MRI scanners, which are not always practical. Predicting the retinotopic mapping using resting-state fMRI instead offers a valuable alternative to retinotopy task-fMRI.

In conclusion, this thesis contributes to the understanding of individual variability in the brain by developing and validating novel MRI-based localisation methods. These approaches have the potential to enhance personalised treatment planning and further our understanding of the underlying neural mechanisms associated with brain variability.

1.2 Thesis Structure

The thesis is organised as follows. Chapter 1 introduces the motivation for this thesis and provides an outline of the subsequent chapters.

Chapter 2 presents an overview of MRI fundamentals and summarises the various MRI modalities employed in this thesis, including structural (T1- and T2-weighted) MRI, functional MRI (fMRI), and diffusion MRI.

Chapter 3 delves into individual anatomical variability. In this chapter, we propose an Image Quality Transfer (IQT) approach to reliably localise a DBS surgical target. Leveraging the publicly-available high-quality HCP dataset, this approach utilises anatomical information from high-quality data to enhance surgical targeting of the structure on clinical-standard lower-quality datasets, outperforming alternative targeting approaches in terms of accuracy and reliability.

Chapter 4 shifts the focus to the functional aspects of individual variability in the brain. In this chapter, we present a framework that uses resting-state spatial configurations to reconstruct profiles of task-evoked brain activity, based on the hypothesis that resting-state spatial processes form a basis set that spans

the subspace of a range of task-evoked activation maps. Specifically, we discuss the importance of "residualisation", a step that removed population-level common patterns from individual spatial maps, leading to improved predictions of individual-specific task activation patterns.

Chapter 5 is closely related to chapter 4 but emphasises the variability of visual processes. In this chapter, we investigate graph embedding methods for reconstructing retinotopic/visuotopic organisation of visual areas, using resting-state timeseries data. This chapter demonstrates that resting-state signals, despite the absence of visual stimuli, can capture individual differences in retinotopic mapping unexplained by population-level retinotopy.

Chapter 6 concludes the thesis by summarising its contributions and discussing the limitations of the presented approaches. Additionally, this chapter briefly explores potential future research directions.

1.3 List of Publications

Journal Articles

- **Ying-Qiu Zheng**, Seyedeh-Rezvan Farahibozorg, Weikang Gong, Hossein Rafipoor, Saad Jbabdi, and Stephen Smith. "Accurate predictions of individual differences in task-evoked brain activity from resting-state fMRI using a sparse ensemble learner." NeuroImage 259 (2022): 119418. (Contributions in Chapter 4.)

Peer Reviewed Conference Proceedings

- **Ying-Qiu Zheng**, Harith Akram, Stephen Smith, and Saad Jbabdi. "A Transfer Learning Approach to Localise a Deep Brain Stimulation Target." MICCAI 2033 accepted. (Contributions in Chapter 3.)

2

Magnetic Resonance Imaging (MRI) Basics

Contents

2.1 Introduction

Magnetic Resonance Imaging (MRI) is a non-invasive imaging technique that harnesses the physics of nuclear magnetic resonance to generate detailed images of the body's internal structures. The basic principle underlying MRI is the behaviour of protons, which is abundant nucleus in the human body, due to the high water content of tissues. When exposed to a strong external magnetic field, these nuclei align with the magnetic field and can be excited with a radio frequency pulse. As

they relax back to their original alignment, the released energy is detected by the MRI scanner and used to create images. MRI enables high-resolution visualisation of soft tissues and boasts a wide range of applications. Since its development in the late 1970s, MRI has become an indispensable tool in clinical diagnosis, medical research, and the study of brain structure and function [41–43]. Common MRI modalities for brain imaging include T1-weighted MRI, T2-weighted MRI, functional MRI (fMRI) and diffusion MRI (also known as dMRI). In this chapter, we will introduce the MRI modalities used in this thesis and briefly discuss their respective principles and applications.

2.2 Structural MRI

Structural MRI focuses on capturing high-resolution images of the brain's anatomy and morphology, such as the grey and white matter, as well as other tissue types. It is the most fundamental modality, serving as a critical tool in both clinical and research settings, allowing for accurate diagnosis and treatment planning, and furthering our understanding of the brain's structural organisation. The most common modalities of structural MRI are T1-weighted and T2-weighted imaging.

2.2.1 T1-weighted MRI

T1-weighted MRI utilises the differences in the longitudinal relaxation time (T1) of protons in various tissues to generate images. T1 relaxation represents the time for the longitudinal magnetisation (parallel to the main magnetic field) of protons to recover to approximately 63% of its initial value after the radio frequency pulse is switched off. In general, tissues with shorter T1 relaxation times appear brighter in T1-weighted MRI, while those with longer T1 relaxation times appear darker. In T1-weighted MRI, grey matter appears darker (lower signal intensity) than white matter, allowing for clear visualisation of the cortical and subcortical structures.

T1-weighted MRI has a wide range of applications in structural analysis of the brain, such as detecting structural abnormalities, assessing brain development, and guiding neurosurgical planning, etc. T1-weighted MRI serves as the basis for

advanced MRI analysis techniques, including voxel-based morphometry and cortical surface reconstruction, which are used to investigate structural changes associated with neurological and psychiatric disorders [44, 45]. Furthermore, T1-weighted MRI plays a vital role in both multi-modal MRI analysis and individual variability studies in neuroimaging. As the anatomical backbone in many neuroimaging studies, it provides a structural reference for aligning and integrating other imaging modalities, such as functional MRI (fMRI) and diffusion MRI. This co-registration process ensures spatial consistency, facilitating the integration of multi-modal data for a more comprehensive understanding of brain structure and function [46]. Simultaneously, T1-weighted MRI offers a common anatomical reference for comparing brain structure and organisation across individuals, allowing for inter-individual comparisons and group-level analyses [44, 47]. To summarise, T1-weighted MRI serves as a critical tool for understanding individual variability in brain structure and its relationship with cognition, behaviour, and clinical outcomes.

2.2.2 T2-weighted MRI

T2-weighted MRI is another type of structural MRI that relies on differences in T2 relaxation times of tissue to create image contrast [43]. T2 relaxation time represents the time for the transverse magnetization (perpendicular to the main magnetic field) of protons to decay to approximately 37% of its initial value. As a result, tissues with longer T2 relaxation times, such as cerebrospinal fluid (CSF), appear brighter, while those with shorter T2 relaxation times, like white matter, appear darker. This contrast is the opposite of what is observed in T1-weighted images, where tissues with shorter T1 relaxation times appear brighter.

T2-weighted MRI has several applications in clinical and research settings. Particularly, T2-weighted MRI has been used to study myelination in the brain, often in combination with T1-weighted MRI. Myelin is a fatty substance that wraps around axons and plays a crucial role in the conduction of nerve impulses. As T2-weighted MRI is sensitive to the water content in tissues, myelinated fibres appear with lower signal intensity (darker) than non-myelinated fibres due to the

lower water content in myelin. T1-weighted MRI also provides complementary information, as myelinated fibres have a higher signal intensity (brighter) than non-myelinated fibres in T1-weighted images. A commonly used approach to study myelination using MRI data is T1w/T2w ratio, which involves calculating the ratio of T1-weighted image intensities to T2-weighted image intensities [48]. This ratio has been shown to be a useful surrogate marker for myelin content in both grey and white matter regions, allowing researchers to investigate myelination patterns across the brain and examine their relationship with age, cognitive function, or neurological disorders [49]. Furthermore, T2-weighted MRI is particularly useful for detecting pathological changes in the brain, such as lesions, tumours, inflammation and demyelination [50] that may be less visible in T1-weighted images.

2.3 Diffusion-weighted MR imaging (diffusion MRI)

Diffusion-weighted MR imaging, also known as diffusion MRI or dMRI, is another specialised type of MRI that focuses on the diffusion of water molecules in different tissues. The principles of diffusion MRI rely on the fact that the diffusion of water molecules is influenced by the microstructure of the surrounding tissue. By applying magnetic field gradients in specific directions, diffusion MRI can detect the random motion of water molecules, providing insights into tissue organisation and structure.

In diffusion MRI, two important parameters are used to describe the acquisition: b-values (also known as bvals) and b-vectors (sometimes called bvecs or gradient directions). The b-values represents the strength and duration of the diffusion-sensitising gradients applied during diffusion MRI acquisition. A higher b-value results in a stronger diffusion weighting, making the images more sensitive to the water molecule diffusion but at the cost of lower signal-to-noise ratio. Common b-values used in clinical settings range from 0 to 1000 s/mm^2, while higher b-values (e.g., 2000-3000 s/mm^2) are more common in advanced research applications. B-vectors represent the direction of the applied diffusion-sensitising gradients. Through different b-values and b-vectors, the MRI scanner acquires images with varying sensitivities to diffusion in various orientations, resulting in volumes where the

signal intensity depends on the rate and direction of water molecule movement. Areas with restricted diffusion in the given direction will have higher signal intensity, while regions with more free diffusion will exhibit lower signal intensity.

Diffusion MRI has numerous applications, including stroke diagnosis [51], tumour characterisation [52], and the investigation of neurodegenerative diseases [53]. More importantly, diffusion MRI is the basis of white matter tractography.

2.3.1 Diffusion Tensor Imaging (DTI)

Diffusion Tensor Imaging (DTI) is a diffusion MRI analysis technique that provides information about the orientation, organisation, and integrity of the white matter fibres by analysing the principal diffusion directions and the degrees of diffusion anisotropy. DTI calculates the diffusion tensor, a mathematical representation of the diffusion process in a voxel, to account for the magnitude, orientation and anisotropy of water diffusion. This calculation is based on multiple diffusion MRI volumes with different diffusion gradients (b-values) and gradient directions (b-vectors). By fitting a diffusion tensor model to each voxel, DTI derives various parameters, such as fractional anisotropy (FA) and mean diffusivity (MD) for the given voxel. Such parameters have numerous applications in both clinical and research settings. For example, they can be used to assess white matter integrity in various neurological and psychiatric disorders [53] and investigate the effects of ageing and neurodevelopment [54, 55]. Furthermore, by identifying the primary direction of diffusion in each voxel (also known as the principal eigenvector), DTI allows for reconstruction of white matter fibre tracts in the brain via fibre tracking (also known as tractography). This information can be used to construct connectomes or structural brain networks, providing information on the anatomical connections between brain regions. In this thesis, DTI was estimated via FSL's DTIFIT tool (`https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT/UserGuide#DTIFIT`). However, most modern tractography methods, such as FSL's PROBTRACKX (`https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT/UserGuide#PROBTRACKX`), do

not over-reduce the data via the DTI tensor model, working instead from more advanced techniques, such as the Ball-and-Stick model.

2.3.2 The Ball-and-Stick model

As DTI models the diffusion properties of water molecules within a voxel using a single tensor, it is usually insufficient to resolve crossing fibres accurately. The ball-and-stick model is another diffusion MRI analysis technique that seeks to address the limitations of DTI in representing complex fibre configurations within a single voxel [56]. Instead, the ball-and-stick model represents the diffusion signals as a combination of multiple compartments, each consisting of a "stick" and a "ball". The "stick" represents the orientation of a major fibre population within a voxel, while the "ball" models the isotropic diffusion component. By using multiple sticks, the ball-and-stick model can accommodate multiple fibre orientations within a single voxel. It is more capable of resolving complex fibre configurations than DTI, leading to more accurate tractography results. In this thesis, the ball-and-stick model was estimated via FSL's BEDPOSTX tool (`https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT/UserGuide#BEDPOSTX`).

2.3.3 Tractography

Another major technique for diffusion MRI analysis used in this thesis is tractography, also known as fibre tracking or streamline tractography, which reconstructs the trajectories of white matter fibre tracts (or streamlines) in the brain using the local diffusion directions [**mori2002fibre**]. To form streamlines or fibre bundles, the fibre tracking algorithm iteratively advances from each seed point along the primary directions within each voxel, connecting neighbouring voxels with similar fibre orientations and finally generating continuous curves that represent the trajectories of the underlying fibre pathways. The process continues until a termination criterion is met, such as reaching a low anisotropy value, which may indicate the termination of a fibre tract, or exceeding a certain angular threshold between neighbouring voxels, which suggests a sharp bend in the tract that is anatomically implausible.

There are two primary types of tractography: deterministic and probabilistic. They differ in how they deal with the uncertainties associated with fibre orientations in diffusion MRI. Deterministic tractography assumes that the estimated fibre orientation in each voxel is accurate and does not account for uncertainty or multiple fibre bundles within a single voxel. Hence, it utilises a single principal orientation for each voxel and usually produces smooth and coherent streamlines. However, it is less effective in resolving complex fibre configurations, such as crossing fibres. Probabilistic tractography addresses the limitations of deterministic tractography by incorporating the uncertainties in fibre orientation estimates [57]. Instead of relying on a single principal orientation, probabilistic tractography typically utilises a distribution of possible fibre orientations for each voxel. At each step, the fibre tracking process samples multiple orientations from these distributions, generating multiple streamlines for each seed point. This process is repeated thousands of times, resulting in a large number of potential streamlines representing the underlying fibre tracts. Probabilistic tractography usually requires more complicated diffusion models, such as the ball-and-stick. Tractography has several applications in studies of individual variability. For example, it may reveal individual differences in the organisation and shape of white matter pathways, linking these image-derived phenotypes with age, cognitive abilities or susceptibilities to neurological disorders [55, 58, 59]. In this thesis, the white matter bundles were estimated using probabilistic tractography via FSL's PROBTRACTX tool (`https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT/UserGuide#PROBTRACKX_-_probabilistic_tracking_with_crossing_fibres`).

2.4 Functional MRI (fMRI)

Functional magnetic resonance imaging (fMRI) is another neuroimaging technique that measures neural activity by detecting changes in blood oxygenation and flow [60–62]. The basic signal underlying fMRI is the blood-oxygen-level-dependent (BOLD) contrast. When a brain region is active, the local neurons consume more oxygen, leading to an increase in blood flow to that region. The increased blood flow

brings in oxygenated hemoglobin, which has different magnetic properties compared to deoxygenated hemoglobin. As a result, the BOLD signal in the active region changes, providing an indirect measure of neural activity. In fMRI scans, a series of images are acquired over time, capturing the dynamic changes in BOLD signal across the entire brain. These images can be analysed to identify brain regions that show correlated activity during rest (resting-state fMRI) or specific tasks (task fMRI), providing insights into the functional organisations underlying cognitive processes.

2.4.1 Resting-state fMRI

Resting-state fMRI (also known as rs-fMRI) measures brain activity while the participant is not performing any specific task, usually with their eyes closed or fixating on a focal point on a blank screen. The primary goal of resting-state fMRI is to examine the intrinsic functional connectivity of the brain, i.e., the co-activation patterns between distinct brain regions that form functional configurations.

Over the years, resting-state fMRI has been widely used in various applications, such the investigation of brain functional organisation, characterisation of alterations in functional connectivity in clinical populations, and examination of the relationship between functional connectivity and individual differences in cognition, behaviour, and personality traits. Resting-state fMRI has been instrumental in identifying intrinsic functional configurations (or functional networks) in the brain, including the default mode network, salience network, and executive control network. These networks have been revealed to play critical roles in various cognitive functions and are altered in neurological and psychiatric disorders [63–66]. More recently, a growing body of studies has focused on the relationship between functional connectivity and individual differences in cognition, behaviour, and personality traits [17, 19].

2.4.2 Task fMRI

Different from resting-state fMRI, task fMRI measures brain activity in response to specific tasks or stimuli. In task fMRI, participants perform a cognitive or sensory task, such as motor movements, visual stimulation, or memory retrieval,

while their brain activity is being recorded. The primary goal of task fMRI is to investigate brain functions or identify the specific brain regions when subjects are performing a particular cognitive, sensory, or motor task. Analysing task fMRI usually involves estimation of task activation maps. Briefly, after initial pre-processing, the BOLD signal changes in task fMRI are modelled using the hemodynamic response function (HRF) to represent the apparent neural response to a task or stimulus [67]. The general linear model (GLM) is employed to fit the predicted BOLD response, derived from convolving the task design matrix with the HRF, to the observed fMRI time series data [68]. This generates beta coefficients as task fMRI activation maps, indicating the magnitude of the BOLD response for each task or stimuli. These resulting activation maps reveal brain regions engaged during the given tasks, revealing functional brain organisation and the neural basis underlying diverse cognitive processes.

Task fMRI activation maps have a wide range of applications in basic and clinical neuroscience. Traditionally, task fMRI activation maps are used in group-level studies to investigate the similarities across individuals in response to specific tasks. These studies reveal brain regions consistently engaged across individuals during a task, providing insights into the neural basis of particular cognitive or behavioural functions [65, 69, 70]. More recent task fMRI studies have focused on individual differences in their task-evoked activation profiles. These studies aim to address the relationship between individual variability in brain activation and variability in cognitive or behavioural performance, shedding light on the neural underpinnings of inter-individual differences in abilities and traits [1]. Furthermore, task fMRI activation maps have also been used to identify patterns of brain activation that are specific to certain disorders [71], demonstrating its potential to provide biomarkers for the diagnosis of various neurological and psychiatric conditions.

2.5 Conclusion

In this chapter, we provided an overview of the MRI modalities used in this thesis, including T1-weighted MRI, T2-weighted MRI, diffusion MRI, and resting-

state/task fMRI. We discussed the principles underlying each modality, along with their unique features and applications, particularly in the context of brain imaging and the study of individual variability. These modalities contribute to a comprehensive understanding of brain structure, function, and connectivity, by offering complementary insights into various aspects of the brain. In the following thesis, each chapter will delve into the specific MRI data and the corresponding analyses for each modality.

3

An Imaging Quality Transfer Technique to Localise Deep Brain Stimulation Targets

Contents

3.1 Introduction

Functional neurosurgical techniques, including deep brain stimulation (DBS) and MR-guided focused ultrasound, have been successfully used for several decades to treat a range of neurological and psychiatric disorders. These techniques target specific neural circuits that are implicated in the pathophysiology of these disorders, allowing for modulation of circuit activity and resulting in the alleviation of symptoms that cannot be managed by medications alone. For instance, the ventral intermediate nucleus of the thalamus (Vim) is a well-established surgical target in DBS and stereotactic ablation for tremor in Parkinson's Disease (PD), essential tremor, and multiple sclerosis [72, 73]. The Vim is a wedge-shaped nucleus located inferiorly within the ventrolateral (VLp) nucleus [74, 75], and is part of the motor thalamic nuclei that play a critical role in tremor circuitry. Classical tract-tracing and immuno-histochemical studies have revealed that the Vim receives efferent fibres from the dentate nucleus of the contralateral cerebellum via the superior cerebellar peduncle at the midbrain level [76], and projects primarily to the primary motor cortex (M1) [77] with minor projections to the supplementary motor area (SMA) and premotor cortex (PMC) [78]. Collectively, the tract connecting the contralateral dentate nucleus to M1 via the Vim in the thalamus forms the dentato-thalamo-cortical pathway (DTCp) and plays a central role in tremor circuitry [76, 79–83].

Accurately targeting the Vim has remained a significant challenge due to the lack of intrinsic contrast in conventional MRI sequences to distinguish the nucleus from neighbouring structures and tissues. Traditional stereotactic targeting of the structure has relied predominantly on standardised atlases/coordinates adapted to individual subjects, aided by visible anatomical landmarks such as the anterior commissure and posterior commissure points [84]. Although stereotactic atlases provide a reproducible way to identify the Vim, they fail to account for the inter-individual anatomical variations, which are not negligible in thalamic nuclei [85–88]. To ensure efficacy of stereotactic surgeries, patients are often required to stay awake in order to allow target confirmation, causing great patient discomfort and potential risks of intracranial haemorrhage leading to neurological deficit or even death [89].

To address this limitation, recent studies have attempted to more accurately identify the Vim *in vivo* by leveraging its anatomical properties [34–36]. Using cutting-edge diffusion MRI (dMRI) techniques and tractography-based methods, these studies reconstruct the DTCp and identify Vim by locating the region of maximum connectivity with both ipsilateral M1 and the contralateral dentate. Diffusion MRI enables the estimation of local fibre orientations, while tractography algorithms generate streamline samples that represent the underlying white matter pathway based on these local fibre orientations. Often referred to as "connectivity-driven ", this approach can better explain the inter-individual anatomical variability of the nucleus, leading to improved efficacy of neurosurgical procedures [34]. However, reconstructing the DTCp usually requires state-of-the-art high angular resolution diffusion imaging (HARDI) and advanced diffusion modelling techniques in order to resolve complex fibre configurations (e.g., crossing fibres), which are impractical in advanced-care clinical settings due to their prolonged acquisition time and higher acquisition/computational cost. As a result, lower-quality diffusion MRI techniques are still commonly used, which affects the reliability of the connectivity-driven approach as a proxy for surgical targeting. Furthermore, even with cutting-edge HARDI and higher-order diffusion modelling techniques, the Vim localised by the connectivity-driven approach has exhibited significant variations across different acquisition protocols (e.g., b-values, spatial and angular resolutions) and processing pipelines (e.g., diffusion signal modelling and tractography algorithm parameters) [36]. Therefore, these methods must be used with caution to ensure that they accurately reflect the true underlying anatomical variability rather than reflecting methodological confounds erroneously interpreted as variability [36].

Accurate localisation of the target nucleus is crucial for the efficacy of stereotactic operations. However, existing surgical targeting approaches have not yet been able to translate into a reliable clinical routine. To address this challenge, we propose a novel Image Quality Transfer (IQT) technique, HQ-augmentation, for reliably localising Vim, particularly on low-quality clinical-like data. IQT is a computational technique in medical imaging that aims to enhance the quality of

low-resolution or low-quality images by transferring information from a set of high-quality reference images [90–93]. This is usually achieved by using machine learning algorithms to learn the mapping between low-quality images and their high-quality counterparts. Following this concept, the HQ-augmentation approach proposed here leverages anatomical information from publicly-available HARDI datasets, such as the Human Connectome Project (HCP), to guide Vim localisation in low-quality diffusion MRI datasets. It also exploits connectivity features with a wide range of brain regions to compensate for the compromised M1 and dentate white matter connectivity features in low-quality diffusion MRI (Figure 3.1). Specifically, we generate approximate locations of the Vim from high-quality diffusion MRI datasets and train the HQ-augmentation model to locate the nucleus in low-quality data that is most equivalent to the counterpart high-quality estimate, given its low-quality anatomical connectivity profiles with a wide range of brain regions. Our results demonstrate that the HQ-augmentation model outperforms existing alternatives, i.e., the atlas-defined approach and the connectivity-driven approach, on surrogate low-quality diffusion MRI data with much lower spatial and angular resolution, suggesting that the HQ-augmentation model is not only capable of accounting for the true anatomical variability but is also robust to the impact of data quality and methodological heterogeneity. Furthermore, the HQ-augmentation model can generalise to unseen low-quality diffusion MRI datasets collected with different acquisition protocols, such as the UK Biobank (UKB) dataset.

Finally, the HQ-augmentation model is not limited to targeting the Vim. Depending on specific symptoms, a range of deep thalamic nuclei have been revealed as effective surgical targets. Within the thalamus, for example, the anterior nucleus has been used as a DBS target to treat epilepsy [94–97]; the medial dorsal nucleus is another DBS target to treat obssessive-compulsive disorder and major depressive disorder [98]. Outside of the thalamus, subthalamic nucleus (STN) [99–101] and globus pallidus internus (GPi) [102–104] are common DBS targets for the treatment of PD; the hypothalamus has been explored for the treatment of cluster headaches, obesity, and other conditions. Our approach can be adapted to many other deep

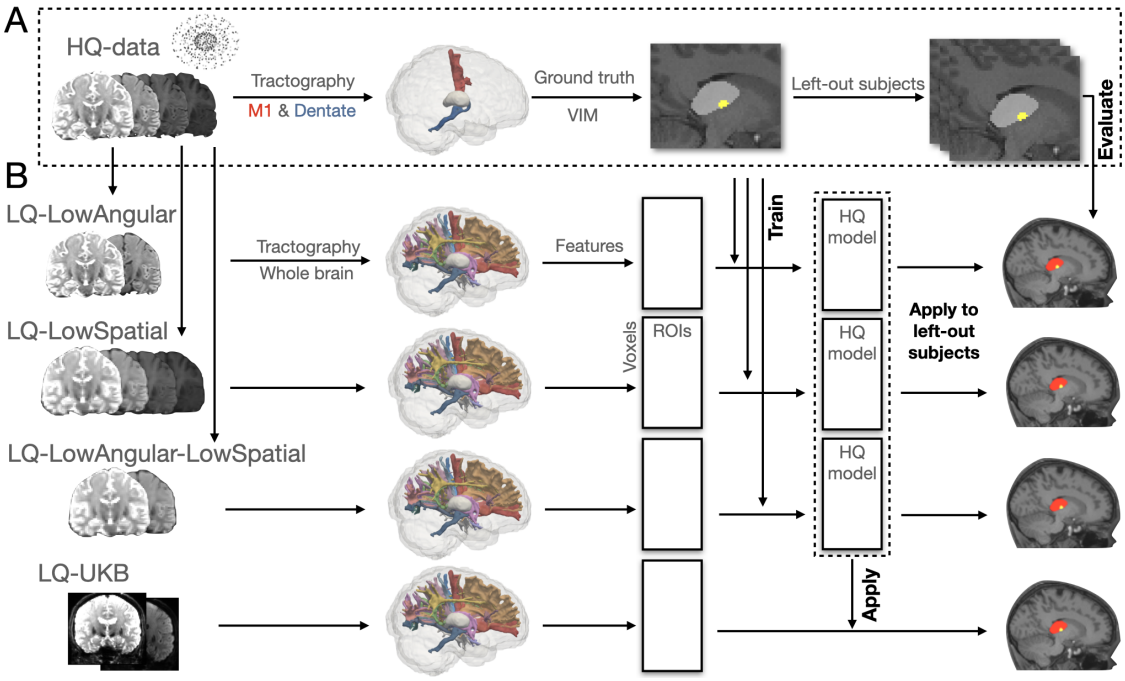

Figure 3.1: Illustration of the HQ-augmentation approach. **(A)** First, the HQ diffusion data were used to create "ground truth" Vim within the thalamic masks, based on its white matter connectivity with M1 and contralateral cerebellum. **(B)** Next, surrogate low-quality diffusion datasets were created by corrupting the HQ datasets. The HQ-augmentation model was trained with the HCP HQ "ground truth" Vim as target labels and an extended set of HCP low-quality connectivity profiles as input features. The trained model was applied to unseen low-quality datasets (surrogate low-quality HCP and UK Biobank diffusion data) to give Vim prediction, and evaluated against the corresponding HQ "ground truth".

brain nuclei, using their literature-based white matter connectivity profiles. Overall, the HQ-augmentation serves as a better proxy for surgical targeting and has the potential to translate into a reliable clinical tool.

3.2 Materials and Methods

3.2.1 Datasets and subjects

3.2.1.1 HCP 3T minimally preprocessed MRI

We used 3T diffusion MRI data from the Human Connectome Project (HCP) [105, 106] as the high-quality dataset. The data were collected by the WU-Minn HCP consortium from a group of healthy young adults (n = 1,062, age range 22-36 years) who consented to participate under the approval of the Washington University in

St Louis Institutional Review Board [105]. The minimally pre-processed T1-, T2-, and diffusion-weighted MRI scans were obtained from the HCP 2017 S1200 release (`https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release`), which included 43 subjects who underwent repeated scanning. T1- and T2-weighted images were acquired with a customised Siemens 3T "Connectome Skyra" scanner at an isotropic spatial resolution of 0.7mm. Diffusion MRI data were collected with a monopolar diffusion-weighted (Stejskal-Tanner) spin-echo EPI sequence using the Siemens 3T "Connectome Skyra" scanner at an isotropic spatial resolution of 1.25mm. The acquisition included three shells (b-values = 1000, 2000, and 3000 s/mm^2) and approximately 90 unique diffusion directions per shell, acquired twice (total scan time 60 min per subject) [106]. As part of the HCP S1200 release, the data were minimally pre-processed [107] and aligned across modalities for each subject. Motion, susceptibility, and eddy current distortions in the diffusion MRI data were corrected [108, 109]. The data were linearly registered to the corresponding T1-weighted images using FSL's FLIRT [110, 111]. Nonlinear transformations to the MNI152 standard space and their inverse warps were obtained using FSL's FNIRT [111, 112]. The complete imaging protocols and pre-processing procedures can be found at `https://www.humanconnectome.org/storage/app/media/documentation/s1200/HCP_S1200_Release_Reference_Manual.pdf`.

3.2.1.2 UK Biobank 3T minimally pre-processed MRI

We also used 3T MRI data from UK Biobank [113] in our analysis, as the overall quality of this data is more representative of what is typically acquired on clinical scanners. Our analysis included a total of 2,560 subjects who had received repeat scanning. T1-weighted images were acquired using a 3D MPRAGE acquisition at 1mm isotropic resolution, while T2-weighted images were acquired using a 3D SPACE sequence at a spatial resolution of 1.05x1x1 mm. Diffusion MRI was performed at isotropic spatial resolution 2mm with two shells (b-values = 1000 and 2000 s/mm^2), and 50 diffusion directions per shell (total scan time of approximately 6 minutes per subject). Additional information on the imaging protocols can be found

in `https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf`. Pre-processing of the UK Biobank data included face removal, brain extraction, and registration across modalities and to the MNI152 standard space [114]. The diffusion MRI images were also minimally pre-processed, corrected for motion, susceptibility and eddy current distortions. [108, 109]. It is important to note that the uncorrupted UK Biobank 3T diffusion MRI still served as a high-quality dataset. In the following analysis, we derived a surrogate version of the low-quality dataset from the UK Biobank dataset.

3.2.2 Low-quality diffusion MRI datasets

To account for the varying data quality typically encountered in clinical contexts, we considered a range of low-quality datasets, including:

1. surrogate HCP low-quality diffusion MRI, obtained by corrupting the original (minimally pre-processed) high-quality HCP 3T diffusion MRI. Three different forms of data corruption were explored: a dataset with reduced angular resolution, referred to as *LQ-LowAngular*; a dataset with decreased spatial resolution, denoted *LQ-LowSpatial*; and a dataset with reductions in both angular and spatial resolution, named *LQ-LowAngular-LowSpatial*.

2. surrogate UKB low-quality diffusion MRI, obtained by corrupting the original (minimally pre-processed) UKB 3T diffusion MRI, which we refer to as *LQ-UKB*;

The HQ-augmentation model was trained using paired high- and low-quality HCP data, specifically, the original HCP 3T diffusion MRI and its various lower-quality counterparts (LQ-LowAngular, LQ-LowSpatial, or LQ-LowAngular-LowSpatial). The paired high- and low-quality datasets provide examples for the model to learn the relationship between the low-quality image features and their corresponding high-quality counterparts, enabling the model to extract and transfer high-quality anatomical information to enhance the target low-quality data. As a result, the trained model is tailored to the specific type of low-quality dataset on which it is trained.

Validation and evaluation were carried out on both HCP and UKB datasets. The Vim derived from low-quality data was assessed against two versions of "ground truth": the connectivity-driven Vim obtained from the high-quality data (referred to as HQ-Vim in the following text), and the atlas-defined Vim adapted to individual native space from the template Vim.

3.2.2.1 LQ-LowAngular

The LQ-LowAngular dataset was designed to resemble properties of single-shell diffusion MRI collected with fewer diffusion directions. To create this dataset, we discarded all volumes corresponding to bvals = 2000 and 3000 s/mm^2 and only considered the single-shell at bvals = 1000 s/mm^2 along with the b0 volumes. We then sampled 32 directions uniformly on a sphere using FSL's GPS tool (`https://git.fmrib.ox.ac.uk/fsl/gps/-/tree/master/`). For each of the sampled directions, we calculated its dot-product with the original 90 directions (at bvals = 1000 s/mm^2) and found its closest equivalent direction (i.e., maximum dot-product) from the actual 90 bvecs, resulting in a total of 32 single-shell volumes selected from the original multi-shell. Finally, we combined the selected 32 single-shell volumes with the bvals = 0 volumes to form the LQ-LowAngular dataset.

3.2.2.2 LQ-LowSpatial

When scanning time is limited, it is possible to achieve better angular resolution by compromising on the spatial resolution for diffusion MRI. The LQ-LowSpatial dataset was designed to imitate this scenario, where the spatial resolution was compromised to achieve higher angular resolution. We downsized the (minimally pre-processed) diffusion data from an isotropic spatial resolution of 1.25mm to 2mm while keeping the original shells and bvals/bvecs, resulting in a surrogate low-spatial-resolution dataset.

3.2.2.3 LQ-LowAngular-LowSpatial

We also created the LQ-LowAngular-LowSpatial dataset, a surrogate low-angular and low-spatial resolution diffusion dataset, to reflect the more extreme poor data

quality when advanced MRI imaging techniques are lacking (or time available for scanning is very short). This dataset was generated by further downsampling the LQ-LowAngular dataset to an isotropic spatial resolution of 2mm. Alternatively, it can also be created by extracting the b0 and bvals = 1000 s/mm^2 single-shell volumes (at the same 32 directions) from the LQ-LowSpatial dataset.

3.2.2.4 LQ-UKB

The UK Biobank diffusion MRI has a spatial resolution of 2x2x2 mm, which is closer to what is typically acquired on clinical scanners. Hence, we did not further degrade this spatial resolution. Instead, we only created its single-shell (low-angular-resolution) counterparts by extracting the b0 and bvals=1000 s/mm^2 volumes, sampled at 32 diffusion directions. The resulting MRI images served as a surrogate UKB low-angular-resolution diffusion MRI dataset, here denoted as LQ-UKB. Similar to the creation of LQ-LowAngular from HCP, the 32 directions of LQ-UKB were selected by first sampling 32 bvecs uniformly on a sphere, and next finding their closest matches respectively from the actual 50 bvecs at bvals = 1000 s/mm^2.

3.2.3 Structural and diffusion data post-processing

3.2.3.1 Segmentation of thalamic masks

Thalamic masks for the left and right hemispheres were segmented for each subject in the HCP and UK Biobank datasets using FSL's MIST (`https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MIST`), a sub-cortical segmentation tool that leverages complementary information from different MRI modalities to achieve accurate segmentation [115]. Although the HCP preprocessed data already included masks for the thalamus, these were not accurate enough and included too much white matter, as they were based on the T1 image only.

Three modalities were used for thalamus segmentation in MIST: T1, T2, and fractional anisotropy (FA) contrast, which was estimated using the diffusion tensor model [116] and FSL's DTIFIT (`https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT/UserGuide#DTIFIT`). To match the spatial resolution of the T1- and

T2-weighted images, the FA images were upsampled from an isotropic resolution of 1.25 mm to 1 mm, while the T1- and T2-weighted images were downsampled from an isotropic resolution of 0.7 mm to 1 mm. The resulting thalamic masks were utilised as seeds regions in tractography to construct white matter tracts projecting from the thalamus.

3.2.3.2 Creation of other anatomical masks

To constrain tractography seeded from the thalamus, we employed various regions-of-interest (ROIs) spanning the brain. These ROIs included:

1. 75 cortical regions and 2 cerebellar parcels (one for cerebellar gray matter and the other for cerebellar white matter) per hemisphere, and the brainstem, segmented using Freesurfer [117–120], released as part of the HCP preprocessed data.

2. 4 ROIs derived from the cerebellothalamic tract of the superior cerebellar peduncle (SCPCT) [121] warped into the individual space, which includes 3 white matter segments situated between the superior cerebellar peduncle (SCP) and the ipsilateral thalamus, as well as an additional parcel overlapping with the brainstem to account for the cerebellothalamic tract decussation. These ROIs were referred to as SCPCT-1, SCPCT-2, SCPCT-3, and SCPCT-brainstem.

3. 29 white matter segments extracted from the major white matter bundles projecting from or passing through the thalamus using the XTRACT atlas [122], including 6 segments from the Superior Thalamic Radiation (STR), 5 from the Acoustic Radiation (AR), 6 from the Anterior Thalamic Radiation (ATR), 6 from the Optic Radiation (OR), and 6 from the Fornix (FX). These white matter ROIs were extracted in the standard space and subsequently warped back into the individual space.

4. 5 white matter ROIs extracted from the fibre bundle joining thalamus and ipsilateral M1. This fiber bundle was created with thalamic voxels as the seed and M1 as both the waypoint and target mask, subsequently warped into the standard space and averaged across subjects. These white matter ROIs were extracted in the standard space and subsequently warped into the individual space. These ROIs were referred to as M1-1, M1-2, M1-3, M1-4, and M1-5.

5. Another 5 white matter segments extracted from the fibre bundle joining thalamus and ipsilateral primary sensory cortex (S1). This fiber bundle was created similarly with thalamic voxels as the seed and S1 as both the waypoint and target mask, subsequently warped into the standard space and averaged across subjects. These white matter ROIs were also extracted in the standard space and subsequently warped into the individual space. These ROIs were referred to as S1-1, S1-2, S1-3, S1-4, and S1-5.

Table A1 provides a summary of the ROIs and details on how they were created.

3.2.3.3 Fibre orientation estimation

Prior to tractography, fibre orientations were estimated for all diffusion MRI data by applying a parametric spherical deconvolution model through FSL's BedpostX (`https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT/UserGuide#BEDPOSTX`). In each voxel, up to three fibre orientations were estimated, along with their respective uncertainties [56, 86, 123, 124]. Two distinct deconvolution models were used to fit the crossing fibres in the multi-shell (HQ-HCP, HQ-UKB, and LQ-LowSpatial) and single-shell (LQ-LowAngular, LQ-LowAngular-LowSpatial, LQ-UKB) diffusion data. The former utilised a Ball-and-Sticks with zeppelins model [123, 124], while the latter used the standard Ball-and-Sticks model [56, 86]. The former approach models the diffusion coefficient using a Gamma distribution, enabling it to more effectively represent multi-shell diffusion MRI data. In contrast, the latter method models the diffusion coefficient with a single value, making it more appropriate for single-shell data.

3.2.3.4 Tractography protocols

Probabilistic tractography was performed using FSL's Probtrackx tool [56, 86, 125] in each dataset's respective individual T1 space (also referred to as native space). For the subjects that underwent repeat scans, the first-visit T1 scans served as the reference native space, to which the second-visit scans were registered. Since all the anatomical masks were in the (first-visit) T1 native space, an affine transformation matrix between the subject's T1 and diffusion space was employed during fibre tracking, allowing the reconstructed streamline distributions to be directly resampled into the respective T1 space.

The anatomical masks included seeds (starting points of the streamlines), waypoints/targets (regions that streamlines must pass through to be valid), exclusion masks (regions that reject streamlines passing through them), and termination masks (regions that serve to stop streamlines running through them). Streamlines were seeded from the thalamic masks using a modified Euler integration with a step length of 0.5mm and 2000 steps per streamline, randomly initialised within a 1mm-radius sphere around each voxel centre. Streamlines were terminated if they entered the exclusion masks, or if the cosine of the angle between two steps exceeded 0.2.

A total of 5000 individual streamlines were drawn for each seed voxel and discarded if they reached the exclusion masks or did not meet the waypoint condition. The choice of waypoints/termination/exclusion masks varied depending on the target masks used in each run of tractography, with additional details available in subsequent sections. The output streamline distributions were corrected for the distance between the target and the seed mask, as the number of streamlines typically decreases with distance from the seed mask. A summary of the tractography options used in this study is provided in Table A2.

3.2.4 Connectivity-driven approach

The connectivity-driven approach used in our study, mirroring from Akram et al. [34] and Bertino et al. [36], resulted in what we refer to as "connectivity-driven" Vim in the subsequent text. This approach identifies the Vim by finding the

maximum probability of connection to M1 and contralateral dentate nucleus within the thalamic mask, generated via probabilistic tractography.

Streamlines seeded from the thalamus, targeting M1 and contralateral cerebellum, were generated via probabilistic tractography, yielding two "tract-density" feature maps for each target. These maps represent the white matter connectivity strength with their corresponding targets. To mitigate bias caused by differences in target volumes, the probability maps were normalised and then thresholded to discard voxels characterised by low tract density (M1 at 50% of the tract density values within the thalamic mask; dentate at 15%). The thresholded M1 map was subsequently multiplied with the thresholded cerebellum map and binarised at a lower 30% threshold, resulting in the connectivity-driven Vim. This process was repeated for each hemisphere and subject for all diffusion MRI scans.

Note that the connectivity-driven approach was either applied to the high-quality data to generate HQ-Vim, serving as approximations of the ground truth location of the Vim, or to the low-quality data to produce low-quality connectivity-driven Vim. These low-quality connectivity-driven Vim segmentations were then compared with alternative approaches on low-quality data.

3.2.5 Atlas-defined approach

The atlas-defined approach identifies the Vim based on the group-average location of the nucleus. To create a group-average Vim atlas, we transformed the binarised connectivity-driven Vim (derived from the high-quality data) to the MNI152 1mm standard space [126] and averaged these across the training subjects to obtain a group-average Vim probability map. The transformation fields were obtained by non-linearly registering the individual T1 to the same standard space. The voxel-wise probabilities in this Vim probability map represent the proportion of subjects that overlap at the given voxel. The group-average Vim probability map was subsequently warped back into the individual T1 space and thresholded at the 50% percentile to define the "atlas-defined" Vim.

It is important to note that we only included a "reliable" subset of subjects to calculate the group-average Vim probability map. As the connectivity-driven approach has limited reliability even with high-quality data, we only included subjects with "trustworthy" high-quality connectivity-driven Vim (HQ-Vim) for calculating the group-average Vim probability map, while discarding subjects whose HQ-Vim were considered unreliable. An HQ-Vim must meet four criteria to be considered reliable: 1) Its volume exceeds 20 mm^3. Given that the size of the Vim is approximately 4x4x6 mm [127], a resulting volume that is too small may suggest unreliable segmentation. 2) It contains only one connected component. This criterion aligns with the anatomical reality, given that the Vim structure is a single connected nucleus. 3) Its center-of-mass is located within 4 mm of the center-of-mass of the Vim cluster in the Thalamic DBS Connectivity Atlas [34], when referenced to native space. 4) Its correlation with the Thalamic DBS Connectivity Vim Atlas exceeds 0.5, again in native space. Criteria 3) and 4) are included to exclude HQ-Vim that deviate excessively from the atlas, thereby being considered unreliable. Nonetheless, these criteria allow for a certain degree of individual variability to be factored in. See Figure A1 for more details on the selection of the "reliable" subset.

3.2.6 HQ-augmented approach

The goal of this approach is to leverage anatomical information in HQ data to infer the likelihood of a voxel belonging to the Vim, given a wide range of tract-density maps (multiple distinct tract bundles) derived from low-quality data as input features. The HQ-augmentation model was trained on the HCP dataset for each type of low-quality dataset (i.e., LQ-LowAngular, LQ-LowSpatial, LQ-LowAngular-LowSpatial). Using the HCP HQ data, we first generated the connectivity-driven Vim (referred to as HQ-Vim) as the "ground truth" location of the nucleus, serving as training labels in the model. Next, for each low-quality counterpart, we generated an extended set of tract-density features, targeting a wide range of ROIs, as the input features of the model. We hypothesise that the richer set of connectivity features serves to compensate for the primary tract-density features (with M1 and

dentate), when those are compromised by insufficient spatial or angular resolution in low-quality diffusion MRI, thus making Vim identification less reliant on the tract-density features used in the connectivity-driven approach, and more robust to variations in data quality. During training, the model learns to use the extended set of low-quality connectivity features to identify the Vim that is closest to the one that can be otherwise obtained from its HQ counterpart.

3.2.6.1 Input features

The extended set of target ROIs comprised: 1) 75 ipsilateral cortical regions and the contralateral cerebellar white matter, derived from the Destrieux atlas [120]; 2) 29 white matter ROIs extracted from the XTRACT atlas [122]; 3) SCPCT-1, SCPCT-2, SCPCT-3, and SCPCT-brainstem, i.e., the 4 ROIs extracted from the cerebellothalamic tract of superior cerebellar peduncle (SCPCT) [121], among which three are white matter segments lying between the thalamus and brainstem (SCPCT-1, SCPCT-2, and SCPCT-3), one overlapping with the brainstem (SCPCT-brainstem); 4) M1-1, M1-2, M1-3, M1-4, and M1-5, i.e., the white matter ROIs joining thalamus with ipsilateral M1; 5) S1-1, S1-2, S1-3, S1-4, and S1-5, i.e., the white matter ROIs joining thalamus with ipsilateral S1.

For each target, we reconstructed streamlines with the thalamus as the seed mask and the selected target as the waypoint mask, producing a tract-density map within the thalamic voxels. The ipsilateral cerebellum and CSF were designated as exclusion masks, whilst the cortex served as termination regions. The streamlines were constrained to be ipsilateral, with the exception of the four SCPCT ROIs and the contralateral cerebellum serving as targets. The resulting 116 tract-density maps were normalised by the maximum density value to correct for overestimation/underestimation biased by target volumes. Finally, six M1-related tract-density maps (targeting at M1 and five white matter segments from M1-1 to M1-5) were each multiplied by the following four tract-density maps (targeting the four SCPCT ROIs and the contralateral cerebellum) respectively, resulting in 6×5 product maps as additional connectivity features. These interaction

features were designed to account for the intersection of the DTCp within the thalamic mask. The whole procedure resulted in a total of 123 features to be fed into the HQ-augmented model. Note that the input features were consistently located within the individual's T1 space. This held true irrespective of the actual spatial resolution of the low-quality diffusion MRI. During the process of fibre tracking, the features were systematically resampled to align with the individual's T1 space. This approach ensured uniform spatial correspondence across all datasets, irrespective of their original resolution.

3.2.6.2 Model setup and implementation

Here we used a Conditional Random Field (CRF) [128] to learn the mapping between the HQ-Vim (training labels) and the low-quality connectivity features (input features), by maximising the likelihood of having the same HQ-Vim label assignments given the set of low-quality connectivity features. CRFs are an extension of Markov Random Fields (MRFs) [129], a technique that models the joint probability distribution of random variables, while taking into account the dependencies between neighbouring variables. While the primary focus of MRFs is to model the joint distribution of the variables, CRFs, on the other hand, are a type of discriminative model and focus on modelling the conditional probability distribution of target variables given the observed input variables, without accounting for the joint distribution of all variables in the system explicitly. Similar to MRFs, CRFs consist of two compartments, one designed to model the likelihood of assigning a label to a given voxel without considering its neighbourhood, the other to account for the fact that neighbouring voxels are more likely than not to have the same label. Specifically, assume $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_V]^T$ is a $V \times d$ connectivity feature matrix for a given subject, where $\mathbf{x}_i$ is a $d \times 1$ vector representing the connectivity features in voxel i, V is the total number of voxels of the thalamus (per hemisphere) for this subject; $\mathbf{y} = [y_1, y_2, ...y_V]^T$ is a $V \times 1$ vector containing the HQ-Vim labels for the V voxels, in which y_i is the label of voxel i. Given the low-quality features $\mathbf{X}$,

we seek to maximise the probability of reproducing the exact same HQ-Vim label assignment $\mathbf{y}$ on its low-quality counterparts, across the training subjects

$$\log P(\mathbf{y}|\mathbf{X}) = \log[\frac{1}{Z(\mathbf{X})}\exp(-E(\mathbf{y}|\mathbf{X}))] \tag{3.1}$$

Here $E(\mathbf{y}|\mathbf{X})$ is the cost of the label assignment $\mathbf{y}$ given the features $\mathbf{X}$, whilst $Z(\mathbf{X})$ serves as an image-dependent normalising term. Maximising the posterior $P(\mathbf{y}|\mathbf{X})$ across subjects is equivalent to minimising the cost of the label assignment $\mathbf{y}$ given the features $\mathbf{X}$. Suppose $\mathcal{N}_i$ is the set of voxels neighbouring voxel i, the cost $E(\mathbf{y}|\mathbf{X})$ is modelled as

$$E(\mathbf{y}|\mathbf{X}) = \sum_i \psi_u(y_i|\mathbf{x}_i) + \sum_i \sum_{j\in\mathcal{N}_i} \psi_p(y_i, y_j|\mathbf{x}_i, \mathbf{x}_j) + \lambda_1||\mathbf{W}||_1 + \lambda_2||\mathbf{W}||_2^2 \tag{3.2}$$

The first component $\psi_u(y_i)$ measures the cost (or inverse likelihood) of voxel i taking label y_i. Here $\psi_u(y_i)$ takes the form $\psi_u(y_i) = \mathbf{w}_{y_i}^T\phi(\mathbf{x}_i)$, where $\phi(\cdot)$ maps a feature vector $\mathbf{x}_i = [x_1, x_2, ...x_d]$ to a further expanded feature space in order to provide more flexibility for the parameterisation. $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2]$ is the coefficient matrix to be learned from the data, each column containing the coefficients for the given class (i.e., belonging to the HQ-Vim or not). Here we chose a series of polynomials along with the group-average Vim probability (registered into native space) to expand the feature space, i.e.,

$$\phi(\mathbf{x}_i) = [x_1, x_2, ...x_d, x_1^{p_1}, x_2^{p_1}, ...x_d^{p_1}, x_1^{p_2}, x_2^{p_2}, ...x_d^{p_2}, x_1^{p_3}, x_2^{p_3}, ...x_d^{p_3}, g_i]$$

where $p_1 = 2, p_2 = 0.5, p_3 = 0.2$ are the power of the polynomials (chosen by testing a range of power values on an independent subset, see Figure A2), and g_i is the group-average probability of voxel i classified as Vim (calculated across the training subjects). The second pairwise cost encourages assigning similar labels to neighbouring voxels, particularly for those sharing similar connectivity features. We modelled this component as $\psi_p(\mathbf{x}_i, \mathbf{x}_j) = \mu(y_i, y_j)\rho k(\phi(\mathbf{x}_i), \phi(\mathbf{x}_j))$. Here $k(\phi(\mathbf{x}_i), \phi(\mathbf{x}_j)) = \exp(-\gamma||\phi(\mathbf{x}_i) - \phi(\mathbf{x}_j)||^2)$ is a kernel function modelling the similarity between voxel i and j in the extended feature space, with length

scale γ, chosen via cross-validation. $\mu(\cdot)$ is a label compatibility function where $\mu(y_i, y_j) = 0$ if $y_i = y_j$ or $\mu(y_i, y_j) = 1$ if $y_i \neq y_j$.

Therefore, in a local neighbourhood, the kernel function penalises inconsistent label assignment of voxels that have similar features, thus allowing modelling local smoothness. ρ controls the relative strength of this pairwise cost weighted by $k(\cdot)$. Lastly, the $L1$ and $L2$ penalty terms serve to prevent overfitting of the model. We used a mean-field algorithm to iteratively approximate the maximum posterior $P(\mathbf{y}|\mathbf{X})$ [128] summed across the subjects. The approximated posterior is maximised via gradient descent in a mini-batch form, where the connectivity feature matrix of each subject serves as a mini-batch, is demeaned and normalised, and sequentially fed into the optimisation problem.

Specifically, we used cross entropy loss to maximise the log-likelihood in Equation (A1). Specifically, suppose $\mathbf{T}$ is the one-hot coding matrix of the HQ-Vim labels $\mathbf{y}$, i.e., y_i is mapped to a binary vector $\mathbf{t}_i = [t_{i1}, ...t_{iK}]$, in which $y_i = k$ corresponds to $t_{ik} = 1$ and other elements in $\mathbf{t}_i$ set to 0. To maximise the log-likelihood in (A1) is equivalent to minimising the cross entropy (negative log likelihood):

$$\min \sum_{i}^{V} \sum_{j}^{K} t_{ij} \log(P(y_i|\mathbf{x}_i)) \tag{3.3}$$

Due to the inter-dependency of neighbouring voxels, the exact minimisation is intractable. Thus, we approximate the CRF distribution $P(\mathbf{y}|\mathbf{X})$ by a simpler function $Q(\mathbf{y})$, and iteratively solve this optimisation problem, explained below.

To set up initialisation, we derived the initial coefficients $\hat{\mathbf{W}}^{(0)} = [\hat{\mathbf{w}}_1^{(0)}, ...\hat{\mathbf{w}}_K^{(0)}]$ by optimising the following

$$\hat{\mathbf{W}}^{(0)} = \arg\min_{\mathbf{W}} \sum_{i}^{V} \sum_{j}^{K} t_{ij} \log(P'(y_i|\mathbf{x}_i)) \tag{3.4}$$

where $P'(y_i|\mathbf{x}_i)$ is the likelihood without considering local smoothness of the label assignment, i.e., ignoring the pairwise loss term in Equation (A2):

$$P'(y_i = l|\mathbf{x}_i) = \frac{\exp(-\psi_u(y_i = l|\mathbf{x}_i))}{\sum_k^K \exp(-\psi_u(y_i = k|\mathbf{x}_i))} \tag{3.5}$$

where, as defined before, $\psi_u(y_i = l|\mathbf{x}_i) = \mathbf{w}_l^T\phi(\mathbf{x}_i)$. The coefficients $\hat{\mathbf{W}}^{(0)}$ were used to initialise $Q(\mathbf{y})$ using Equation (A5), i.e., $Q_i(y_i = l) \leftarrow P'(y_i = l|\mathbf{x}_i)$ evaluated at $\hat{\mathbf{W}}^{(0)}$.

After initialisation, $Q_i(y_i = l)$ is updated as the weighted sum of its neighbouring Q values, $\tilde{Q}_i(y_i = l) \leftarrow \rho\sum_{j\in\mathcal{N}_i} k(\phi(\mathbf{x}_i), \phi(\mathbf{x}_j))Q_j(y_j = l)$, where $\tilde{Q}_i(y_i = l)$ is the updated Q value. This is the message-passing step and is equivalent to applying M image-dependent Gaussian filters on the Q values. After message passing, label incompatibility was calculated as a penalty to encourage local smoothness. The incompatibility for label l at a given voxel i, denoted by $\hat{Q}_i(y_i = l)$, was calculated as the sum of the updated $\tilde{Q}_i$ that take a different label, i.e., $\hat{Q}_i(y_i = l) \leftarrow \sum_{l'} \mu(l, l')\tilde{Q}_i(y_i = l')$. Next, this penalty incurred by incompatibility was subtracted from the unary inputs $\psi_u(y_i = l|\mathbf{x}_i)$, i.e., $Q_i(y_i = l) \leftarrow \frac{1}{Z_i'}\exp(-\psi_u(y_i = l|\mathbf{x}_i) - \hat{Q}_i(y_i = l))$, where $Z_i' = \sum_k^K \exp(-\psi_u(y_i = k|\mathbf{x}_i) - \hat{Q}_i(y_i = k))$ is the normalisation constant. The above steps were repeated until Q converges. The resulting $Q_i(y_i|\mathbf{x}_i)$ is an approximation of the likelihood $P(y_i|\mathbf{x}_i)$, and was used when calculating the cross entropy in Equation (A3). This cross entropy (A3) was minimised in a mini-batch style via an ADAM optimiser [130] with learning rate 0.01, in which the connectivity features $\mathbf{X}$ of each subject served as a mini-batch. The model was trained using only the reliable subjects. To clarify, "reliable" refers to the subjects that passed the selection criteria described in Chapter 3.2.5. These reliable subjects were further divided into two separate groups or "folds". The model was trained using one fold at a time, with the other fold and "unreliable" subjects (those who did not pass the quality control) used for testing. This approach ensured that our model was exposed to and learned from the most accurate data available.

The pseudo code of the above steps is summarised in Algorithm 1.

3.2.7 Evaluation of Vim localisation

The three approaches for Vim localisation on low-quality data were evaluated separately on the two subsets of subjects: the reliable subset of subjects, in which the HQ-Vim passed the selection criteria (described in Chapter 3.2.5), and the

Algorithm 1 Mean-field iteration in CRF

$Q_i(y_i = l) \leftarrow \exp(-\psi_u(y_i = l|\mathbf{x}_i))/\sum_k \exp(-\psi_u(y_i = k|\mathbf{x}_i))$ ▷ Initialisation
while not converged **do**
  $\tilde{Q}_i(y_i = l) \leftarrow \rho \sum_{j \in \mathcal{N}_i} k(\phi(\mathbf{x}_i), \phi(\mathbf{x}_j)) Q_j(y_j = l)$ ▷ Message Passing
  $\hat{Q}_i(y_i = l) \leftarrow \sum_{l'} \mu(l, l') \tilde{Q}_i(y_i = l')$ ▷ Compatibility with neighbours
  $Q_i(y_i = l) \leftarrow \frac{1}{Z_i'} \exp(-\psi_u(y_i = l|\mathbf{x}_i) - \hat{Q}_i(y_i = l))$ ▷ Approximate $P(y_i = l|\mathbf{x}_i)$
end while

unreliable subset, in which the HQ-Vim did not meet the criteria. On the reliable subset, the corresponding HQ-Vim served as the ground truth (i.e., HCP HQ-Vim as ground truth for HCP; UKB HQ-Vim as ground truth for UKB); on the unreliable subset, the atlas-defined Vim served as the ground truth.

Two metrics, Dice coefficient and centroid displacement, were employed to assess the correspondence between the low-quality Vim localisation and the ground truth. The Dice coefficient measures the extent of overlap between two segmentations relative to their combined size. Specifically, for two segmentations, A and B, the Dice coefficient can be expressed as $2|A \cap B|/(|A| + |B|)$, where $|A \cap B|$ represents the number of voxels classified as Vim in both A and B, whilst $|A| + |B|$ represents the total number of Vim voxels in A and B. Since the HQ-augmentation approach and the atlas-based approach produce continuous Vim probability maps, calculating their Dice coefficients with the "ground truth" requires thresholding to binarise these maps *a priori*.

The other metric, centroid displacement, provides an alternative means of evaluating segmentation similarity that does not rely heavily on precise thresholding. It measures the Euclidean distance between the centroids of two Vim clusters, thus offering a measure of the spatial displacement between predicted and true Vim locations. The centroid for the HQ-augmented Vim is calculated as the weighted average coordinates, where the weights correspond to the estimated posterior probability of a voxel being classified as Vim. A low threshold of 0.1 is applied to the output posterior probability map when calculating the centroid coordinate, which eliminates voxels that have a low likelihood of being classified as Vim. The centroid for the atlas-based Vim is computed in a similar fashion, using the group-average

Vim probability as weights (once warped into native space) and also applying a 0.1 threshold. For the connectivity-driven Vim, the centroid coordinate is derived from the average coordinates of the binarised map.

3.3 Results

3.3.1 Accuracy of the HQ-augmentation model on HCP surrogate low-quality data

As the connectivity-driven approach may even fail on high-quality data, producing HQ-Vim (i.e., the connectivity-driven Vim derived from high-quality HCP data) that are unreliable as "ground truth", we split the evaluation subjects into two subsets based on the reliability of the HCP HQ-Vim and made evaluations separately, against different ground truth. The first subset consisted of subjects with trustworthy HQ-Vim, where the HQ-Vim served as the ground truth location of the Vim. The second subset consisted of subjects with unreliable HQ-Vim, where the atlas-defined Vim served as the ground truth instead. The HQ-Vim has to pass four criteria to serve as a trustworthy ground truth: 1) its volume exceeds 20 mm^3; 2) it contains only one blobs; 3) its center-of-mass locates within 4 mm from the center-of-mass of the Thalamic DBS Connectivity Vim Atlas [34] in native space; 4) its correlation with the Thalamic DBS Connectivity Vim Atlas exceeds 0.5 in native space. These loose criteria were chosen to reject the HQ-Vim that was too far away from the previously-validated Vim atlas to be considered as reliable, while preserving the inter-individual anatomical variability by allowing for some extent of deviation from the standard atlas (see Figure A1A for the histograms of Euclidean distances and correlations between HQ-Vim and the Thalamic DBS Connectivity Vim Atlas across HCP subjects). A total of 459 HCP subjects out of 1063 passed the four criteria, thus selected as the reliable subset (see Figure 3.3 for the example contours of reliable HQ-Vim, and Figure 3.4C, 3.4F and 3.4I for contours of unreliable HQ-Vim). A group-average Vim probability map was calculated across these 459 subjects' HQ-Vim (warped into standard space), and subsequently transformed back into the individual space and thresholded at 50% as the ground truth for the unreliable

subset of HCP subjects. To further validate our choices of ground truth, we assessed the spatial proximity between this group-average Vim probability map (in 1mm MNI standard space) and three sets of literature-based optimal stimulation points targeting the Vim for the treatment of tremor [131] in left hemisphere (because, the three sets of coordinates were provided in left hemisphere only). The group-average map overlapped with all three sets of optimal stimulation points (TABLE 3.1). Particularly, it overlapped with the peak location of an unweighted volume of tissue activated (VTA) frequency map with high probability.

We compared three Vim localisation approaches against the "ground truth" for each type of HCP surrogate low-quality data (LQ-LowAngular, LQ-LowSpatial, or LQ-LowAngular-LowSpatial). The three approaches were: 1) the atlas-defined Vim, created by transforming the HCP group-average Vim probability map into the individual native space; 2) the connectivity-driven approach, derived from the respective low-quality M1 and dentate tract-density features; 3) the HQ-augmented Vim, obtained by applying the corresponding HQ-augmentation model to the extended set of low-quality connectivity features.

When evaluated against the respective HQ-Vim, the HQ-augmented Vim outperformed the atlas-defined Vim and the low-quality connectivity-driven Vim for all three types of HCP low-quality data (Figure 3.2), producing higher Dice coefficient with the respective HQ-Vim and smaller Euclidean distance from the HQ-Vim's center-of-mass. In particular, on LQ-LowAngular (Figure 3.2A, 3.2B and 3.2C) and LQ-LowAngular-LowSpatial data (Figure 3.2G, 3.2H and 3.2I), the connectivity-driven approach often failed to generate any clusters due to the degraded angular resolution, resulting in "zero" identified Vim volume and thus having zero Dice coefficient with the HQ-Vim. In this case, its Euclidean distance from the HQ-Vim's center-of-mass was set to Infinity. In contrast, the HQ-augmentation model invariably succeeded in finding a coherent Vim cluster that was close to the respective HQ-Vim, suggesting that the HQ-augmentation model is more capable of reproducing Vim localisation than the connectivity-driven approach when high-quality data is unavailable. Furthermore, the HQ-Vim approach produced higher

Literature-based stimulation points	I: classical target center of gravity	II: clinically-weighted hotspot	III: unweighted frequency map peak
MNI coordinates (x, y, z)	(-13.1, -19.3, -2.1)	(-17.3, -13.9, 4.2)	(-16.5, -17.0, 1.2)
Overlap probability (HCP)	0.51	0.37	0.96
Overlap probability (UKB)	0.62	0.23	0.95

Table 3.1: Spatial proximity between the group-average Vim probability map and the literature-based optimal stimulation points (left hemisphere). The group-average Vim probability map calculated from the reliable subset overlapped with three sets of literature-based optimal stimulation points. The overlap probability (in third and fourth row) represents the proportion of subjects (of the respective dataset) that overlap at the given target location (in second row). The classical target center of gravity was defined by Horn et al. based on literature-reported effective contact location. The other two sets were defined by Elias et al. 2021 on the basis of volume of tissue activated (VTA) stimulation maps, either weighted by the corresponding VTA percentage change from baseline (clinically-weighted hotspot coordinates), or unweighted but thresholded at top 10% (unweighted frequency map peak coordinates).

Dice coefficient and smaller centroid displacement with the HQ-Vim than the atlas-defined approach, suggesting that the HQ-augmented Vim preserved more inter-individual variations of the Vim than a rigid standardised atlas. Obviously, the subset of HQ-Vim as evaluation "ground truth" was defined to resemble the standardised Vim atlas (i.e., excluding those having a large mismatch with the atlas), which thus limits inter-individual anatomical variability in the first place. It is notable that the HQ-augmentation approach produced results that were closer to the HQ-vim than the atlas-defined Vim, even within this limited margin of anatomical variability (see Figure 3.3 for contours of reliable HQ-Vim, atlas-defined, HQ-augmented and connectivity-driven Vim).

On the subset for which the HQ-Vim were regarded as unreliable, we evaluated the accuracy of the HQ-augmented Vim against the atlas-defined Vim (Figure 3.4). Although the atlas-defined Vim cannot account for individual variability, it may still serve as a version of the "gold standard," as the unknown ground truth location of

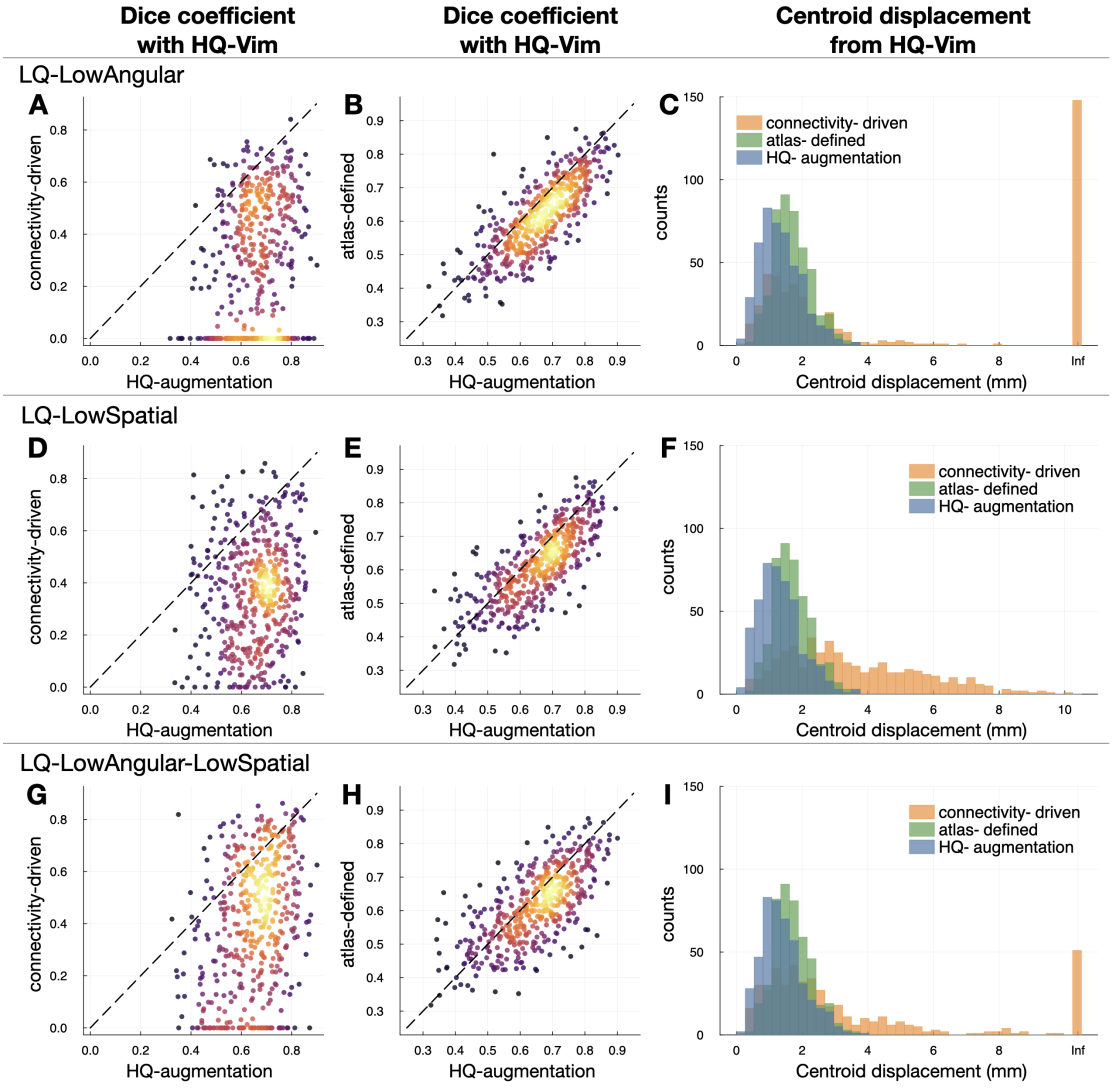

Figure 3.2: Accuracy of the HQ-augmentation model on LQ-LowAngular, LQ-LowSpatial, and LQ-LowAngular-LowSpatial, evaluated against HQ-Vim. On each type of HCP surrogate low-quality data, we evaluated the HQ-augmentation, connectivity-driven, and atlas-defined approach against the HQ-Vim. **(A)** On LQ-LowAngular data, the HQ-augmentation model produced higher Dice coefficient (X-axis) with the respective HQ-Vim than the connectivity-driven approach (y-axis). Warmer colours are indicative of areas where data points are more densely populated. **(B)** On LQ-LowAngular data, the HQ-augmentation model also gave higher Dice coefficient with the respective HQ-Vim than the atlas-defined vim (y-axis). **(C)** On LQ-LowAngular data, the HQ-augmentation model (blue histogram) gave smallest centroid displacement from the HQ-Vim than the connectivity-driven (orange) and atlas-defined (green) approach. **(C), (D) and (F)** Equivalent plots of **(A), (B) and (C)**, on LQ-LowSpatial. **(G), (H) and (I)** Equivalent plots of **(A), (B) and (C)**, on LQ-LowAngular-LowSpatial. For a more direct representation of the summary statistics, please refer to Figure A4 which displays boxplots of the Dice coefficients.

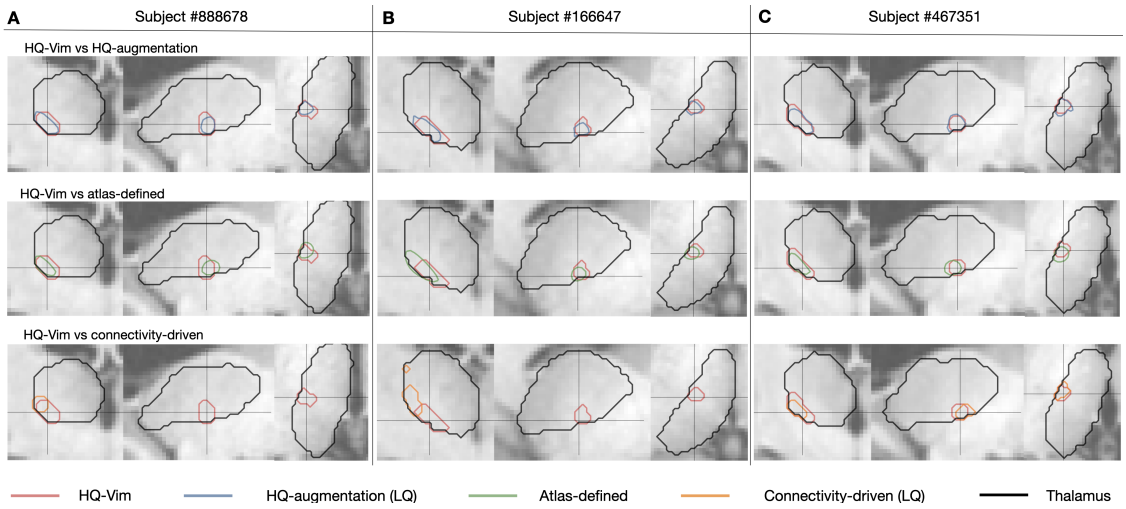

Figure 3.3: Example contours of HQ-Vim (ground truth), overlayed with the HQ-augmented or connectivity-driven Vim from HCP low-quality data, or by the atlas-defined Vim. (A) Contours of HQ-vim (red) versus LQ-LowAngular Vim (blue: HQ-augmentation; orange: connectivity-driven), or the atlas-defined Vim (green), of an example subject. **(B)** Contours of HQ-vim (red) versus LQ-LowSpatial Vim or the atlas-defined Vim (green), of an example subject. **(C)** Contours of HQ-vim (red) versus LQ-LowAngular-LowSpatial Vim or the atlas-defined Vim (green), of an example subject.

the Vim is not expected to be far from its atlas position. For all three types of low-quality data, the HQ-augmentation approach on the low-quality data gave higher Dice coefficient and smaller centroid displacement with the atlas-defined Vim, even outperforming the connectivity-driven approach applied on high-quality HCP. This suggests that the HQ-augmentation approach is more robust to corruptions of data quality. Similarly, this evaluation (on unreliable subjects) is obviously somewhat circular, in the sense that these subjects are defined by a large mismatch between HQ-Vim and atlas-based Vim, but it is nevertheless notable that the HQ-augmentation results (which are driven by tractography) are closer to the atlas results than the HQ-Vim (see Figure 3.4C, 3.4F and 3.4I for example contours of unreliable HQ-Vim, the HQ-augmented Vim on each low-quality dataset, and the atlas-defined Vim).

Overall, the above results demonstrate that the HQ-augmentation model is superior to existing approaches in its robustness to data quality, as well as ability to preserve individual variability. Its reliability against data corruptions is primarily achieved via the extended set of connectivity-features with the wide range of brain regions. Instead of merely relying on the M1 and dentate tract-density features,

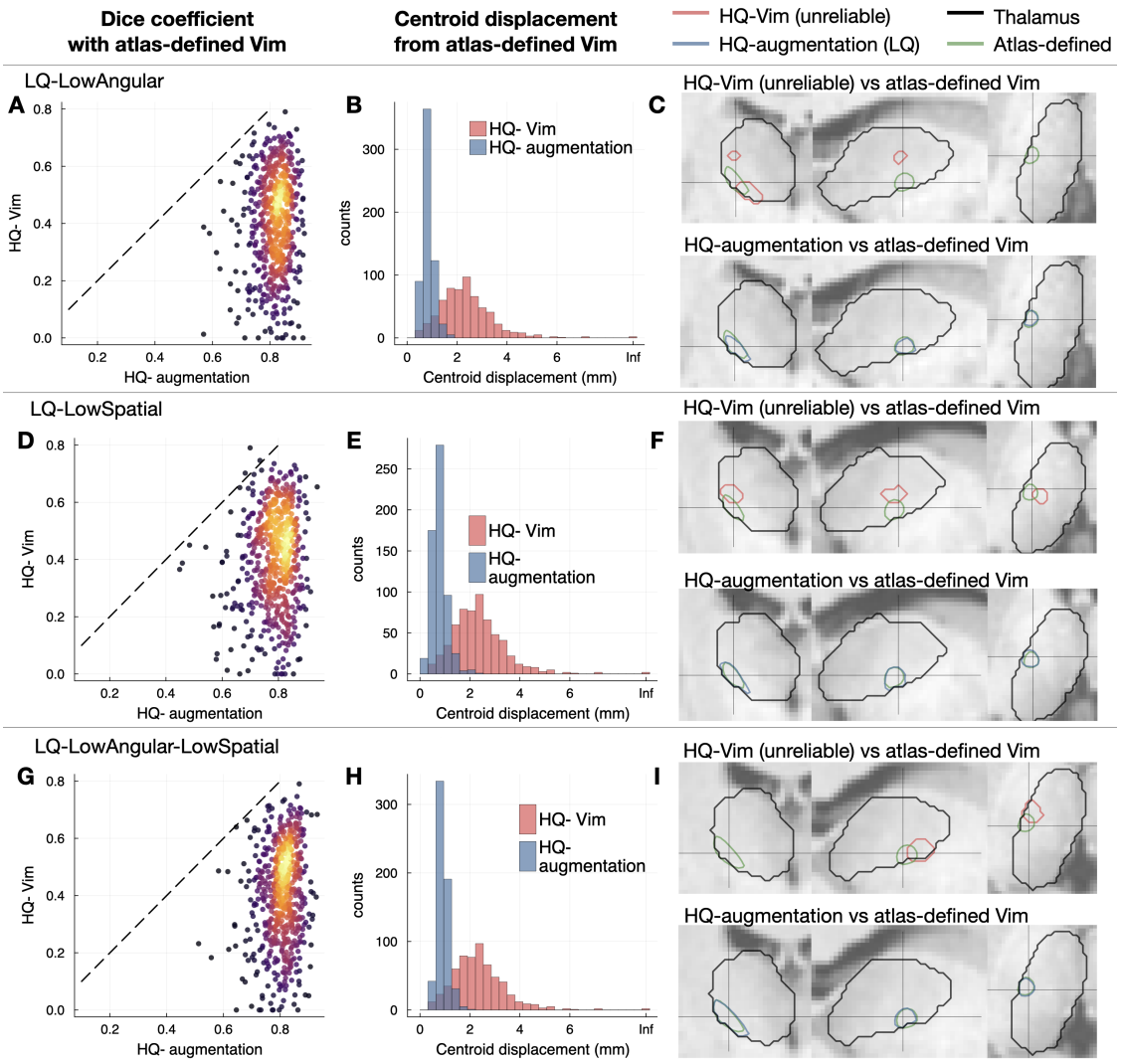

Figure 3.4: Accuracy of HQ-augmentation on LQ-LowAngular, LQ-LowSpatial, LQ-LowAngular-LowSpatial, evaluated against the atlas-defined Vim. For the left-out subjects in which HQ-Vim is untrustworthy, the atlas-defined Vim served as "ground truth" instead. **(A) and (B)** The HQ-augmentation approach (on LQ-LowAngular) gave higher Dice coefficient and smaller centroid displacement with the atlas-defined Vim even than the HQ-Vim. Warmer colours are indicative of areas where data points are more densely populated. **(C)** Example contours of HQ-Vim (unreliable), atlas-defined Vim (green) and HQ-augmented Vim (blue). The HQ-augmentation approach on LQ-LowAngular produced Vim closer to the group-average location of Vim. **(D), (E) and (F)** Equivalent plots of **(A), (B) and (C)**, on LQ-LowSpatial data. **(G), (H) and (I)** Equivalent plots of **(A), (B) and (C)**, on LQ-LowAngular-LowSpatial data.

the HQ-augmentation model gathers evidence from the richer set of connectivity features, together with the group-average "prior" location to make decisions; when the M1 and dentate features are no longer trustworthy (e.g., giving contrasting evidence with respect to the location of Vim or tending to make predictions that deviate too much from the group-average location), the HQ-augmentation model may still obtain sufficient evidence from the rest of the connectivity features to produce a coherent Vim cluster that roughly aligns with the atlas-defined Vim. Even if the data quality is too bad, the model can still rely on the group-average prior, producing baseline Vim localisation that approximates the atlas-based Vim (see Figure 3.4 for more discussion). This also explains why the HQ-augmentation model may give Vim closer to the atlas even when the HQ-Vim fails.

Another interesting result is that the LQ-LowAngular-LowSpatial dataset, which was created by spatially downsampling the LQ-LowAngular dataset, does not appear to have an actual lower quality than the latter. Specifically, the connectivity-driven approach encountered fewer failures on the LQ-LowAngular-LowSpatial dataset than on the LQ-LowAngular dataset. A potential explanation for this phenomenon is that the reduced spatial resolution of the LQ-LowAngular-LowSpatial dataset results in a higher signal-to-noise ratio (SNR). This higher SNR might improve estimation of fibre orientations and streamline tracking, resulting in better connectivity-driven Vim localisation.

3.3.2 Generalising the HQ-augmentation model to the UKB surrogate low-quality dataset

We also tested whether the HQ-augmentation models trained on HCP were generalisable to other datasets collected under different protocols. It is crucial for a model to be generalisable to unseen protocols, as collecting large datasets for training purposes in clinical contexts can often be impractical. Here we leveraged the UKB diffusion MRI dataset to test the generalisability of the HQ-augmentation model trained on HCP. The UKB diffusion MRI were acquired at isotropic spatial resolution 2 mm, with two shells bvals = 1000 s/mm^2 and 2000 s/mm^2, sampled

at 50 directions per shell. We similarly corrupted the UKB diffusion MRI data by extracting the b0 and bvals=1000 s/mm^2 volumes, the latter sampled uniformly at 32 directions, as the surrogate UKB low-quality data, referred to as LQ-UKB. We applied the HQ-augmentation model trained on LQ-LowAngular-LowSpatial, which roughly matched the spatial and angular resolution of LQ-UKB, to the low-quality connectivity features derived from LQ-UKB, producing a predicted Vim probability map per UKB subject.

The evaluations on the UKB dataset were similar to those on HCP, where the evaluation subjects were divided into a reliable subset with UKB HQ-Vim as ground truth and an unreliable subset with atlas-defined Vim as ground truth. The UKB HQ-Vim were derived from the original multi-shell UKB dataset via the connectivity-driven approach. The UKB atlas-defined Vim was derived from the UKB version of group-average Vim probability map, created by averaging the reliable UKB HQ-Vim, as it is more representative of the age population in UKB than the HCP group-average Vim probability map. The selection criteria were the same as HCP (Figure A1B), i.e., the UKB HQ-Vim has to satisfy the following four conditions to be regarded as trustworthy "ground truth": 1) its volume is greater than 20 mm^3; 2) it contains only one cluster; 3) its correlation with the Thalamic DBS Connectivity Atlas is greater than 0.5; 4) its centroid displacement from the Thalamic DBS Connectivity Atlas is smaller than 4 mm. The group-average Vim probability map calculated from the subset of reliable HQ-Vim, though slightly different from the one calculated from the HCP subjects, also overlappped with the literature-based optimal stimulation points for tremor (see TABLE 3.1, the fourth row, for the overlap probabilities).

Similarly to HCP, the performance of the HQ-augmentation model was compared with the connectivity-driven Vim on LQ-UKB and the atlas-defined Vim. Although purely trained on HCP low-quality features, the HQ-augmentation model again produced higher Dice coefficient and smaller centroid displacement than the connectivity-driven and atlas-defined approach (Figure 3.5A) with the UKB HQ-Vim, suggesting its robustness to variations in data heterogeneity as well as its

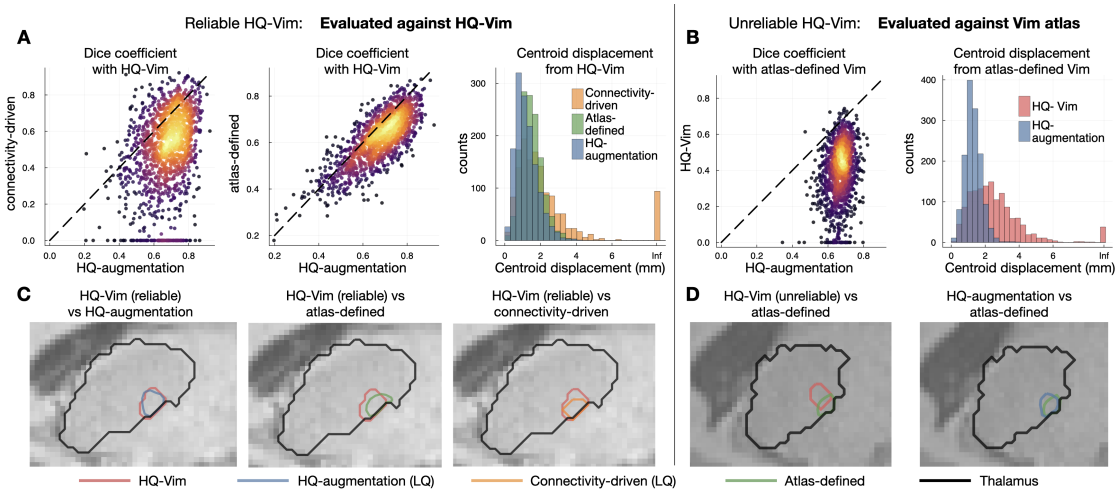

Figure 3.5: Generalisability of HQ-augmentation to UKB. The HQ-augmentation model trained on HCP was applied to the UKB low-quality features (blue) as HQ-augmented Vim. **(A)**. When using the UKB HQ-Vim as ground truth, the HQ-augmentated Vim has higher Dice coefficient and smaller centroid displacement with the UKB HQ-Vim, than the atlas-defined Vim (green) and the connectivity-driven Vim using low-quality features (orange). **(B)** When using the atlas-defined Vim as ground truth, the HQ-augmentation model using low-quality features even gave more reliable Vim than the UKB HQ-Vim (red), which used HQ features to target Vim. **(C)** and **(D)**. Contours of Vim, identified by each approach.

ability to capture individual variability. Furthermore, when evaluated against the atlas-defined Vim, this HCP-transferred HQ-augmentation model generated Vim estimates that were much closer to the atlas than the UKB HQ-Vim, indicating again that the HQ-augmentation approach, even trained on a different dataset, was even more reliable than the connectivity-driven approach applied to high-quality data (Figure 3.5B). Overall, these results demonstrate that the HQ-augmentation approach trained on HCP data can indeed generalise to other low-quality datasets acquired using different protocols. A possible reason for this generalisability is that the HQ-augmentation model trained on HCP has already learned some general features and patterns that can be useful for predicting Vim on UKB data; hence, leveraging the anatomical knowledge gained from HCP improves its performance on another related dataset, e.g., LQ-UKB. The HCP dataset itself contains a diverse range of subjects, and thus the model has learned to handle the variability in the LQ-UKB to some extent.

3.3.3 Reliability analysis of the HQ-augmentation approach on HCP and UKB

We conducted a comprehensive reliability analysis for the three approaches at the individual level to assess their consistency in producing Vim localisation despite variations in data quality and across scanning sessions. It is crucial for a localisation approach to produce consistent results not only across different scanning sessions but also across varying data quality, as its clinical value lies in capturing the true underlying anatomical variability unique to each individual, rather than reflecting variabilities arising from noise or methodological factors. Our reliability assessment consisted of two dimensions: 1) consistency of Vim localisations across scans with different quality levels for a given subject (referred to as across-quality reliability); and 2) correspondence of Vim localisations across different scanning sessions of the same quality level for a given subject (referred to as test-retest reliability). Prior to conducting the reliability analysis, we trained an HQ-augmentation model specifically for the high-quality data, using the HQ-Vim as training labels and the extended set of connectivity profiles derived from high-quality data instead as input features. The HQ-augmented model tailored to high-quality data was applied to the left-out subjects to produce a "high-quality version" of HQ-augmented Vim per subject. This version was then compared with the HQ-augmentation outputs from low-quality datasets (to measure its across-quality reliability) or retest sessions (to measure its test-retest reliability).

To assess across-quality consistency of the HQ-augmentation approach, we calculated the Dice coefficient and centroid displacement between the HQ-augmented Vim derived from each low-quality dataset (LQ-LowAngular, LQ-LowSpatial, LQ-LowAngular-LowSpatial) and its high-quality counterpart (i.e, the high-quality version of HQ-augmented Vim) per subject, resulting in six measurements per subject (2 metrics $\times$ three pairs of low-versus-high comparisons). Similarly, we also assessed the across-quality consistency of the connectivity-driven approach, by comparing the connectivity-driven Vim derived from each low quality dataset against the respective HQ-Vim, again resulting in six measurements. On both UKB

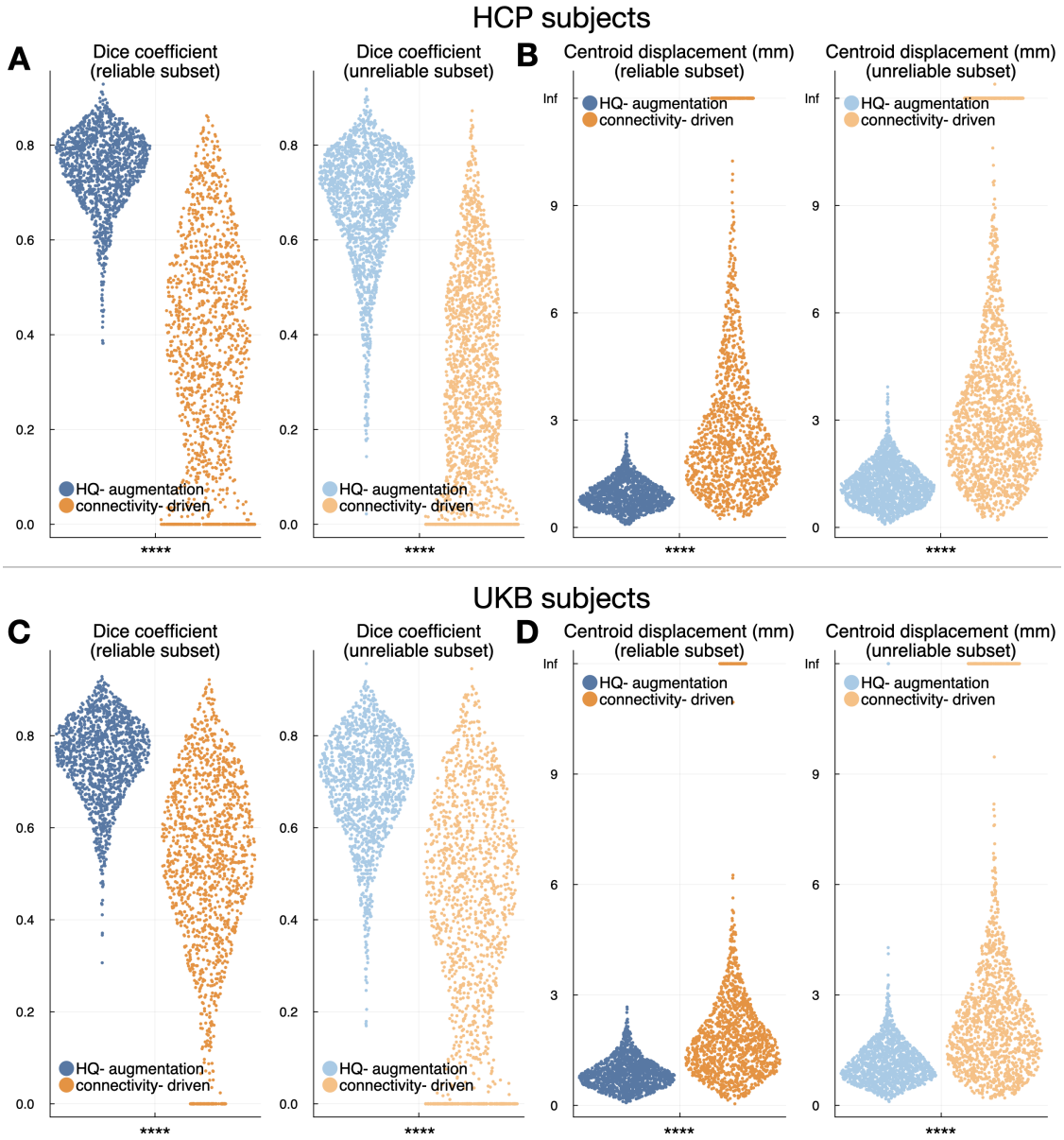

Figure 3.6: Across-quality consistency of the HQ-augmentation and connectivity-driven approach. The HQ-augmentation approach is significantly more consistent than the connectivity-driven approach (paired t-tests, Bonferroni-corrected by 8 tests in total), producing higher Dice coefficient and smaller centroid displacement between the outputs derived from different data quality, both across the reliable subset (dark blue and orange) and unreliable subset (light blue and orange). **(A)** Blue: Dice coefficient between the HQ-augmentation results on high-quality data (HQ-HCP) and each of the low-quality data (LQ-LowAngular, LQ-LowSpatial, LQ-LowAngular-LowSpatial), pooled together (i.e., each dot represents a single high-vs-low comparison); orange: equivalent plots for the connectivity-driven approach. **(B)** Equivalent plots of (A), using centroid displacement as the consistency metric. **(C)** Dice coefficient between the HQ-augmentation results on HQ-UKB and LQ-UKB (blue) and between the connectivity-driven results on HQ-UKB and LQ-UKB (orange). **(D)** Equivalent plots of (C), using centroid displacement as the consistency metric. Dark colours indicate reliable subjects (i.e., their HQ-Vim are trustworthy), while light colours indicate the opposite.

and HCP, the HQ-augmentation approach produced substantially more consistent Vim localisation results than the connectivity-driven approach across datasets of different quality (Figure 3.6). Both approaches had compromised "across-quality" consistency on the unreliable subset, demonstrated by the smaller Dice coefficient and higher centroid displacement (Figure 3.6, light colours).

The test-retest reliability of the two approaches displayed a similar pattern. To assess the test-retest reliability of the HQ-augmentation approach, we applied the HQ-augmentation model trained on the multi-shell first-visit data to the second-visit connectivity features, producing the HQ-augmented Vim for the repeat sessions, and calculated the Dice coefficient and centroid displacement between first-visit and second-visit HQ-augmented Vim. The test-retest reliability of the connectivity-driven approach was likewise measured by the similarity between the connectivity-driven Vim derived from the first-visit and second-visit M1/dentate connectivity features respectively. The HQ-augmentation approach outperforms the connectivity-driven approach in providing substantially higher Dice coefficient and smaller centroid displacement between sessions (Figure 3.7). Similarly, both approaches had compromised test-retest reliability on the unreliable subset (Figure 3.7, light colours).

3.4 Discussion

We proposed an Image Quality Transfer (IQT) approach, HQ-augmentation, for improving the accuracy of localising the ventral intermediate nucleus of the thalamus (Vim), a common target for deep brain stimulation and MR-guided ultrasound. This approach leverages the anatomical information (i.e., approximate ground truth location of the nucleus) obtained from high-quality HCP to augment Vim identification on the low-quality data, by learning to produce Vim localisation as close as possible to the high-quality "ground truth" using a wide range of white matter connectivity features derived from the low-quality data. Our results demonstrate that the HQ-augmentation approach is more accurate in producing Vim and more robust to data corruptions than its predecessor, the connectivity-driven approach, especially when high-quality data is unavailable; moreover, it proves more

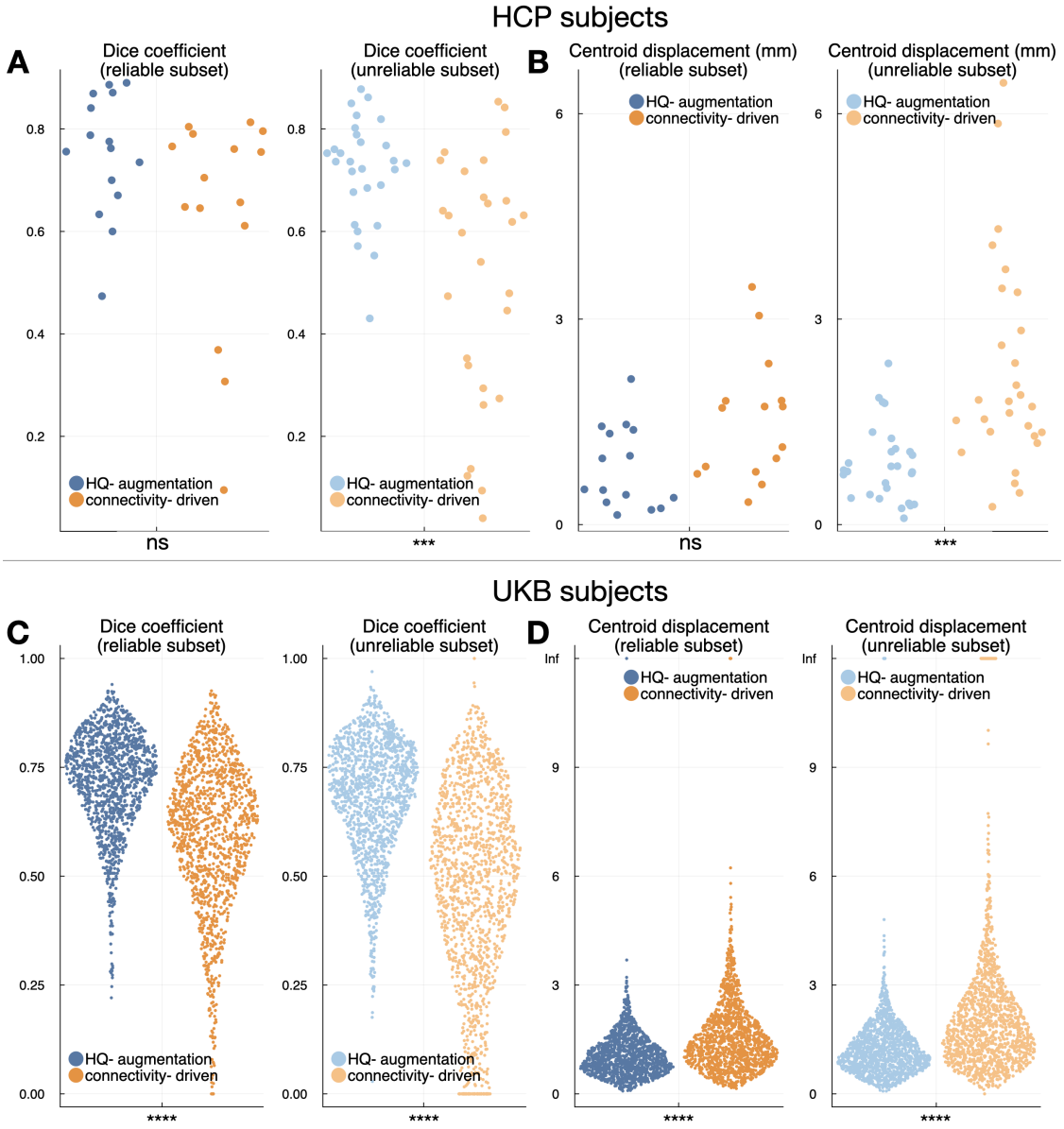

Figure 3.7: Test-retest reliability of the HQ-augmentation and connectivity-driven approaches. The HQ-augmentation approach is significantly more consistent than the connectivity-driven approach (8 paired t-tests, Bonferroni-corrected by 8 tests in total) across scanning sessions, producing higher Dice coefficient and smaller centroid displacement between the outputs derived from the first-visit and repeat scans, both across the reliable subset (dark blue and orange) and unreliable subset (light blue and orange). **(A)** Blue: Dice coefficient between the HQ-augmentation outputs on first-visit and repeat scans across 43 HCP subjects that had received retest sessions; orange: equivalent plots for the connectivity-driven approach. **(B)** Equivalent plots of (A), using centroid displacement as the consistency metric. **(C)** and **(D)** Equivalent plots of (A) and (B) for 2,760 UKB subjects, using centroid displacement as the consistency metric. Dark colours indicate reliable subjects (i.e., their HQ-Vim are trustworthy), while light colours indicate the opposite.

reliable than the connectivity-driven approach even when the latter uses high-quality information to localise Vim. The HQ-augmentation approach performs consistently well across different data qualities and scanning sessions, surpassing the performance of the connectivity-driven approach. Furthermore, the HQ-augmentation approach preserves more inter-individual anatomical variability than the atlas-based approach. Importantly, our HQ-augmentation model trained on HCP surrogate low-quality data is generalisable to UKB low-quality data, outperforming the connectivity-driven approach in accuracy and reliability while preserving the anatomical variability unique to UKB individual subjects. Being a reliable and generalisable tool, our approach is especially valuable in clinical contexts where large samples of data collection and retraining are impractical.

Why is the HQ-augmentation approach capable of reproducing HQ-Vim on low-quality data, while the connectivity-driven approach using M1 and dentate tract-density features struggles to do so in isolation? The HQ-augmentation model leverages connectivity features with other target ROIs, which may also contain information about the anatomical location of Vim. For instance, Vim is anterior to thalamic clusters exhibiting high diffusion connectivity to S1, the primary sensory area. Consequently, tract-density with S1 serves as a "negative" feature, providing evidence against the likelihood of a given voxel being classified as Vim. Connectivity targeting at white matter segments can also be informative, particularly when diffusion data has limited angular resolution. This is because middle point targets are more accessible for streamlines to reach compared to more distant cortical ROIs. For example, streamlines originating from a seed thalamic voxel might not reach M1 but could still reach white matter segments between the ipsilateral M1 and the thalamus. In such cases, tract-density features with these white matter ROIs may compensate for the inadequate M1 connectivity feature, resulting in Vim segmentation that remains close to the one defined by high-quality M1 and dentate tract-density features. In essence, the HQ-augmentation model synthesises evidence from an extended set of connectivity features to determine the most probable location of the "true" Vim. When primary connectivity features defining Vim fail to provide

reliable identification due to limited data quality, the remaining connectivity features may still guide the model to produce Vim segmentation close to its high-quality counterpart. This makes the model more robust against data quality degradation. It is worth noting that the HQ-augmentation model generates Vim segmentation closer to the atlas than the HQ-Vim, which is defined by high-quality M1 and dentate tract-density features. This is due to the incorporation of a "prior" feature representing the group-average location of Vim. The HQ-augmentation approach may converge to the baseline, i.e., the atlas-defined Vim, when the entire extended set of connectivity features is too noisy to provide reliable Vim identification.

Numerous techniques exist for deep nuclei segmentation, including atlas based segmentation, image intensity based segmentation, and machine/deep learning based segmentation. Atlas-based segmentation has the advantage of being the most reproducible, as the only factor that may affect its segmentation performance is suboptimal nonlinear registration of the atlas image to a new image set. By employing an average atlas derived from the same image population, this technique can often enhance its performance through improved alignment between the atlas and individual image sets. However, atlas-based segmentation has limitations in accounting for individual variability inherent in the data, which is essential for personalising treatments in clinical practice. The connectivity-driven approach is an example of image intensity based segmentation. The product (or intersection) of M1 and dentate tract-density feature maps creates an intensity contrast, offering probabilistic information about the likelihood of a voxel belonging to the Vim nucleus. A related method is Markov Random Field (MRF), which models image intensity as Gaussian mixtures and classifies voxels based on their posterior probability of belonging to respective mixture components, while promoting similarity in membership assignments for neighbouring voxels. The HQ-augmentation approach, on the other hand, is a machine learning-based segmentation technique. It shares similarities with MRF, as it contains two compartments for modelling the posterior probability of voxel label assignment. One of the compartments usually represents the likelihood of a voxel belonging to a given class without considering its neighbours.

The other compartment is usually a pairwise term, akin to MRF, penalising neighbouring voxels with incompatible label assignments. The HQ-augmentation approach outperforms MRF in this context due to its discriminative power, which is achieved by learning the mapping between low-quality image features and high-quality labels. This optimisation of the likelihood of observing high-quality labels given low-quality features results in a stronger classifier than Gaussian mixtures in MRF. In the future, it may be worthwhile to explore other machine learning or deep learning techniques, as well as transfer learning methods, to evaluate their potential for further enhancing deep nuclei segmentation/localisation tailored to low-quality data.

The proposed HQ-augmentation model falls within the scope of Image Quality Transfer (IQT), a new computational imaging technique aimed at propagating information from rare or expensive high-quality data to more prevalent low-quality datasets. This transfer enhances the image resolution or information content of the latter, facilitating subsequent analysis. IQT typically relies on matched pairs of high- and low-quality data to establish a mapping from low-quality features to the corresponding high-quality information [90, 91, 132]. Although closely related to these previous IQT techniques, the HQ-augmentation approach differs slightly as it does not primarily focus on enhancing lower-quality images to augment subsequent analysis. Instead, it directly learns the mapping between high-quality Vim labels and lower-quality connectivity features. Once learned and applied to unseen image sets, the mappings "augment" the low-quality connectivity features to produce predicted Vim probability maps as if they were derived from their "imaginary" high-quality counterparts. An alternative option for augmenting Vim localisation on low-quality data involves enhancing the low-quality image content via IQT first, followed by identifying Vim using the connectivity-driven approach on the enhanced image. However, this route was not chosen because the connectivity-driven approach may fail even on high-quality data. Therefore, in addition to high-quality image transfer, it is equally important to develop an approach that is more reliable and flexible than existing methods, even when high-quality data

is available. Nevertheless, it is worth exploring whether enhancing low-quality image content via IQT first, followed by the HQ-augmentation model, might further improve Vim localisation on low-quality data.

The HQ-augmentation approach has several limitations. First, it relies on the availability of high-quality data to train the model, which may not always be accessible in certain research or clinical settings. Although our model demonstrates generalisability from HCP to UKB datasets, it is important to assess its performance on other low-quality datasets and populations to ensure broad applicability. Moreover, given the variation in MRI acquisition protocols and hardware, the model may require retraining or fine-tuning for specific clinical or research contexts. Second, the current model is based on a machine learning approach, which may be more sensitive to the choice of features and the representation of the data. Recent advancements in deep learning, specifically in the domain of medical image analysis, may offer alternative solutions with improved performance and adaptability to various imaging contexts. Investigating the integration of deep learning methods, such as convolutional neural networks (CNNs), with the HQ-augmentation framework, could lead to further advancements in deep nuclei segmentation on low-quality data. Lastly, while the HQ-augmentation approach demonstrates improved performance compared to existing techniques, there is still room for refinement. Exploring additional image features, optimising the model architecture, and incorporating advanced regularisation techniques may further improve the segmentation accuracy and robustness of the model, ultimately benefiting clinical decision-making and personalised patient care.

Note that this approach is not exclusively applicable to the Vim nucleus. Rather, it can be generalised to any brain region for which detailed connectivity information is available. This adaptability is particularly suitable for subcortical regions, for which the connectivity in animal models, such as macaques (where we can use tracers), is more likely preserved in humans.

4

Accurate predictions of individual differences in task-evoked brain activity from resting-state fMRI using a sparse ensemble learner

Contents

4.1 Introduction

Studying individual differences in brain activity and how they relate to cognitive and genetic traits is an important area of research in basic and clinical neuroscience. Traditionally, functional Magnetic Resonance Imaging (fMRI) analysis has primarily been concerned with group-average inference. While averaging data across individuals substantially improves signal-to-noise (SNR) ratio and has proved fruitful in identifying common patterns across subjects, this approach treats unexplained individual variations as noise, discarding unique attributes of brain activity specific to a particular subject. Individual variations in neuroanatomical or functional activity often carries valuable information. For example, if a small number of subjects in a large cohort has a rare disease, an indiscriminate data reduction prior to the analysis will very likely hide this information.

The rapid development of cutting-edge neuroimaging techniques in recent decades has led to substantial improvements in the reliability and validity of blood-oxygen-level-dependent (BOLD) measurements, providing an unprecedented opportunity to investigate individualised patterns of brain activity. Moreover, emerging "big data" projects, such as the Human Connectome Project [105] and UK Biobank [133], have collected multi-modal neuroimaging data on very large samples, enabling researchers to more closely examine individual variations in neuroanatomical patterns and functional activities with enhanced statistical power. Among previous fMRI studies of individual variabilities, an active line of research focuses on understanding how individual brains vary in response to external cognitive tasks. Following the work of [19, 37], a number of studies [38, 39, 134, 135] have since shown that spatial patterns of task-evoked activation form a stable trait marker, encoded in resting-state brain activity, i.e., in the absence of any explicit task. In contrast to previous studies that mostly rely on correlation analysis (in-sample inference) to investigate individual differences, these works adopt predictive frameworks that allow for out-of-sample inference and greatly improved generalisability of these investigations of individual variability.

Why is task-free fMRI predictive of task-evoked activation? Previous studies have suggested that resting-state networks and task networks may share the same intrinsic architecture [37, 65, 136–138]. Therefore, a reasonable corollary is that resting-state heterogeneity should inform on variability of task-evoked brain activity. Typically, resting-state data are summarised as spatially continuous parcels distributed across the brain [139–144]. These spatial maps are often referred to as "functional modes", characterising functionally unified sub-processes underlying human cognition. Among the approaches of finding functional modes to predict task-fMRI, dual-regression [145, 146] is a widely-used algorithm, showing ability to predict individual idiosyncrasies in their response profiles [19, 38–40]. Although these previous attempts have successfully characterised individual-unique patterns of task-evoked brain activity, there are a few limitations yet to be accounted for. For example, these approaches either focused on cortical regions [40] or relied on pre-determined brain parcellations to make predictions within small patches of brain and concatenate the results [38, 39]. Here, we are interested in activations across the whole brain, not necessarilly limited to the cortex. Furthermore, our model uses global features, because parcellating the brain into small patches *a priori* may introduce more free-parameters and increase the risk of over-fitting. More importantly, these approaches did not attempt to explicitly model cross-subject variability of the rest and task states *per se*, and thus may be sub-optimal to capture cross-subject variations. In contrast, Ngo et al., 2021 introduced a contrastive loss in combination with the common loss to maximise inter-individual differences in cortical regions. However, in practice, such loss functions are often non-convex and may have complicated behaviours (e.g., multiple local minima) rendering optimisation difficult [147]. To fully account for inter-individual variations, an alternative is to explicitly train on residualised data, i.e., residuals where group-average information has been regressed out. The data obtained in this way has minimal shared variance with the group-level information, thus serves as a cleaner description of individual-level differences.

Accurate prediction of individual differences in task-fMRI response using resting-state fMRI has a number of potential applications. In a clinical setting, considerable variability across patients makes it difficult for surgeons to rely on anatomy (or population-averaged task activation maps) when finding crucial surgical targets. Task-fMRI can be used to localise surgical targets more accurately for individual patients, often by asking them to perform a specific cognitive/sensory-motor task. However, in practice, task-fMRI localisers are often limited by the patients' poor performance (or, the patients are unable to perform the task) and measurement noise. As a stable trait marker, resting-state fMRI has proved a good alternative to task-fMRI localisers, without the need of task execution prior to an operation. Accurate predictions of task-fMRI localisers using resting-state fMRI can therefore greatly benefit surgical targetting. Furthermore, having a better rest-predict-task model may improve our understanding of the mechanisms underlying individual differences in resting-state and task-fMRI.

Here we propose a framework that explicitly models individual variations in task-evoked brain activity using the resting-state variability, the latter profiled by a set of common spatial modes derived from a recently developed technique, Stochastic Probabilistic Functional Modes (sPROFUMO). We show that, consistent with previous studies [148–150], sPROFUMO provides better sets of "bases" (later referred to as PFMs) to reconstruct the variations in task-evoked activation patterns than the widely used approach of group-ICA followed by dual-regression. Additionally, we show that an ensemble learner that combines global and local bases has improved capacity of not only reproducing typical activation patterns but also preserved patterns unique to individuals. We demonstrate that modelling of individual-level task contrast maps comprises the modelling of two separate sources of variability, shape of activations and the overall activation strength. Considering these two aspects separately in task prediction is at least as effective as or even more desirable than simply modelling the original task contrast maps. Furthermore, the proposed model can recapitulate the spatial patterns of inter-individual variability, recovering regions that are more variable at the group-level. The model achieves state of

the art prediction accuracy for both datasets, and is also on par with task test-retest reliability. These results demonstrate the potential of resting-state features to reproduce task-fMRI features, and serve as a supplement to task localisers in pre-surgical plannings.

4.2 Materials and Methods

4.2.1 UK Biobank data

UK Biobank (UKB) is a large national project that collects a wide range of health-related measures for over 500,000 subjects, initially aged between 40 and 69. We used the resting-state and task functional MRI data from a total of 17,560 subjects. The acquisition parameters and processing details can be found in [113, 114]. Briefly, all resting-state fMRI scans were acquired with identical scanners (3T Siemens Skyra) with a TR of 735ms for a total of 490 time points for each individual. After the initial preprocessing, the data were ICA-FIX cleaned to remove structured artefacts [151], and then registered to the standard MNI space. Next, each individual's resting-state 4D time series were further spatially smoothed with a Gaussian kernel of sigma 3mm. The task used is the Hariri faces/shapes "emotion" task [70, 152], scanned and processed under the same protocols as the resting-state data (except that the task-fMRI data is not ICA-denoised). Individual as well as group-average activation z-statistic maps of three contrasts (faces, shapes, and faces-shapes) were estimated from the task fMRI scans using FEAT [153, 154]. Additionally, 473 subjects in this 17,560 subset received second-time scanning (mean test-retest-interval 2.25 years, std 0.12). These second-time scans provided test-retest reliability scores as a benchmark for our model performance.

4.2.2 Human Connectome Project data

We used the MSMAll-registered data provided by the Human Connectome Project (HCP), S1200 Release (https://www.humanconnectome.org/study/hcp-young-adult). Details on the acquisition protocols and processing pipelines can be found in [105, 107, 155]. Resting-state and task fMRI data from 991 subjects, aged 22 to 35

years, were used in the analysis. Each individual had four runs of resting-state scans with a TR of 0.72s for a total of 1,200 time points per run. The data were ICA-FIX denoised to remove structured artefacts automatically, then resampled onto the "32k_fs_LR" grayordinates space (surface vertices + subcortical voxels) and minimally-smoothed by 2mm FWHM. All subjects were MSMAll-registered to improve functional and structural alignment [155]. To further increase the signal-to-noise ratio, an additional smoothing of 4mm FWHM was applied to the MSMAll-registered data (with subcortical structures smoothed within parcel boundaries, and cortical data smoothed in 2D on the surface) using the Connectome Workbench (https://www.humanconnectome.org/software/connectome-workbench). The task fMRI scans were acquired and pre-processed in the same way (though without FIX). We used the MSMAll-registered individual and group-average contrast maps with 4mm FWHM smoothing in the analysis, including 47 contrasts across seven task domains [70]. In contrast to the UKB dataset, which is presented in a volumetric format, the HCP data is represented in the 32k fs_LR standard greyordinates space. As described by Glasser et al. [107], this space consists of 91,282 vertices in total, including both surface vertices and sub-cortical voxels. Another difference between the two datasets is how they were registered to the standard space. The UKB spatial maps were non-linearly registered to the MNI152 standard space via FSL's FNIRT [111, 112], while the HCP spatial maps were aligned into the standard grayordinates space via "MSMAll", an aggressive registration that not only considers cortical folding patterns but also other modalities such as myelination and resting-state connectivity [155, 156].

Similarly to the UKB dataset, we used the HCP repeat scans as the reliability benchmark for the predictions. Among the 991 subjects, 43 have received second-time scanning under the same 3T imaging and behaviour protocols with test-retest-interval ranging from 18 to 328 days (mean 134.78; std 62.49).

4.2.3 Generation of resting-state functional modes

We used resting-state functional modes to predict individual task-fMRI. Functional modes are typically modelled as parcel-like spatial configurations of unified functional processes distributed across brain, each characterised by a summary time course that captures mode activity over time. Here we explored two approaches of generating individual resting-state modes, group-ICA followed by Dual-Regression (DR-ICA) and Stochastic Probablistic Functional Modes (sPROFUMO). DR-ICA is a conventional group-average algorithm to estimate individual "versions" of group-level spatial configurations, using a set of common spatial modes as templates [157]. In DR-ICA, group-PCA was carried out on each dataset (UKB and HCP) by MELODIC's Incremental Group-PCA [158] on the resting-state time series of all subjects (temporally demeaned and variance normalised), producing 1,200 weighted spatial eigenmaps for UKB and 4,500 eigenmaps for HCP. These eigenmaps were subsequently fed into ICA using FSL's MELODIC tool to generate group-ICA spatial maps at multiple ICA-dimensions (i.e., the number of distinct ICA components). To obtain dual-regression maps for a specific subject at a given ICA-dimension k, we first regressed the corresponding k-dimensional group-ICA spatial maps into the individual 4D time-series data, yielding a set of k time courses per subject. The resulting time courses were subsequently regressed into the same 4D time-series, generating k dual-regression spatial maps for each subject. However, a major limitation of DR-ICA is that it only allows unidirectional flow between group and individuals, i.e., the estimated individual modes cannot in turn drive the refinement of group-average modes, and may have limited ability to cope with individual deviations from the group-average [159, 160]. A recently developed technique, sPROFUMO, uses a Bayesian model that simutaneously estimates functional modes both at group- and individual-level, and is scalable to large datasets [150]. In sPROFUMO, individual resting-state time-series are factorised into a set of spatial modes and their summary time courses (one per mode), together with the time course amplitudes. The group-level parameters constrain the estimation of (the posteriors over) individual-level parameters, of which the posterior evidence is accumulated

across individuals to in turn infer the group-level parameters. The bidirectional information flow between the group and individuals aims to result in improved subject-specific spatial alignments. Below, we refer to the resting-state feature maps as either DR-ICA maps or PFMs depending on the approach used to derive them.

4.2.4 Residualisation of the resting-state and task contrast maps

Our aim is to derive a model that can predict task activation in individuals given their resting state modes. One of the innovations in this chapter is to try to explicitly capture individual variations in our model. We propose that training and evaluating the model on residualised data (i.e., data and features where the group-averaged maps have been regressed out) would be of more value than training a model on the original resting-state and task contrast spatial maps.

To understand this, consider each individual task contrast map. This can be written as the sum of the group-average activation map (scaled by some factor) and a spatial residual map specific to the given individual. The resting-state feature maps can also be similarly decomposed. Thus, the mapping between resting-state features and target maps contains both desirable and undesirable mappings, in the sense that we do not want to optimise, e.g., for resting-state residual features to map onto the group-average activation map. Residualisation separates the group-average components from the residual components, thereby the two desirable mappings are learned separately as opposed to jointly (Figure 4.1a). Furthermore, using the residualised data has another added benefit in model evaluation. The test-retest reliability of the variability (with the group-mean effects removed) between different scanning sessions is an estimate of noise ceiling, which theoretically bounds the performance a model can achieve [161]. Consider $aX + b + \epsilon_1$ and $cX + b + \epsilon_2$ as the activation maps measured at the first and the repeat session respectively, where X is the group-average activation, b the individual-unique features and ϵ_1, ϵ_2 the measurement noise (and session variability) of the two sessions. The test-retest reliability of the residuals $b + \epsilon_1$ and $b + \epsilon_2$ is $\text{corr}(b + \epsilon_1, b + \epsilon_2) \approx \sigma_b^2/(\sigma_b^2 + \sigma_\epsilon^2)$,

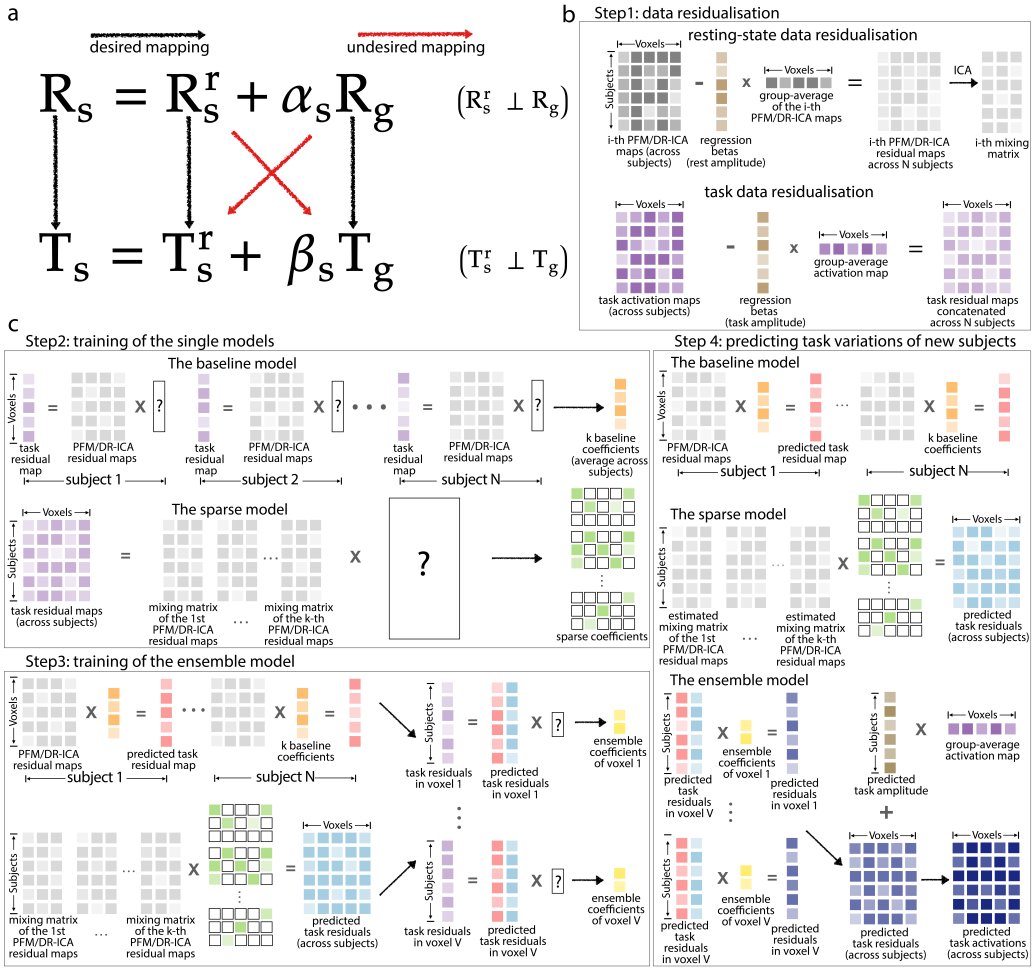

Figure 4.1: Model illustration. **(a)** The residualisation separates the features and targets into two (the group maps, denoted by subscript g and the subject residual, denoted by superscript r). The mapping between rest R_s and task T_s for subject s contains desirable and undesirable mappings, in the sense that we do not want to optimise, e.g., for the individual subject features to map onto the group maps. **(b)** Step 1. Residualisation of the resting-state modes and task contrast maps. The residualised resting-state maps were further ICA-reduced as the input of the sparse model. **(c)** Step 2. Training of the baseline and sparse model: per training subject, the baseline model yielded k reconstruction coefficients (one coefficient per map), which were averaged across subjects as the final baseline coefficients (orange). Next, the resting-state residual maps and the task residual maps were concatenated across subjects accordingly and then reduced to lower dimensions via ICA. The sparse model was trained on the (ICA-reduced) across-subject residual matrices to give the sparse regression coefficients (green). Step 3. the estimated baseline coefficients and sparse coefficients were applied to the training subjects to get the baseline-model-fitted (pink) and sparse-model-fitted (light blue) task residual maps. Next, for each voxel across subjects, we estimated the ensemble coefficients (yellow) by fitting another linear regression model with the baseline-model-fitted activations and the Lasso-fitted activations (in the given voxel) as the two regressors. Step 4. These three sets of coefficients were finally applied to the test subjects to make new predictions (dark blue).

where σ_b corresponds to the variability of the signal and σ_ϵ corresponds to the variability due to the measurement noise. Suppose the predicted activation is $\hat{a}X+\hat{b}$; the prediction accuracy of the residuals $\text{corr}(b+\epsilon_1, \hat{b})$, similarly, can be written as $\text{cov}(b+\epsilon_1, \hat{b})/\text{std}(\hat{b})\sqrt{\sigma_b^2+\sigma_\epsilon^2}$. It can be easily seen that when the prediction $\hat{b}$ is perfect (i.e., $b=\hat{b}$), the accuracy can be higher than the test-retest reliability due to the denoising effect. Therefore, by comparing prediction accuracy against its upper bound, we can show whether the model can fundamentally capture the individual differences in activation. However, if we use the un-residualised data to evaluate accuracy and reliability, the bias due to the group-average would make it difficult to establish a straightforward comparison between the test-retest reliability and prediction accuracy.

Hence, we built the model entirely on residualised data. The residualised resting-state functional modes and the residualised task contrast maps are also referred to as "resting-state variation maps" and "task variation maps" in the remainder of this chapter. To residualise the resting-state data, each of the k group-average (across training subjects) ICA spatial maps were regressed out from the corresponding individual DR-ICA maps for all subjects (i.e., a one-variable linear spatial regression per subject per dual-regression map), and similarly, each sPROFUMO group-level spatial map was regressed out from the same mode's individual-level sPROFUMO spatial maps (PFMs). These residualised spatial maps represent individual variations in resting-state activity, serving as features to predict individual variations in task-fMRI.

The task activation maps were residualised similarly for each individual, to give task variation maps. For a given task contrast, the group-average activation map was regressed out from the individual contrast maps (i.e., a simple linear regression per subject per contrast, with the group-average activation map as the regressor). These task variation maps describe individual differences in task-evoked brain activity that deviate from the typical activation patterns. Therefore the mapping between the rest and task states is entirely based on the variations rather

than on the original resting-state and task-evoked activity. See Figure 4.1b for an illustration of residualisation.

Finally, to compare this residualised model with the model trained on the original resting-state and task contrast maps (i.e., un-residualised), the task activations are predicted as a combination of the modelled variations and the average activation patterns. For both the rest and task data, we record the regression parameters as part of regressing out group-mean maps; these measures of overall "amplitude" are used later in the work, described below, and can of course be used (multiplied by the group-mean maps) to add the group-mean contribution back in where desired.

4.2.5 The ensemble learner

Our overall ensemble approach combines two separate models, "baseline" and "sparse". We start by describing these two individual models, and then go on to describe the ensemble method.

4.2.5.1 The baseline model

The baseline model assumes that, for a given task contrast, the individual task variations (i.e., residualised activation maps) can be represented by a linear combination of the variations in resting-state functional modes (i.e., residualised DR-ICA maps or PFMs). In this sense, the resting-state modes serve as a set of "bases" that span the task space. To obtain the reconstruction coefficients for each subject, we regressed the subject's resting-state bases into its spatial activation map (i.e., a multiple-regression per subject per task contrast, with resting-state variation maps as regressors and the task variation map as response). More specifically, suppose the number of voxels is V, and each individual has k number of bases (i.e., there are k group-average ICA spatial maps); to find the reconstruction coefficients of a specific task contrast map $\mathbf{y}_j$ (a $V \times 1$ vector) for a given subject j, we regress the given subject's k resting-state variation maps, denoted by a $V \times k$ matrix $\mathbf{X}_j = [\mathbf{x}_j^1, \mathbf{x}_j^2, ..., \mathbf{x}_j^k]$, into the task variation map $\mathbf{y}_j$ of this subject. As a standard

linear regression problem, the reconstruction coefficients $\boldsymbol{\beta}_j$ (a $k \times 1$ vector) of subject j minimise the following loss function

$$\hat{\boldsymbol{\beta}}_j = \underset{\boldsymbol{\beta}_j \in \mathbb{R}^k}{\operatorname{argmin}} ||\mathbf{y}_j - \mathbf{X}_j \boldsymbol{\beta}_j||_2^2 \tag{4.1}$$

for $j \in \mathcal{S}$, where $\mathcal{S}$ is the set of training subjects. The estimated reconstruction coefficients $\hat{\boldsymbol{\beta}}_j$ are given as $(\mathbf{X}_j^T \mathbf{X}_j)^{-1} \mathbf{X}_j^T \mathbf{y}_j$, where $\mathbf{X}_j$ is subject j's resting-state variation maps. These coefficients were averaged across the training subjects to give the final estimates of the reconstruction coefficients, i.e., $\hat{\boldsymbol{\beta}} = \frac{1}{N} \sum_{j \in \mathcal{S}} \hat{\boldsymbol{\beta}}_j$, where $|\mathcal{S}| = N$ is the number of training subjects.

To predict the activation map of an unseen subject l in a test set $\mathcal{T}$, we applied the reconstruction coefficients averaged from the training set to the subject's own resting-state variation maps $\mathbf{X}_l$, i.e.,

$$\hat{\mathbf{y}}_l = \mathbf{X}_l \hat{\boldsymbol{\beta}} \tag{4.2}$$

for $l \in \mathcal{T}$. Note that the baseline model is different from [19, 38, 39] in two ways. First, their models were primarily local, i.e., one linear regression per brain region, rather than a global linear regression for the whole brain. Second, with the group-average content regressed out from both the resting-state dual-regression maps and task activations, our baseline model aims to establish linear relationships between the variations of the two states (relative to the group-average) rather than the original resting-state and task activity (which is possibly dominated by the group average).

4.2.5.2 The sparse model

The baseline model has a few limitations. First, it has very few free parameters, resulting in one reconstruction coefficient per basis, which is then pooled (averaged) across all subjects. Crucially, each feature (spatial map) is associated with a single regression coefficient, regardless of which part of the brain is being modelled. Second, the coefficients learned from each training subject are estimated separately, which ignores patterns of between-subject variations. We re-formulated the problem in a more flexible, highly-parameterised framework, referred to below as the sparse model, with appropriate regularisation techniques to protect against too much flexibility.

First, to create feature maps that contain information of cross-subject variability, each of the k resting-state variation maps is first concatenated across the set of training subjects $\mathcal{S}$, yielding one $N \times V$ resting-state matrix per group-average spatial map, with a total of k such matrices. Denoting the i-th matrix $\tilde{\mathbf{X}}^i_{\mathcal{S}} = [\mathbf{x}^i_1, \mathbf{x}^i_2, ..., \mathbf{x}^i_N]^T$, where $\mathbf{x}^i_j$ is the i-th resting-state variation map of subject j, we then dimensionality-reduce these matrices into a set of d components using ICA: $\tilde{\mathbf{X}}^i_{\mathcal{S}} = \mathbf{A}^i_{\mathcal{S}}\mathbf{S}^i$ for $i = 1, 2, ..., k$. Following this decomposition, $\mathbf{A}^i_{\mathcal{S}}$ is an $N \times d$ mixing matrix and $\mathbf{S}^i$ is a set of d independent components representing common spatial variations across the training subjects $\mathcal{S}$. The mixing matrices of each ICA contains "coordinates" of each individual in the resting-state space spanned by these common modes, providing profiles of the resting-state variabilities of these individuals. The k mixing matrices are concatenated to give a single reduced variation matrix $\mathbf{A}^{\text{rest}}_{\mathcal{S}}$ as the final predictors, where $\mathbf{A}^{\text{rest}}_{\mathcal{S}} = [\mathbf{A}^1_{\mathcal{S}}, \mathbf{A}^2_{\mathcal{S}}, ..., \mathbf{A}^k_{\mathcal{S}}]$ is an $N \times dk$ matrix.

Likewise, the task variation maps (residualised activation maps) are concatenated across the training subjects $\mathcal{S}$, resulting in an $N \times V$ task variation matrix $\mathbf{Y}_{\mathcal{S}} = [\mathbf{y}_1, \mathbf{y}_2, ...\mathbf{y}_N]^T$ per contrast. The reduced resting-state variation matrix ($\mathbf{A}^{\text{rest}}_{\mathcal{S}}$) will be used to predict the concatenated task variation matrix. Under this formulation, the model has a large number of potential predictors. To prevent over-fitting, we enforce sparsity on the prediction regression coefficients, to enable selection of the subset of features that are most desirable for prediction. In addition, given that predictions made on the original task matrix are not only computationally expensive but also involve many redundant and noisy features (which will likely compete with the "real" features in the training), we also consider to similarly decompose the task matrix into a set of p independent components, i.e., $\mathbf{Y}_{\mathcal{S}} = \mathbf{A}^{\text{task}}_{\mathcal{S}}\mathbf{S}^{\text{task}}$, where $\mathbf{A}^{\text{task}}_{\mathcal{S}}$ is the $N \times p$ mixing matrix, and $\mathbf{S}^{\text{task}}$ is the set of p independent components. Thus, both the features matrix ($\mathbf{A}^{\text{rest}}_{\mathcal{S}}$) and the regression target used in training ($\mathbf{Y}_{\mathcal{S}}$) are sparse, low-rank versions of their original versions (through ICA), and contain information on individual variations (through concatenation of subjects).

To find the sparse coefficients $\mathbf{W}$, we solve the following regularised regression problem on the ICA-reduced task matrix $\mathbf{A}_{\mathcal{S}}^{\text{task}}$

$$\hat{\mathbf{W}} = \underset{\mathbf{W}=[\mathbf{w}_1,\mathbf{w}_2,...,\mathbf{w}_p]\in\mathbb{R}^{dk\times p}}{\operatorname{argmin}} \{||\mathbf{A}_{\mathcal{S}}^{\text{task}} - \mathbf{A}_{\mathcal{S}}^{\text{rest}}\mathbf{W}||_F^2 + \sum_{i=1}^{p}\lambda_i||\mathbf{w}_i||_1\} \tag{4.3}$$

or on the original task maps $\mathbf{Y}_{\mathcal{S}}$

$$\hat{\mathbf{W}} = \underset{\mathbf{W}=[\mathbf{w}_1,\mathbf{w}_2,...,\mathbf{w}_V]\in\mathbb{R}^{dk\times V}}{\operatorname{argmin}} \{||\mathbf{Y}_{\mathcal{S}} - \mathbf{A}_{\mathcal{S}}^{\text{rest}}\mathbf{W}||_F^2 + \sum_{i=1}^{V}\lambda_i||\mathbf{w}_i||_1\} \tag{4.4}$$

where the Lasso penalty is univariately applied to columns of $\mathbf{W}$ (with different hyper-parameters to allow differential amounts of regularisation), encouraging it to be element-wise sparse. Note that an alternative way of introducing sparsity is to use an $L_{1,2}$ penalty on $\mathbf{W}$ that enforces row-wise sparsity, as commonly applied in the grouped Lasso and the multivariate Lasso. That strategy would permit simultaneous use of all outputs to estimate a sole regularisation parameter. It implicitly assumes that predictions of different outputs (columns of $\mathbf{A}_{\mathcal{S}}^{\text{task}}$ or $\mathbf{Y}_{\mathcal{S}}$) tend to require the same set of features. This underlying assumption of row-sparsity penalty is not very appropriate and tends to require heterogeneous feature selection. Other alternatives that simultaneously use all outputs include Partial Least Squares (PLS), Canonical Correlation Analysis (CCA), and their variants, as well as a range of multi-task learning approaches. Given that multi-task learning approaches with sparsity regularisations usually have more complex behaviours than the pure Lasso, we simply choose the Lasso penalty, which is also particularly easy to parallelise across columns of $\mathbf{A}_{\mathcal{S}}^{\text{task}}$ or $\mathbf{Y}_{\mathcal{S}}$ (i.e., across task voxels).

To predict task variation maps for a set of unseen subjects, denoted by $\mathcal{T}$, we first need to translate the subjects' resting-state variations into the subspace spanned by the resting-state common modes (decomposed from the training subjects). This is conducted by regressing each across-subject basis matrix onto the corresponding set of resting-state common modes. Again, suppose the i-th (across-subject) resting-state variation matrix of the test subjects is denoted by an $n \times V$ matrix $\tilde{\mathbf{X}}_{\mathcal{T}}^{i}$,

where $\mathcal{T}$ is the test set, and $n = |\mathcal{T}|$ the number of test subjects. We seek to solve the linear regression problem

$$\hat{\mathbf{A}}^i_{\mathcal{T}} = \underset{\mathbf{A}^i_{\mathcal{T}} \in \mathbb{R}^{n \times d}}{\operatorname{argmin}} ||\tilde{\mathbf{X}}^i_{\mathcal{T}} - \mathbf{A}^i_{\mathcal{T}} \mathbf{S}^i||^2_F \tag{4.5}$$

where $\hat{\mathbf{A}}^i_{\mathcal{T}}$ is the estimated "mixing matrix" of the i-th resting-state variation matrix across the test subjects $\mathcal{T}$, and $\mathbf{S}^i$ is the independent components calculated from the training subjects for $i = 1, 2, ..., k$. Next, the sparse coefficients $\hat{\mathbf{W}}$, estimated via (3) or (4), are applied onto the concatenated variability profiles $\hat{\mathbf{A}}^{\text{rest}}_{\mathcal{T}} = [\hat{\mathbf{A}}^1_{\mathcal{T}}, \hat{\mathbf{A}}^2_{\mathcal{T}}, ..., \hat{\mathbf{A}}^k_{\mathcal{T}}]$ (an $n \times dk$ matrix), to give predictions for the set of unseen subjects $\mathcal{T}$

$$\hat{\mathbf{Y}}_{\mathcal{T}} = \hat{\mathbf{A}}^{\text{rest}}_{\mathcal{T}} \hat{\mathbf{W}} \mathbf{S}^{\text{task}} \tag{4.6}$$

if $\hat{\mathbf{W}}$ is solved via (3) or

$$\hat{\mathbf{Y}}_{\mathcal{T}} = \hat{\mathbf{A}}^{\text{rest}}_{\mathcal{T}} \hat{\mathbf{W}} \tag{4.7}$$

if $\hat{\mathbf{W}}$ is solved via (4).

This completes the specification of the sparse model. To summarise the approach, we use concatenation of training subjects to incorporate information on subject variability in the training, we apply ICA to sparcify this data to help with fitting, and we employ further regularisation via the Lasso cost function on the regression coefficients. For UKB, we chose to reduce each across-subject resting-state matrix to 3,000 independent components and the task matrix to 4,000 independent components (however, reducing the resting-state and task matrices to 1,000 independent components yields comparable results, see Figure B1). For HCP, in contrast, we chose to reduce each resting-state matrix to its full rank (i.e., number of the training subjects, around 900 in each fold) but kept the original spatial dimension of each task matrix (i.e., no ICA on the task matrix), which yielded the best performance on a left-out subset (see Figure B2).

4.2.5.3 The ensemble model

A single model usually represents a single hypothesis space of the particular prediction problem. Although the single models may contain the hypothesis space already well-suited for a specific problem, combining multiple hypotheses allows for more flexible structures to exist between predictors and response variables, and can potentially improve model performance (again, as long as over-fitting is avoided through correct use of, for example, cross-validation or left-out data). Here the two single models are tailored to different aspects of the underlying hypothesis that variations in resting-state activity can inform task variations. As mentioned above, the baseline model treats the resting-state variation maps as a set of "bases" that spans the task variation map for each individual. It is obvious that the baseline model assumes that the mapping between resting-state and task space is within-subjects and thus ignores between-subject patterns of variations which might also be useful for the predictions. The underlying hypothesis of the sparse model captures a different aspect, though closely-connected with the baseline model hypothesis. With the variation maps reduced to the corresponding subspaces, the sparse model assumes that the "coordinates" of the subjects in resting-state space can be translated into their "coordinates" in task space.

Here we aggregate the predictions of each single model, to give the final prediction for unseen subjects using simple linear regression. Suppose $\hat{\mathbf{Y}}_{\mathcal{S}}^{\text{baseline}}$ is the $N \times V$ baseline-model-fitted activations of the training subjects $\mathcal{S}$, and $\hat{\mathbf{Y}}_{\mathcal{S}}^{\text{sparse}}$ is the sparse-model-fitted maps. Particularly, we use $\hat{\mathbf{y}}_{\cdot i}^{\text{baseline}}$ and $\hat{\mathbf{y}}_{\cdot i}^{\text{sparse}}$ to denote the fitted activations in voxel i across subjects (i.e., each is an $N \times 1$ vector). At the ensemble stage, we aim to find the coefficients for each constituent model by column-wisely fitting a simple linear regression on the task matrix of training subjects $\mathbf{Y}_{\mathcal{S}}$, i.e.,

$$\hat{\theta}_i^{(1)}, \hat{\theta}_i^{(2)} = \underset{\theta_i^{(1)}, \theta_i^{(2)}}{\operatorname{argmin}} ||\mathbf{y}_{\cdot i}^{\mathcal{S}} - \theta_i^{(1)} \hat{\mathbf{y}}_{\cdot i}^{\text{baseline}} - \theta_i^{(2)} \hat{\mathbf{y}}_{\cdot i}^{\text{sparse}}||_2^2 \tag{4.8}$$

for the "true" activations in voxel i across the N training subjects, denoted by $\mathbf{y}_{\cdot i}^{\mathcal{S}}$, for $i = 1, 2, ..., V$. The two coefficients, $\theta_i^{(1)}$ and $\theta_i^{(2)}$, will then be applied to the baseline-model-predicted and sparse-model-predicted maps to yield predictions of

task variations for the unseen subjects. See Figure 4.1 for an illustration of model training and Table B2 for a summary of the notations.

For UKB, the ensemble model and its constituent base models were trained and tested on 17,560 subjects (3-fold cross-validation); for HCP, the models were trained and tested on 991 subjects (10-fold cross-validation). The hyper-parameter of the L_1 penalty was optimised within each fold's training data via nested cross-validation (3-fold). The other free parameters (e.g., the number of resting-state bases and the number of independent components in the sparse model) were determined on a different subset of 4,700 subjects for the UKB dataset (trained on 4,000 and tested on 700). Due to the limited number of HCP subjects, we randomly selected 10% of the HCP subjects and investigated how the choice of these parameters would affect the model.

4.2.5.4 The amplitude model

The amplitude model aims to predict the task activation amplitude for each individual (i.e., the beta coefficients from regression against the group-average activation map, as recorded during residualisation, one scalar value per subject per contrast) using the resting-state amplitude (i.e., the beta coefficients from regression of resting maps against the group-mean dual-regression maps, one scalar value per subject per basis). There are a few reasons for incorporating a separate amplitude estimation. First, one important source of individual variabilities in task-evoked activity is the (overall) activation amplitude. Explicitly predicting this information may help capture a different kind of individual variability that cannot be fully modelled by the aforementioned spatial models (indeed we would not expect to capture this from the residualised predictions). Second, the final predictions for test subjects are ideally given as a combination of modelled residual variations and the typical activation patterns. In order to recover the activation maps from the variation maps for each individual, the group-average activations are yet to be added back in appropriately, scaled by the activation amplitude of the specific individual. However, the activation maps of the test subjects are of course not seen

during training. Therefore, we are not able to estimate the activation amplitude (i.e, betas from task residualisation) by simply regressing the group-average activations into the inidividual activations. As an alternative, the resting-state amplitude may be predictive of the overall activation amplitude (Figure B3 and B4) and thus may serve as a substitute for this information. The surrogate activation amplitude was generated as follows. Remembering that each subject has k resting-state amplitude values, corresponding to each of the k group-average spatial maps (i.e., one amplitude value per map): for a given contrast, a multiple linear regression model with the activation amplitude as the response and the resting-state amplitude as the predictors was trained across subjects (3-fold cross validation on UKB; 10-fold cross validation on HCP). These surrogate activation amplitude are subsequently applied to the predicted variation maps as the new beta coefficients, such that the re-scaled group-average effects can be added back in accordingly.

The other hypothesis about the overall activation amplitude is that it serves as another important source of individual variabilities. To explore this possibility, we also consider to incorporate the amplitude information into the ensemble stage to test whether it can further improve model performance. Given that the k resting-state amplitude values of the k sets of dual-regression maps are correlated (across subjects), we reduce the k amplitude features into a few principal components, the number of which are determined via cross-validation. These components are included in the ensemble model as additional predictors to predict each column (voxel) of $\mathbf{Y}$.

4.2.6 Measures of model performance

Assessment of model performance is primarily based on Pearson correlations between predicted maps and the actual maps (in subjects left out of the training process). Apart from the standard MNI152 brain mask applied at the beginning of all the analysis, we choose not to apply further thresholding of the resulting maps. Although further masking of the images may emphasise certain regions that are more of interest, the choice of thresholds can have a complex impact on evaluation and requires caution.

For a given task contrast, the predicted maps are correlated with the actual maps for all subjects, yielding a "subject by subject" correlation matrix, where the entry in the i-th row and j-th column corresponds to the correlation between subject i's predicted map and subject j's actual map. The mean of the diagonal elements measures the overall prediction accuracy, i.e., how well the model can reproduce the spatial patterns of activation for each subject, averaged across subjects. However, this measure cannot fully quantify model performance because the overall model accuracy can be boosted by simply reproducing the group-average activation, particularly when most subjects are "normal", having activations patterns close to the group-average. Therefore, it is also important to make differentiated predictions, i.e., how well the model can capture atypical variations that deviate from the group-average activations. This necessitates measuring the extent to which, for a specific subject, the model can make predictions that are closer to the subject's own activation maps than to the others. This is of course particularly relevant if doing non-residualised prediction.

The new evaluation measure is calculated as follows: after the correlation matrix (between predicted maps and the actual maps for all subjects) is normalised via Fisher's transformation, for each subject, we calculate the difference between two values: (i) correlation between the subject's predicted map and the subject's actual map; (ii) mean of the correlations between the subject's predicted maps and other subjects' actual maps. The difference between (i) and (ii) provides a quantitative evaluation of the model's capability of predicting individual differences distinct from the group mean. In the following text, the first measure is referred to as "prediction accuracy", and the second one is referred to as "prediction discriminability".

Additionally, we calculated the between-subject standard deviation map of the actual task variations (as a measure of inter-individual voxel-wise variability) and also of the predicted variations (as a measure of predicted variability) for each contrast. We then correlated the predicted variability maps against the actual variability map as a third measure of model performance. A higher correspondence

between the two standard deviation maps indicates better ability to reproduce the spatial pattern of between-subject variability.

4.3 Results

4.3.1 The ensemble model outperforms its constituent single models

To compare DR-ICA maps with PFMs, we chose the optimal dimensionality of each method, DR-ICA25 and PFM50 for UKB, and DR-ICA50 and PFM150 for HCP, respectively. The fact that PFM optimal dimensions were found to be higher than those of DR-ICA suggests that the former yielded more reliable functional modes particularly at higher dimensions. (However, note that PFMs consistently outperformed DR-ICA across all dimensions. See Figure B5). In the baseline model, overall, most variation maps contributed to the predictions (Figure B6), suggesting that these resting-state variation modes did capture a significant proportion of the variance in task variation maps.

We found that the sPROFUMO modes had overall higher accuracy in predicting task variations than the DR-ICA maps, consistently across the baseline, sparse, and ensemble model (Figure 4.2a and 4.2b). Compared with predictions based on DR-ICA, the biggest improvement introduced by sPROFUMO modes was evident from the baseline model, suggesting that sPROFUMO provides a fundamentally better set of resting-state bases to reconstruct task variations than DR-ICA. This corresponds with previous evidence that sPROFUMO better accounts for cross-subject misalignment and accommodates higher predictive power of population heterogeneity [150]. Additionally, sPROFUMO modes also exhibited higher prediction accuracy for the sparse and ensemble model. Interestingly, the baseline and sparse model based on DR-ICA had very distinct performance on the two datasets. For HCP, the baseline model yielded higher prediction accuracy than the sparse model (Figure 4.2b, blue and green), while for UKB, this relationship was entirely reversed (Figure 4.2a, blue and green). Here we provide a possible explanation for this discrepancy. The two single models are tailored to different data scenarios. If the resting-state modes form

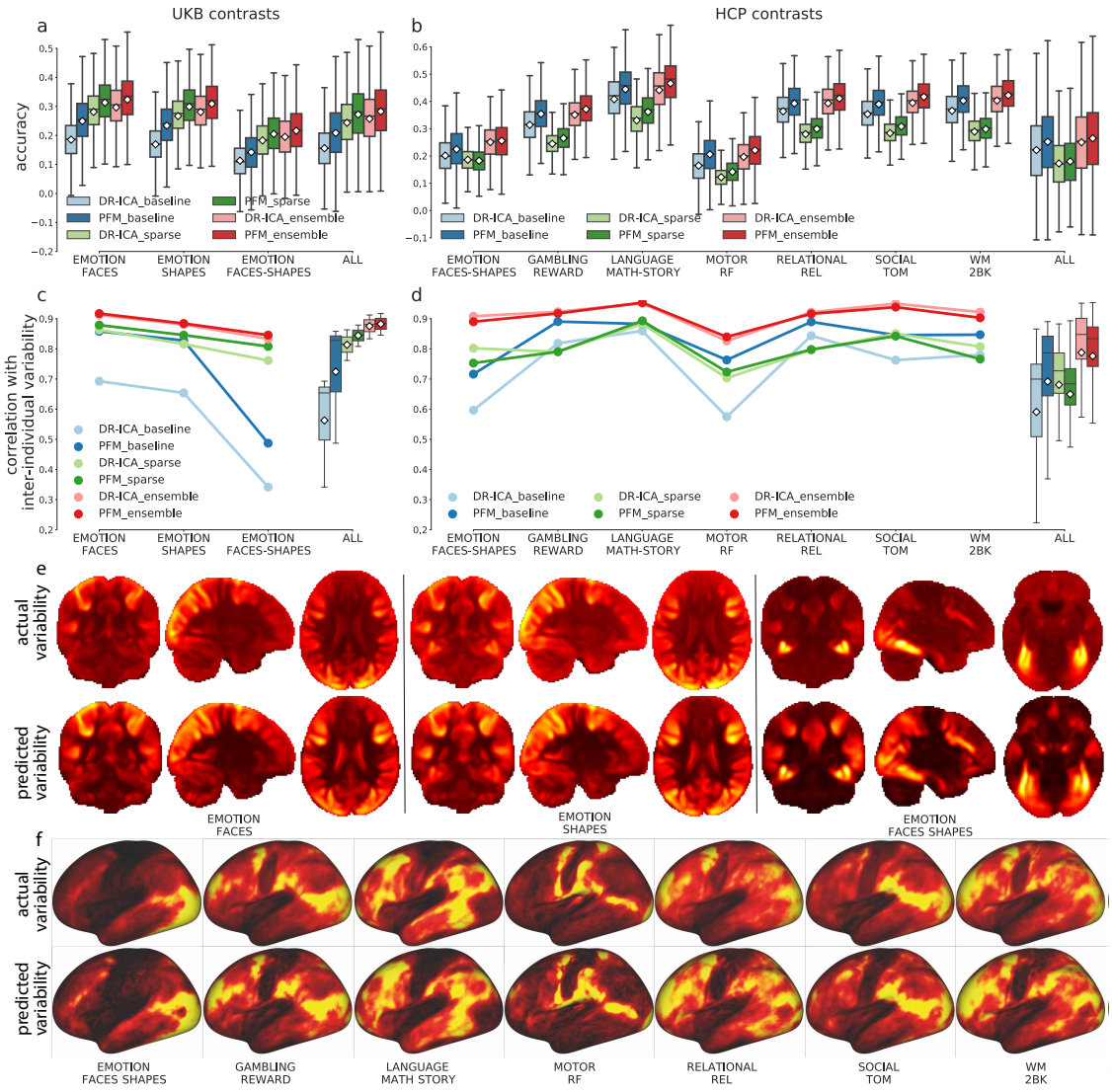

Figure 4.2: Prediction accuracy of the individual task variations and of the inter-individual variability. PFM better captures task variations than DR-ICA maps (dark colors vs light colors); the ensemble model outperformed its constituent single models in predicting individual task variations and reproducing inter-individual variability patterns (blue, green, and red). **(a)** Prediction accuracy of the baseline, sparse, and ensemble models for 17,560 UKB subjects across the three contrasts, the last columns showing all contrasts pooled together. White diamonds show the means along with the boxplots showing the medians and quartiles. **(b)** Equivalent plots of 991 HCP subjects across seven representative contrasts, the last column showing all 47 contrasts pooled together. The statistical tests for (a) and (b) can be found in Table B3, all significant after being Bonferroni-corrected. **(c)** Correlations between the predicted and the actual inter-individual variability maps calculated across 17,560 UKB subjects. Overall, ensemble trained on PFM yielded the highest correspondence with the inter-individual variability. **(d)** Equivalent plots across 991 HCP subjects. See Figure B7 for all HCP contrasts. The statistical tests for (c) and (d) can be found in Table B4. **(e)** The actual (first row) and the predicted (second row) inter-individual variability across 17,560 UKB subjects of the three contrasts, shown volumetrically. Warmer colors indicate higher variability with the maximum normalised to 1. **(f)** The actual (first row) and the predicted (second row) inter-individual variability calculated across 991 HCP subjects of the seven representative contrasts, shown on the cortical surface.

the set of bases that do fundamentally have the ability to predict the task maps, then the baseline model should suffice, i.e., we do not need the sparse model to emphasise specific spatial features. In practice, however, DR-ICA maps are not the perfect sets of individual "versions" of the group-average modes, containing many noisy voxels irrelevant to task-fMRI prediction. A major difference between the two datasets is that the UKB data we used to train the model is volumetric while the HCP data is grayordinates. As a consequence, there is more functional spatial variability (misalignment) in the UKB data [162] and hence more "errors" in its individual dual-regression maps. In addition, HCP data is MSMAll-aligned and UKB is not. On the other hand, sPROFUMO better accounts for cross-subject misalignment and allows more fine-grained delineation of individual differences in resting-state data, thus it has improved ability to capture variations in task data. Furthermore, due to the shorter scanning sessions, the resting-state and task-fMRI scans in UKB have higher noise than in HCP, requiring additional benefits of identifying which voxels/spatial features are more desirable in the modelling. Hence, UKB requires greater spatial modelling complexity as well as greater spatial smoothing, provided by the sparse model (note that conducting ICA on the resting-state and task matrices across subjects in the sparse model may serve as a kind of de-noising).

For both datasets, overall, the ensemble model outperformed its constituent single models. Remember that the task variations are the residuals of regressing the group-average activations into the individuals, thus they are orthogonal to the group-mean by design. This also implies that these task variation maps have minimal overall cross-subject similarity, i.e., the spatial correlations between pairs of subjects fluctuate around zero. Therefore, the plots of prediction accuracy and of discriminability will look almost identical, because the predicted maps will have near-to-zero correlations with the maps of the other subjects, i.e., the off-diagonals of the (subject by subject) correlation matrices (between the predicted maps and the actual maps) are all close to zero (Figures B9 and B10). The discrimination metric for prediction accuracy on the residuals can be found in Figure B8.

In addition to predicting the individual variations in task activity, all three models could reproduce the spatial pattern of inter-individual variability (standard derivation maps across subjects) for both datasets (Figure 4.2c and 4.2d). Similar to the previous scenario, using sPROFUMO modes as bases improved the prediction of inter-individual variability for the baseline model on both datasets (Figure 4.2c and 4.2d, blue), corroborating the conclusion that sPROFUMO better aligns the subjects, refines the spatial details of cross-subject heterogeneity, and thus provides a better set of bases to reconstruct task variation space. In terms of the sparse and ensemble model, DR-ICA and sPROFUMO yielded comparable correspondence with the true inter-individual variability.

Finally, we also calculated the subjects' identification accuracy, i.e. the probability that predicted maps had the highest correlation with the subjects' own residual maps, for each task contrast. The subject identification accuracies are all close to 100% (Figure B9 and B10), suggesting that the predicted variability for a given subject best corresponds with the subject's own variability patterns than with the others. These results demonstrate that resting-state variations can capture well the inter-subject differences in task-evoked brain activity. These actual and predicted (via the ensemble model) inter-individual variability maps are shown in Figure 4.2e and 4.2f. Regions of higher variability across subjects are those more involved in the corresponding task execution. For example, somato-sensory and motor regions are more variable across subjects in the motor contrasts; fronto-parietal regions exhibit higher variability in more cognitive contrasts; the visual areas tend to be more variable in general, for all contrasts. In summary, all three models are able to capture individual-unique activation patterns that deviate from the typical activation patterns as well as recapitulating the spatial pattern of inter-individual variability. In the subsequent analysis, we used PFM50 for UKB, and PFM150 for HCP.

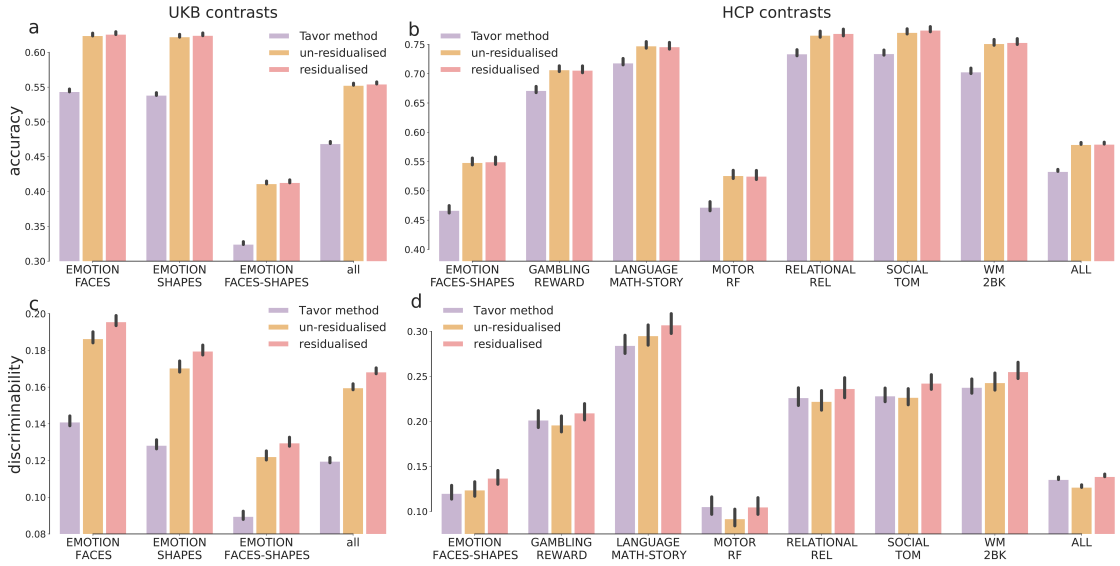

Figure 4.3: Comparison between the Tavor model and the ensemble models. Overall, the ensemble model trained on variation maps (residualised maps) outperformed the other two options; error bars show the 95% CI of the means. (a) Prediction accuracy across 17,560 UKB subjects of the three contrasts, the last column showing all contrasts pooled together. (b) Equivalent plots across 991 HCP subjects of seven representative contrasts, the last column showing all 47 contrasts pooled together. (c) Prediction discriminability in UKB. (d) Prediction discriminability in HCP in a subset of task contrasts (see Figure B13 for all HCP contrasts). The statistical tests can be found in Table B5.

4.3.2 Training on the un-residualised data is suboptimal to capture individual differences

The next question we asked is whether the model can recover individual idiosyncrasies in task-fMRI, if trained on the un-residualised resting-state spatial modes and the task activations, as opposed to the residualised data (i.e., variation maps). Having close-to-zero shared variance with the group-average, the residuals more accurately profile the individual differences by design; we posit that training on residuals avoids the contamination of group-level information and thus may potentially facilitate capturing individual-unique patterns. To fairly compare the two options requires recovering the actual task-evoked responses (as opposed to the residuals) from the predicted variations for each individual. To explore this, we next generated the surrogate activation amplitude using the PFMs' amplitude for each individual, then added the group-average activation map (scaled by the resting-state-

predicted amplitude) back to the predicted variation maps. These predictions with group-average added back in were correlated against the actual (un-residualised) activations for all subjects, again yielding a subject by subject correlation matrix per contrast. We calculated the prediction accuracy and discriminability from these correlation matrices, and compared them with the model trained on the un-residualised PFMs and task activations.

Overall, both options manifested considerable predictive power of individual activations, as suggested by the overall accuracy and discriminability (Figure 4.3, red and orange). Additionally, we found that although training on variations exhibited little improvement on the actual prediction accuracy (Figure 4.3a and 4.3b), it tended to improve prediction discriminability (Figure 4.3c and 4.3d). This suggests that it is more desirable to establish a mapping between the variations in rest and task data *per se* than simply use the original data with group-average effects present. This is probably because residualisation orthogonalises the individual maps with respect to the group-average maps and prevents the dominance of the typical activation patterns. Furthermore, this shows that separating out the modelling of overall amplitude from (group-mean-removed) map variability, and then recombining these parts of the model later, is at least as effective as predicting raw task from raw resting maps. This is valuable, as it does suggest that these different data aspects can indeed be considered separately. The subject identification accuracies based on residualised predictions (with group-average effects added back in for evaluation) are again close to 100% for most contrasts, shown in Figure B11 and Figure B12.

We also benchmark our model against previous GLM-based methods [19] using the same subjects. The Tavor method is based on multiple GLMs, essentially very similar to the baseline model, except for a few differences: (1) instead of training a global GLM for the whole brain between the resting-state and the task maps (as in our baseline model), the Tavor model seeks to fit multiple "local" GLMs within each of the pre-determined parcels; (2) the features of the Tavor model are seed-based connectivity maps, while our baseline model uses the dual-regression maps (i.e., multiple regression against the many "seed" timeseries output

by the first stage of dual-regression). The ensemble model, trained either on the un-residualised data or on the variation maps, yielded higher prediction accuracy than the Tavor method. On the UKB dataset, the ensemble model substantially improved prediction accuracy and discriminability; on the HCP dataset, the Tavor method and the ensemble model trained on variations manifested comparable discriminability, both superior to the ensemble model trained on un-residualised data (see Figure B13 for all HCP task contrasts). We also compared our approach with other studies [38, 39] that also utilised the ensemble technique. Cohen et al. 2020 showed that using Random forest bootstrap aggregation enhanced prediction accuracy, benchmarked against the linear regression approach used in Tavor et al. 2016. Dohmatob et al. 2021 used random parcellations to improve prediction accuracy. Instead of using a single parcellation scheme to predict local patches of activations (one ridge regression per patch, concatenated afterwards), they averaged the predictions based on random parcellations as the final prediction. We show that our ensemble approach outperforms the others, particularly on the UKB dataset (Figure B14). Note that, among the HCP contrasts, motor-tasks exhibited weak prediction discriminability. A possible explanation for this is that the individual response profiles to motor-related stimulus had little cross-subject variation, such that the model was not able to extract sufficient information to discriminate between subjects. The relatively lower prediction accuracy of motor tasks is, on the other hand, unexpected, especially considering the strong activations in cortical regions that are supposed to enable the model to learn the mapping between resting-state and motor tasks. Understanding this discrepancy between motor tasks and resting-state activity requires future investigations and would be important to understand the ongoing interplay of resting-state networks in task execution. Here we provide possible explanations for the poor accuracy of motor-related tasks. Motor task activations often pose unique challenges due to their inherently dynamic nature compared to more static cognitive tasks, making prediction from resting-state data, which is relatively static, more complex. Additionally, the influence of non-neural factors such as muscle activity, motion artifacts, and physiological responses,

which might not be adequately captured or differentiated in the resting-state data, could further complicate the prediction accuracy. These elements contribute to the observed lower prediction performance for Motor network-related tasks.

The fact that the model trained on the variations *per se* (with an explicit and separate amplitude prediction) can better capture patterns unique to individuals than its un-residualised counterpart corroborates the assumption that, in addition to the spatial layout (shapes) of activations, the overall activation intensity may also contribute to the variability of task-evoked brain activity. Following this, we also tested whether incorporating resting-state amplitude as additional predictors explicitly at the ensemble stage would further facilitate capturing individual-unique patterns for the un-residualised model. We found that, though having little effect on the actual prediction accuracy (Figure B15a), including the PFMs' amplitude as explicit predictors (in addition to the other two predictors, the baseline-model predicted and sparse-model-predicted values in the corresponding voxel) did further improve discriminability (Figure B15b), particularly on UKB. This improved discriminability for the un-residualised model (Figure B15b, gray bars), however, is still not as good as the discriminability of the residualised predictions (Figure B15e, orange bars). For the ensemble model trained on the residualised data, regressing out the group-average response "removes" the overall activation intensity relative to the group-average activations for each individual. Therefore, introducing resting-state amplitude to the residualised ensemble model, in theory, should have little effect on model performance. However, in practice, we found that incorporating resting-state amplitude as additional features in the ensemble stage also increased prediction discriminability for the residualised ensemble model (Figure B15e). There are a few possible explanations for this discrepancy. One possible explanation is that the group-average activation patterns were not entirely removed particularly from the subjects that are very atypical, probably due to GLM's sensitivity to outliers or noise in the fitting (e.g., related to regression dilution). In this sense, including resting-state amplitude as additional features thus accounted for the remnants of the amplitude information particularly for those atypical subjects, and thus

increased the overall prediction discriminability. Another possibility is that the overall activation intensity may still inform the (strength of the) variabilities of the shape of activations. This possibility can be partially validated by the finding that it further improved the fit with the spatial pattern of inter-individual variability by including resting-state amplitude as additional features at the ensemble stage (Figure B15c and B15f). Note that, however, the resting-state amplitude is not expected to be a perfect surrogate of the task amplitude, as R^2 between the actual and the predicted task amplitude is actually small (Figure B3 and B4).

4.3.3 Prediction accuracy paralleled test-retest reliability

To evaluate whether the predicted task contrast maps can reliably capture individual differences in task-evoked brain activity, we compared the accuracy of task predictions against their corresponding test-retest reliability, leveraging the repeat fMRI scans in UKB and HCP datasets. The test-retest reliability of task-fMRI was defined as the spatial correlation between the first-visit and repeat-session task contrast maps, for each subject and each contrast. For both datasets, the PFM-predicted contrast maps yielded higher overall accuracy than the repeat scans, consistent across all task contrasts (Figure 4.4b and 4.4c, light blue and light red), suggesting that resting-state predicted activations can surpass task-fMRI test-retest reliability. This agrees with previous studies that resting-state features serves as a reliable trait marker and may even be more heritable than task-fMRI phenotypes [12]. Note that, the accuracy of PFM-predicted task maps, which is on par with the test-retest reliability, is not a result of over-fitting to the first-visit task-fMRI data. To illustrate this, we also correlated the predicted task maps against the second-visit task contrast maps as "second-visit prediction accuracy" (Figure 4.4a). If the model is over-fit to the noise component or simply reflects some analytic circularity in the first-visit sessions, this second-visit prediction accuracy should be much lower than the first-visit accuracy and far below the test-retest reliability benchmark. However, we found that the second-visit accuracies were very close to the first-visit accuracies (Figure 4.4b and 4.4c, blue vs green). This suggests

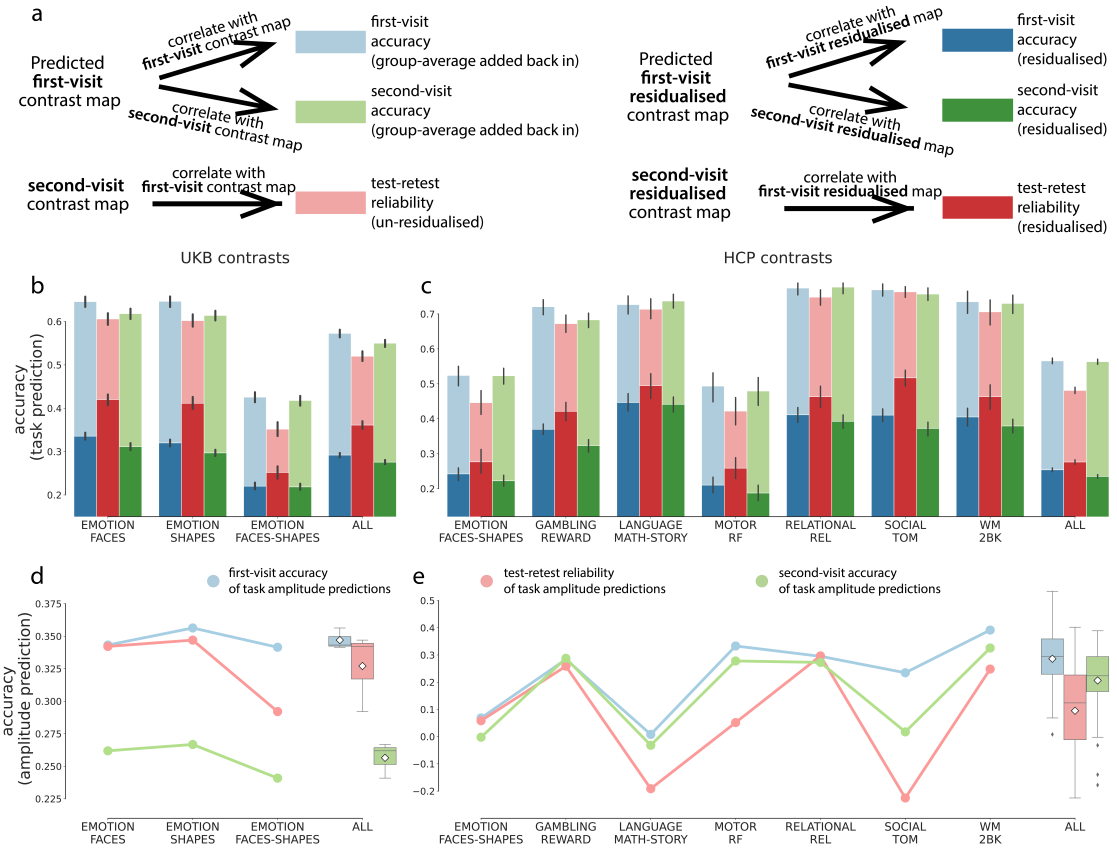

Figure 4.4: Test-retest reliability of PFM-predicted task maps. (a). Illustration of how first- and second-visit accuracy and test-retest reliability were calculated for the residualised and "group-average-added-back-in" predictions. **(b)** and **(c)**. Blue: prediction accuracy of first-visit contrast maps. Green: prediction accuracy of second-visit contrast maps. Red: test-retest reliability. Light colours denote the prediction accuracy of the un-residualised/group-average-back-in maps, dark colours the accuracy of the residualised maps. The last columns show all contrasts pooled together. Error bars show 95% CI of the means. **(d)** and **(e)**. For both datasets, the PFM-predicted task amplitude for the first-visit scans gave higher accuracy than the repeat scans (blue vs red). On HCP, the predicted amplitude also gave higher accuracy for the second-visit task amplitude than the corresponding test-retest reliability (red vs green). However, on UKB, this accuracy was worse (green vs red in Figure 4.4d), again suggesting that the UKB resting-state sessions are less reliable, possible due to the much shorter scanning time. The statistical tests can be found in Table B6 and B7. See Figure B18 and B19 for equivalent plots of (c) and (e) of all HCP contrasts.

that the model trained on a single session generalises well to the repeat sessions and is not over-fitting to in the first-visit sessions.

As mentioned in previous sections, predicting residuals is of more interest. The residualised task activation maps, with the group-average removed, consists of two components that explain the across-session variability, one the measurement

noise and the other the true individual-unique features. Thus, the test-retest reliability (correlation between sessions) of residuals measures the proportion of the true signal's variance relative to the noise variability, and is a fundamental criterion for the model performance [161]. A model that yields prediction accuracy comparable with this reliability is fundamentally capable of capturing the inter-individual differences in activation. On the HCP dataset, the prediction accuracy of residualised activations was close to the test-retest reliability of task residuals (Figure 4.4c) for most contrasts, and even gave higher accuracy for a few contrasts (GAMBLING_REWARD, GAMBLING_PUNISHMENT, SOCIAL_MATCH-REL, etc., see Figure B18). On the UKB dataset, however, the (residualised) repeat-session task-fMRI scans still yielded much higher accuracy than the PFM-predicted task variations (Figure 4.4b), possibly because of the much shorter resting-state scanning sessions. The repeat scans also had higher prediction discriminability than did the group-average-back-in predictions (Figure B17), which is un-surprising due to the noise ceiling effect.

We also investigated the reliability of task activation amplitude predictions by comparing the accuracy of predicted amplitude against the test-retest reliability of the activation amplitude. Similarly, the test-retest reliability of activation amplitude is defined as the correlation between first- and second-visit regression betas (i.e., the coefficients obtained by regressing the group-average task map out from the individual task contrast maps) across subjects. We tested whether resting-state-predicted amplitude is more robust than that measured directly in repeat-session task-fMRI scans. The PFM-predicted activation amplitude (using the first-visit resting-state scans) indeed proved more reliable than task-fMRI scans in replicating the actual activation amplitude (Figure 4.4d and 4.4e, blue vs red) for both datasets. Similarly, to exclude the possibility of over-fitting for the amplitude model, we also correlated this PFM-predicted amplitude against the second-visit actual amplitude (Figure 4.4d and 4.4e, green). For HCP, the predicted amplitude still gave higher accuracy for the second-visit data than the test-retest reliability (Figure 4.4e, green vs red). However, this no longer held for UKB (Figure 4.4d, red vs green), possibly

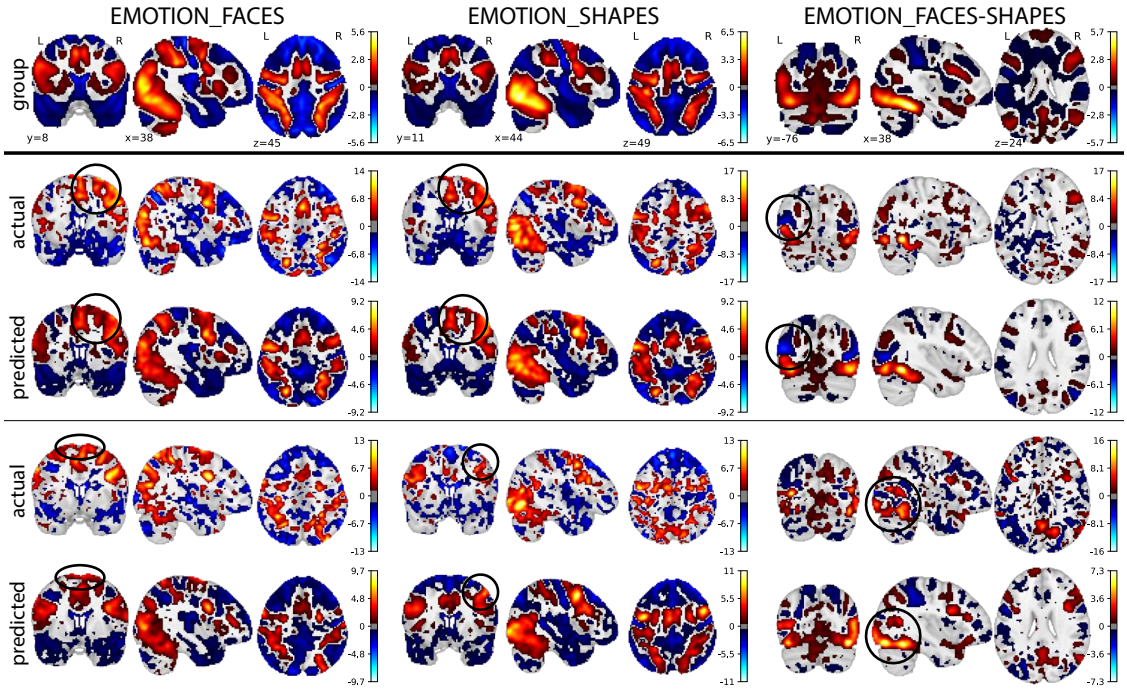

Figure 4.5: Predicted, actual, and group-average activations of 6 example UKB subjects. The predicted activations captured the atypical activations in individual subjects (with group-mean-related components added back in). The subjects ranked between 50% to 75% according to their correlations with the corresponding group-average activations. See Figure B20 for the plots of the predicted and the actual task variation maps of the same example subjects.

also due to the much shorter scanning time. Overall, these results did suggest that resting-state data is potentially a more stable trait marker than task-fMRI features, but this depends on the reliability of the resting-state scans.

Figures 4.5 and 4.6 show the comparison between the predicted, actual, and group-average activations volumetrically (for UKB) and on the surface (for HCP). It can be seen that the predicted activations provide a "smoothed" estimation of the individual activations, while preserving the unique patterns in individual subjects (for the actual and the predicted task variation maps of the same example subjects, see Figure B20 and B21).

4.4 Discussion

In this chapter, we extended previous GLM-based approaches [19, 38, 39] and proposed an ensemble learner to model individual variations in task activations

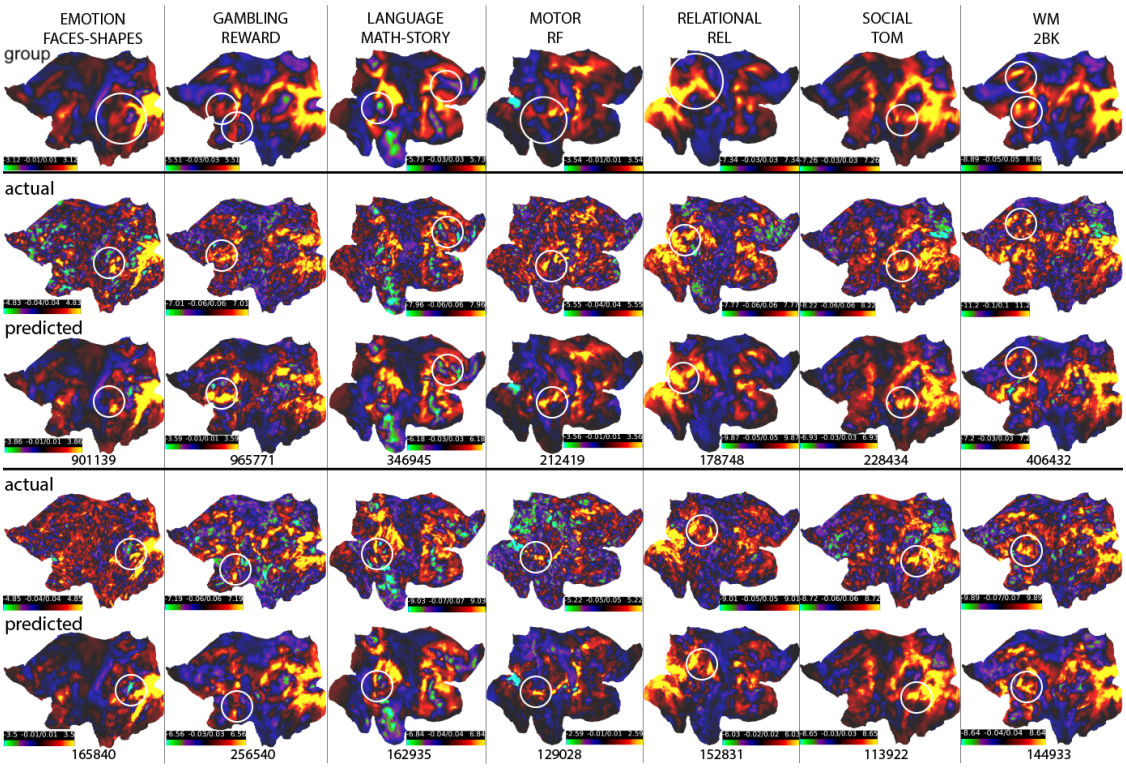

Figure 4.6: Predicted, actual, and group-average activations of example HCP subjects. The predicted activations captured the atypical activations in individual subjects; these subjects ranked in the lower 50% percentile according to their correlations with the corresponding group-average activations. See Figure B21 for plots of the actual and the predicted task variations maps of the same example subjects.

on two large datasets, UKB and HCP. Enabled by a recently developed technique, sPROFUMO, we exploited the richness of individual variability in resting-state to reproduce task-evoked activation patterns unique to individuals. We demonstrated that sPROFUMO can accommodate higher predictive power than DR-ICA, especially in terms of the overall capacity of reproducing between-subject differences. This added advantage of sPROFUMO arises from its enhanced ability to depict fine-grained resting-state variability in rich detail due to its bidirectional and hierarchical architecture between the group-average and individual, in contrast to the unidirectional group-average algorithms (e.g., DR-ICA). Furthermore, we showed that modelling the individual activation profiles as a combination of the group-average and predicted variations can be more desirable than simply modelling the raw task map, suggesting that two sources of task variability, shape and amplitude, factorise into different compartments and can be modelled separately. Characterising

different aspects of task variability is important to understanding the sources of these cross-subject differences. Overall, resting-state functional modes serve as a set of bases that can not only sufficiently reconstruct individual task-fMRI space but also yield more reliable localisation of individual task-evoked response profiles.

Our ensemble framework consists a baseline model and a sparse model, each tailored to a different scenario. In the baseline model, for each individual, the resting-state modes span the space of the task activation maps and thus, in theory, can reproduce task-fMRI in itself. In practice, however, more spatial complexity is often required to select local features that are "cleaner" or of more interest. The sparse model largely accounts for this limitation. For example, the motor network in resting-state modes contains components that are often in sync with each other and are part of the same spatial basis. The baseline model cannot split them, while the sparse model may select the components more desirable for prediction. However, the sparse model has a limitation. Despite the existing rescaling techniques (e.g., fitting another OLS on top of the selected features; introducing a re-scale factor), the Lasso penalty often introduces too much shrinkage, particularly when the prediction involves too many candidate features. As a result, the predicted response may become too biased towards zero, thus degrading model discriminability. The ensemble model, by fitting another OLS on top of each voxel, de-biases the over-shrinkage of the sparse model.

It is important to consider our approach in the context of other frameworks, for example, the activity flow approach [37, 136, 163], which also investigated the idea of residualisation. It is a different type of framework in that it did not establish a supervised model trained to relate resting-state data to task-fMRI; instead, it sought to explicitly build the underlying mechanisms that link functional connectivity to task activations by testing whether such a mechanistic construct can capture task activations better than chance. The two types of frameworks are of different nature but both are able to reproduce individual task-fRMI patterns. They are tailored to different purposes. The activity flow approach provided an explanation for the observed rest-state and task-fMRI relationships and tested the

corresponding hypothesis against chance level. It may help us better understand the intrinsic mechanism or biology underlying functional connectivity and task activations. Our framework is a supervised model, which may give more accurate predictions of individual activations. Note that the activity flow approach also investigated the idea of residualisation. There are major differences between their residualisation and ours. For example, the resting-state data were residualised differently in the two studies. In our analysis, the component to be regressed out is the group-average spatial map, while in their model, this component is the subject-general voxel-to-region FC patterns. Furthermore, when predicting residual activations for a given subject, they used the subject-general activation patterns (calculated from other subjects) in combination with the residualised FC patterns of the given subject, whereas we did not utilise this information of subject-general task activations at all. Both approaches exhibited the benefits of residualisation. Their approach showed that the individualised functional connectivity routes can predict individual-specific activations better than chance. This provides further evidence for the utility of residualisation in making more individualised activation patterns for individual subjects.

Note that the group-average activation patterns alone can have considerable overlap with individual activation maps. Thus one can obtain moderate prediction accuracy by simply reproducing the group-average. Hence, the accuracy of residualised predictions, or the discriminability of the group-average-added-back predictions, are more informative on the model's ability to make individualised predictions. This is of particular importance, because many existing algorithms tend to push predictions towards the mean. In a higher-dimensional setting, the relation between the two measures becomes complicated, but it is not difficult to see that the improvement of discriminability may degrade accuracy a little. Training and evaluating the model on residualised resting-state and task data thus have more desirable properties, not only to simplify the assessment of model performance but also to maximise ability to capture inter-individual differences. Other approaches to improve prediction discriminability include introducing a contrastive loss term

to push between-subject differences to be large [40]. It is yet to be investigated whether the two approaches are comparable. However, introducing extra terms may complicate the loss function (for example, turn a convex loss function into a non-convex one) and thus may be less stable. Training on residualised data keeps the original loss function structure and is usually simpler to train.

It is also important to note that the accuracy of such predictions is not by itself a good indicator of how well the model captures the underlying individual differences. The prediction accuracy is not only limited by inaccuracies of model assumptions or inappropriate specification (i.e. mismodelling) but also by other sources of variability such as measurement noise. In general, computational models do not account for the variability due to measurement noise, thus this imposes a bound to the model's ability to capture the true signal of interest. This effect is usually referred to as noise ceiling, and can be estimated by calculating the correlation between two repetitions of the test set (i.e. split-half estimator). If a model can fundamentally capture the variations in brain response due to individual or stimulus differences, the accuracy of prediction should at least verge on the test-retest reliability, or even be higher. It has previously been recommended in the neuroimaging community to report the performance of a model with respect to the noise ceiling [161]. In our results, the prediction accuracies on the HCP dataset almost paralleled the test-retest reliability for most task contrasts, suggesting that our model can fundamentally capture the individual variations in activation. On the UKB dataset, the accuracies are less encouraging, which is unsurprising due to the shorter scanning sessions and higher noise level.

In addition to predicting individual-unique activations, it is also of value to investigate the causes of the variations in task-evoked activations, particularly, what information in resting-state activity drives the individual differences in task activity. For example, do variations in peak activation patterns correspond to the changes in resting-state activity in the same location, or is it actually driven by more complicated configuration of the dense connectivity pattern? Such investigations would help us understand the nature of the inherent resting-state

features that characterise variations in task activity. For example, these features can be “structurally” inherent (characterised by brain organisation and connectivity) or “functionally inherent” (related to the cognitive state of subjects during the resting-state scan) [19], both of which may cause the re-configuration or re-allocation of peak activation patterns. Note that individual differences in task-evoked activations may be partially due to inter-subject misalignment. Indeed, registration remains an empirical question and may be sub-optimal in practice. However, it is very unlikely that our results only account for misalignment between subjects, as the model can capture variations not only in shape and position but also in topology of the activations. Indeed, it is likely that the relatively state-of-the-art alignments used here in preprocessing reduced intersubject variability, rather than increased it.

Using resting-state fMRI scans to infer individualised task-evoked responses has a wide range of implications in translational and clinical neuroscience. One potential application of the proposed model is to infer individualised functional localisers based on resting-state fMRI scans. This is important because task-fMRI scans are often of limited accuracy and reliability [164], possibly due to poor task performance and noise that is hard to remove in pre-processing. Such a framework can supplement task localisers, potentially improving the prediction of individual functional mapping (and for multiple networks and regions) and facilitating investigations of individualised response profiles of localised brain regions. Although our results primarily focus on the residualised results, this does not conflict with application targets such as pre-surgical planning. One of the biggest challenges in surgical targeting is to account for individual differences in the location of a target structure, which may vary considerably across subjects. Therefore, it is more desirable to have a model better capable of capturing the underlying inter-individual differences, not just reproducing the typical patterns shared by all subjects. The model trained on the residualised data, with group-average added back in afterwards, yielded more differentiated or "individualised" predictions, which matches this goal. The next question arising is: how many subjects are required in the training to create such a model and make predictions useful, given that in a surgical context, it is

impractical to collect a substantial number of subjects as in the UKB and HCP datasets. Firstly, it is worth noting that even though the numbers of training subjects are large in this study, neither of the two datasets requires a big "N" to achieve good prediction accuracy (Figure B22 and B23). In fact, prediction accuracy as a function of training size converges very quickly. Secondly, one does not necessarily have to re-collect training data using the same scanner or task for a given clinical dataset, because the model trained on one dataset may be transferable to another if aided by transfer learning or data harmonisation techniques. For example, Jones et al., 2017 applied the GLM model [19] trained on the HCP healthy cohort to a disease cohort and successfully learned the variations in their language maps. Another possibility is to generalise the resting-state features across different datasets, instead of transferring the mapping between resting-state and task data from one dataset to another. More specifically, one can generate matched dual-regression or sPROFUMO maps for different datasets. The coefficients thereby learned on one dataset may also be useful to another, as they refer to the same spatial configurations. Of course, to accurately generalise such predictions requires more complicated modelling techniques; developing and validating such approaches is outside the scope of this chapter. As numerous multi-site multi-scanner consortia emerge, in future, it is important to develop a model that is capable of learning features invariant across scanners and insensitive to confounds. Once generalisable to other populations, such a model can be used to localise regions of interest for those who cannot perform tasks, such as paralysed patients and infants.

There are a few limitations in this study. First, the ensemble model is a linear combination of two single (largely) linear models and thus has limited ability to capture higher-order non-linear relationships within and between the resting-state and task-evoked brain activity. Second, the decompositions of common modes of variations are unsupervised. In the future, more complex modelling could be adopted to simultaneously estimate the common modes of variations and the reconstruction coefficients. Third, the rich information derivable from T1 and diffusion MRI scans

may further aid the predictions of individual differences in task-evoked activity, and this model is yet to be adapted into a multi-modal framework.

5

Resting-state neural activity predicts individual differences in retinotopic organisation in early visual areas

5.1 Introduction

Contents

A large proportion of human cerebral cortex is dedicated to visual processing. The large-scale spatial organisation of visual cortex that corresponds to the spatial

arrangement of the retina is known as “retinotopic mapping” or "retinotopy". In essence, it represents the mapping of the visual field onto the cortical surface, via topography-preserving anatomical connections.

Studying retinotopy has important implications for both basic and clinical neuroscience. As a fundamental feature of the visual cortex, retinotopic mapping provides insights into mechanisms underlying visual processing. Studies on retinotopic organisation have illuminated the specialisation of different cortical regions for processing specific aspects of visual information, such as colour, motion, and object recognition [165, 166]. Moreover, understanding retinotopic organisation can facilitate the diagnosis and treatment of neurological disorders. For example, retinotopic mapping can help localise lesions in the visual pathway or assess the functional integrity of the visual cortex in patients with visual impairments, such as those caused by stroke or neurodegenerative diseases [167, 168]. Additionally, knowledge of retinotopic organisation can facilitate the development of visual prosthetic devices or rehabilitation strategies aimed at restoring or enhancing visual function [169].

Retinotopic mapping is typically acquired using task-fMRI while subjects view visual stimuli specifically designed to evoke responses in the visual cortex that correspond to specific locations in the visual field. Common stimulus types include expanding rings and rotating wedges/semicircles, which are designed to vary spatiotemporally in a repetitive manner, covering the visual field over time. The patterns of BOLD signal changes in response to these stimuli are used to calculate spatial maps in the visual cortex that represent different locations in the visual field, as "retinotopic maps". Frequently studied retinotopic maps include polar angle and eccentricity, which describe the location of receptive fields in the visual cortex, as well as the size of receptive fields, which measures the spatial extent of the region responsive to the given stimuli [165]. Despite being a well-established approach, deriving retinotopic mapping from dedicated task-fMRI can be impractical, as it typically requires high-field MRI scanners and long scanning time (usually between 15-60 min). Furthermore, it may be impossible to apply this approach to certain

populations, such as individuals with cognitive or sensory impairments or young children. Therefore, it is important to develop alternative methods for estimating retinotopic maps at the individual level.

While limited, previous efforts to characterise retinotopic mapping in vivo without using dedicated task-fMRI have typically turned to other MRI modalities, such as structural MRI and resting-state fMRI. The anatomical structure of the visual cortex, given its fundamental role in visual processing, serves as a valuable indicator of retinotopic organisation in early visual areas [170–174]. Traditional methods for predicting retinotopic maps involve using purely geometric templates (e.g., iso-eccentric bands as an eccentricity template) adjusted to an individual's cortical anatomy, allowing for estimation of individual variations in the functional organisation of the early visual cortex [171, 173]. Such template-based approaches can only account for gross individual differences. More recently, Ribeiro et al. [175] achieved more individualised prediction of retinotopic mapping using the anatomical features of early visual cortex, via deep learning techniques. However, relying solely on anatomical information, these anatomy-based predictions of retinotopic mapping overlook the functional aspect of visual processing and thus cannot account for all unique differences concerning the structure-function relationship in visual areas. Moreover, while cortical folding predicts the border of V1 well [176, 177], higher visual areas are not well correlated with the location of the folds.

As demonstrated in Chapter 4, resting-state fMRI possesses an inherent ability to reconstruct task-evoked brain responses, capturing intrinsic functional connectivity between brain regions even in the absence of a specific task. This raises an intriguing question: can resting-state BOLD signal also predict more fine-grained task-elicited retinotopic mapping? This is based on the hypothesis that adjacent regions in the visual cortex, which represent nearby locations in the visual field, will likely exhibit correlated spontaneous activity patterns. A growing body of research has explored this hypothesis, characterising the retinotopic organisation of visual areas using resting-state brain activity [174, 178–184]. For example, Haak et al. [183, 184] characterised retinotopic mapping using resting-state connectopy, which is the

topographically organised resting-state brain connectivity that changes spatially and gradually according to an orderly organisation, wherein nearby locations connect with nearby locations elsewhere in the brain; Heinzle et al. and Bock et al. [178, 180] utilised cortico-cortico receptive field modelling, also revealing that locations in visual areas that tend to co-activate are responsive to similar or nearby visual fields. Despite these advances, existing resting-state studies on retinotopic mapping have primarily focused on either a single highly-sampled individual or identifying common patterns of retinotopic mapping across a population. As a result, there is still a lack of methods for individualised prediction of retinotopy using resting-state fMRI data. Recent work by Watson et al. [185] has attempted to address this challenge by evaluating the accuracy of individualised resting-state retinotopy predicted by different manifold learning approaches. However, this study was confined to the primary visual cortex (V1) and did not investigate higher order visual areas. More research is needed to refine and validate these approaches for broader application.

In this chapter, extending the connectopy-based approach adopted in Haak et al. [183, 184], we proposed a framework for individualised retinotopic mapping in early visual areas, ranging from primary visual cortex (V1) up to extrastriate cortex (V4), using voxel-wise functional "fingerprints" of these areas derived from resting-state fMRI. We focused on two properties of retinotopy: polar angle and eccentricity, as they are the most fundamental attributes of the visual field. Polar angle corresponds to the angular position of a point in the visual field relative to the center of gaze, while eccentricity represents the radial distance from the center of gaze. These two attributes together provide a precise description of a point's position in the visual field. Moreover, early visual areas are known to exhibit a topographic organisation based on the polar angle and eccentricity of the visual field. Focusing on polar angle and eccentricity is not only more straightforward and practical than other retinotopic properties (e.g. receptive field size), but also more interpretable.

The proposed framework consists of two steps. First, informed by the brain connectivity elicited by specific visual stimuli, we exploit the rich information

contained in resting-state fMRI timeseries to generate a set of voxel-wise resting-state "fingerprints" related to the functional processing of visual information. From these fingerprints, we derive pairwise similarities among voxels in the visual cortex, constructing a "resting-state graph" that reflects resting-state similarity pertinent to visual processing. This graph is optimised to capture the functional connectivity that arises from visual processing, and contains less of the gross connectivity patterns found in conventional resting-state functional connectivity. Second, we apply a modified graph embedding method to uncover the topographic organisation of this "optimised" resting-state graph, hidden within its low-dimensional representation. Graph embedding methods usually transform the high-dimensional graph into a lower-dimensional representation (or 'embedding') in such a way that important characteristics of the graph, such as distance relationships between nodes (i.e., voxels), are maintained in the lower-dimensional space. This step serves to uncover spatially varying patterns of functional connectivity between voxels underlying the high-dimensional complex resting-state graphs. We demonstrate that this resting-state retinotopic mapping not only reproduces the task-elicited retinotopy, previously obtained using 7T task fMRI with dedicated visual tasks, but also captures individual differences in retinotopic organisation in V1-V4. Additionally, the proposed framework more accurately delineates borders of the early visual areas, V1-V4, compared to group-average parcellation of the visual cortex. We demonstrate the effectiveness of our approach using both high-field 7T and more common 3T resting-state fMRI data. Overall, this approach presents a potential alternative to resource-intensive task-fMRI visual experiments. It has the potential to be applied to diverse populations, such as infants or patient cohorts, characterising retinotopic organisation of visual areas when dedicated visual experiments are impractical. Such applications will potentially provide valuable insights into the neural mechanisms underlying neurodevelopment or pathology.

5.2 Materials and Methods

5.2.1 The Human Connectome Project MRI data

5.2.1.1 Subjects

A total of 181 participants (109 females, 72 males) between the ages of 22 and 35, obtained from the Young Adult Human Connectome Project (HCP), were included in the analysis (`https://www.humanconnectome.org/study/hcp-young-adult/data-releases`). They underwent multimodal MRI scans and received behavioral and demographic assessments. All participants had normal or corrected-to-normal visual acuity. The dataset included 53 pairs of identical twins, 34 pairs of fraternal twins, two pairs of non-twin siblings, and three individuals whose twins/siblings were not included. The family structure of this dataset can be found in "Restricted Data" on ConnectomeDB (`https://www.humanconnectome.org/study/hcp-young-adult/document/wu-minn-hcp-consortium-restricted-data-use-terms`).

5.2.1.2 Structural image acquisition and preprocessing

The T1-weighted (T1w) and T2-weighted (T2w) structural scans were acquired at 0.7-mm isotropic resolution on a customized Siemens 3T Connectom scanner at Washington University [105]. They served as as the anatomical substrate for the retinotopy data. White and pial cortical surfaces were reconstructed using the HCP preprocessing Pipelines [107]. The surface were aligned across subjects to the HCP 32k fs_LR standard surface space using two registration stages: a gentle folding-based registration called "MSMSulc", driven by sulcal and gyral patterns, followed by a more aggressive registration called "MSMAll", driven not only by cortical folding patterns but also by other modalities such as myelination and resting-state connectivity [155, 156, 186]. As the MSMAll registration method partly involves information of resting-state networks and resting-state-based estimates of visuotopic organisation, it likely improves alignment of population receptive field (pRF) solutions across individuals compared to MSMSulc, which is solely based on cortical folding. Subcortical volume data were aligned to MNI152 standard space using FNIRT.

5.2.1.3 fMRI acquisition and preprocessing

In brief, fMRI data were acquired on a Siemens 7T Magnetom scanner. Whole-brain fMRI data were acquired at 1.6mm isotropic resolution and 1s TR (multiband acceleration 5, in-plane acceleration 2, 85 slices). The fMRI data were corrected for head motion and EPI spatial distortion, and were aligned with the HCP standard surface space described above using the HCP preprocessing pipelines [107]. The preprocessed fMRI data are in the "CIFTI" format (a file format that merges fMRI data from the gray matter surface and subcortical structures into one file), including 91,282 grayordinates in the 32k fs_LR standard space, covering both cortical and subcortical brain regions with approximately 2-mm spatial resolution [107]. Next, the fMRI data were denoised for spatially specific structured noise via multi-run sICA+FIX [151, 187, 188]. Finally, the dimensions of the preprocessed data are 181 subjects $\times$ 91,282 grayordinates $\times$ 6 runs $\times$ 300 timepoints. Full details on fMRI data acquisition and preprocessing can be found elsewhere [107, 189]. The preprocessed fMRI data are available from ConnectomeDB (`https://db.humanconnectome.org/`).

5.2.1.4 Stimuli

The retinotopic mapping stimuli used in this study were generated by slowly moving apertures containing a range of colourful visual objects. Both the apertures and textures were generated at a resolution of 768 x 768 pixels, and were confined to a circular region with a diameter of 16.0°. Beyond the circular region, the display was uniform gray.

The retinotopy task-fMRI experiments consisted of six runs, RETCCW, RETCW, RETEXP, RETCON, RETBAR1, and RETBAR2. Each run lasts 300.0 seconds, during which participants underwent three different types of apertures: wedges (RETCCW and RETCW, designed to elicit polar angle mapping), rings (RETEXP and RETCON, designed to elicit eccentricity), and bars (RETBAR1 and RETBAR2, designed to elicit both polar angle and eccentricity). These apertures slowly moved across the visual field and were periodically interrupted by blank periods, designed

to help differentiate between non-visual responses and responses from neurons with very large receptive fields [190]. More specifically, the ring stimuli comprised of a ring either expanding from or contracting towards the visual field center with a period of 32 seconds. The wedges stimuli consisted of 90°wedges rotating either counterclockwise or clockwise across the visual field with a period of 32 seconds. The bar stimuli consisted of bars with different orientations (i.e., four orientations) moving across various directions in the visual field. Throughout each experiment, participants were asked to focus on a semitransparent dot located at the center of the display and report any changes in colour of the dot by pressing a button to encourage fixation and allocation of attention to the center of the display. For further details on the aperture and texture design, please refer to Benson et al. [191].

5.2.2 Population receptive field (pRF) analysis

The retinotopic maps were constructed using population receptive field (pRF) modelling [191, 192]. The receptive field of a neuron is the region in the visual field within which a visual stimulus can influence its activity. The pRF model aims to describe the receptive field properties of a population of neurons within a specific region of the visual cortex (Figure 5.1A). The pRF model represents the receptive fields as 2-D isotropic Gaussian functions, characterised by three parameters, polar angle and eccentricity (which indicates the position of the receptive field) and a size parameter (which represents the spatial extent of the receptive field).

To perform pRF modelling, the stimulus-related neural response is predicted as the dot product between the stimulus aperture time series and the 2-D Gaussian, summed sublinearly according to a static power-law nonlinearity with exponent 0.05 [192] and scaled by a gain factor, finally convolved with a canonical hemodynamic response function (HRF) [191]:

$$r(t) = (g \times (S(t) \cdot G)^n) * h(t) \tag{5.1}$$

where $r(t)$ is the predicted stimulus-elicited fMRI time series, g is a scaling factor, $S(t)$ is the stimulus aperture at time t, G is the 2-D isotropic Gaussian, n is an

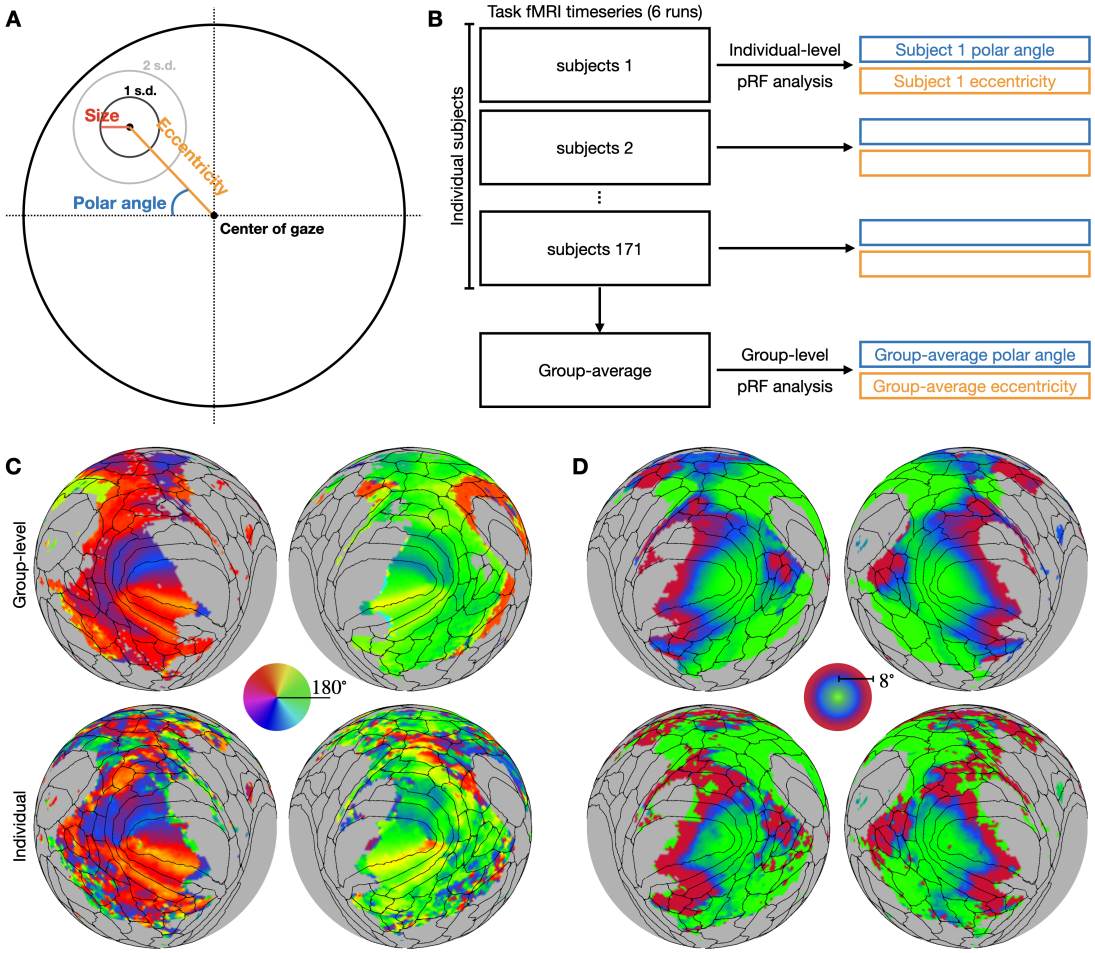

Figure 5.1: Illustration of task-elicited retinotopy and pRF analysis. **(A)** Illustration of the three properties of a receptive field, polar angle (blue), eccentricity (orange), and receptive field size (red). Polar angle and eccentricity describe the location of a point in the visual field relative to the center of gaze, while pRF size describes the spatial extent that a group of neurons is reponsive to. **(B)** The individual polar angle and eccentricity spatial maps were derived from individual task fMRI times series as "ground truth" via individual-level pRF analysis. The group-average polar angle and eccentricity maps were derived from group-average task fMRI timeseries via the group-level pRF analysis as the baseline prediction. **(C)** Examples of group-average polar angle map (top) and individual polar angle map (bottom), across the brain. **(D)** Examples of group-average eccentricity map (top) and individual eccentricity map (bottom), across the brain.

exponent parameter (here $n = 0.05$), and $h(t)$ is the canonical HRF. The pRF model's parameters are adjusted iteratively until the best-fitting model is found via grid search, yielding the pRF properties (polar angle, eccentricity and pRF size) for each voxel. See [191] for more details on pRF model fitting. In this chapter, our focus is limited to polar angle and eccentricity.

5.2.2.1 Group-level pRF analysis

Group-level pRF models were fitted on the group-average task-fMRI (Figure 5.1B). As the visual stimuli are synchronised across the subjects, the MSMAll-registered retinotopy task-fMRI timeseries were simply averaged across the 181 subjects for each run (e.g., RETCCW, RETCW, RETEXP, RETCON, RETBAR1, and RETBAR2), giving a single 91282 $\times$ 300 timeseries per run. The fitted pRFs that best explained the group-average task-fMRI timeseries, six runs concatenated together, produced the group-level spatial maps of polar angle, eccentricity, and pRF size, as well as the variance explained as a measure of goodness-of-fit of the pRF models. The group-level polar angle and eccentricity maps serve as the baseline prediction of retinotopic mapping. The map of variance explained by group-level pRF modelling was used to exclude noisy vertices, i.e., the neural activity of which cannot be explained by the visual stimuli.

5.2.2.2 Individual-level pRF analysis

In contrast to the group-level pRF analysis, we chose MSMSulc-registered retinotopy task-fMRI data for individual-level analysis. This is because MSMAll registration utilised visuotopic information to align the cortical surface across subjects, consequently minimising individual variability in retinotopic mapping [186]. The individual-level polar angle maps were obtained by fitting pRF models to the concatenated RETCCW and RETCW task-fMRI time series (i.e., rotating wedge). Meanwhile, eccentricity maps were generated by fitting pRF models to the concatenated RETCON and RETEXP time series (i.e., expanding/contracting ring). The resulting individual version of polar angle and eccentricity maps serves as "ground truth" for resting-state retinotopic mapping.

5.2.3 Definition of visual ROIs

We restricted the analysis to early visual areas, from V1 to V4. The V1, V2, V3, and V4 ROIs were obtained from a multimodal group-average template [186] in the HCP 32k fs_LR standard surface space [107]. Vertices for which the fitted

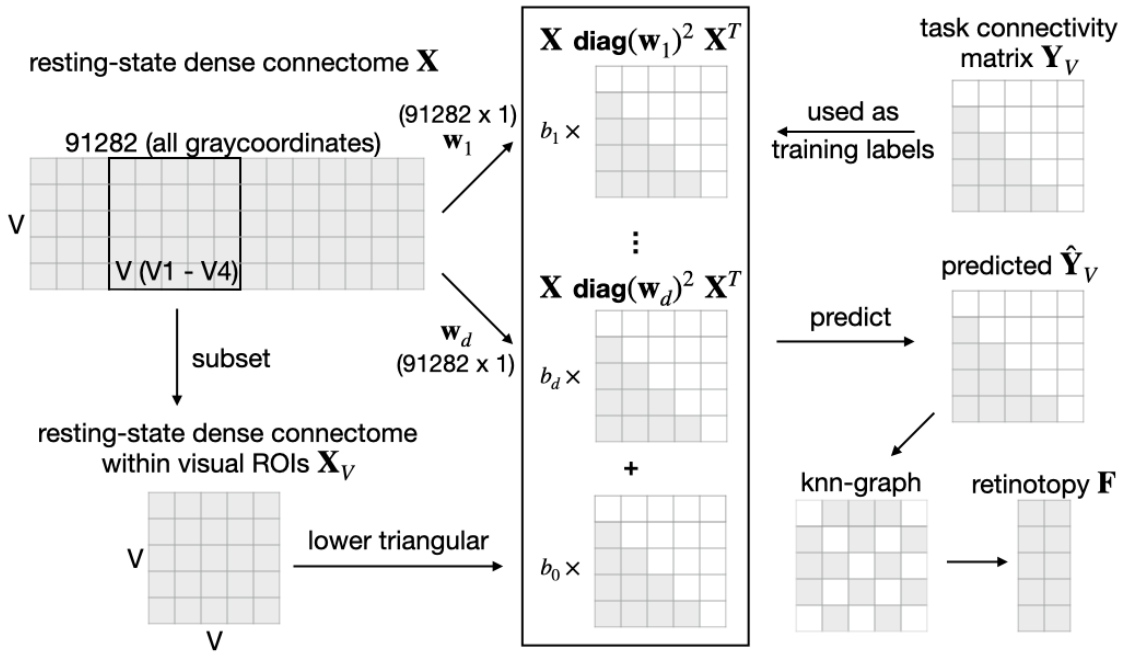

Figure 5.2: Illustration of the retinotopy prediction framework. First, a $V \times 91282$ resting-state dense connectome $\mathbf{X}$ is constructed for the V vertices within the selected ROIs (V1-V4), serving as resting-state "fingerprints" for these vertices. The dense connectome within the ROIs, denoted by $\mathbf{X}_V$, is a subset of $\mathbf{X}$. A set of $V \times V$ similarity matrices are subsequently constructed from the dense connectome $\mathbf{X}$, each emphasising the fingerprints with respect to a specific functional configuration. The similarity matrices were combined with $\mathbf{X}_V$ to predict the task connectivity matrix $\mathbf{Y}_V$ elicited by visual stimuli. The predicted task connectivity matrix, $\hat{\mathbf{Y}}_V$, is referred to as the "optimised resting-state graph". A knn-graph $\mathbf{D}$ is next constructed from $\hat{\mathbf{Y}}_V$. Finally, the resting-state retinotopy F is solved from the knn-graph $\mathbf{D}$ via the modified Laplacian Eigenmap algorithm.

pRF model at group-level yielded less than 55% explained variance were excluded from the V1-V4 parcels, as the "ground truth" polar angle and eccentricity of these vertices are regarded unreliable as they can't be explained well by the pRF model.

5.2.4 Framework setup and implementations

In this study, we used resting-state fMRI to predict task-based retinotopy by focusing on enhancing the specificity of the resting-state graphs to visual-elicited functional connectivity (see Figure 5.2 for an illustration of the workflow). The analysis was conducted for each hemisphere separately.

5.2.4.1 Optimising the resting-state graph

The original resting-state graph, calculated as the correlations of voxel-wise resting-state timeseries (i.e., dense connectome), is likely biased by functional connectivity patterns unrelated to visual processing, as the timeseries in resting-state fMRI inevitably reflects not only the neural activity underlying visual processing but also fluctuations modulated by other functional networks. Therefore, by incorporating connectivity information from task-fMRI, it is possible to improve the "specificity" of the resting-state graphs to visual-related functional connectivity. We hypothesise that this specificity can be achieved by accounting for the voxel-wise (or vertex-wise) similarities with respect to the functional spatial configurations that contain visual information. To illustrate, the functional connectivity of a given pair of vertices (i.e., an edge in the resting-state graph) can be attributed to two sources: one due to the fluctuations of neural activity normally elicited by visual stimuli, and the other due to the low-frequency continuous fluctuations arising from intrinsic functional organisation. By reducing the latter, we can guide the resting-state graph to better align with visual functional networks, thus allowing for a more accurate prediction of task-based retinotopy. To achieve this, we calculate a set of resting-state similarity matrices and use linear combinations of these matrices to approximate the task-elicited connectivity. The hypothesis is that this approximation can better capture the neural fluctuations specific to visual processing, compared to the original resting-state functional connectivity.

Specifically, suppose $\mathbf{X}$ is the $V \times 91282$ dense connectome matrix between the V visual vertices and the 91282 graycoordinates across the brain (both cortical vertices and subcortical voxels). The original resting-state graph $\mathbf{X}_V$, i.e., the dense connectome within the visual ROIs, is a $V \times V$ subset of $\mathbf{X}$. To calculate the task connectivity graph elicited by a given task stimulus, we took correlations between the task-fMRI timeseries of the given task (e.g., the task connectivity graphs elicited by rotating wedges were constructed by calculating the task-fMRI time series of RETCW and RETCCW tasks, temporally concatenated) and constructed the corresponding k-nearest-neighbour (knn) graph, removing the weakly correlated

edges in response to the stimuli. Here we choose $k = 50$ to construct the knn task connectivity graph for the rotating wedge task, and $k = 200$ for the ring task. The resulting task-fMRI connectivity graph is denoted by $\mathbf{Y}_V$. Suppose $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, ...\mathbf{w}_d] \in \mathbf{R}^{91282 \times d}$ is a set of d spatial configurations. We seek to find a set of linear coefficients $\mathbf{b} = [b_0, b_1, b_2, ...b_d]$ that minimises the following loss function:

$$\min_{\mathbf{b}} ||(b_0\mathbf{X}_V + \sum_{i}^{d} b_i \mathbf{X}\text{diag}^2(\mathbf{w}_i)\mathbf{X}^T - \mathbf{Y}_V) \odot \mathbf{M}||_F^2 + \lambda_2||\mathbf{b}||_2^2 + \lambda_1||\mathbf{b}||_1 \qquad (5.2)$$

where M is a $V \times V$ matrix that has ones below the diagonal and zeros on and above the diagonal, and $\odot$ is the element-wise multiplication. Briefly, we used a set of spatial configurations to modulate the voxel-wise similarity derived from the resting-state dense connectome and then used a linear combination of these resting-state similarities to approximate the task connectivity graph. Here we used the group-average ICA50 spatial maps as the spatial weights $\mathbf{W}$ (see Chapter 4.2.3 for more details on the ICA maps).

The coefficients $\mathbf{b}$ signify the weights assigned to each of the resting-state similarity matrices in their linear combination. These weights help determine the contribution of each individual similarity matrix to the overall approximation of task-elicited connectivity. They were solved for the rotating wedge (RETCW and RETCCW) and expanding/contracting ring (RETCON and RETEXP) task per subject, resulting in two sets of coefficients for each task per subject (see Figure C1 and Figure C2 for the corresponding coefficients). The coefficients were averaged across the training set, yielding a single set of coefficients for the given task, and were subsequently applied to the left-out subjects to produce the "optimised" resting-state graph $\hat{\mathbf{Y}}_V$. These "optimised" resting-state graphs would ideally better describe the co-activation patterns elicited by visual stimuli, compared to the original resting-state graphs. We next constructed the knn resting-state graph from the predicted $\hat{\mathbf{Y}}_V$, denoted as $\mathbf{D}$ in the following analysis.

5.2.4.2 Uncovering retinotopy from the resting-state graph

The resting-state retinotopy maps were uncovered as the low-dimensional representations of the optimised resting-state graph. We adopted a modified Laplacian Eigenmap (LE) to calculate these low-dimensional representations. Specifically, for a given graph $\mathbf{A}$ (in the form of a $V \times V$ connectivity matrix, here the optimised resting-state graph $\hat{Y}_V$), its retinotopic low-dimensional representations at the individual level, $\mathbf{F} = [\mathbf{f}_{\text{angle}}, \mathbf{f}_{\text{ecc}}] \in \mathbb{R}^{V \times 2}$, minimises

$$\min_{\mathbf{F}} \text{Tr}(\mathbf{F}^T \mathbf{L} \mathbf{F}) + \lambda_1 ||\mathbf{F} - \mathbf{G}||_2^2 + \lambda_2 ||\mathbf{F}^T \mathbf{F} - \mathbf{I}||_2^2 \tag{5.3}$$

where $\mathbf{G} = [\mathbf{g}_{\text{angle}}, \mathbf{g}_{\text{ecc}}] \in \mathbb{R}^{V \times 2}$ is the group-average retinotopic mapping, $\mathbf{g}_{\text{angle}}$ the group-level polar angle map and $\mathbf{g}_{\text{ecc}}$ the group-average eccentricity (which are derived from the group-level pRF analysis, described in Chapter 5.2.2.1). $\mathbf{L} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$ is the normalised graph Laplacian of the resting-state graph $\mathbf{A}$, where $\mathbf{D} = \text{diag}(\mathbf{d})$ is its degree matrix (i.e., $\mathbf{d} \in \mathbf{R}^{V \times 1}$ is the node degree vector of $\mathbf{A}$). The optimal $\mathbf{f}_{\text{angle}}$ and $\mathbf{f}_{\text{ecc}}$ that give the smallest Laplacian loss are solved as the individual polar angle and eccentricity maps. The modified Laplacian map differs from the classical one in two significant ways. First, it includes an additional penalty term, as indicated by the second term in Equation (5.3), designed to align the low-dimensional representations, denoted as $\mathbf{F}$, with group-level retinotopy $\mathbf{G}$. This additional penalty term is absent in the conventional Laplacian Eigenmap. Second, it relaxes the orthogonality constraint present in the traditional Laplacian Eigenmap, transforming it into a softer constraint, illustrated by the third term in the equation. Despite these modifications, this Laplacian Eigenmap remains an unsupervised method as it doesn't involve pairing input and output data for training.

5.2.5 Delineating borders of V1-V4 from retinotopy

Typically, the polar angle map reverses (i.e. its gradients flip direction) at borders of the visual areas. For example, from the lower to upper bank of the calcarine sulcus, the polar angle gradually changes from the upper vertical meridian to the lower vertical meridian, and reverses at the lips of the calcarine sulcus in the bordering

dorsal and ventral V2 maps [191]. Where the polar angle map reverses, or where its gradients change direction, indicates transitions between adjacent visual areas. We derived the borders between V1 to V4 from the predicted and group-average polar angle respectively for each individual, and compared them against the ground truth borders, derived from individual polar angle.

Characterising the borders based on a given polar angle map involves two steps. First, the gradients of the given polar angle map are calculated on the surface. Second, the dot product between the gradient vectors and the direction vectors with respect to a reference point (e.g., the centroid of V1) are calculated for each location. This dot product map reflects whether the gradient change is aligned with the expected direction of the polar angle reversal, and thus contains information of the transitions between visual areas (see Figure 5.3B for an illustration).

5.2.5.1 Calculating retinotopic gradients on the cortical surface

Estimating gradients on a cortical surface involves determining the rate of change in a particular feature at a given vertex on the surface. We employed the method described by Glasser et al. [186] to compute gradients of polar angle maps on individual surface meshes.

Here we briefly recap the gradient calculation procedure. For each vertex on the surface mesh, the normals of the vertex's associated mesh triangles were averaged to obtain the vertex's normal vector. Next, the vertex and its neighbours were unrolled onto a plane orthogonal to the vertex's normal that passes through the vertex by conducting the following steps:

1. Between the centre vertex and a neighbouring vertex, draw a circular arc that is tangent to the plane at the centre vertex;

2. Calculate the arc length $L_{\text{arc}} = L_{\text{Euclid}} * \sin^{-1}(L_{\text{opposite}}/L_{\text{Euclid}}) * L_{\text{Euclid}}/L_{\text{opposite}}$, where L_{opposite} is the dot product of the vector representing the edge (between the centre vertex and the neighbour) and the normal vector of the centre vertex (see Figure 5.3A for an illustration);

3. Project the neighbouring vertex onto the plane, adjusting its position to maintain the projected direction from the centre vertex while preserving a distance equal to the circular arc length from the centre vertex;

4. Perform linear regression within the plane, $f(t, u) = at + bu + c$, where t and u correspond to distances along orthogonal axes within the plane, using the values and positions of the centre vertex and the unrolled neighbouring vertices. The gradient vector comprises the spatial coefficients, projected into 3-D space by the unit vectors of the plane's coordinate system.

The process is repeated for each subject, on the individual 32k sphere mesh. We calculated gradients of different versions of polar angle maps, including the ground truth polar angle (as defined by individual-level pRF), the group-average polar angle (as defined by the group-level pRF), and the predicted polar angle. Note that since gradients are calculated on individual-specific surface meshes, the group-average polar angle map yields slightly different gradients across individuals.

5.2.5.2 Deriving borders from retinotopic gradients

To better characterise where the polar angle gradients change direction, we computed the dot product between the 3-D gradient vectors and the position vectors pointing towards the V1 centroid. Specifically, we first averaged the 3-D coordinates (determined by the individual sphere mesh) of all V1 vertices (per hemisphere) to obtain the centroid coordinates of V1, denoted as (x_0, y_0, z_0). Next, for a given vertex v, we computed a direction vector by subtracting the V1 centroid coordinates from the vertex's coordinates (x_v, y_v, z_v), resulting in a direction vector $(x_v - x_0, y_v - y_0, z_v - z_0)$, representing the direction towards V1 at location v. Suppose the gradient vector at vertex v is $(g_{x_v}, g_{y_v}, g_{z_v})$; we then took the dot product between this direction vector and the gradient vector at the given vertex, $(x_v - x_0, y_v - y_0, z_v - z_0) \cdot (g_{x_v}, g_{y_v}, g_{z_v})$. This step was repeated for every vertex within the ROIs (V1-V4), resulting in a dot product map. We then smoothed the dot product maps with a kernel of FWHM 4mm. At each vertex, we assigned

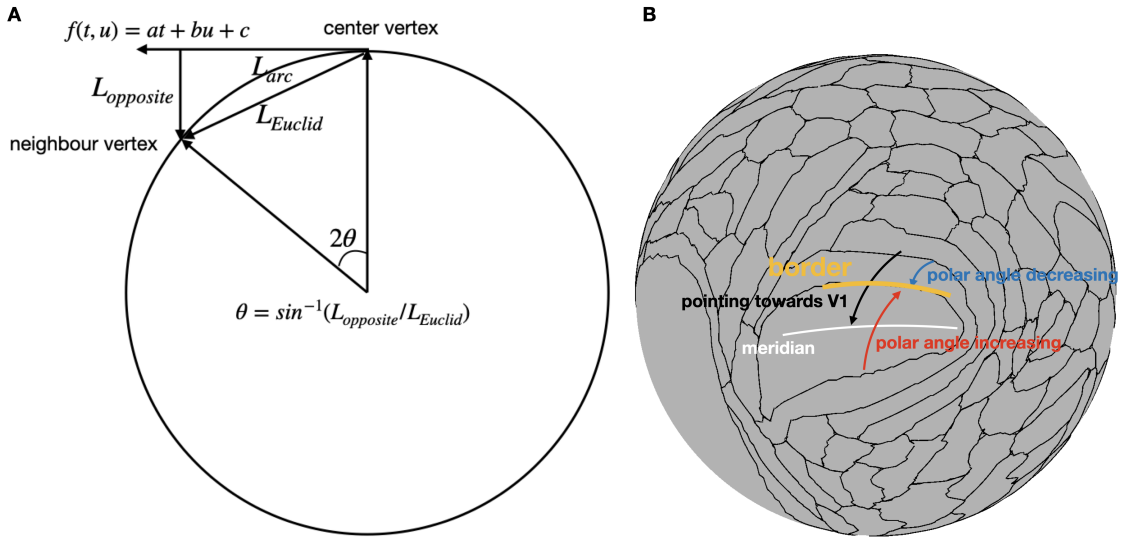

Figure 5.3: Illustration of the calculation of gradients and borders. **(A)** Calculation of the gradients on the surface. The arc length between the centre vertex and the neighbour is calculated as $L_{\text{arc}} = L_{\text{Euclid}} * \sin^{-1}(L_{\text{opposite}}/L_{\text{Euclid}}) * L_{\text{Euclid}}/L_{\text{opposite}}$, where L_{opposite} . The neighbour vertex is projected onto the plane tangent to the arc and elongated from the centre vertex to match the distance of L_{arc}. A linear regression is subsequently fitted at the centre vertex within the plane $f(t,u) = at + bu + c$ for the given feature $f(\cdot,\cdot)$ (e.g., polar angle), giving the spatial coefficients as the gradients. See [186] for more details. **(B)** Calculation of the borders. The dot product map between the gradient vectors and direction vectors reflects whether the polar angle changes align with the expected direction (e.g. pointing towards V1). Hence, where the dot product map flips the sign indicates where the polar angle reverses.

$+1$ to it if the dot product is positive, otherwise -1, finally resulting a map of ± 1s. The locations where the gradients change sign along the directions towards V1 indicate reversals in the polar angle map, suggesting the boundaries of the visual areas. For a visual representation of the steps taken to derive the borders, please refer to Figure 5.3B.

5.3 Results

5.3.1 Resting-state fMRI better predicts individual differences in task connectivity graph

We first tested whether the optimised resting-state graph can better capture individual differences in task connectivity structure, elicited by the dedicated retinotopy task stimuli. Here we briefly recap how these graphs were calculated.

The individual task connectivity graphs were constructed from the temporal correlations of 7T retinotopy task-fMRI times series, either from the rotating wedge visual task (RETCCW and RETCW, concatenated temporally), or from the expanding/contracting ring visual task (RETCON and RETEXP, concatenated temporally). To remove the weakly connected edges, we constructed the respective knn-graphs from the task connectome, with $k = 80$ for the rotating wedge task and $k = 200$ for the ring task. The group-average task connectivity graphs were constructed likewise, but from the group-average retinotopy task-fMRI time series. The original resting-state graph were simply the dense connectome calculated from the original resting-state fMRI time series (with the average signal across the brain regressed out from the time series of each individual voxel), while the "optimised" resting-state graphs were constructed using the approach described in Chapter 5.2.4.1. We hypothesised that the original resting-state functional connectivity between two visual vertices comprises two sources, one arising from the functional configuration dedicated to visual processing, the other due to the gross connectivity patterns and/or other non-visual functional configurations. The "optimised" resting-state graph was designed to approximate the task connectivity structure elicited by dedicated visual stimuli by reducing the contribution of the latter source, leveraging resting-state "fingerprints" with the rest of the brain.

For each retinotopy task (rotating wedge or expanding/contracting ring), the lower triangular elements of the optimised and original resting-state graphs were respectively correlated with those of the actual task connectivity graphs at the individual level. Such correlations were also calculated for the group-average task connectivity graphs, i.e., between the group-average and individual-level lower triangular elements, as the baseline. Both the optimised and original resting-state graphs produced higher correlations with the actual task connectivity graphs than the respective group-average task connectivity graph (Figure 5.4A and 5.4B, orange/red vs blue), suggesting that resting-state fMRI, even in the absence of dedicated visual-focused analysis, can indeed capture the individual differences in co-activation patterns of neural activity elicited by specific visual stimuli. Furthermore,

the optimised resting-state graphs gave higher correlations than the original resting-state graphs (Figure 5.4A and 5.4B, orange vs red). This corroborates the hypothesis that resting-state "fingerprints" in the rest of the brain can indeed inform the co-activation structure in visual processing. The results obtained from 3T resting-state fMRI exhibited a similar pattern (Figure 5.4C and 5.4D).

5.3.2 The optimised resting-state graph better captures individual-specific retinotopy

We subsequently investigated whether the resting-state retinotopy, comprising the polar angle and eccentricity maps derived from the optimised resting-state graphs, could more accurately capture individual differences in task-elicited retinotopy (i.e., polar angle and eccentricity maps obtained from individual-level pRF analysis) compared to group-average task retinotopy (i.e., polar angle and eccentricity maps obtained from group-level pRF analysis). Resting-state retinotopy represents the low-dimensional embeddings of resting-state graphs, which were determined using the modified Laplacian Eigenmap algorithm detailed in Chapter 5.2.4.1. These embeddings were designed to capture the gradual changes in resting-state functional organisation across two primary, orthogonal directions: one representing changes in the group-average polar angle, and the other reflecting changes in group-average eccentricity. Thus, each resting-state graph gave two low-dimensional embeddings, one supposed to approximate polar angle, the other eccentricity. Note that, however, the resting-state graphs optimised for the wedge task connectivity inevitably suppress eccentricity information while enhancing the polar angle information, and vice versa. Hence, we discarded the eccentricity embeddings from the resting-state graphs optimised for the wedge task, and only used the polar angle embeddings as the predicted polar angle maps; similarly, we also discarded the polar angle embeddings from the resting-state graphs optimised for the ring task, and only adopted the eccentricity embeddings as the predicted eccentricity maps.

Evaluations were conducted separately for each hemisphere. The predicted eccentricity maps, based on either the optimised or original resting-state graphs,

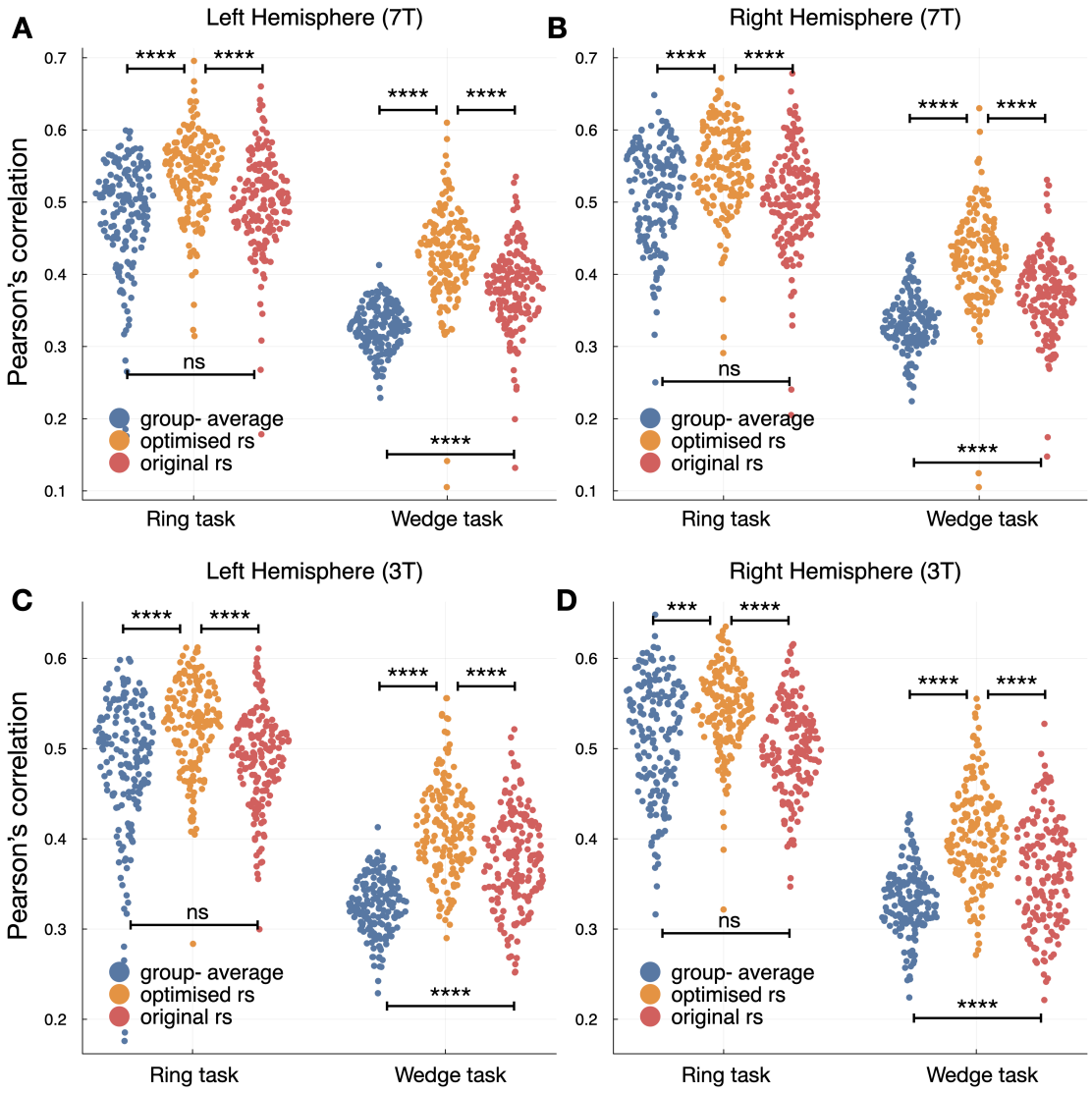

Figure 5.4: Correlations with the task connectivity graph edges. The edges of the optimised resting-state graph (orange), the original resting-state graph (red), and the group-average task connectivity graph (blue), were correlated with those of the individual task-fMRI connectivity graphs of the respective task, rotating wedge or expanding/contracting ring. Both the original and optimised resting-state graphs gave higher correlations with the individual task connectivity graphs than the group-average baseline. **(A)** Such correlations in left hemisphere, where the resting-state graphs were calculated using 7T timeseries. **(B)** Equivalent plots of (A) in right hemisphere. **(C)** and **(D)** Equivalent plots of (A) and (B), where the resting-state graphs were based on 3T data. The task connectivity graphs, either at the individual- or group-level, were calculated from 7T retinotopy task-fMRI times series. All tests (paired t-tests) were Bonferroni-corrected (24 tests in total).

were correlated with the actual eccentricity maps obtained via individual-level pRF analysis to assess prediction accuracy. Additionally, the group-average eccentricity, calculated using group-level pRF analysis, was correlated with individual eccentricity to establish a baseline. Similar correlation calculations were performed for polar angle. It is important to note that polar angle values are cyclical (i.e., $360^o = 0^o$). To circumvent discontinuity issues in polar angle values, especially at the horizontal meridian in the left hemisphere, both individual and group-average polar angle maps (output by pRF analysis) were transformed into their corresponding sinusoid maps.

Overall, the 7T resting-state retinotopy yielded high correspondence with actual task-elicited polar angle and eccentricity maps (Figure 5.5A and 5.5C). Compared to the group-average polar angle and eccentricity, the resting-state retinotopy gave higher correspondence with the actual retinotopy derived from individual-level pRF (Figure 5.5A and 5.5C). This is unsurprising, given that the resting-state graphs better captured the task-elicited co-activation patterns from V1-V4 (see Figure 5.6 for visualisations of the task-elicited, group-average, and resting-state-predicted polar angle maps on the surface). We also calculated the discrimination metric as described in Chapter 4.2.6. This metric measures, in addition to the overall accuracy, the extent to which the predicted retinotopy is more close to the subject to be predicted than to the others. The group-average prediction, while exhibiting high correspondence with the individual ground truth, does not discriminate between subjects, i.e., lacking the ability to tell subjects apart. The resting-state retinotopy produced much higher discriminability than the group-average patterns, suggesting that the model is indeed capable of capturing individual differences in task-elicited retinotopy (Figure 5.5B and 5.5D). The results based on 3T resting-state data produced a similar pattern (Figure C5), with the optimised resting-state retinotopy yielded the highest accuracy and discriminability.

Overall, the prediction accuracies for eccentricity are higher than those for polar angle. This discrepancy may have several possible interpretations. First, the organisation of eccentricity in the visual cortex might be simpler and more consistent across individuals compared to the organisation of polar angle. This could make

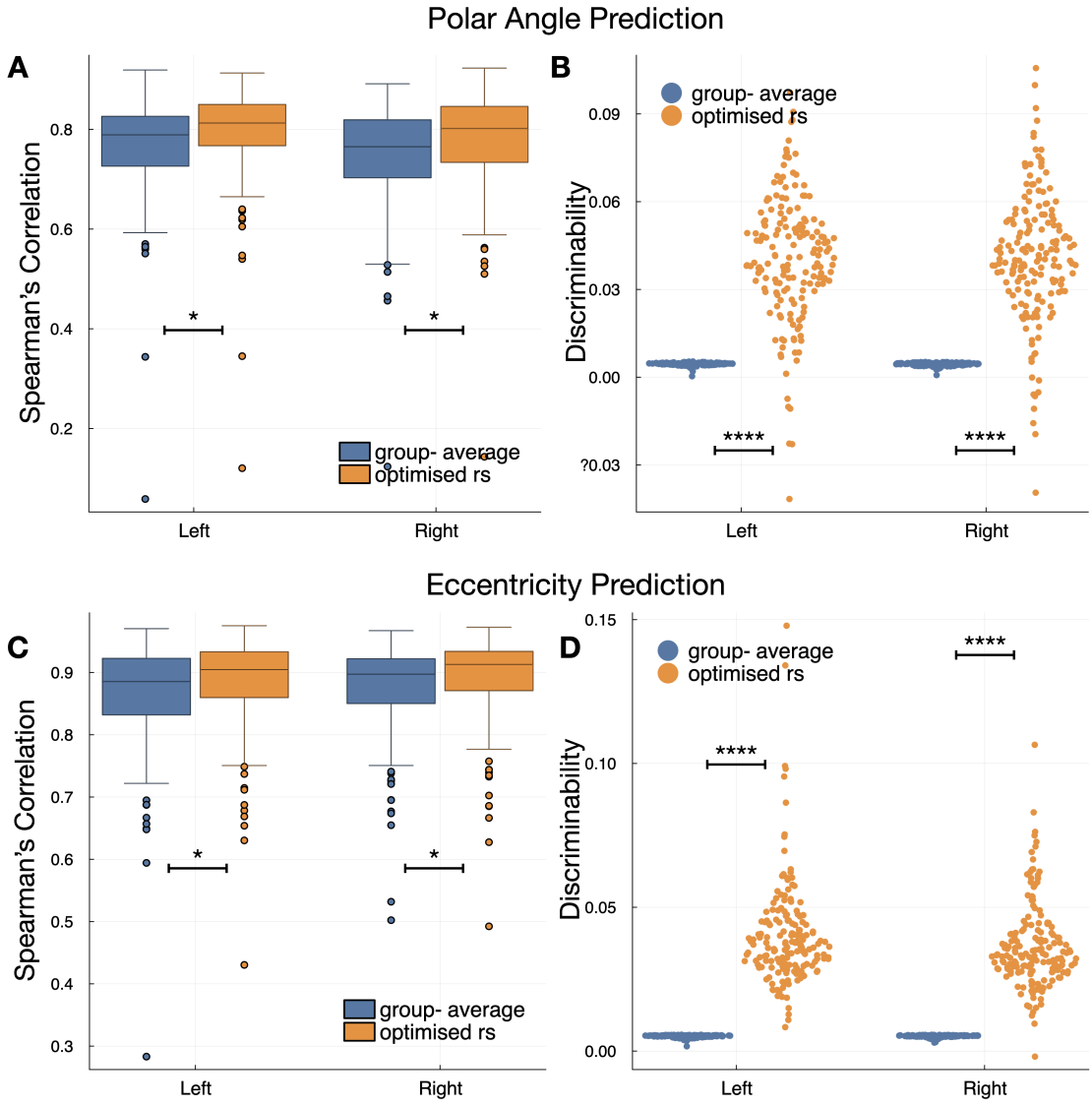

Figure 5.5: Prediction accuracy of 7T resting-state retinotopy. **(A)** 7T resting-state polar angle gave higher correlations with individual-level task-elicited polar angle than the group-average polar angle map. **(B)** 7T resting-state polar angle gave much higher discriminability than the group-average polar angle. **(C)** and **(D)** Equivalent plots of (A) and (B), for eccentricity prediction.

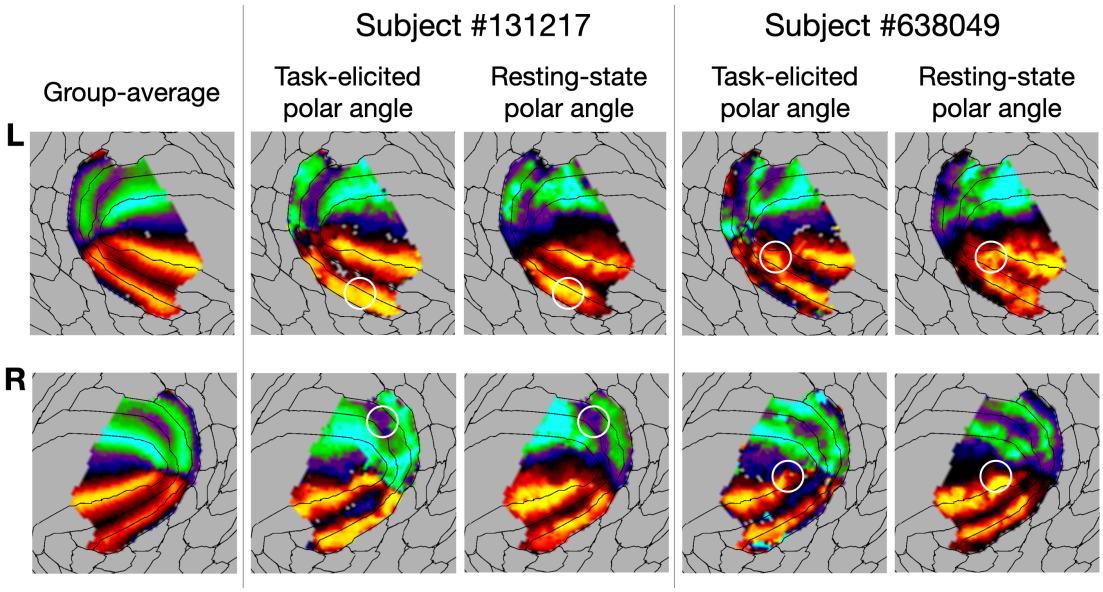

Figure 5.6: Task-elicited, group-average, and resting-state polar angle, shown on the sphere surface. In task-elicited polar angle, the patterns distinct from the group-average are highlighted by the white circles. These "abnormal" patterns are however captured by resting-state polar angle. Here the resting-state polar angle were based on 7T data; see Figure C3 for the 3T resting-state counterparts. The polar angle was converted to sinusoid maps so that it ranges from -1 to 1.

it easier to predict and capture eccentricity representations accurately. Second, the neural activity associated with eccentricity may be stronger or more robust than those associated with polar angle. As a result, the signals for eccentricity might be easier to detect and predict.

5.3.3 Delineation of borders between V1-V4 via the optimised resting-state graph

One of the important applications of polar angle is to delineate transitions between adjacent visual areas, particularly areas V1, V2, V3, and V4. This is because the borders of the early visual cortex are characterised by reversals in the polar angle representation; by identifying the locations where the polar angle reverses, it is possible to delineate the borders of these early visual areas.

Characterising where polar angle reverses typically involves calculating its changes (i.e., its gradients) on the surface and identify where the gradients flip their sign. To delineate the transitions of visual areas, it is usually required to

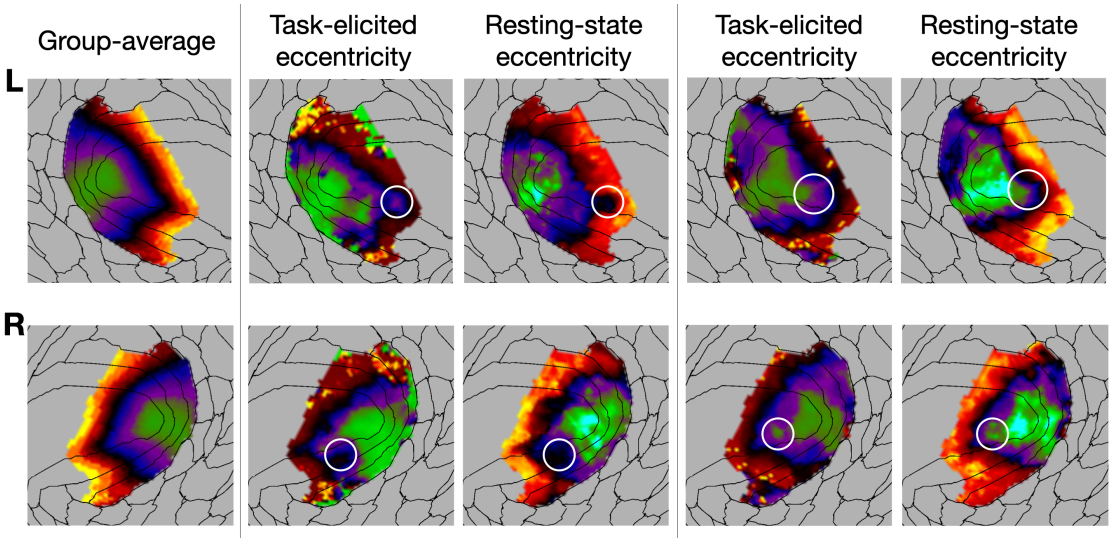

Figure 5.7: Task-elicited, group-average, and resting-state eccentricity, shown on the sphere surface. Equivalent plots of Figure 5.6, for eccentricity. The abnormal patterns in individual task-elicited eccentricity are highlighted by the white circles. These "abnormal" patterns are however captured by resting-state eccentricity. Here the resting-state eccentricity were based on 7T data; see Figure C4 for the 3T resting-state counterparts. Warmer colours indicate larger eccentricity, and vice versa.

further characterise the reversal locations with respect to a reference point, e.g., the centroid of V1, to help determine if the gradient change is aligned with the expected direction of the polar angle reversal. To illustrate, from the lower to upper bank of the calcarine sulcus, the polar angle gradually changes from the upper vertical meridian to the lower vertical meridian, and reverses at the lips of the calcarine sulcus in the bordering dorsal and ventral V2 maps [191]. By identifying whether the gradients flip sign along the expected direction of the polar angle change (i.e., from the lower to upper bank of calcarine sulcus), we can determine the reversal locations of the given polar angle map as the borders of adjacent visual areas.

Therefore, on the individual cortical surface, we calculated the gradients of three types of polar angle maps, the ground truth polar angle derived from individual-level pRF analysis, the resting-state polar angle derived from the optimised resting-state graphs, and the group-average polar angle derived from the group-level pRF analysis. For each type of polar angle gradient map, we calculated the dot product between the gradient vectors and position vectors pointing towards the centroid of V1 at each location, resulting in a map of dot products. Next, we assigned a positive or

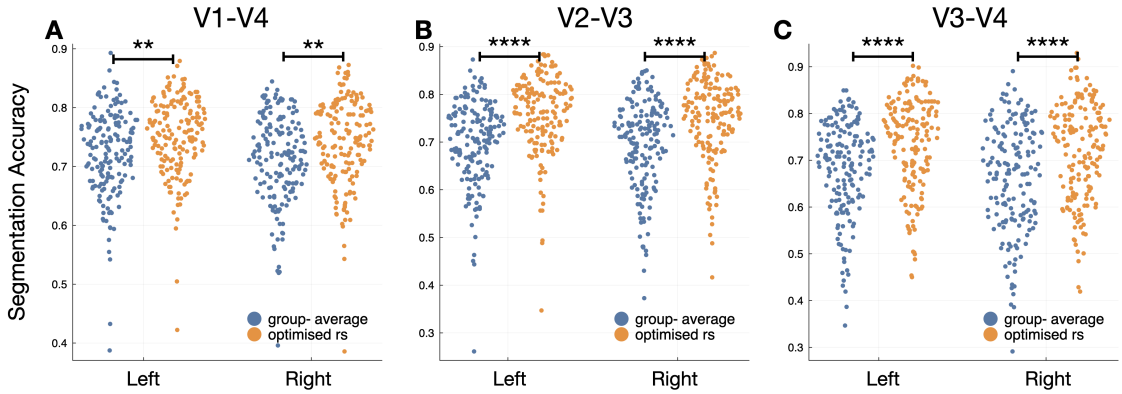

Figure 5.8: Segmentation accuracy of the 7T resting-state polar angle. **(A)** Across V1, V2, V3, and V4, the 7T resting-state retinotopy (orange) gave marginally higher segmentation accuracy than the baseline group-average accuracy (blue). **(B)** Across V2 and V3, however, the 7T resting-state retinotopy better captured transitions between visual areas. **(C)** Across V3 and V4, the 7T resting-state polar angle gave significantly higher segmentation accuracy than the group-average segmentation. See Figure C6 for results based on 3T resting-state polar angle.

negative sign to each dot product based on whether the dot product is positive or negative, resulting in a map of positive and negative values. The positive values represent where the change of polar angle aligned with the expected direction of the polar angle change, while the negative values represent the opposite; the transitions between the positive and negative values suggest the reversal of polar angle along the reference direction. Hence, the resulting positive/negative assignments naturally provide the segmentation from V1 to V4.

Such positive/negative-assigned dot product maps, derived from the individual task-elicited polar angle maps, were adopted as the ground truth segmentation in the following evaluation. To quantify whether the resting-state polar angle maps (either from the optimised or original resting-state graphs) can characterise the borders between V1-V4, we calculated the percentage of correctly assigned vertices in the resting-state dot product maps, against this ground truth, as a measure of prediction accuracy. Similarly, we also calculated the baseline accuracy, i.e., the percentage of correctly assigned vertices in the group-average dot product maps. Again, the resting-state-based segmentation between V1-V4 turned out to be slightly more accurate than the group-average segmentation (Figure 5.8A), particularly

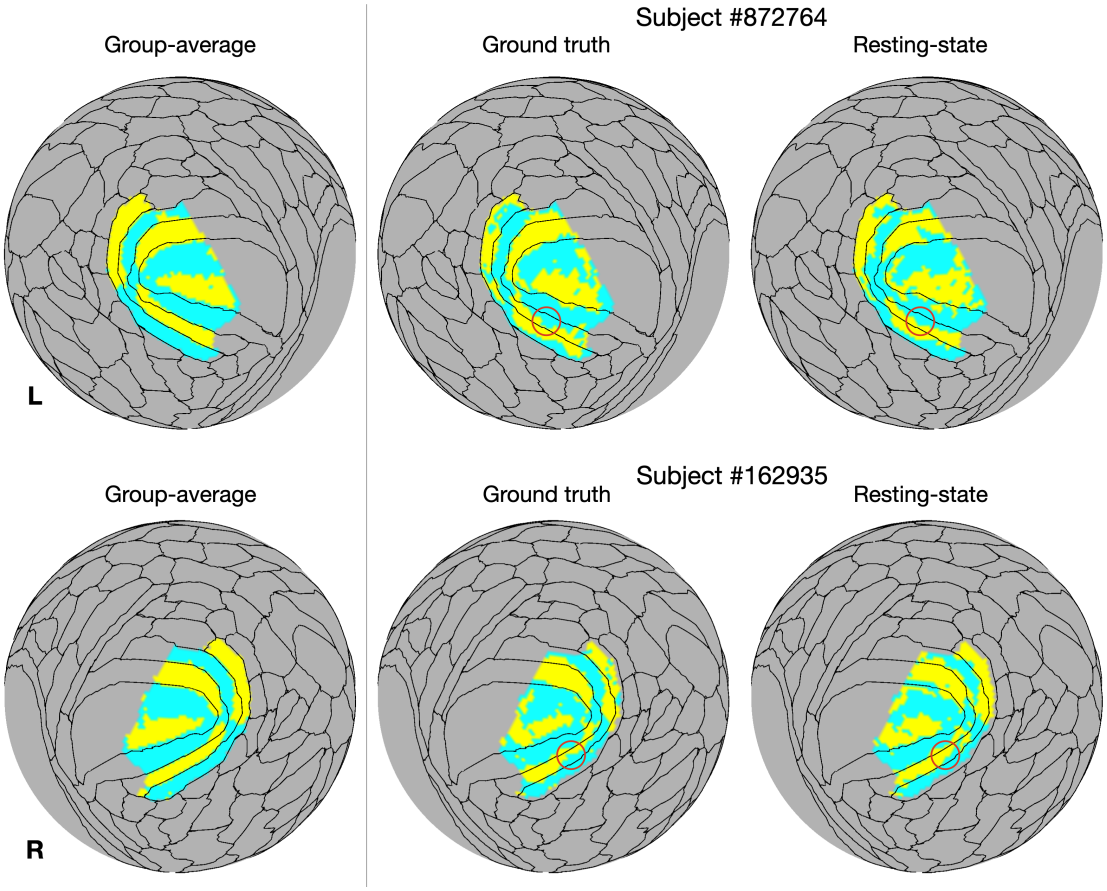

Figure 5.9: Segmentation between V1-V4 using the group-average polar angle (left), the individual task-elicited polar angle (middle), and the 7T resting-state polar angle (right). The segmentation derived from individual task-elicited polar angle served as the ground truth. Resting-state segmentation can capture the individual-unique patterns, highlighted by red circles.

in higher order visual areas (Figure 5.8B and 5.8C). These results suggest that more individual variability exists in higher order visual areas, which is in line with previous studies [193]. Results based on 3T resting-state followed a similar pattern with those based on 7T (Figure C6) though with degraded segmentation accuracy. Visualisation of the segmentations between V1-V4 can be found in Figure 5.9 for 7T resting-state, and in Figure C7 for 3T.

5.4 Discussion

In this chapter, we proposed a two-step supervised Laplacian Eigenmap (LE) approach to reconstruct retinotopic mapping (i.e., polar angle and eccentricity

of visual receptive fields) in early visual areas from resting-state fMRI. While the original LE algorithm is an unsupervised method for finding low-dimensional representations of the data, our proposed framework leveraged the retinotopic information modulated by dedicated task stimuli, and better captured the topographically varying retinotopic patterns in the lower-dimensional embedding at the individual level. Specifically, in the first step, the construction of the resting-state graph is optimised to approximate the co-activation patterns in the retinotopy task-fMRI. By incorporating this task information, the resting-state graph is better suited to characterising the gradual change of functional organisation that arises from processing of visual information, rather than from the global connectivity or irrelevant functional configurations, thus producing more visual-specific embeddings. In the second step, the dimensionality reduction is guided by the group-average retinotopic embeddings. This is achieved by introducing an additional term to the traditional Laplacian Eigenmap loss function, encouraging the low-dimensional embedding of individual resting-state graphs to resemble the overall retinotopic organisations. This additional penalty further forces the resting-state embeddings to "focus" more on the functional connectivity patterns relevant to visual processing, while preserving sensitivity to individual variability in the resting-state graph. The proposed approach leads to better performance in characterising individual-specific retinotopy in early visual areas from resting-state fMRI. Furthermore, based on the gradients of the resulting resting-state embeddings, the proposed approach yielded more accurate delineations of the borders of the early visual area than its unsupervised counterpart and the group-average parcellation. This is more true for higher-order areas, such as V3 and V4, than for V1, as the latter has less individual variability and can be well predicted by group-average folding patterns. Overall, these results demonstrate the potential of resting-state fMRI for replacing task-fMRI to characterise retinotopic organisations in the visual cortex and serve a reliable functional localiser, particularly when the dedicated task-fMRI is impractical in infants or disease cohorts.

Why is resting-state fMRI capable of predicting task retinotopy, even in the absence of a specific visual stimulus? During resting-state fMRI, spontaneous neural activity continues to occur. This intrinsic organisation reflects the underlying structure and connections of the neural networks responsible for processing specific types of sensory information, such as visual input. In the case of retinotopy, the visual cortex is organised in a topographically specific manner, with nearby neurons tending to respond to similar or adjacent regions of the visual field. This organisation is preserved in the resting-state functional connectivity patterns, as regions that share similar functional properties are more likely to be interconnected. The common anatomical substrate underlies the correspondence between resting-state and task-elicited retinotopy.

The proposed approach, while promising, remains a preliminary framework with several methodological limitations that warrant further investigation. The first major limitation is underfitting. In optimising the resting-state graph, we applied a set of spatial weights to emphasise specific connectivity "fingerprints" to create more visually-specific vertex-wise similarities. These spatial weights were precomputed as group-average spatial ICA maps, which may not be the optimal set of spatial weights for generating a more visually-specific resting-state graph. The actual number of free parameters (i.e., the number of coefficients corresponding to the set of spatial weights) is substantially smaller than the sample size (i.e., the number of edges). It is worth investigating whether optimising the spatial weights simultaneously with the coefficients will lead to better model performance. Second, the overall prediction accuracy of the framework needs further improvement. While we demonstrated that the proposed approach can capture individual differences in task retinotopy, its prediction accuracy is only marginally better than group-average retinotopy. Moreover, the predicted retinotopy does not correctly scale with the actual retinotopy, which should uniformly range from 0 to 2π in polar angle and from 0 to 8 in eccentricity within the selected ROIs. The distributions of predicted retinotopy exhibit bimodal patterns, complicating direct comparisons with the ground truth retinotopy (see Figure C8 for more details and discussions on this

point). Additional efforts are necessary to predict biologically meaningful retinotopy with accurate scales and distributions. Furthermore, the current framework has been trained and validated on the high-quality 7T HCP dataset, while using 3T fMRI data results in reduced model performance. Investigating advanced machine learning techniques to transfer the high-quality information from 7T fMRI to 3T fMRI could enhance retinotopy prediction on more commonly used low-field MRI systems. A third limitation is that the current work focused solely on polar angle and eccentricity. While these properties provide valuable information about the organisation of the visual cortex, other aspects, such as pRF size, are also essential to understanding the complete functional organisation of the visual system. pRF size refers to the spatial extent of the visual field that a group of neurons responds to. It varies across different regions of the visual cortex and is crucial for understanding how visual information is processed at various spatial scales. In future, it is also worth exploring frameworks that can predict individual differences in pRF size. Lastly, the framework has yet to be adapted for diverse populations, such as infants and blind patients, who may have different functional organisations that could affect model performance. Investigating the retinotopy of these populations is crucial for understanding the mechanisms underlying neurodevelopment and pathology. To translate the proposed framework into a clinically reliable routine, it must account for the differences in functional organisation among various populations.

6 Conclusion

Contents

6.1 Contributions

This thesis delved into three unique facets of individual variability in the human brain *in vivo*, encompassing both anatomical distinctions and functional variations, using MRI analysis and machine learning techniques. In this chapter, we will recapitulate the primary contributions of this thesis and discuss its limitations. Lastly, we will concisely outline potential future research directions stemming from this work.

- **An Imaging Quality Transfer (IQT) Technique to Localise a Deep Brain Stimulation (DBS) Target**

  In Chapter 3, we proposed an IQT technique for robustly localising a DBS target, the Vim, on low-quality clinical-like data by leveraging anatomical information from high-quality data sources. We demonstrated that the HQ-augmentation approach outperforms its predecessor, the connectivity-driven

approach, in terms of accuracy, robustness against data corruption, and consistency across varying data qualities and scanning sessions. We also demonstrated that the HQ-augmentation approach preserves more inter-individual anatomical variability compared to the atlas-based approach. Furthermore, we presented the generalisability of the HQ-augmentation model across different datasets, such as the HCP and UKB low-quality data, making it more valuable in clinical contexts where large samples of data collection and retraining may be impractical. Finally, we showed the versatility of the approach by indicating its potential to be adapted for targeting other deep brain structures.

- **Accurate predictions of individual differences in task-evoked brain activity from resting-state fMRI using a sparse ensemble learner**

  In Chapter 4, we proposed a framework to improve state-of-the art prediction accuracy of individual differences in task-fMRI activation profiles using resting-state fMRI. Based on the assumption that resting-state functional modes are the underlying substrate of task-evoked activity, we demonstrated that features extracted using Stochastic Probabilistic Functional Modes (sPROFUMO) outperform the commonly-used dual-regression spatial modes. We showed that the shape and overall intensity of individualised task activations can be modelled separately and explicitly. Furthermore, we discussed the importance of "residualisation" and presented evidence that training the model on predicting residual differences in brain activity further boosts individualised predictions. The framework's state-of-the-art prediction accuracy was validated on both HCP data (surface-based analyses) and UKB data (volumetric analyses) and on par with the test-rest reliability of task-fMRI, suggesting its potential to supplement traditional task localisers.

- **Resting-state neural activity predicts individual differences in retinotopic organisation in early visual areas**

In Chapter 5, we presented a two-step supervised Laplacian Eigenmap (LE) approach for reconstructing retinotopic mapping in early visual areas using resting-state fMRI. In this framework, we optimised the construction of resting-state graphs by incorporating co-activation patterns from retinotopy task-fMRI, resulting in more visual-specific embeddings. We introduced an additional term to the Laplacian Eigenmap loss function that guides dimensionality reduction using group-average retinotopic embeddings, thereby encouraging individual resting-state graphs to resemble overall retinotopic organisations. We showed that the framework had improved performance in characterising individual-specific retinotopy in early visual areas (V1 to V4) from resting-state fMRI compared to unsupervised methods and group-average parcellation, demonstrating its potential to replace task-fMRI for characterising retinotopic organisations in the visual cortex and serving as a reliable functional localiser, particularly in cases where task-fMRI is impractical, such as with infants or disease cohorts.

6.2 Limitations

In this section, we will recapitulate the limitations of each project, as previously discussed in Chapters 3.4, 4.4, and 5.4, and provide a comprehensive overview of these constraints in a broader context.

- In Chapter 3, we proposed an IQT technique to reliably localise a DBS target, the Vim, leveraging high-quality data resources. There are several limitations. First, its reliance on high-quality data for model training may render it impractical in specific research or clinical settings. While the model exhibits generalisability from HCP to UKB datasets, it is crucial to evaluate its performance on other low-quality datasets and diverse populations to ensure widespread applicability. Additionally, due to variations in MRI acquisition protocols and hardware, the model might necessitate retraining or fine-tuning for particular clinical or research scenarios. Second, the model did not fully

explore other advanced techniques in machine learning. Recent advances in deep learning, particularly in medical image analysis, could provide alternative solutions with enhanced performance and adaptability to different datasets. For example, integrating deep learning methods like convolutional neural networks (CNNs) with the HQ-augmentation framework may lead to further progress in deep nuclei segmentation for low-quality data. Lastly, there is potential for further refinement by examining additional image features, optimising the model architecture, and incorporating advanced regularisation techniques to improve segmentation accuracy and robustness.

- In Chapter 4, we proposed an ensemble framework that achieves state-of-the-art prediction accuracy of individual differences in task-evoked brain activity. Several limitations exist. First, the approach relies on unsupervised decomposition of common modes of variations. Future work could adopt more sophisticated modelling strategies for simultaneous estimation of variations and reconstruction coefficients. Second, the model has not yet been extended to include the valuable information available from T1 and diffusion MRI scans, which could enhance the predictions of individual differences in task-evoked activity. Developing a multi-modal framework that integrates data from various imaging modalities would likely provide a more comprehensive and accurate representation of individual brain activity patterns, further improving the predictive capabilities of the model. Finally, the generalisability of the model to other datasets remains an area for further exploration. Evaluating the model's performance on diverse datasets, e.g., in populations, as well as refining the model to account for variations in MRI acquisition protocols and hardware, will be crucial to ensuring the broad applicability of the ensemble framework in both research and clinical settings.

- In Chapter 5, the proposed framework that predicts task retinotopy from resting-state fMRI has several methodological limitations. First, there is potential underfitting in the resting-state graph optimisation, where spatial

weights were precomputed using group-average spatial ICA maps. Optimising these weights simultaneously with coefficients may improve model performance. Second, the framework's prediction accuracy needs enhancement. While capturing individual differences in task retinotopy, its accuracy is only slightly better than group-average retinotopy. Predicted retinotopy distributions exhibit a Gaussian pattern, complicating comparisons with ground truth retinotopy. Further efforts are needed to achieve accurate retinotopy predictions. Additionally, the framework has been trained and validated on high-quality 7T HCP datasets, while using 3T fMRI data results in reduced performance. Advanced machine learning techniques may help transfer information from 7T fMRI to more common 3T hardware. Furthermore, the current framework focused solely on polar angle and eccentricity. While they are the most important properties regarding the organisation of the visual cortex, incorporating other aspects of retinotopy, such as pRF size, may provide a more comprehensive understanding of the functional organisation of the visual cortex. Finally, the framework needs adaptation to diverse populations, such as infants and blind patients, whose functional organisations may differ from the young and healthy HCP subjects. Understanding retinotopy in these populations is crucial, and the framework should account for such differences to become a clinically reliable tool.

Although each chapter varies in terms of the methodology and applications, they share several common limitations in the analysis of individual variability. The first shared limitation is restricted generalisability. As the chapters are all aimed at developing clinically reliable tools, it is crucial to ensure their applicability to diverse datasets collected in clinical settings, potentially from different disease/age population and different sites with varying MRI acquisition protocols and hardware configurations. Although the IQT technique presented in Chapter 3 demonstrates its generalisability from HCP to UKB datasets, the remaining two chapters do not investigate the feasibility of extending the models trained on one dataset to others with different MRI protocols or populations. To enhance the applicability

of the proposed approaches across a wider context, it is essential to examine advanced techniques that can accommodate the heterogeneity in MRI acquisition protocols, hardware, data populations, and other factors. For example, incorporating domain adaptation techniques could further improve the applicability of these approaches in various settings. The second notable limitation pertains to scalability. The proposed methods may not scale well to larger datasets or higher-resolution imaging data, which could impact their computational efficiency and effectiveness in real-world applications. For example, while the IQT technique introduced in Chapter 3 employed a mini-batch method that substantially reduced RAM requirements, its scalability for high-resolution data might still be limited due to the enforcement of local smoothness, which typically involves complex mathematical operations on voxel-by-voxel matrices. In Chapter 4, both the calculation of functional modes and the sparse model is computationally demanding due to the large size of voxels. Evaluating and optimising the scalability of these methods would be essential for their successful implementation in research and clinical settings. Finally, the current chapters may not have explored the full range of available analytical techniques, which could potentially enhance the predictive accuracy, generalisability, and robustness of the models. For example, incorporating advanced deep learning techniques, such as CNNs, could capture more complex and often complementary patterns in the MRI data and lead to better performance. In summary, while the current studies have demonstrated promising results, there is still room for improvement in model performance by considering more advanced machine learning techniques and continuously refining the models based on the latest developments in the field.

6.3 Future Directions

There are several future directions that can be pursued to address the aforementioned limitations. To tackle with restricted generalisability, it is necessary to investigate transfer learning or domain adaptation techniques and conduct validation studies on more diverse datasets to ensure the models' applicability to various clinical and

research contexts. Transfer learning and domain adaptation are machine learning techniques that leverage knowledge gained from one task or domain to improve the performance of a model in another related task or domain. The idea behind these techniques is that features or patterns learned from the source task (source domain) can inform the target task (target domain). In our context, MRI data acquired at different sites (or with different protocols and hardware) or from different population may have different data properties (e.g., having different SNRs). Transfer learning or domain adaptation techniques may help models to adapt to differences in data properties, making them more robust and applicable across various contexts.

Another future direction to consider is the integration of multimodal data. While each chapter primarily focuses on a single MRI modality to investigate specific aspects of individual variability, combining information from multiple imaging modalities could potentially enhance the performance of the proposed models. For example, localising the DBS target in Chapter 3 relies on diffusion MRI, while predicting the task-evoked brain activity in Chapter 4 and 5 utilises resting-state fMRI. Leveraging complementary information in multimodal MRI may lead to better performance of these models and provide better understanding of the underlying neural mechanisms that may not be fully captured by a single modality. An potential example is the integration of structural features (e.g., cortical thickness and myelination) into the retinotopy-prediction framework proposed in Chapter 5, as previous evidence has revealed that these structural features are also predictive of the retinotopic organisation in the visual cortex. Furthermore, incorporating multimodal MRI data could also facilitate the discovery of novel biomarkers, better prediction of clinical outcomes, and more effective individualised interventions.

A third future direction to pursue is the validation of the models on diverse populations. Since this thesis primarily focuses on the publicly available HCP and UKB datasets, the proposed methods have not yet been tested on a wide range of populations, such as specific age groups, genders, or individuals with particular health conditions who may have limited sample sizes. Conducting validation studies on diverse populations would be crucial in ensuring the broader

applicability and relevance of the developed methods. For example, it would be valuable to explore whether the retinotopy-prediction method presented in Chapter 5 can be applied to infants or individuals with visual impairments, such as the blind. Investigating the applicability of the proposed models in these populations could provide essential insights into the neural mechanisms underlying neurodevelopment or the pathology of the visual cortex. Moreover, by encompassing a wider range of populations, researchers can better understand the potential impact of factors such as age, sex, and health status on the performance of the proposed models. This broader validation will not only help refine the models but also contribute to a more comprehensive understanding of individual variability in brain function and structure, ultimately promoting the development of personalised treatments for diverse populations.

Appendices

A

Additional Results for Chapter 3

A.0.1 Selection criteria of reliable HQ-Vim

The subjects were split into two subsets, depending on the reliability of HQ-Vim. A subject's HQ-Vim has to pass 4 criteria in order to be accepted as "reliable":

1. the HQ-Vim's volume exceeds 30mm^3;
2. the HQ-Vim contains one blob;
3. Its correlation with the Vim from Thalamic DBS Connecitivity Atlas (Akram et al., 2018) is larger than 0.5;
4. Its center-of-mass is within 4mm from the center of mass of the Thalamic DBS Connecitivity Dentate Atlas [34];

The thresholds in 3. and 4. are shown as vertical lines in Figure A1. 459 out of 1063 HCP subjects and 1445 out of 2760 UKB subjects passed all four criteria. These selection criteria exclude subjects whose HQ-Vim clusters locate too far away from the atlas to be considered as trustworthy, while preserving the inter-individual anatomical variability of the structure as much as possible. Note that, however, passing all the selection criteria does not necessarily guarantee the selected HQ-Vim as the perfect "ground truth". Instead, this only suggests that the HQ-Vim may

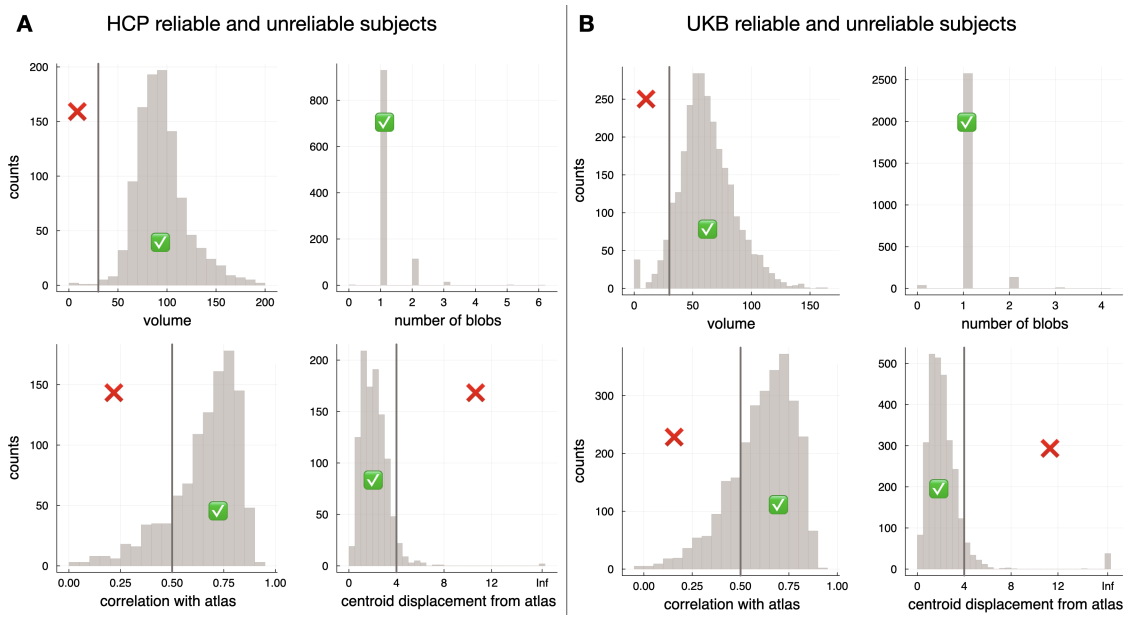

Figure A1: Split between the reliable subset and unreliable subset. **(A)** Histograms of HQ-Vim's volume (mm^3), number of blobs, correlations and centroid displacement with the Thalamic DBS Connecitivity Atlas, across HCP subjects. **(B)** Equivalent plots of (A), across UKB subjects. Ticks suggest that the subjects fell above the respective criteria, while crosses suggest their HQ-Vim were rejected as untrustworthy.

serve as the ground truth Vim with relatively high confidence, as it is the best available estimate of the ground truth location of the Vim.

A.0.2 List of ROIs

We provide a table listing the ROIs used in this study, as well as the respective tractography protocols and the procedures to obtain them.

ROI name (per hemisphere)	**Procedures to obtain it**
Thalamus	Obtained using T1, T2 and FA via FSL's MIST [115], a multimodal subcortical segmentation tool
Brainstem (both hemisphere)	Obtained via Freesurfer's subcortical segmentation tool, as released in the *aparc.a2009s.aseg.nii.gz files
Cerebellum grey matter (WM)	
Cerebellum grey matter (GM)	
Caudate	
Pallidum	
Putamen	
Cerebrospinal fluid (CSF)	
Ventrical	
AR-1, AR-2, AR-3, AR-4, AR-5, and AR-6: six Acoustic Radiation fiber segments	Manually extracted from the Acoustic Radiation fibers (XTRACT atlas), and warped into individual space
STR-1, STR-2, STR-3, STR-4, STR-5, and STR-6: six Superior Thalamic Radiation WM segments	Manually extracted from the Superior Thalamic Radiation fibers (XTRACT atlas), and warped into individual space
OR-1, OR-2, OR-3, OR-4, OR-5, and OR-6: 6 Optic Radiation fiber segments	Manually extracted from the Optic Radiation fibers (XTRACT atlas), and warped into individual space
ATR-1, ATR-2, ATR-3, ATR-4, ATR-5, and ATR-6: 6 Anterior Thalamic Radiation fiber segments	Manually extracted from the Anterior Thalamic Radiation fibers (XTRACT atlas), and warped into individual space
FX-1, FX-2, FX-3, FX-4, and FX-5: five Fornix ROIs	Manually extracted from the Fornix (XTRACT atlas), and warped into individual space
SCPCT-brainstem, SCPCT-1, SCPCT-2, and SCPCT-3: four segments of the SCPCT tract, three lying between the brainstem ROI and ipsilateral thalamus ROI (SCPCT-1, SCPCT-2, SCPCT-3), one overlapping with the brainstem ROI (SCPCT-brainstem)	Extracted from the cerebellothalamic tract [121], and warped into individual space

4 white matter segments on the tracts that connect M1 and thalamus	manually extracted from the tract that joins M1 and thalamus
4 white matter segments on the tracts that connect S1 and thalamus	manually extracted from the tract that joins S1 and thalamus
G_and_S_frontomargin: Fronto-marginal gyrus (of Wernicke) and sulcus	Obtained from the Destrieux Atlas [120], via Freesurfer's cortical parcellation tool (`https://surfer.nmr.mgh.harvard.edu/fswiki/CorticalParcellation`) [118, 119]
G_and_S_occipital_inf: inferior occipital gyrus (O3) and sulcus	
G_and_S_paracentral: Paracentral lobule and sulcus	
G_and_S_subcentral: Subcentral gyrus (central operculum) and sulci	
G_and_S_transv_frontopol: Transverse frontopolar gyri and sulci	
Transverse frontopolar gyri and sulci	
G_and_S_cingul-Mid-Ant: Middle-anterior part of the cingulate gyrus and sulcus	
G_and_S_cingul-Mid-Post: Middle-posterior part of the cingulate gyrus and sulcus	
G_cingul-Post-dorsal: Posterior-dorsal part of the cingulate gyrus	
G_cingul-Post-ventral: Posterior-ventral part of the cingulate gyrus	
G_cuneus: Cuneus	
G_front_inf-Opercular: Opercular part of the inferior frontal gyrus	
G_front_inf-Orbital: Orbital part of the inferior frontal gyrus	
G_front_inf-Triangul: Triangular part of the inferior frontal gyrus	
G_front_middle: Middle frontal gyrus	
G_front_sup: Superior frontal gyrus	
G_Ins_lg_and_S_cent_ins: Long insular gyrus and central sulcus of the insula	
G_insular_short : Short insular gyri	
G_occipital_middle: Middle occipital gyrus	
G_occipital_sup: Superior occipital gyrus	
G_oc-temp_lat-fusifor: Lateral occipito-temporal gyrus	
G_oc-temp_med-Lingual: Lingual gyrus, ligual part of the medial occipito-temporal gyrus	
G_oc-temp_med-Parahip: Parahippocampal gyrus, parahippocampal part of the medial occipito-temporal gyrus	
G_orbital: Orbital gyri	

G_pariet_inf-Angular: Angular gyrus
G_pariet_inf-Supramar: Supramarginal gyrus
G_parietal_sup: Superior parietal lobule
G_postcentral: Postcentral gyrus
G_precentral: Precentral gyrus
G_precuneus: Precuneus
G_rectus: Straight gyrus
G_subcallosal: Subcallosal area, subcallosal gyrus
G_temp_sup-G_T_transv: Anterior transverse temporal gyrus
G_temp_sup-Lateral: Lateral aspect of the superior temporal gyrus
G_temp_sup-Plan_polar: Planum polare of the superior temporal gyrus
G_temp_sup-Plan_tempo : Planum temporale or temporal plane of the superior temporal gyrus
G_temporal_inf: Inferior temporal gyrus
G_temporal_middle:
Lat_Fis-ant-Horizont: Horizontal ramus of the anterior segment of the lateral sulcus
Lat_Fis-ant-Vertical: Vertical ramus of the anterior segment of the lateral sulcus
Lat_Fis-post: Posterior ramus (or segment) of the lateral sulcus
Medial_wall:
Pole_occipital: Occipital pole
Pole_temporal: Temporal pole
S_calcarine: Calcarine sulcus
S_central: Central sulcus (Rolando's fissure)
S_cingul-Marginalis: Marginal branch (or part) of the cingulate sulcus
S_circular_insula_ant: Anterior segment of the circular sulcus of the insula
S_circular_insula_inf: Inferior segment of the circular sulcus of the insula
S_circular_insula_sup: Superior segment of the circular sulcus of the insula
S_collat_transv_ant: Anterior transverse collateral sulcus
S_collat_transv_post: Posterior transverse collateral sulcus
S_front_inf: Inferior frontal sulcus
S_front_middle: Middle frontal sulcus
S_front_sup: Superior frontal sulcus

S_interm_prim-Jensen: Sulcus intermedius primus
S_intrapariet_and_P_trans: Intraparietal sulcus (interparietal sulcus) and transverse parietal sulci
S_oc_middle_and_Lunatus: Middle occipital sulcus and lunatus sulcus
S_oc_sup_and_transversal: Superior occipital sulcus and transverse occipital sulcus
S_occipital_ant: Anterior occipital sulcus and preoccipital notch (temporo-occipital incisure)
S_oc-temp_lat: Lateral occipito-temporal sulcus
S_oc-temp_med_and_Lingual: Medial occipito-temporal sulcus (collateral sulcus) and lingual sulcus
S_orbital_lateral: Lateral orbital sulcus
S_orbital_med-olfact: Medial orbital sulcus (olfactory sulcus)
S_orbital-H_Shaped: Orbital sulci (H-shaped sulci)
S_parieto_occipital: Parieto-occipital sulcus (or fissure)
S_pericallosal: Pericallosal sulcus (S of corpus callosum)
S_postcentral: Postcentral sulcus
S_precentral-inf-part: Inferior part of the precentral sulcus
S_precentral-sup-part: Superior part of the precentral sulcus
S_suborbital: Suborbital sulcus (sulcus rostrales, supraorbital sulcus)
S_subparietal: Subparietal sulcus
S_temporal_inf: Inferior temporal sulcus
S_temporal_sup: Superior temporal sulcus
S_temporal_transverse: Transverse temporal sulcus

Table A1: The list of anatomical ROIs used in Chapter 3.

A.0.3 Tractography protocols

Target mask (per hemisphere)	**Waypoint mask**	**Exclusion mask**	**termination mask**
Contralateral cerebellar WM (dentate)	SCPCT-brainstem, contralateral cerebellar WM	Ipsilateral cerebellum WM and GM, CSF, Putamen, Caudate, contralateral thalamus	cortex
SCPCT-brainstem (the segment of SCPCT that overlap with the brainstem ROI)	SCPCT-brainstem	Ipsilateral cerebellum WM and GM, CSF, Putamen, Caudate, contralateral thalamus	cortex
SCPCT-1, SCPCT-2, and SCPCT-3	Same as the respective target mask	Ipsilateral cerebellum WM and GM, CSF, Putamen, Caudate, contralateral thalamus	SCPCT-brainstem, cortex
G_and_S_frontomargin	same as target mask	Ipsilateral cerebellum, CSF	cortex
G_and_S_occipital_inf			
G_and_S_paracentral			
G_and_S_subcentral			
G_and_S_transv_frontopol			
G_and_S_cingul-Mid-Ant			
G_and_S_cingul-Mid-Post			
G_cingul-Post-dorsal			
G_cingul-Post-ventral			
G_cuneus			
G_front_inf-Opercular			
G_front_inf-Orbital			
G_front_inf-Triangul			
G_front_middle			
G_front_sup			
G_Ins_lg_and_S_cent_ins			
G_insular_short			
G_occipital_middle			
G_occipital_sup			
G_oc-temp_lat-fusifor			

G_oc-temp_med-Lingual
G_oc-temp_med-Parahip
G_orbital
G_pariet_inf-Angular
G_pariet_inf-Supramar
G_parietal_sup
G_postcentral
G_precentral (M1)
G_precuneus
G_rectus
G_subcallosal
G_temp_sup-G_T_transv
G_temp_sup-Lateral
G_temp_sup-Plan_polar
G_temp_sup-Plan_tempo
G_temporal_inf
G_temporal_middle
Lat_Fis-ant-Horizont
Lat_Fis-ant-Vertical
Lat_Fis-post
Medial_wall
Pole_occipital
Pole_temporal
S_calcarine
S_central
S_cingul-Marginalis
S_circular_insula_ant
S_circular_insula_inf
S_circular_insula_sup
S_collat_transv_ant
S_collat_transv_post
S_front_inf
S_front_middle
S_front_sup
S_interm_prim-Jensen
S_intrapariet_and_P_trans
S_oc_middle_and_Lunatus
S_oc_sup_and_transversal
S_occipital_ant
S_oc-temp_lat
S_oc-temp_med_and_Lingual
S_orbital_lateral
S_orbital_med-olfact
S_orbital-H_Shaped

S_parieto_occipital			
S_pericallosal			
S_postcentral			
S_precentral-inf-part			
S_precentral-sup-part			
S_suborbital			
S_subparietal			
S_temporal_inf			
S_temporal_sup			
S_temporal_transverse			
AR1-AR6			
STR1-STR6			
ATR1-ATR6			
OR1-OR6			
FX1-FX5			
S1-1-4			
M1-1-4			

Table A2: The list of seed, waypoint, exclusion and termination masks used in tractography.

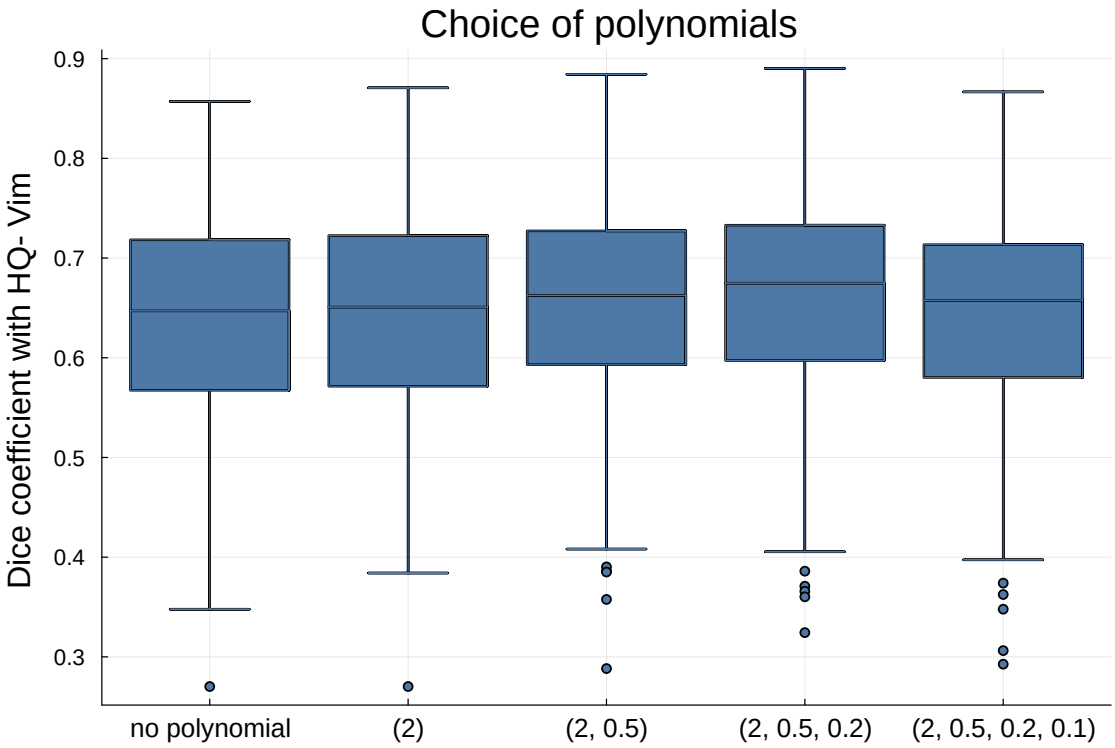

Figure A2: Choices of polynomial features. In addition to the original features (no polynomial), we tested a range of polynomial features on a smaller independent subset of subjects. The numbers in the x-axis denote the powers of polynomials used in the HQ-augmentation model, e.g., (2, 0.5, 0.2, 0.1) indicates that four additional polynomial features were included, with $p_1 = 2, p_2 = 0.5, p_3 = 0.2, p_4 = 0.1$ where p_1, p_2, p_3, p_4 were the powers. Overall, the choices of polynomials did not affect the Dice coefficient with HQ-Vim very much. However, (2, 0.5, 0.2), i.e., $p_1 = 2, p_2 = 0.5, p_3 = 0.2$, appeared to be slightly better than the other choices.

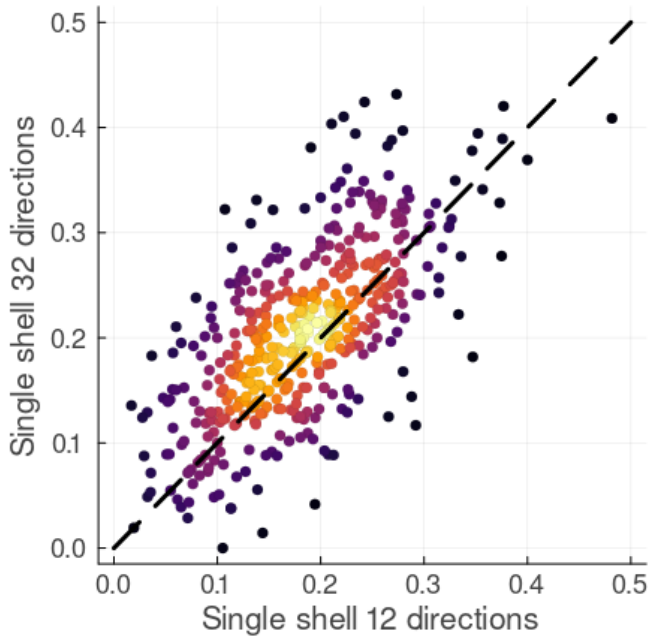

Figure A3: Dice coefficient with the HQ-Vim that excluded the atlas-defined Vim content, for the HQ-augmentation approach on two types of low-quality data, single shell sampled at 32 directions and 12 directions. We further degraded the LQ-LowAngular data (sampled at 32 directions) by extracting volumes that correspond to 12 directions. Therefore, "Single shell 12 directions" is a low-quality version of "Single shell 32 directions". We trained a HQ-augmentation model tailored to the "single-12" dataset as we did for the LQ-LowAngular, and generated the corresponding HQ-augmented Vim. For each dataset ("single-12" or LQ-LowAngular, a.k.a "single-32"), we calculated the Dice coefficient between the HQ-augmented Vim and the HQ-Vim that excluded the atlas-defined Vim (i.e., the voxels belonging to the HQ-Vim subtract the voxels belonging to the atlas-defined Vim). Hence, higher Dice coefficient here suggests less similarity to the atlas. Such Dice coefficients on "single-32" are higher than those on "single-12", suggesting that the HQ-augmented Vim on more corrupted dataset relies more on the group-average information.

A.0.4 Mean-field approximation of the CRF distribution

As mentioned in Chapter 3.2.6.2, we seek to maximise the probability of reproducing the exact same HQ-Vim label assignment $\mathbf{y}$ on its low-quality counterparts, across the training subjects

$$\log P(\mathbf{y}|\mathbf{X}) = \log[\frac{1}{Z(\mathbf{X})}\exp(-E(\mathbf{y}|\mathbf{X}))] \tag{A1}$$

where the cost (energy) of label assignment $\mathbf{y}$ given the connectivity features $\mathbf{X}$ is modelled as

$$E(\mathbf{y}|\mathbf{X}) = \sum_i \psi_u(y_i|\mathbf{x}_i) + \sum_i \sum_{j\in\mathcal{N}_i} \psi_p(y_i, y_j|\mathbf{x}_i, \mathbf{x}_j) + \lambda_1||\mathbf{W}||_1 + \lambda_2||\mathbf{W}||_2^2 \tag{A2}$$

Here we briefly re-explain these terms. The first component $\psi_u(y_i)$ measures the cost (or inverse likelihood) of voxel i taking label y_i. Here $\psi_u(y_i)$ takes the form $\psi_u(y_i = k) = \mathbf{w}_k^T\phi(\mathbf{x}_i)$ for $k = 1, 2$, where $\phi(\cdot)$ maps a feature vector $\mathbf{x}_i = [x_1, x_2, ...x_d]$ to a further expanded feature space in order to provide more flexibility for the parameterisation. $\mathbf{W} = [\mathbf{w}_1, ...\mathbf{w}_K]$ is the coefficient matrix ($K = 2$), each column containing the coefficients for the given class (i.e., belonging to the HQ-Vim or not). Here we chose a series of polynomials along with the group-average Vim probability (registered into native space) to expand the feature space, i.e.,

$$\phi(\mathbf{x}_i) = [x_1, x_2, ...x_d, x_1^{p_1}, x_2^{p_1}, ...x_d^{p_1}, x_1^{p_2}, x_2^{p_2}, ...x_d^{p_2}, x_1^{p_3}, x_2^{p_3}, ...x_d^{p_3}, g_i]$$

where $p_1 = 2, p_2 = 0.5, p_3 = 0.2$ are the power of the polynomials, and g_i is the group-average probability of voxel i classified as Vim (calculated across the training subjects). The second pairwise cost encourages assigning similar labels to neighbouring voxels, particularly for those sharing similar connectivity features. We modelled this component as $\psi_p(\mathbf{x}_i, \mathbf{x}_j) = \mu(y_i, y_j)\rho_m k_m(\phi(\mathbf{x}_i), \phi(\mathbf{x}_j))$. Here $k_m(\phi(\mathbf{x}_i), \phi(\mathbf{x}_j)) = \exp(-\gamma_m||\phi(\mathbf{x}_i) - \phi(\mathbf{x}_j)||^2)$ is a kernel function modelling the similarity between voxel i and j in the extended feature space, with length scale γ_m, chosen via cross-validation. $\mu(\cdot)$ is a label compatibility function where $\mu(y_i, y_j) = 0$ if $y_i = y_j$ or $\mu(y_i, y_j) = 1$ if $y_i \neq y_j$.

Here we used cross entropy loss to maximise the log-likelihood in Equation (A1). Specifically, suppose $\mathbf{T}$ is the one-hot coding matrix of the HQ-Vim labels $\mathbf{y}$, i.e., y_i is mapped to a binary vector $\mathbf{t}_i = [t_{i1}, ...t_{iK}]$, in which $y_i = k$ corresponds to $t_{ik} = 1$ and other elements in $\mathbf{t}_i$ set to 0. To maximise the log-likelihood in (A1) is equivalent to minimising the cross entropy (negative log likelihood):

$$\min \sum_{i}^{V} \sum_{j}^{K} t_{ij} \log(P(y_i|\mathbf{x}_i)) \tag{A3}$$

Due to the inter-dependency of neighbouring voxels, the exact minimisation is intractable. Thus, we approximate the CRF distribution $P(\mathbf{y}|\mathbf{X})$ by a simpler function $Q(\mathbf{y})$, and iteratively solve this optimisation problem, explained below.

To set up initialisation, we derived the initial coefficients $\hat{\mathbf{W}}^{(0)} = [\hat{\mathbf{w}}_1^{(0)}, ...\hat{\mathbf{w}}_K^{(0)}]$ by optimising the following

$$\hat{\mathbf{W}}^{(0)} = \arg\min_{\mathbf{W}} \sum_{i}^{V} \sum_{j}^{K} t_{ij} \log(P'(y_i|\mathbf{x}_i)) \tag{A4}$$

where $P'(y_i|\mathbf{x}_i)$ is the likelihood without considering local smoothness of the label assignment, i.e., ignoring the pairwise loss term in Equation (A2):

$$P'(y_i = l|\mathbf{x}_i) = \frac{\exp(-\psi_u(y_i = l|\mathbf{x}_i))}{\sum_k^K \exp(-\psi_u(y_i = k|\mathbf{x}_i))} \tag{A5}$$

where, as defined before, $\psi_u(y_i = l|\mathbf{x}_i) = \mathbf{w}_l^T \phi(\mathbf{x}_i)$. The coefficients $\hat{\mathbf{W}}^{(0)}$ were used to initialise $Q(\mathbf{y})$ using Equation (A5), i.e., $Q_i(y_i = l) \leftarrow P'(y_i = l|\mathbf{x}_i)$ evaluated at $\hat{\mathbf{W}}^{(0)}$.

After initialisation, $Q_i(y_i = l)$ is updated as the weighted sum of its neighbouring Q values, $\tilde{Q}_i(y_i = l) \leftarrow \rho_m \sum_{j \in \mathcal{N}_i} k_m(\phi(\mathbf{x}_i), \phi(\mathbf{x}_j)) Q_j(y_j = l)$, where $\tilde{Q}_i(y_i = l)$ is the updated Q value. This is the message-passing step and is equivalent to applying M image-dependent Gaussian filters on the Q values. After message passing, label incompatibility was calculated as a penalty to encourage local smoothness. The incompatibility for label l at a given voxel i, denoted by $\hat{Q}_i(y_i = l)$, was calculated as the sum of the updated $\tilde{Q}_i$ that take a different label, i.e., $\hat{Q}_i(y_i = l) \leftarrow \sum_{l'} \mu(l, l') \tilde{Q}_i(y_i = l')$. Next, this penalty incurred by incompatibility was subtracted from the unary inputs $\psi_u(y_i = l|\mathbf{x}_i)$, i.e., $Q_i(y_i = l) \leftarrow \frac{1}{Z_i'} \exp(-\psi_u(y_i =$

	LQ-LowAngular		LQ-LowSpatial	
	Dice coefficient	Centroid displacement	Dice coefficient	Centroid displacement
HQ-augmentation vs connectivity-driven	<1e-99	<1e-54	<1e-99	<1e-89
HQ-augmentation vs atlas-based	<1e-10	<1e-17	0.0043	<1e-17
	LQ-LowAngular-LowSpatial		LQ-UKB	
	Dice coefficient	Centroid displacement	Dice coefficient	Centroid displacement
HQ-augmentation vs connectivity-driven	<1e-72	<1e-38	<1e-96	<1e-66
HQ-augmentation vs atlas-based	<1e-4	1e-12	<1e-19	<1e-12

Table A3: P values of paired t-tests comparing the HQ-augmentation against the alternative methods, based on their correspondence with HQ-Vim. Metrics used for comparison include the Dice coefficient and centroid displacement. P values were Bonferroni corrected (12 tests in total).

$l|\mathbf{x}_i) - \hat{Q}_i(y_i = l))$, where $Z_i' = \sum_k^K \exp(-\psi_u(y_i = k|\mathbf{x}_i) - \hat{Q}_i(y_i = k))$ is the normalisation constant. The above steps were repeated until Q converges. The resulting $Q_i(y_i|\mathbf{x}_i)$ is an approximation of the likelihood $P(y_i|\mathbf{x}_i)$, and was used when calculating the cross entropy in Equation (A3). This cross entropy (A3) was minimised in a mini-batch style via an ADAM optimiser [130] with learning rate 0.01, in which the connectivity features $\mathbf{X}$ of each subject served as a mini-batch.

The pseudo code of the above steps is summarised in Table 2.

Algorithm 2 Mean-field iteration in CRF

$Q_i(y_i = l) \leftarrow \exp(-\psi_u(y_i = l|\mathbf{x}_i))/\sum_k \exp(-\psi_u(y_i = k|\mathbf{x}_i))$ ▷ Initialisation
while not converged **do**
  $\tilde{Q}_i(y_i = l) \leftarrow \rho_m \sum_{j \in \mathcal{N}_i} k_m(\phi(\mathbf{x}_i), \phi(\mathbf{x}_j)) Q_j(y_j = l)$ ▷ Message Passing
  $\hat{Q}_i(y_i = l) \leftarrow \sum_{l'} \mu(l, l') \tilde{Q}_i(y_i = l')$ ▷ Compatibility with neighbours
  $Q_i(y_i = l) \leftarrow \frac{1}{Z_i'} \exp(-\psi_u(y_i = l|\mathbf{x}_i) - \hat{Q}_i(y_i = l))$ ▷ Approximate $P(y_i = l|\mathbf{x}_i)$
end while

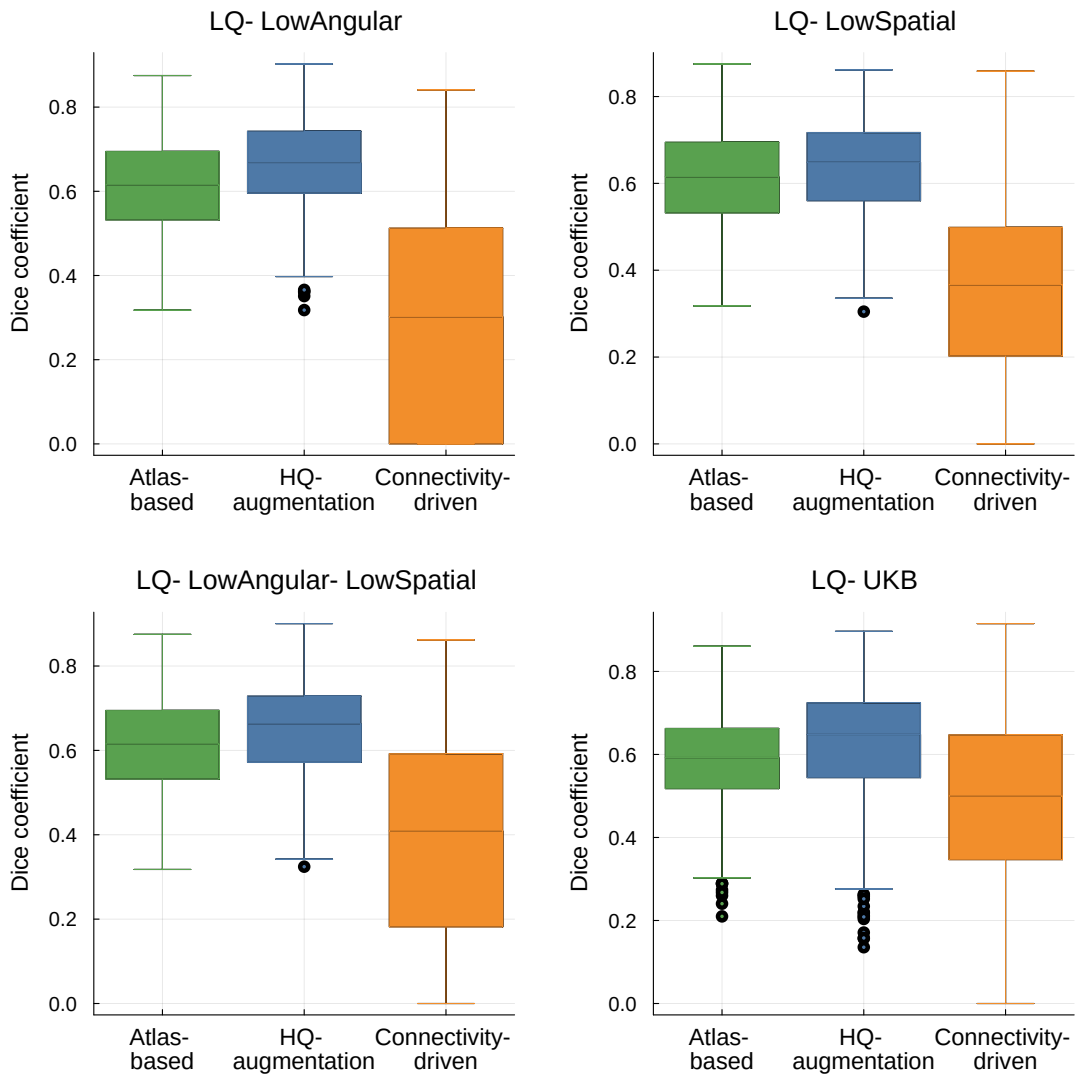

Figure A4: Boxplots of the Dice coefficient with HQ-Vim on surrogate low-quality datasets. **(A)**. Dice coefficient with the HQ-Vim, produced by the atlas-based (green), HQ-augmentation (blue), and connectivity-driven (orange) approach, on LQ-LowAngular. **(B)**. Equivalent plots of **(A)**, on LQ-LowSpatial. **(C)**. Equivalent plots of **(A)**, on LQ-LowAngular-LowSpatial. **(D)**. Equivalent plots of **(A)**, on LQ-UKB (where the HQ-augmentation model was trained on HCP and applied on LQ-UKB).

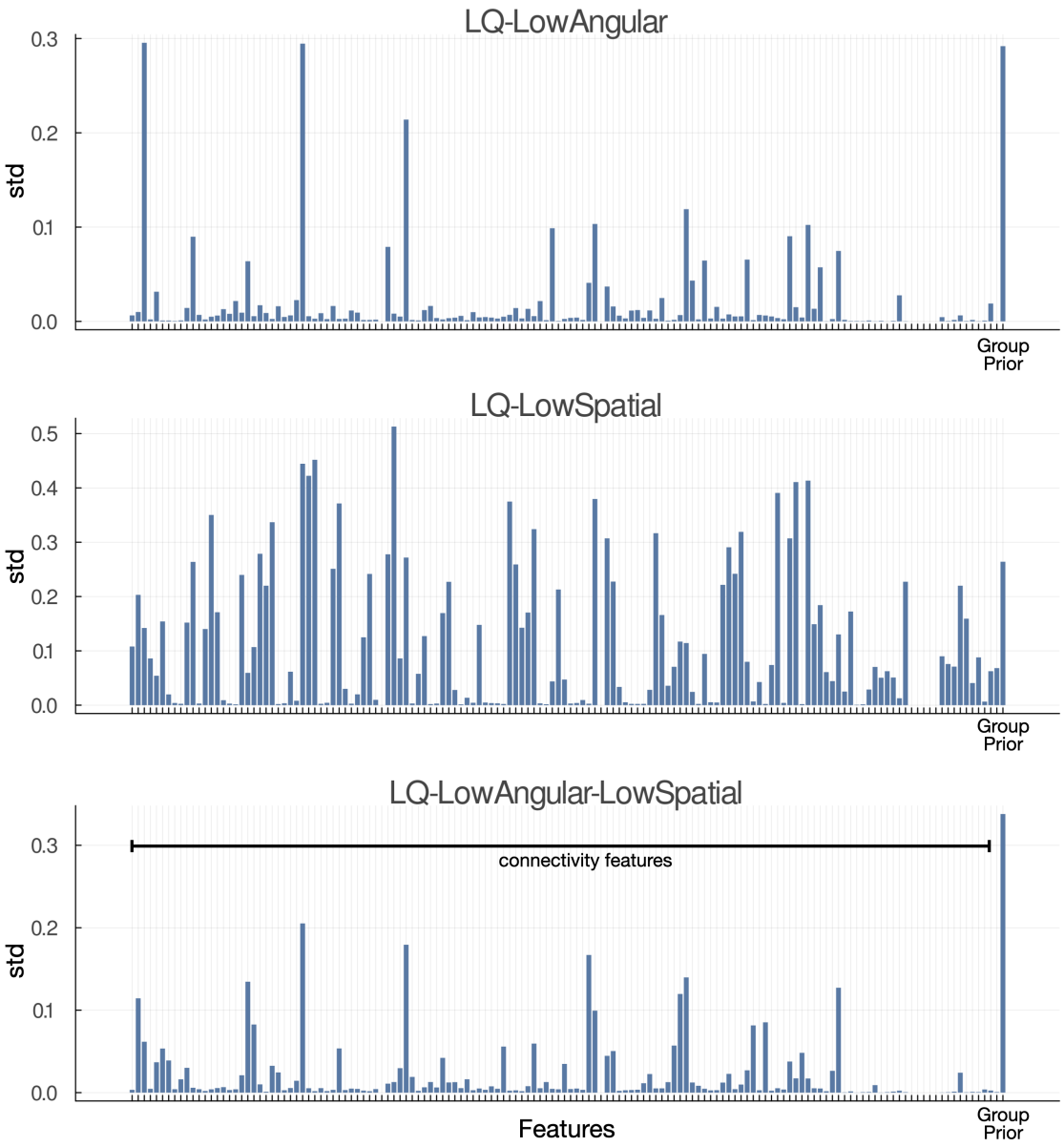

Figure A5: Relative contributions of the features in HQ-augmentation model for LQ-LowAngular, LQ-LowSpatial, LQ-LowAngular-LowSpatial. We plotted the std. of each feature contribution (i.e., the feature times its coefficient) across 50 subjects. The group-average prior feature contributed the most to Vim prediction, which is unsurprising given that the ground truth HQ-Vim labels were confined to be close to the atlas-defined Vim. For the purpose of visualisation, we summed the std. across the polynomials of a given connectivity feature. For example, the std. for M1 shown here was obtained by summing the std. of M1's polynomial features. Specifically, suppose β_1 is the coefficient for $\mathbf{x}_{\text{M1}}$, and β_2 for $\mathbf{x}_{\text{M1}}^2$, β_3 for $\mathbf{x}_{\text{M1}}^{0.5}$, and β_4 for $\mathbf{x}_{\text{M1}}^{0.2}$; then the std for M1 shown here is $std(\beta_1\mathbf{x}_{\text{M1}}) + std(\beta_2\mathbf{x}_{\text{M1}}^2) + std(\beta_3\mathbf{x}_{\text{M1}}^{0.5}) + std(\beta_4\mathbf{x}_{\text{M1}}^{0.2})$ across 50 subjects.

B

Additional Results for Chapter 4

HCP contrast category	**HCP contrast index and name**
EMOTION	01_EMOTION_FACES 02_EMOTION_SHAPES 03_EMOTION_FACES-SHAPES
GAMBLING	07_GAMBLING_PUNISH 08_GAMBLING_REWARD 09_GAMBLING_PUNISH-REWARD
LANGUAGE	13_LANGUAGE_MATH 14_LANGUAGE_STORY 15_LANGUAGE_MATH-STORY
MOTION	19_MOTOR_CUE 20_MOTOR_LF 21_MOTOR_LH 22_MOTOR_RF 23_MOTOR_RH 24_MOTOR_T 25_MOTOR_AVG 26_MOTOR_CUE-AVG 27_MOTOR_LF-AVG 28_MOTOR_LH-AVG 29_MOTOR_RF-AVG 30_MOTOR_RH-AVG 31_MOTOR_T-AVG
RELATIONAL	45_RELATIONAL_MATCH 46_RELATIONAL_REL 47_RELATIONAL_MATCH-REL
SOCIAL	51_SOCIAL_RANDOM 52_SOCIAL_TOM 53_SOCIAL_RANDOM-TOM
WORKING MEMORY	57_WM_2BK 58_WM_2BK 59_WM_2BK 60_WM_2BK 61_WM_0BK 62_WM_0BK 63_WM_0BK 64_WM_0BK 65_WM_2BK 66_WM_0BK 67_WM_2BK-0BK 71_WM_BODY 72_WM_FACE 73_WM_PLACE 74_WM_TOOL 75_WM_BODY-AVG 76_WM_FACE-AVG 77_WM_PLACE-AVG 78_WM_TOOL-AVG

Table B1: List of the 47 HCP contrasts. We used the 47 unique contrast maps for HCP, excluding all redundant contrasts.

Notation	Explanation	Notation	Explanation
V	the number of voxels	k	the number of dual-regression maps per subject
$\mathcal{S}$	the training set	$\mathcal{T}$	the test set
N	the number of training subjects	n	the number of test subjects
$\mathbf{x}_j^i \in \mathbb{R}^V$	the i-th resting-state variation map of subject j	$\mathbf{X}_j \in \mathbb{R}^{V\times k}$	the resting-state variation maps of subject j
$\boldsymbol{\beta}_j, \hat{\boldsymbol{\beta}}_j \in \mathbb{R}^k$	the (estimated) baseline coefficients for subject j	$\hat{\boldsymbol{\beta}} \in \mathbb{R}^k$	the baseline coefficients (averaged across the training set)
$\tilde{\mathbf{X}}_{\mathcal{S}}^i \in \mathbb{R}^{N\times V}$	the i-th across-subject resting-state variation matrix (of the training set)	$\mathbf{Y}_{\mathcal{S}}, \hat{\mathbf{Y}}_{\mathcal{S}} \in \mathbb{R}^k$	the (predicted) across-subject task variation matrix
$\mathbf{A}_{\mathcal{S}}^i, \hat{\mathbf{A}}_{\mathcal{S}}^i \in \mathbb{R}^{N\times d}$	the (estimated) mixing matrix of the i-th across-subject resting-state variation matrix	$\mathbf{S}^i \in \mathbb{R}^{d\times V}$	the independent components of the i-th across-subject resting-state variation matrix
$\mathbf{A}_{\mathcal{S}}^{\text{rest}}, \hat{\mathbf{A}}_{\mathcal{S}}^{\text{rest}} \in \mathbb{R}^{N\times dk}$	the concatenated k mixing matrices of the resting-state variation matrices	d	the number of modes/independent components of each resting-state variation matrix
$\mathbf{A}_{\mathcal{S}}^{\text{task}} \in \mathbb{R}^{N\times p}$	the mixing matrix of the across-subject task variation matrix	$\mathbf{S}^{\text{task}} \in \mathbb{R}^{p\times V}$	the independent components of the task variation matrix
$\mathbf{y}_j \in \mathbb{R}^V$	the task variation map of subject j	p	the number of modes/independent components of the task variation matrix
$\mathbf{W}, \hat{\mathbf{W}} \in \mathbb{R}^{dk\times p}$ or $\mathbb{R}^{dk\times V}$	the (estimated) sparse coefficients	$\mathbf{w}_i, \hat{\mathbf{w}}_i \in \mathbb{R}^{dk}$	the i-th column of the estimated sparse coefficients
λ_i	the hyper-parameter of the L_1 penalty for the i-th column of $\mathbf{W}$	$\hat{\mathbf{Y}}_{\mathcal{S}}^{\text{baseline}} \in \mathbb{R}^{N\times V}$	the baseline-model-fitted task variation matrix for the training subjects
$\hat{\mathbf{Y}}_{\mathcal{S}}^{\text{sparse}} \in \mathbb{R}^{N\times V}$	the sparse-model-fitted task variation matrix for the training subjects	$\hat{\mathbf{y}}_{\cdot i}^{\text{baseline}}, \hat{\mathbf{y}}_{\cdot i)}^{\text{sparse}} \in \mathbb{R}^N$	the i-th column/voxel of the baseline- / sparse-model fitted task variation matrix
$\theta_i^{(1)}$	ensemble coefficient for the i-th voxel of the baseline mode	$\theta_i^{(2)}$	ensemble coefficient for the i-th voxel of the sparse mode

Table B2: List of the notations used in Chapter 4.

Prediction accuracy of the residualised activations	UKB	HCP
DR-ICA baseline vs PFM baseline	p<1e-30	p<1e-30
DR-ICA sparse vs PFM sparse	p<1e-30	p<1e-30
DR-ICA ensemble vs PFM ensemble	p<1e-30	p<1e-30
DR-ICA baseline vs DR-ICA sparse	p<1e-30	p<1e-30
DR-ICA baseline vs DR-ICA ensemble	p<1e-30	p<1e-30
DR-ICA sparse vs DR-ICA ensemble	p<1e-30	p<1e-30
PFM baseline vs PFM sparse	p<1e-30	p<1e-30
PFM baseline vs PFM ensemble	p<1e-30	p<1e-30
PFM sparse vs PFM ensemble	p<1e-30	p<1e-30

Table B3: Two-sided paired t-test between prediction accuracies of the baseline, sparse and ensemble model (residualised), based on DR-ICA or PFM repsectively. The task contrasts were pooled together, corresponding to the last columns of Figure 4.2a and 4.2b (UKB three contrasts and HCP 47 contrasts, across all subjects). The p-values were all highly significant, after being Bonferroni-corrected (18 tests in total).

Correlation with inter-individual variability	UKB	HCP
DR-ICA baseline vs PFM baseline	0.0243	4.5176e-12
DR-ICA sparse vs PFM sparse	0.6724	0.0003
DR-ICA ensemble vs PFM ensemble	1.1192	0.0218
DR-ICA baseline vs DR-ICA sparse	0.8775	5.7470e-5
DR-ICA baseline vs DR-ICA ensemble	0.6582	1.4465e-16
DR-ICA sparse vs DR-ICA ensemble	0.1026	1.7612e-25
PFM baseline vs PFM sparse	3.2002	0.1466
PFM baseline vs PFM ensemble	2.3015	5.4127e-8
PFM sparse vs PFM ensemble	0.0005	5.5898e-23

Table B4: Two-sided paired t-test between correlations with the inter-individual variability (std map) of the baseline, sparse and ensemble model (residualised), based on DR-ICA or PFM repsectively. The task contrasts were pooled together, corresponding to the last columns of Figure 4.2c and 4.2d (UKB three contrasts and HCP 47 contrasts). The p-values were Bonferroni-corrected (18 tests in total). Overall, UKB results were non-significant, which is un-surprising given that UKB has only three task contrasts (i.e., three samples).

Accuracy & discriminability		UKB	HCP
Tavor method vs un-residualised	accuracy	p<1e-30	p<1e-30
	discriminability	p<1e-30	p<1e-30
Tavor method vs residualised	accuracy	p<1e-30	p<1e-30
	discriminability	p<1e-30	p<1e-30
un-residualised method vs residualised	accuracy	0.4105	3.7932
	discriminability	p<1e-30	p<1e-30

Table B5: Two-sided paired t-test between model performance of different approaches. The task contrasts were pooled together, corresponding to the last columns of Figure 4.3 (UKB three contrasts and HCP 47 contrasts, across all subjects). The p-values were Bonferroni-corrected (12 tests in total).

Comparison with test-retest reliability (task prediction)		UKB	HCP
first-visit prediction vs reliability	un-residualised	p<1e-30	p<1e-30
	residualised	p<1e-30	p<1e-30
second-visit prediction vs reliability	un-residualised	p<1e-30	p<1e-30
	residualised	p<1e-30	p<1e-30

Table B6: Two-sided paired t-test between model performance (with group-average added back in) and test-retest reliability. The task contrasts were pooled together, corresponding to the last columns of Figure 4.4a and 4.4b (UKB three contrasts and HCP 47 contrasts, across all subjects). The p-values were all highly significant, after being Bonferroni-corrected (8 tests in total).

Comparison with test-retest reliability (amplitude prediction)	UKB	HCP
first-visit prediction vs reliability	1.2543	6.5663e-11
second-visit prediction vs reliability	0.0734	2.9706e-6

Table B7: Two-sided paired t-test between amplitude prediction accuracy and the corresponding test-retest reliability. The task contrasts were pooled together, corresponding to the last columns of Figure 4.4c and 4.4d (UKB three contrasts and HCP 47 contrasts). The p-values were Bonferroni-corrected (4 tests in total). UKB results were non-significant, which is un-surprising given that UKB has only three task contrasts (i.e., three samples).

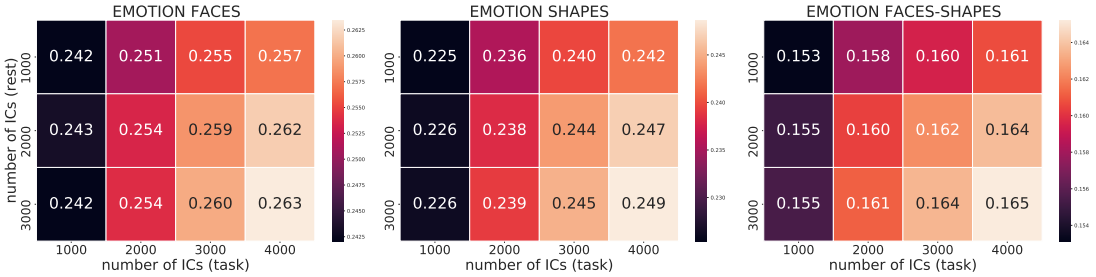

Figure B1: Prediction accuracy of the sparse model at a range of PFM dimensions, trained on a subset of 4,000 UKB subjects and tested on 700. Overall, prediction accuracy increases with the number of functional modes. Note that the results were based on residualised data. The un-residualised data exhibited similar accuracy patterns, though with smaller differences between the choices of dimensions (not shown here).

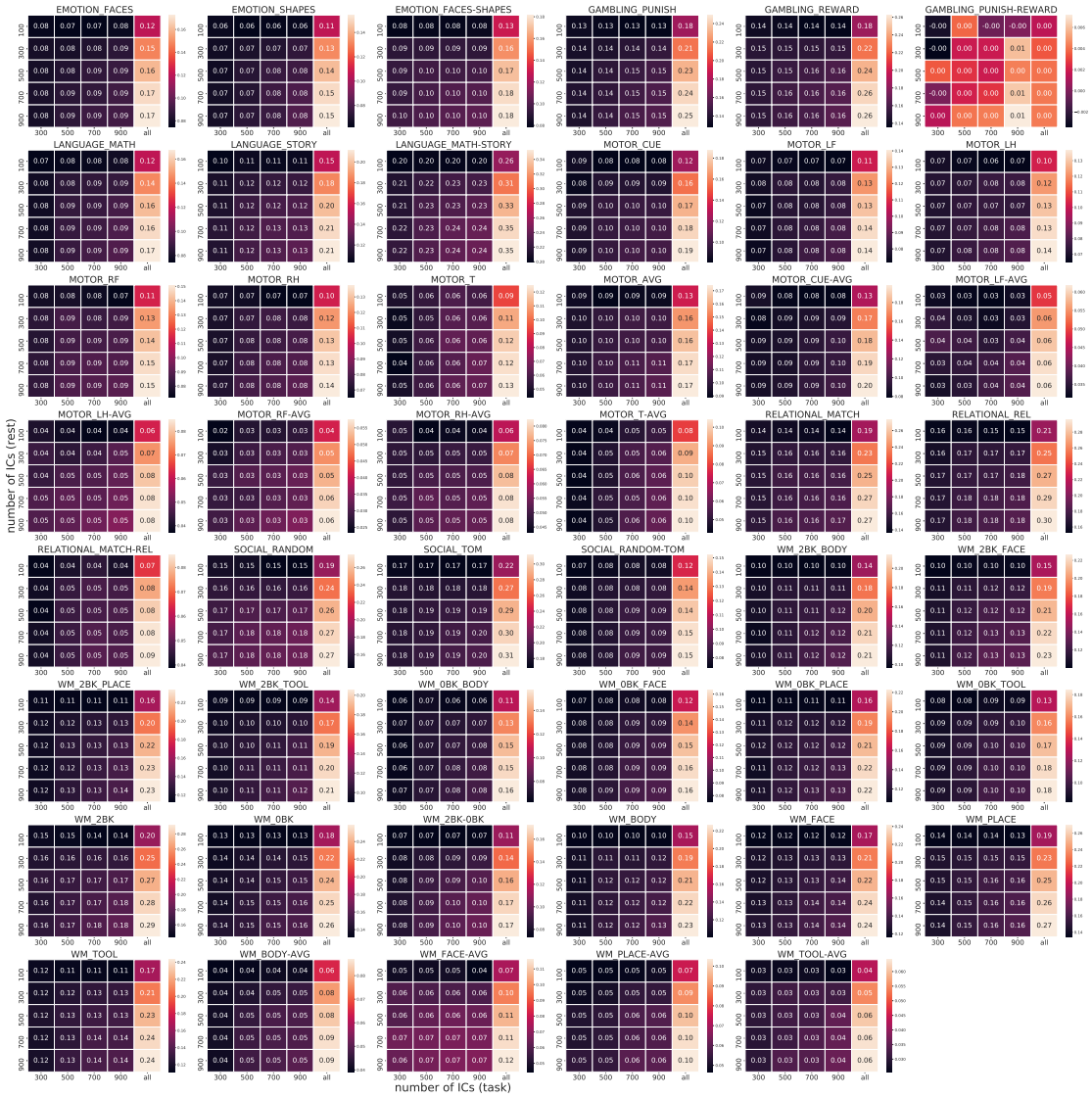

Figure B2: Prediction accuracy of the sparse model at a range of PFM dimensions, trained on 891 HCP subjects and tested on 98. Overall, accuracy increases with the number of functional modes. The un-residualised data exhibited similar accuracy patterns, though with smaller differences between the choices of dimensions (not shown here).

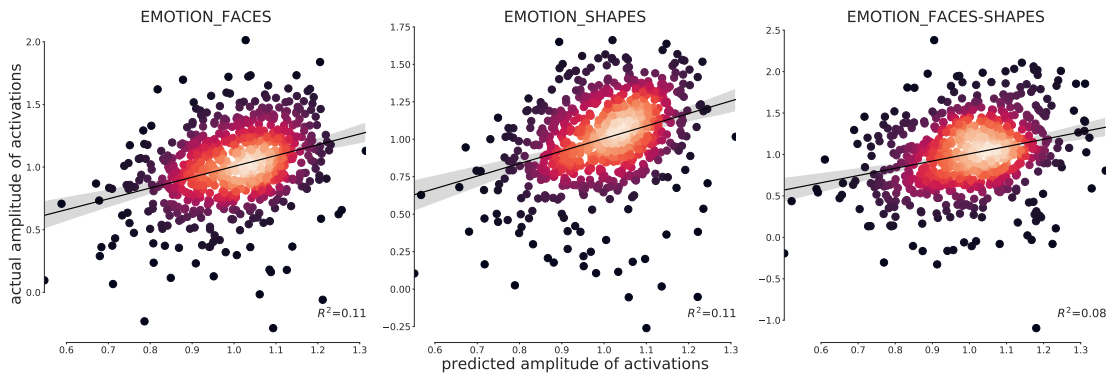

Figure B3: Using resting-state amplitude to predict activation amplitude (UKB). For each task contrast, the activation amplitude was predicted using the amplitude of the 50 PFMs (700 subjects shown).

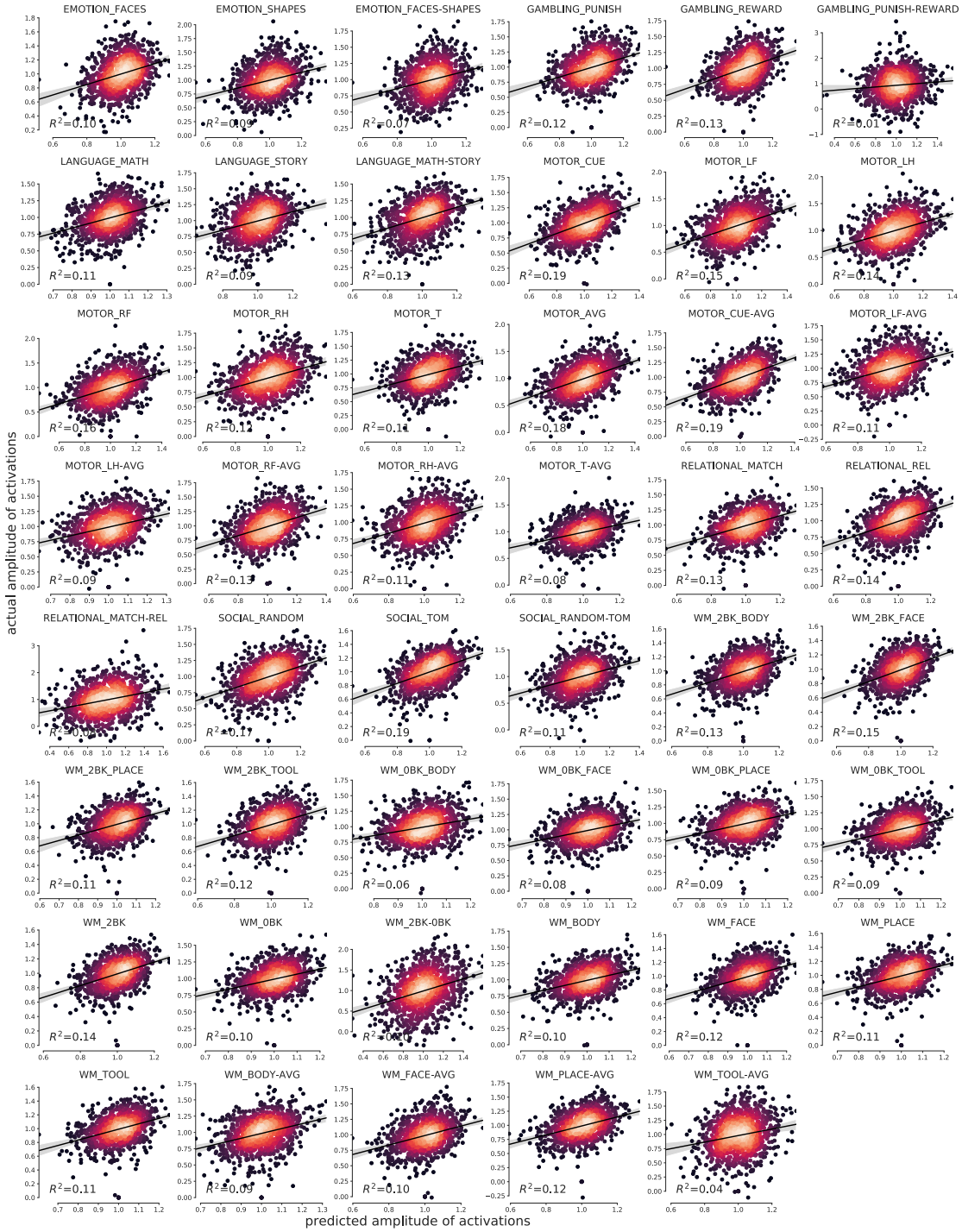

Figure B4: Using resting-state amplitude to predict activation amplitude (HCP). For each task contrast, task amplitude was predicted using the amplitude of 150 PFMs via 10-fold cross-validation (i.e., trained on 9 folds and predicted on the rest).

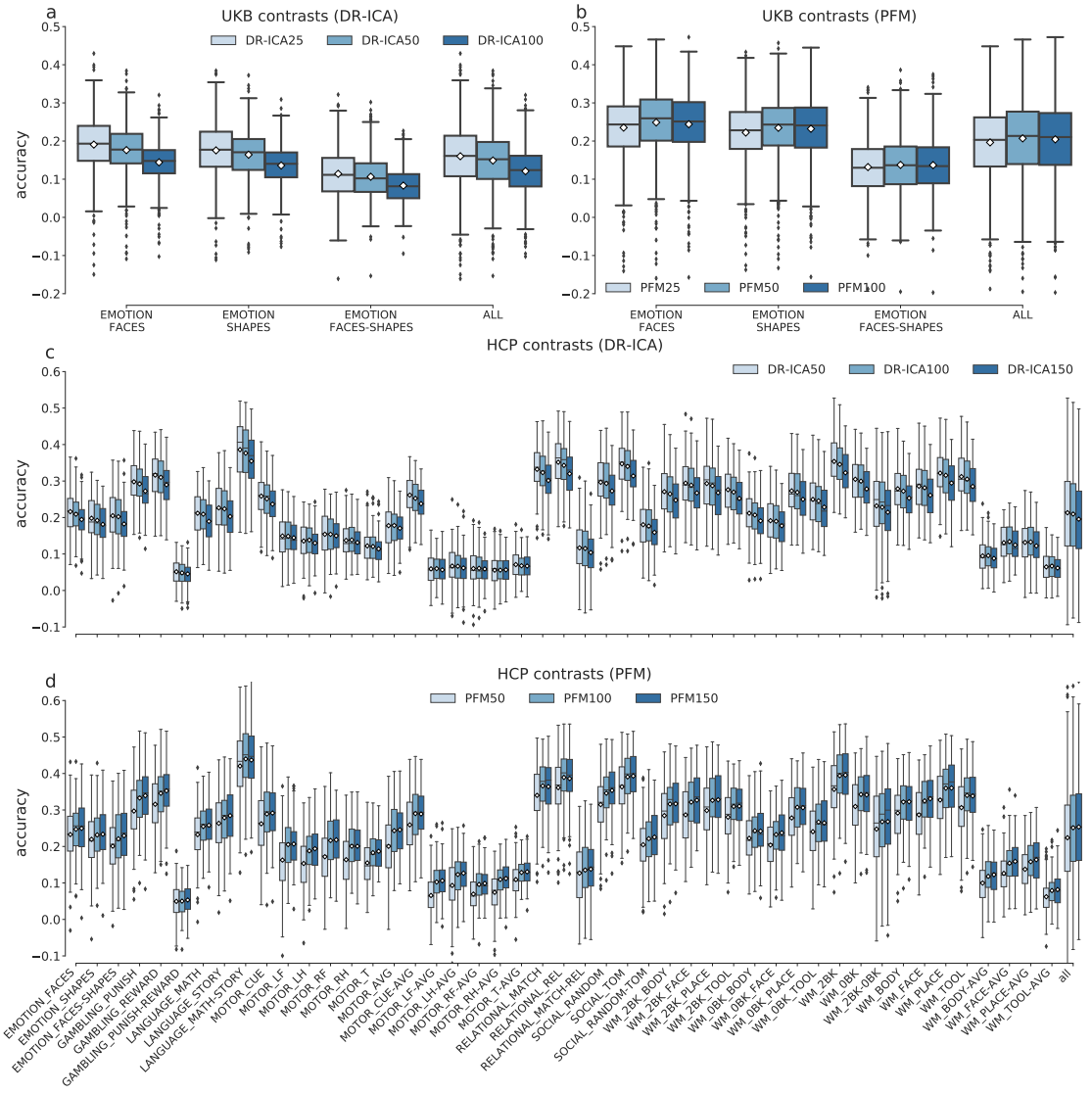

Figure B5: Choices of the functional modes' dimensions. Lower dimensions often result in larger parcels and tend to reflect whole brain resting-state networks. Higher dimensions, in contrast, tend to break down the parcels into more fine-grained functional subprocesses. We argue that the resolution of functional parcellations may non-trivially impact prediction of individual variations in task-evoked activations. The number of resting-state modes must be optimised in the first place for the the subsequent analysis. White diamond shows the mean. (a) Prediction accuracy of DR-ICA25, DR-ICA50, and DR-ICA100 across a subset of 700 UKB subjects. (b) Equivalent plots of PFM25, PFM50, and PFM100. (c) and (d) Equivalent plots of 98 HCP subjects at 50, 100, and 150 modes. The results were based on the residualised data; for the un-residualised data, different dimensions had similar effects on the accuracy, though with smaller differences (not shown here).

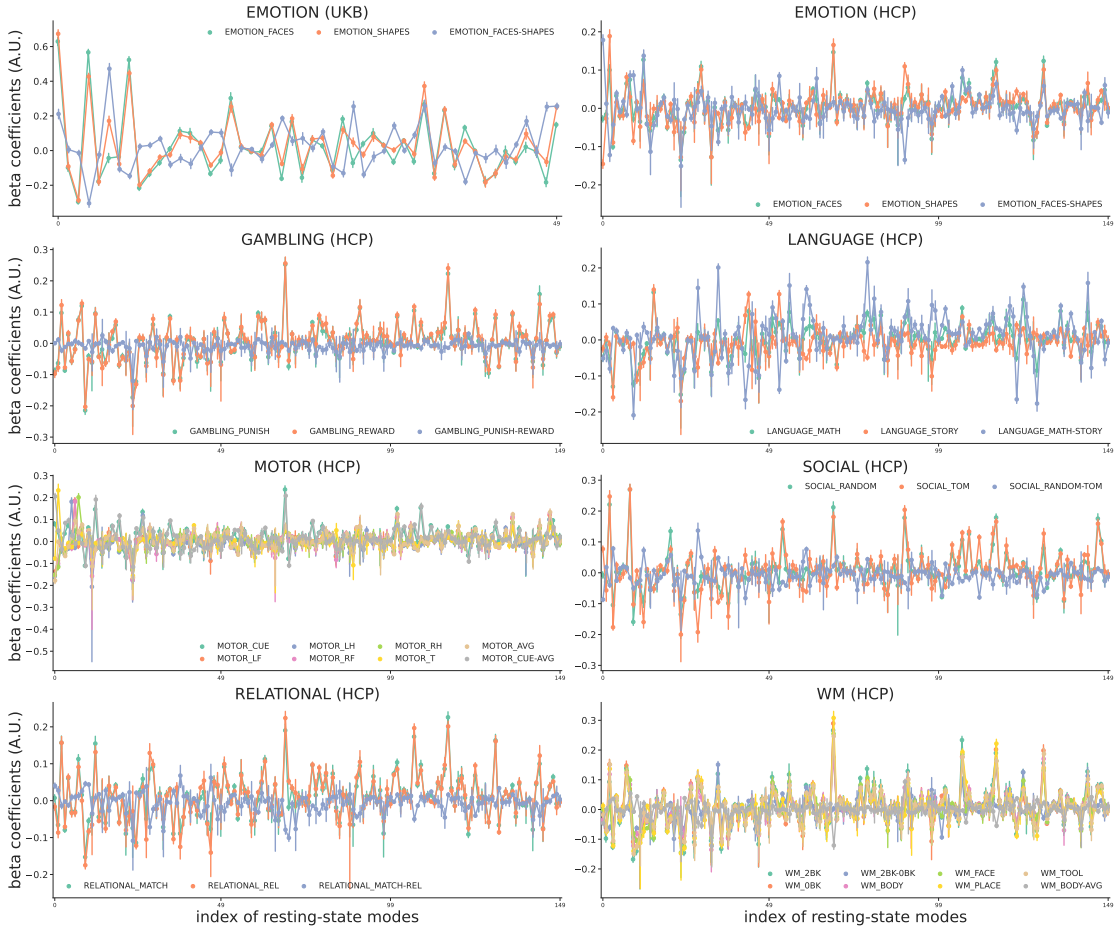

Figure B6: Baseline coefficients (betas of the baseline model) of the residualised data. Error bars showing 95% CI of the mean beta values (calculated across 1,000 UKB subjects and 891 HCP subjects). For each subject, the coefficients were divided by the maximum beta value within the given contrast. Overall, the non-differential functional modes exhibited consistent patterns within each task domain, while the differential contrasts showed different patterns.

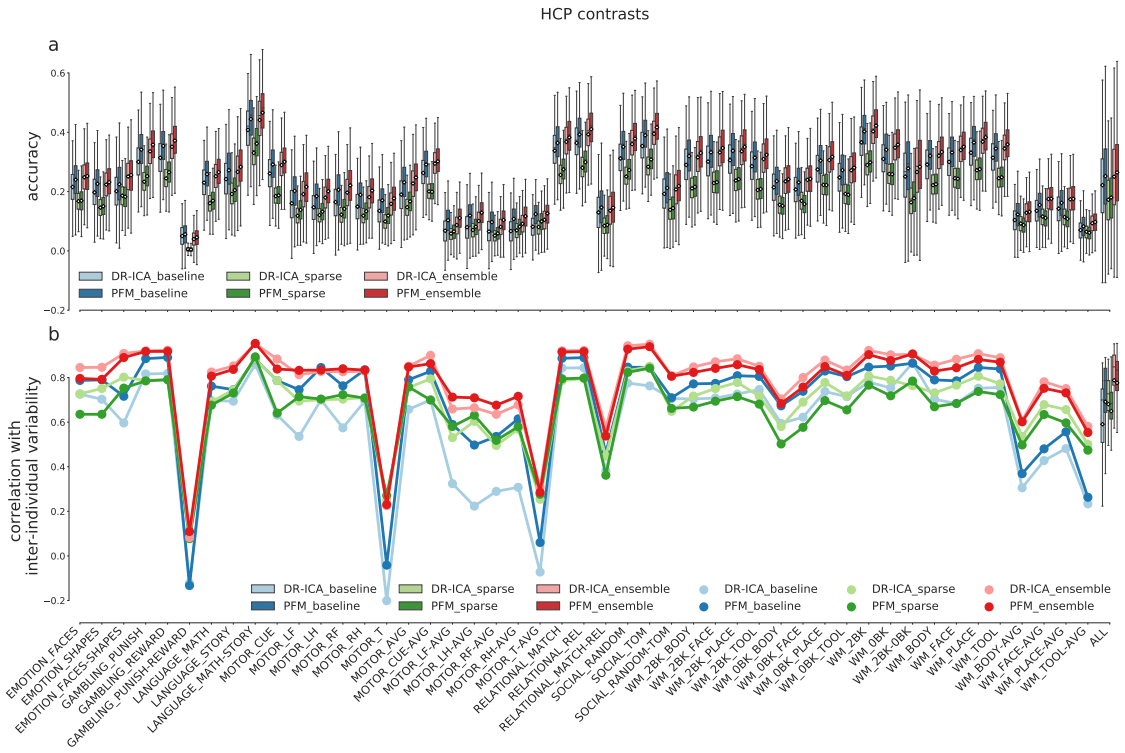

Figure B7: Comparison between the ensemble model and the single models, shown across all 47 HCP task contrasts. (a) Equivalent plots of Figure 4.2b. (b) Equivalent plots of Figure 4.2d.

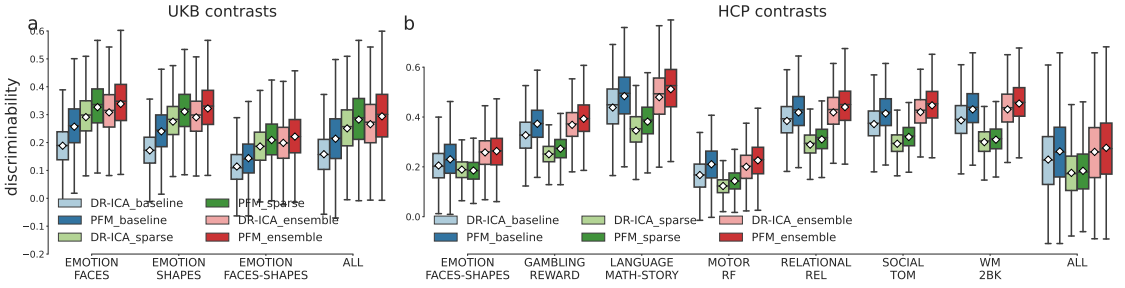

Figure B8: Discrimination metric for the residualised contrast maps. (a) and **(b)**: equivalent plots of Figure 2a and 2b., showing discriminability for the same sets of subjects. The discriminability values are close to the accuracy (correlation) values, because the "null models" mostly have zero accuracy.

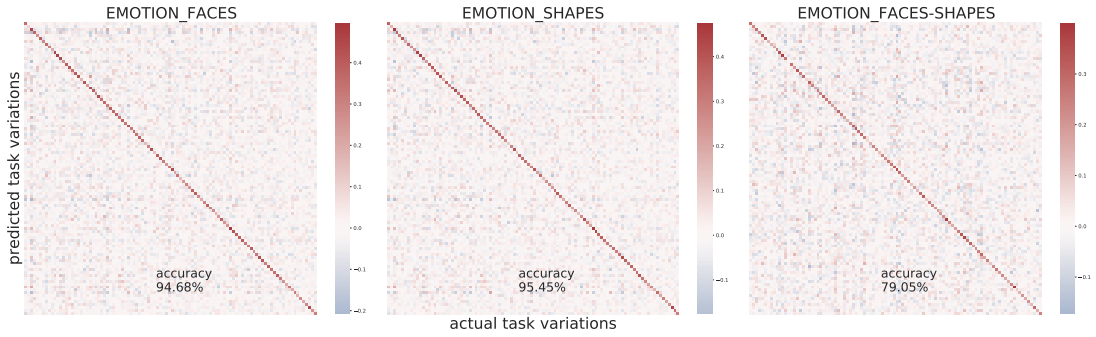

Figure B9: UKB subjects identification accuracy (based on residualised data). Accuracy (pearson's correlation between the predicted and actual variations) was calculated across all subjects; for illustration purpose, only 100 subjects were shown above. The off-diagonal elements fluctuate around zero, i.e., accuracy and discriminability calculated on residualised predictions are almost identical.

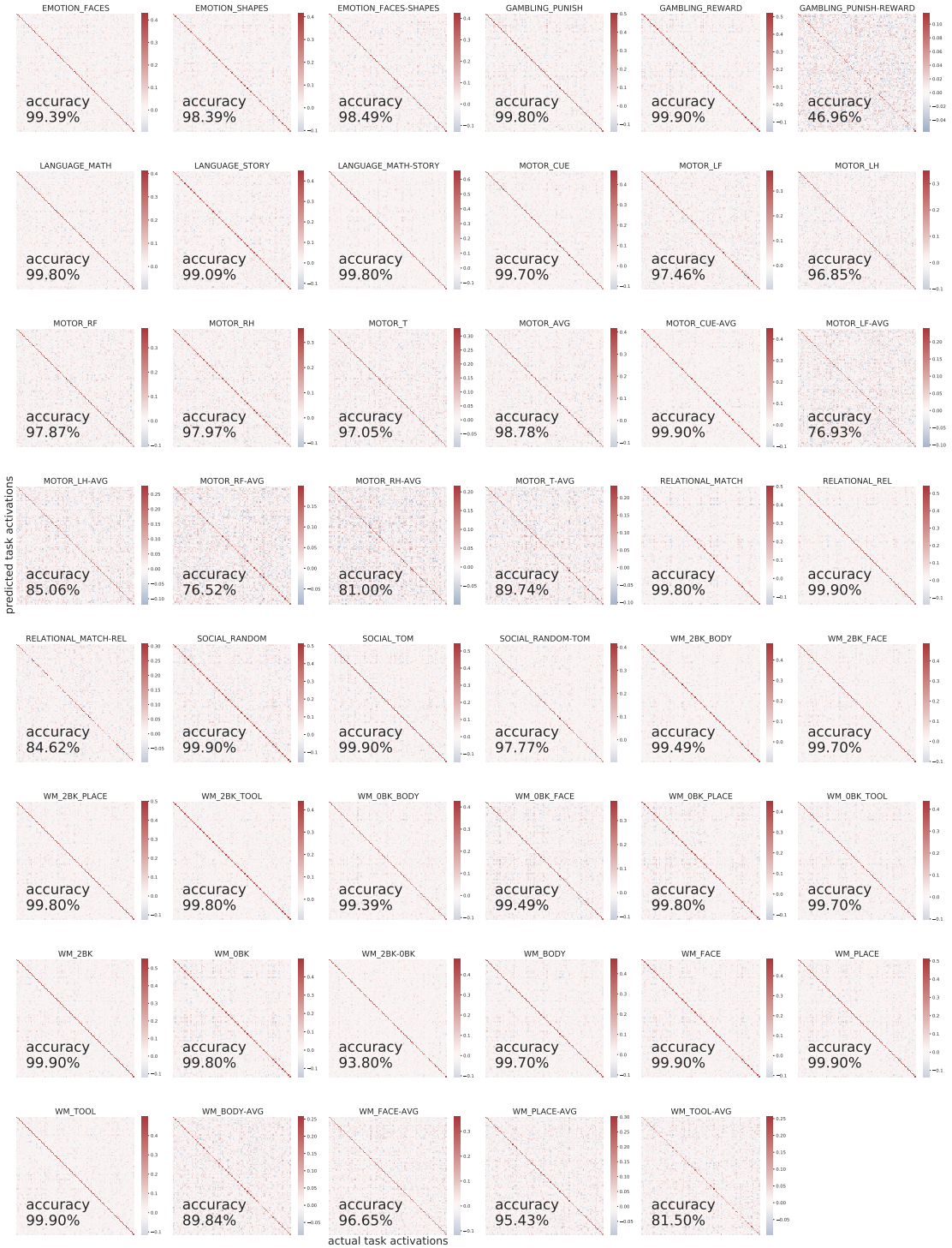

Figure B10: HCP subjects identification accuracy (based on residualised data). For illustration purpose, only 100 subjects were shown above.

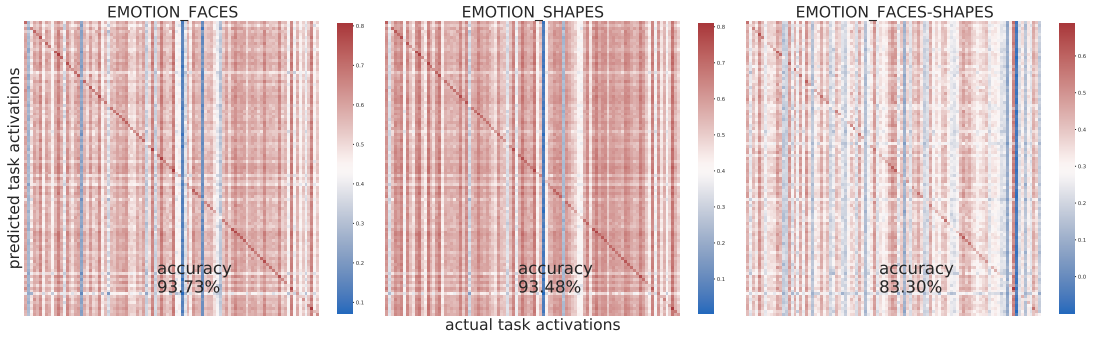

Figure B11: UKB subjects identification accuracy (with group-average activations added back in). The off-diagonal values no longer fluctuate around zero. The subject identification accuracy remains high. For illustration purpose, only 100 subjects were shown above.

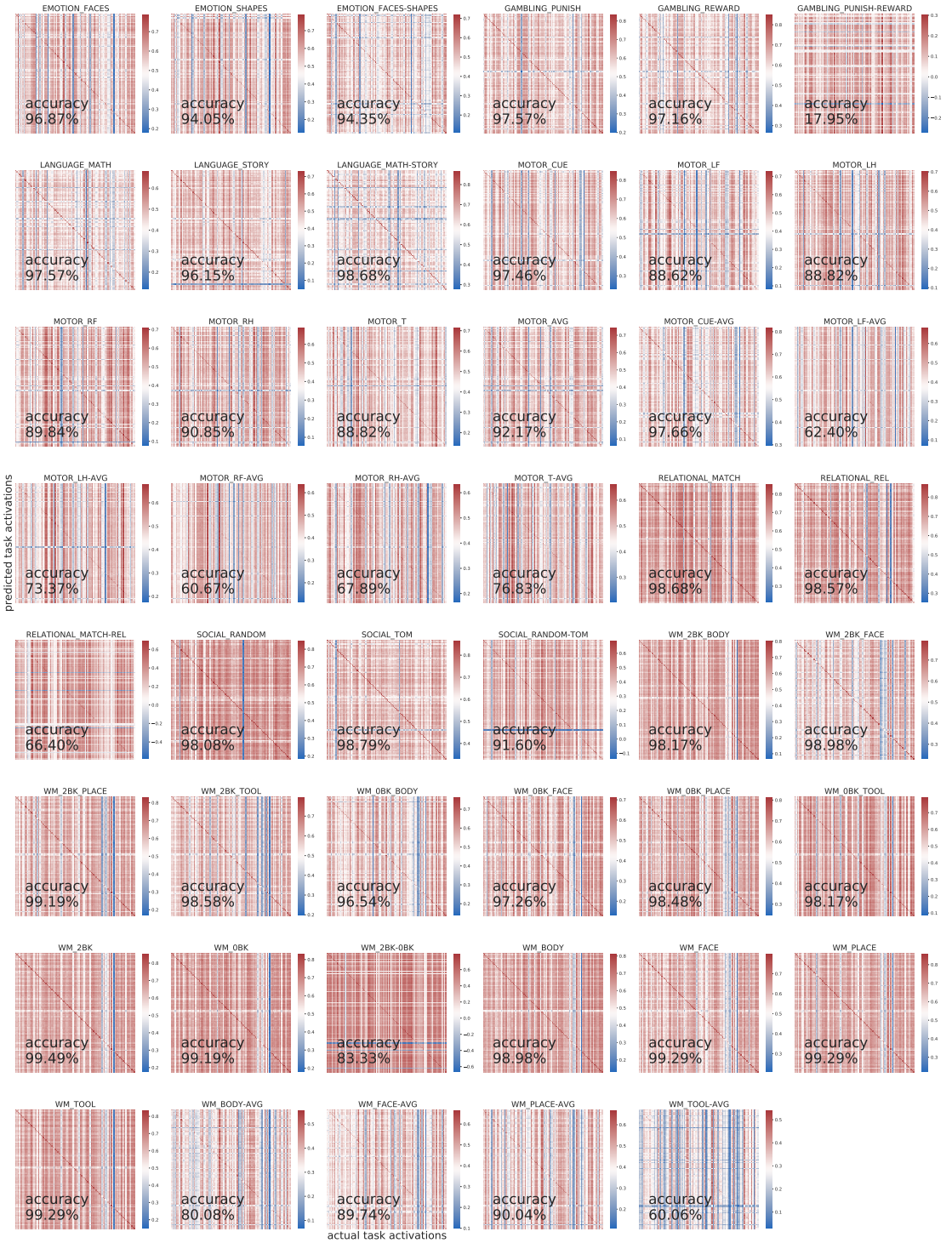

Figure B12: HCP subjects identification accuracy (with group-average activations added back in). For illustration purpose, only 100 subjects were shown above.

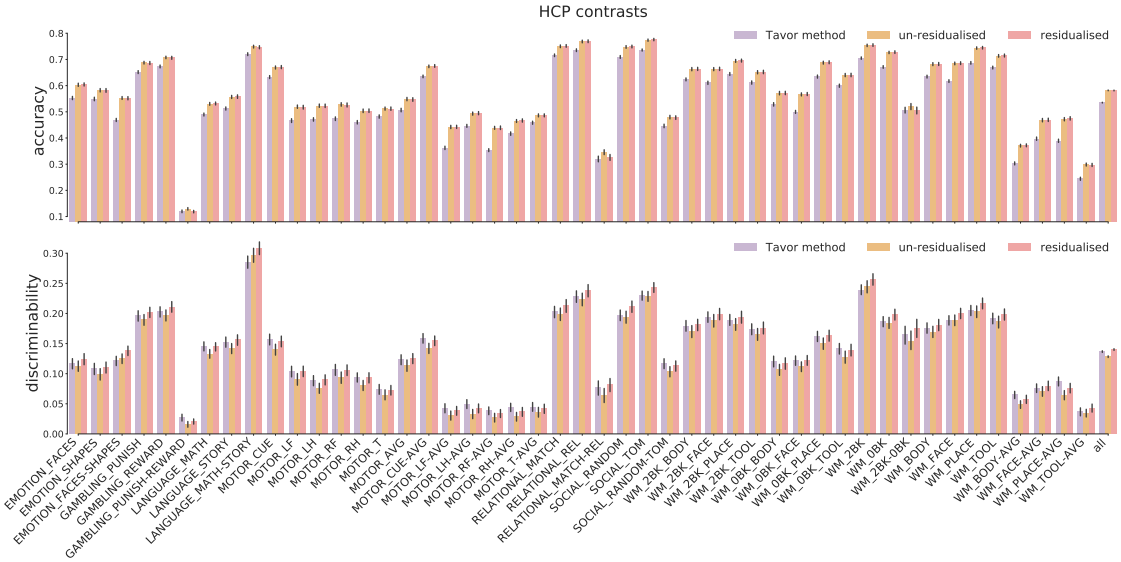

Figure B13: Comparison of the Tavor model and the ensemble model (both un-residualised and residualised) across 991 HCP subjects for all 47 task contrasts. The ensemble model (either residualised or not) outperformed the Tavor model in terms of the actual prediction accuracy; however, the Tavor model could make more individualised predictions than the ensemble model if both trained on un-residualised data. The residualised ensemble model outperformed the other two both in accuracy and discriminability, except for the motor task domain.

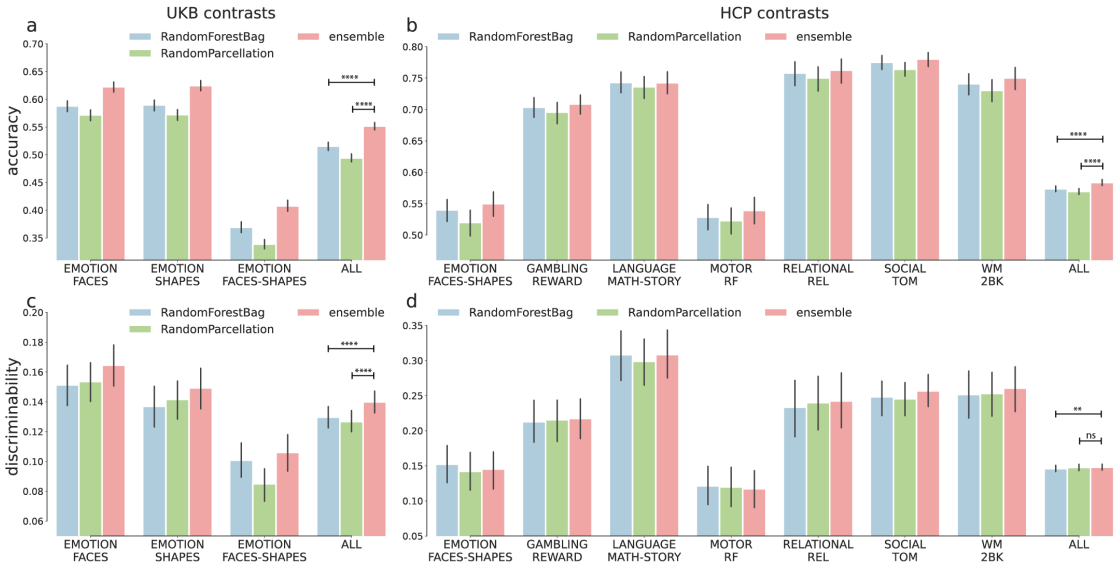

Figure B14: Prediction accuracy and discriminability of previous methods. The hyper-parameters of Random Forest Bagging were optimised on a different subset. P-values of the two-sided paired t-tests are Bonferroni-corrected (8 tests in total). ns: $P > 0.05$; **: $P < 0.01$; ****: $P < 0.0001$. Blue: the Random Forest Bagging approach used in Cohen et al., 2020; Green: the Random Parcellation approach used in Dohmatob et al. 2021; Red: Our ensemble approach. (a) Prediction accuracy across 700 UKB subjects, trained on 4,000 subjects. (b) Prediction accuracy across 98 HCP subjects, trained on 991 subjects. (c) Equivalent plot of (a), showing prediction discriminability. (d) Equivalent plot of (b), showing prediction discriminability.

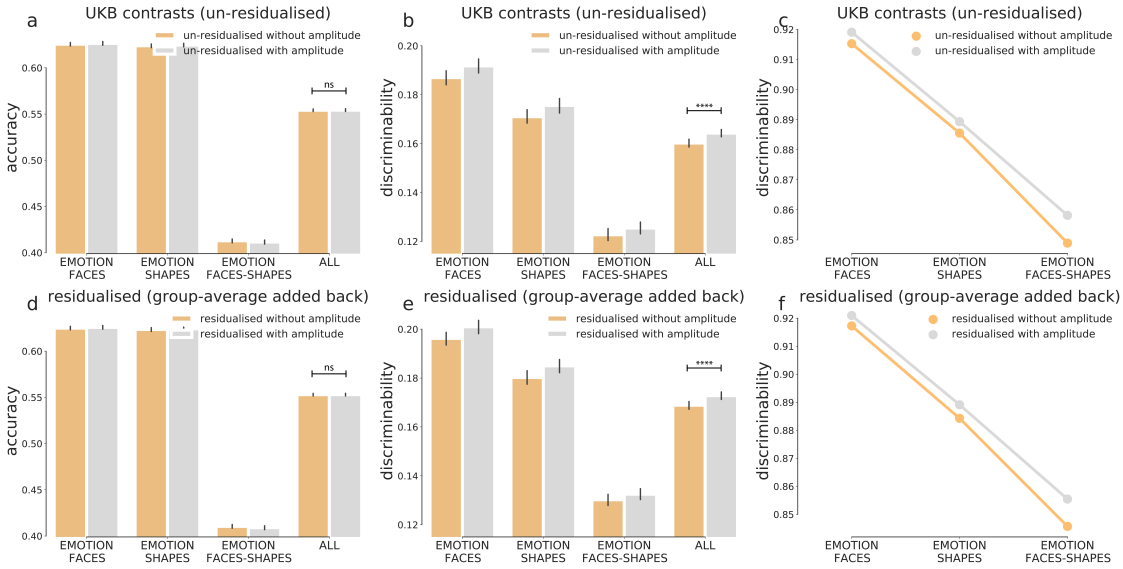

Figure B15: Prediction accuracy and discrimininability of the ensemble model with or without PFM amplitude as additional features, calculated across 17,560 UKB subjects Although incorporating amplitude did not further increase the overall accuracy for UKB, it did marginally improve prediction discriminability. This coincides with (c) and (d), which shows that the std. maps of predicted activations (across subjects) exhibited higher correspondence with the actual inter-individual variability. P-values of the two-sided paired t-tests are Bonferroni-corrected (4 tests in total). ns: P $>$ 0.05; **: P $<$ 0.01; ****: P $<$ 0.0001.

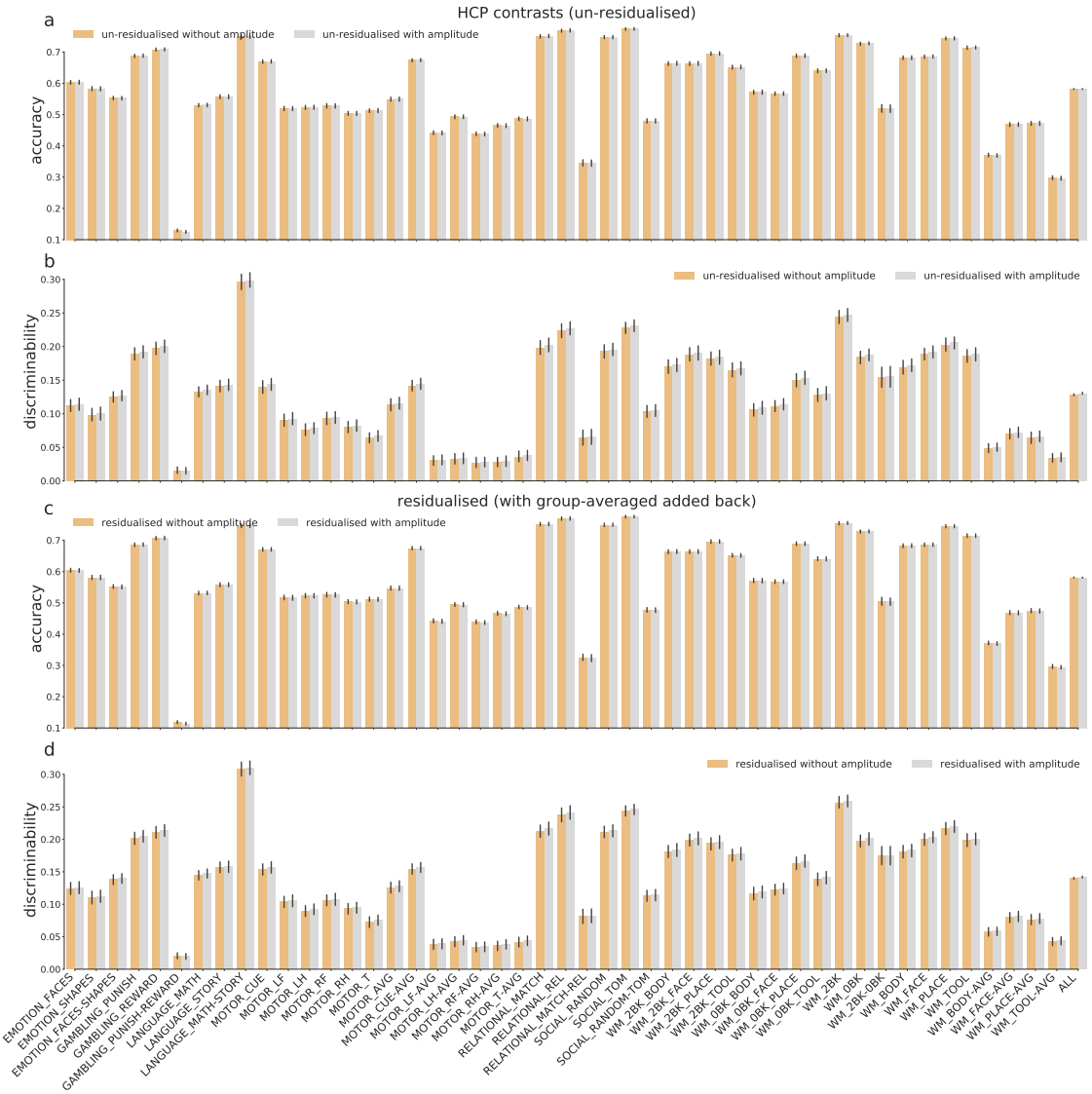

Figure B16: Prediction accuracy and discrimininability of the ensemble model with and without resting-state amplitude as additional features for all 47 HCP contrasts. For HCP, however, including PFM amplitude as additional features at the ensemble stage only marginally improved prediction discriminability.

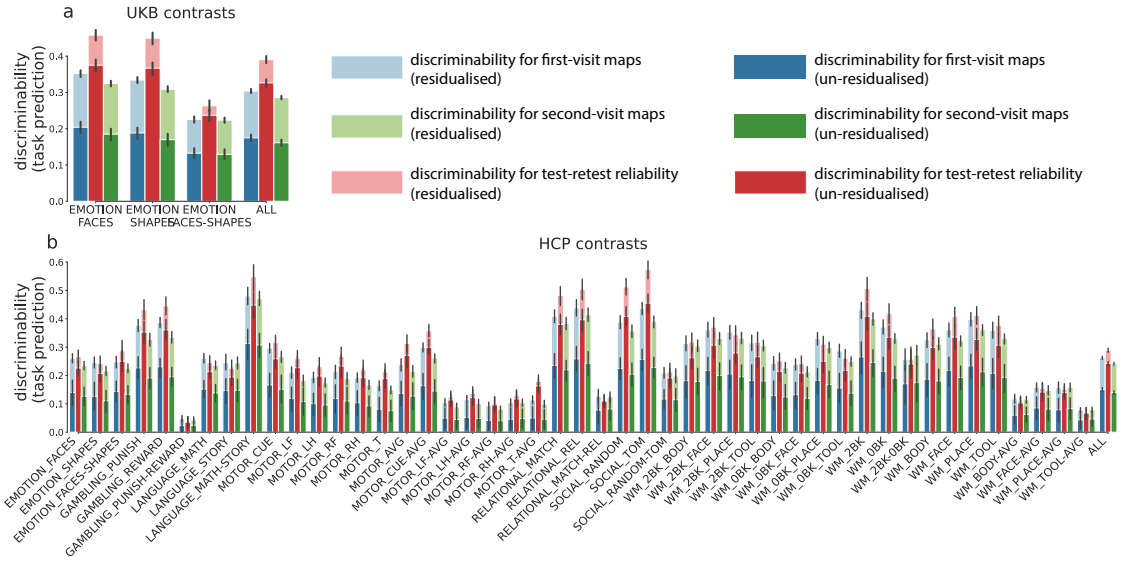

Figure B17: Discrimination metric for the predictions and the test-retest reliability. Blue: prediction discriminability for the first-visit task contrast maps. Green: prediction discriminability for the second-visit task contrast maps. Red: test-retest discriminability of task contrast maps. Note that opposite to Figure 4.4b and 4.4c, here light colours denote the prediction discriminability of the residualised maps, dark colours the discriminability of the un-residualised maps.

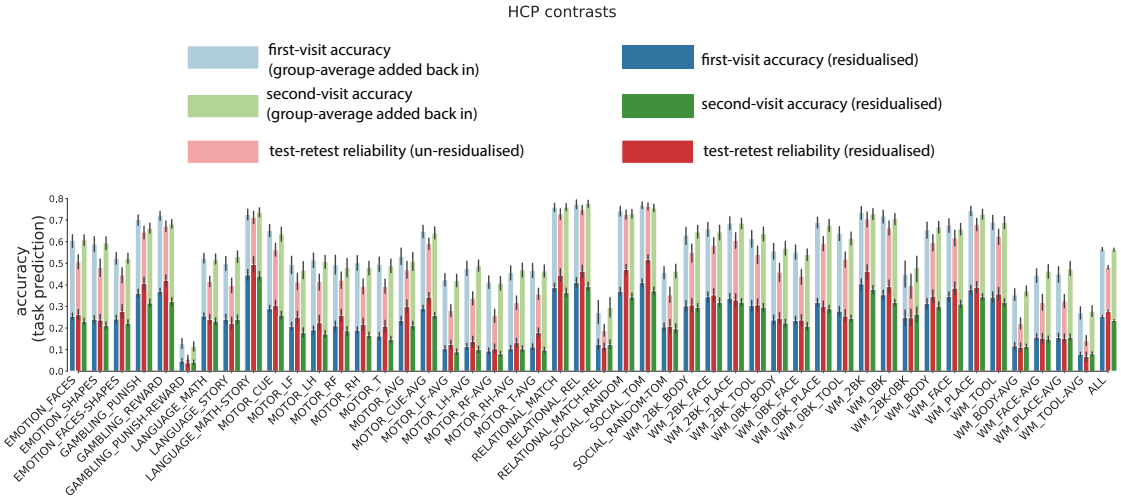

Figure B18: Prediction accuracy of first-visit and second-visit contrast maps and test-retest reliability for all HCP contrasts. Equivalent plots of Figure 4.4c, for all HCP task contrasts.

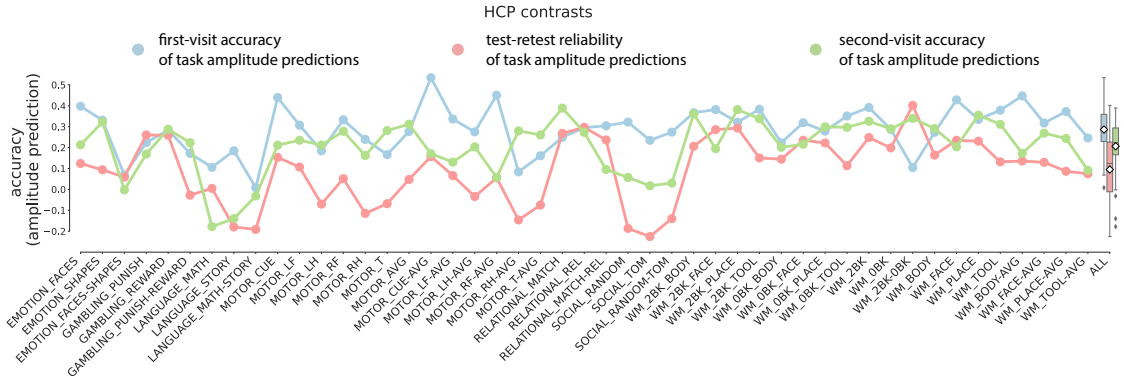

Figure B19: Prediction accuracy of first-visit and second-visit amplitude and test-retest reliability for all HCP contrasts. Equivalent plots of Figure 4.4e, for all HCP task contrasts.

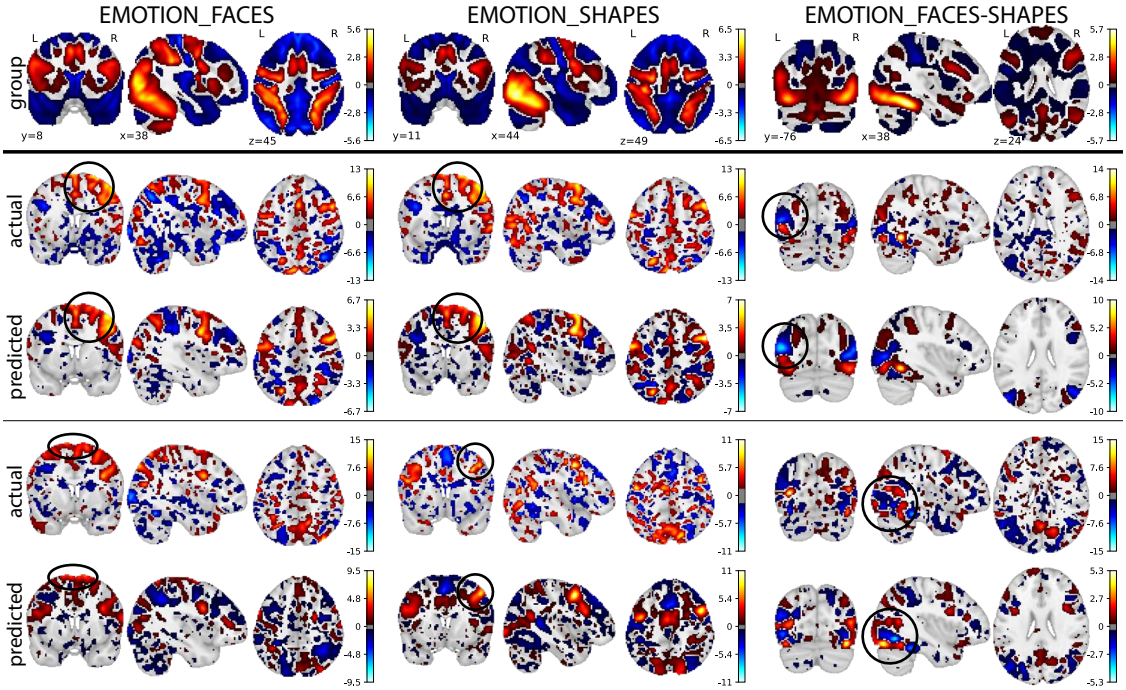

Figure B20: The actual and the predicted task variations (residuals) of the example UKB subjects, shown on the brain.

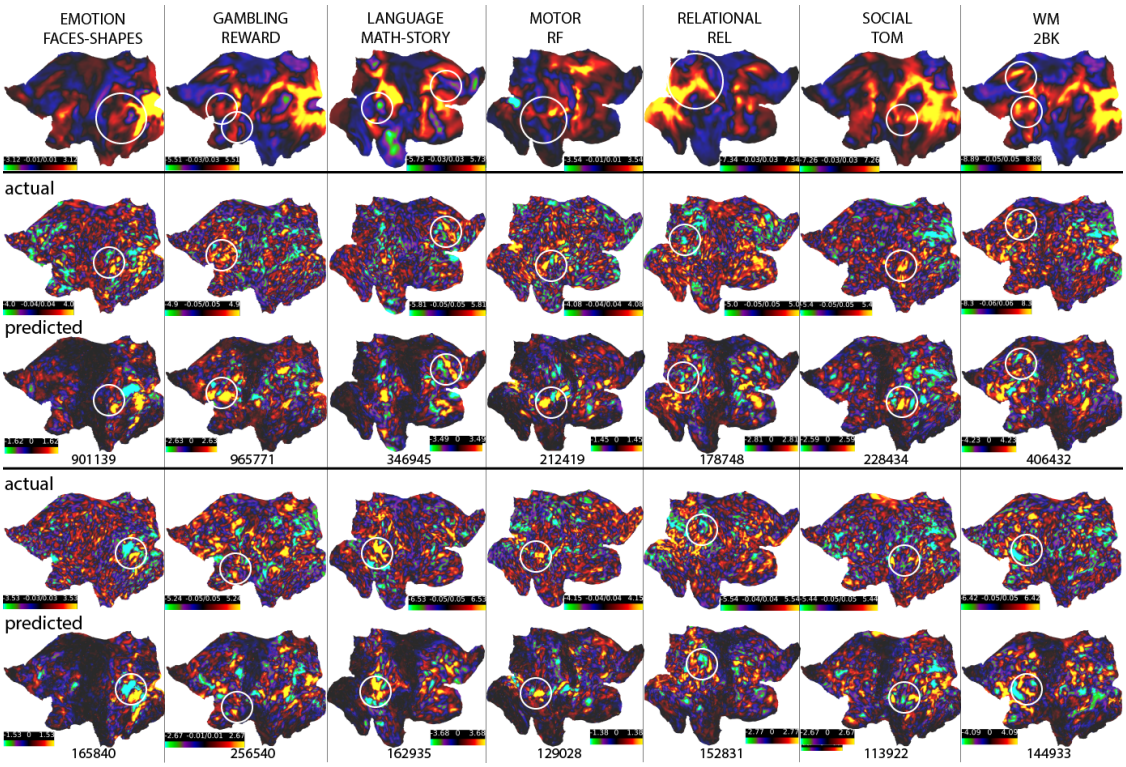

Figure B21: The actual and the predicted task variations (residuals) of the HCP subjects, shown on the surface.

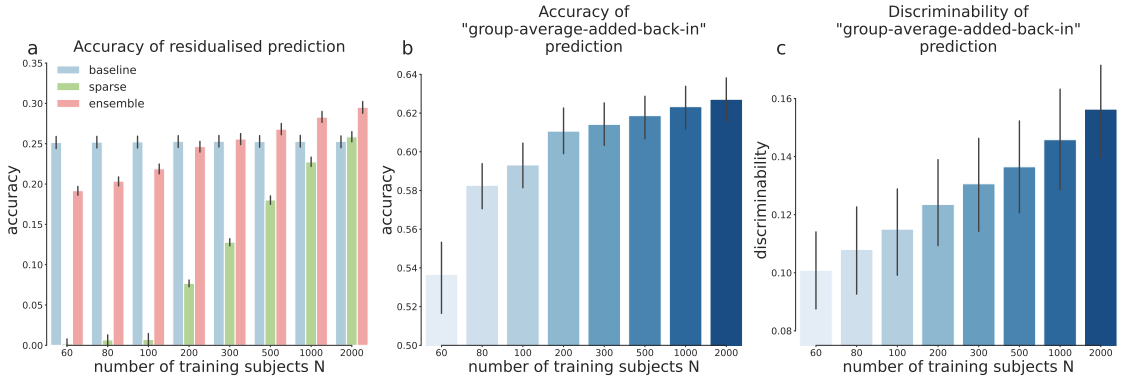

Figure B22: Prediction accuracy as a function of number of training subjects (UKB). (a) Prediction accuracy of the baseline, sparse and ensemble model for the residualised maps. The baseline and ensemble model converges very quickly; the sparse model requires a larger number of training subjects. (b) Prediction accuracy of the ensemble model, with group-average added back in. (c) Equivalent plot of (b), showing prediction discriminability.

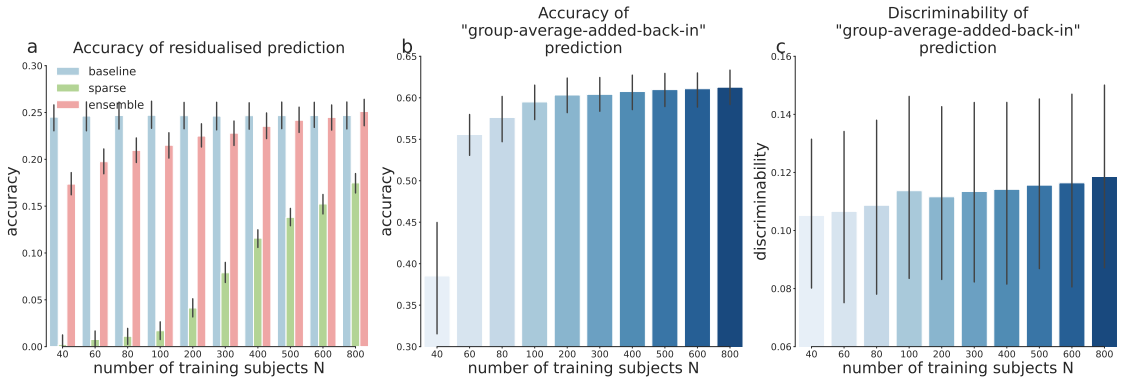

Figure B23: Prediction accuracy as a function of number of training subjects (HCP). (a) Prediction accuracy of the baseline, sparse and ensemble model on the residualised maps for 98 HCP subjects. The baseline and ensemble model gave good accuracy at a small N; the sparse model requires a larger number of training subjects. (b) Prediction accuracy of the ensemble model, with group-average added back in. (c) Equivalent plot of (b), showing prediction discriminability.

C

Additional Results for Chapter 5

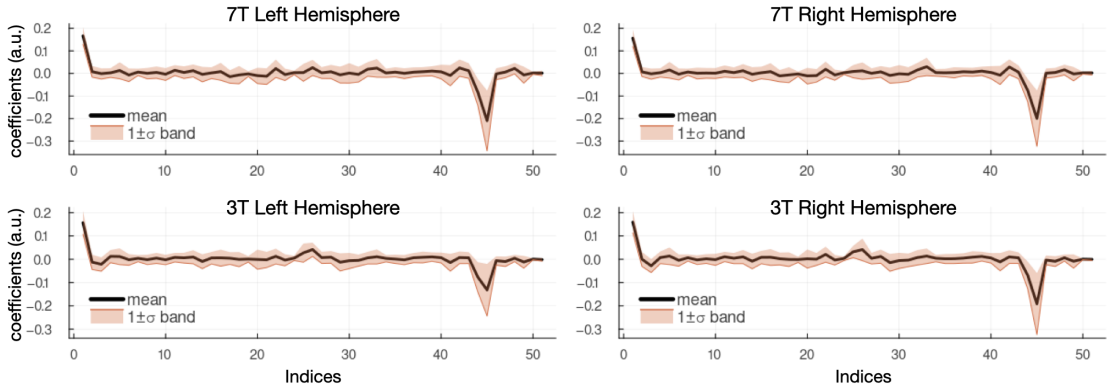

Figure C1: Coefficients for the wedge task (RETCCW and RETCW). The first coefficient corresponds to the contribution of the original resting-state connectivity matrix (i.e., dense connectome). The other coefficients correspond to the contributions of the other resting-state smilarities, each derived from the resting-state fingerprints weighted by group-average ICA50.

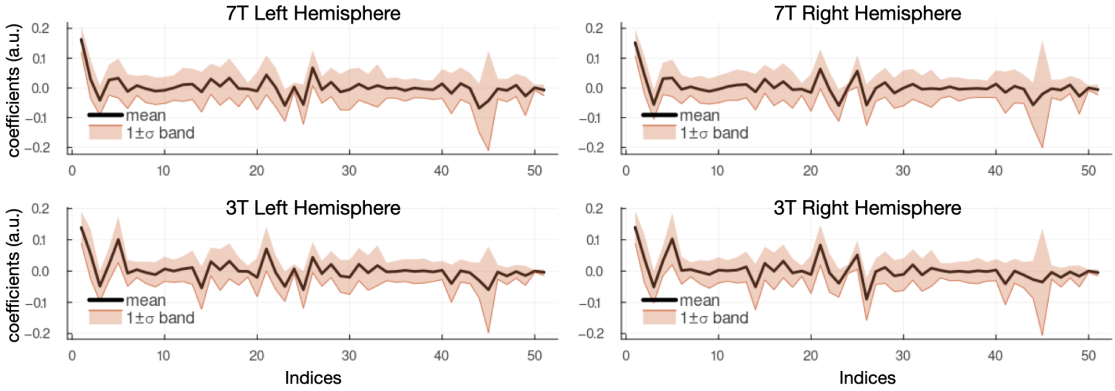

Figure C2: Coefficients for the ring task (RETEXP and RETCON). The first coefficient corresponds to the contribution of the original resting-state connectivity matrix (i.e., dense connectome). The other coefficients correspond to the contributions of the other resting-state smilarities, each derived from the resting-state fingerprints weighted by group-average ICA50.

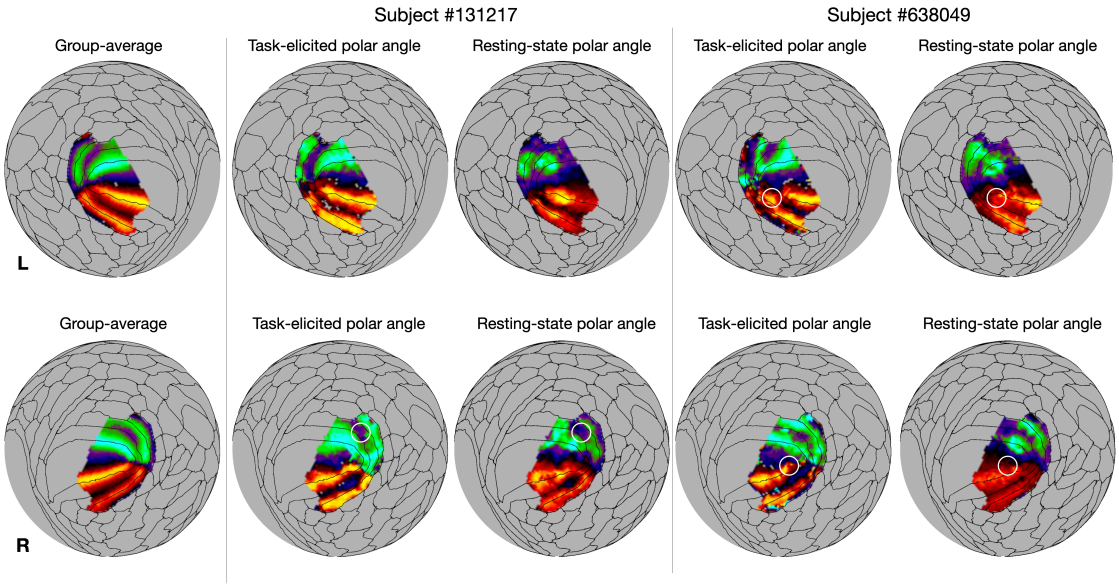

Figure C3: Task-elicited, group-average, and 3T resting-state polar angle. Equivalent plots of Figure 5.6, based on 3T resting-state.

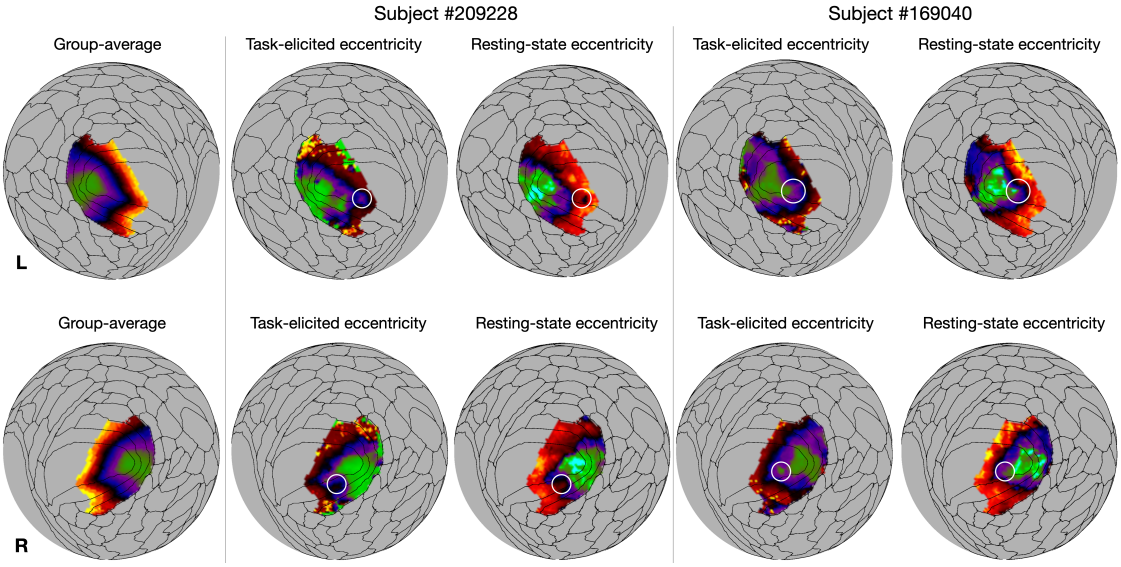

Figure C4: Task-elicited, group-average, and 3T resting-state eccentricity. Equivalent plots of Figure 5.7, based on 3T resting-state.

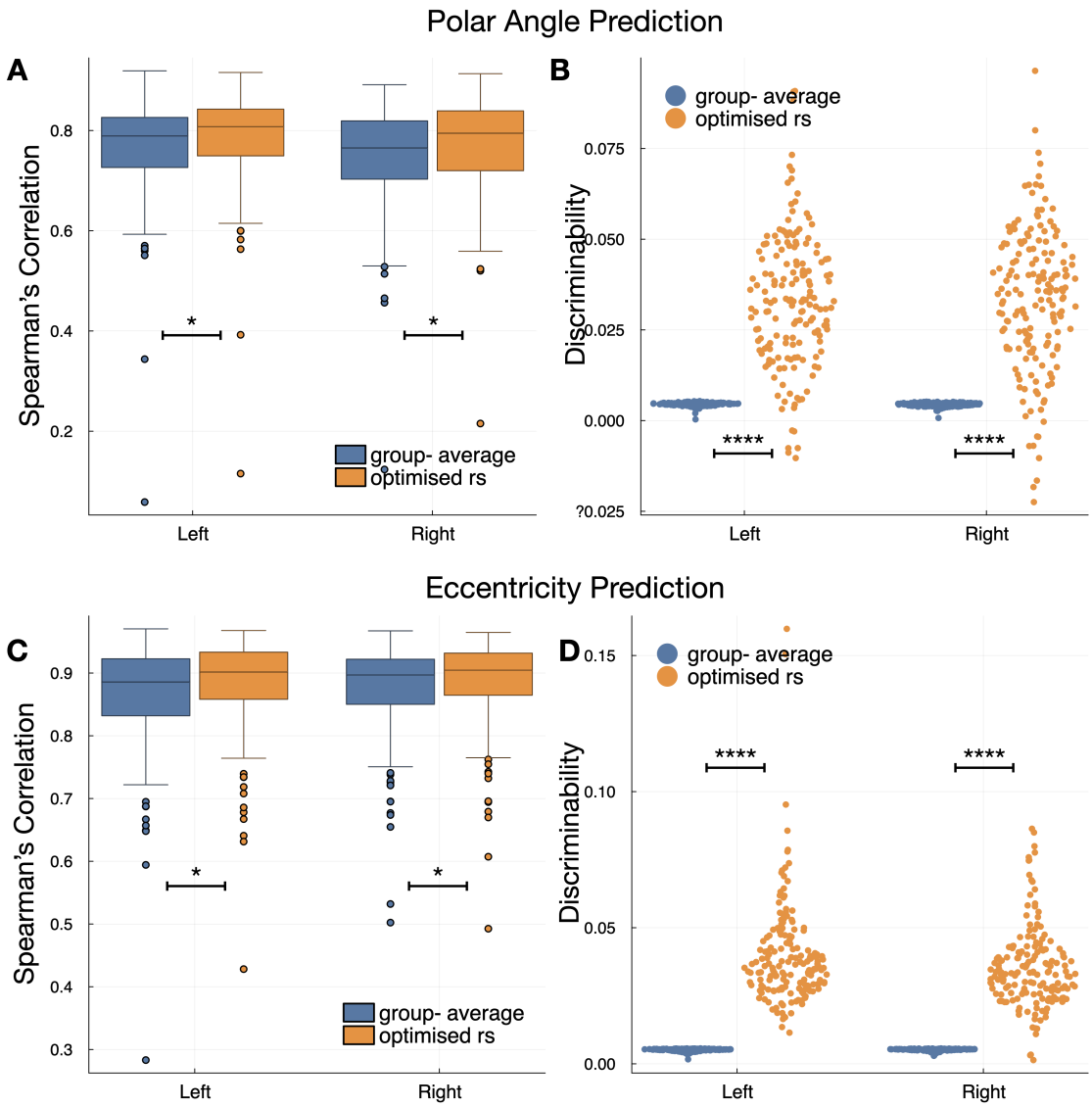

Figure C5: Prediction accuracy of 3T resting-state retinotopy. Equivalent plots of Figure 5.5, using 3T resting-state fMRI.

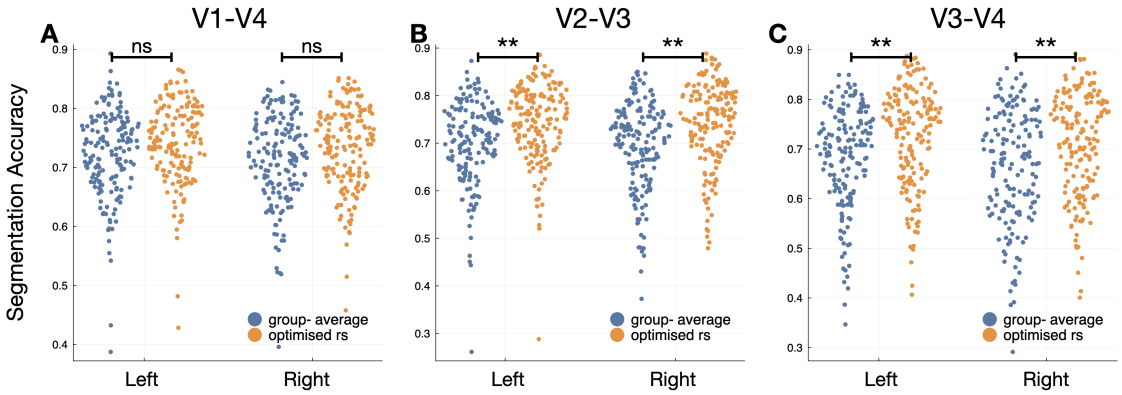

Figure C6: Segmentation accuracy of the 3T resting-state polar angle. Equivalent plots of Figure 5.8, based on 3T resting-state polar angle.

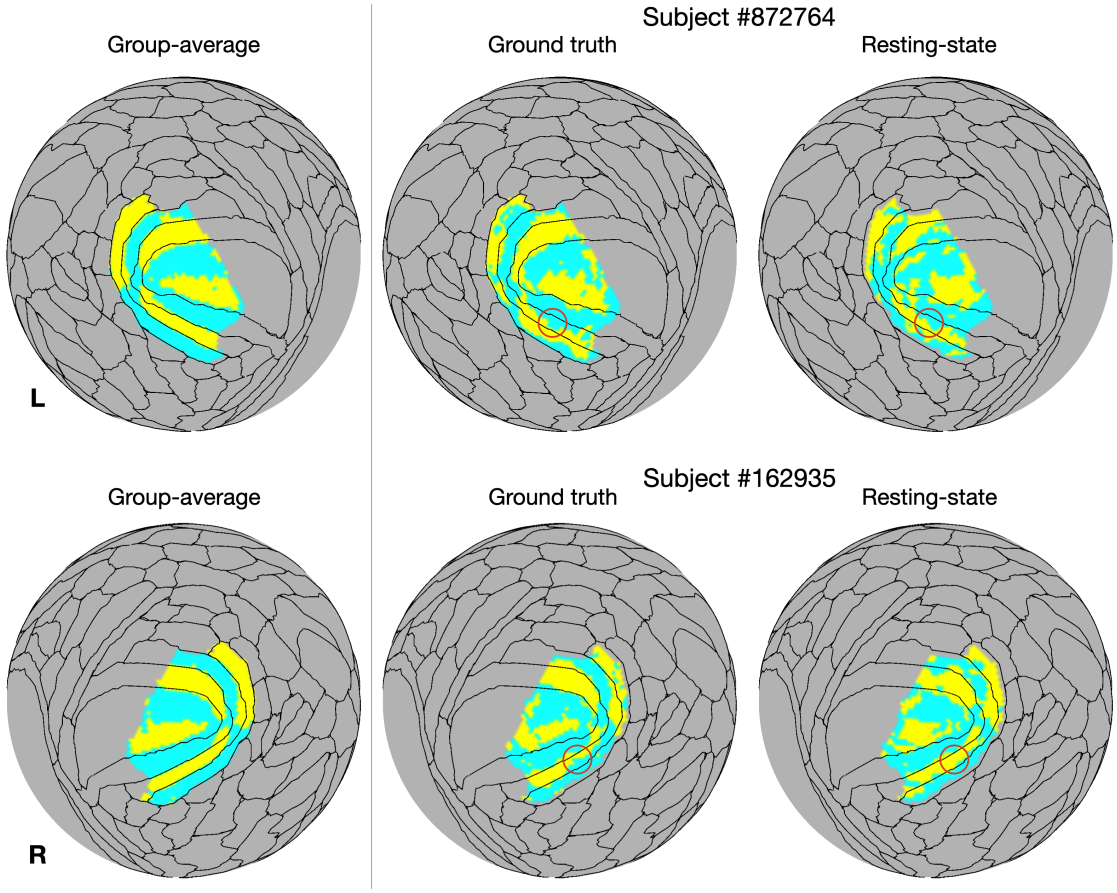

Figure C7: Segmentation between V1-V4 using the group-average polar angle (left), the individual task-elicited polar angle (middle), and the 3T resting-state polar angle (right). Equivalent plots of Figure 5.9, while the resting-state segmentation were based on 3T resting-state polar angle maps.

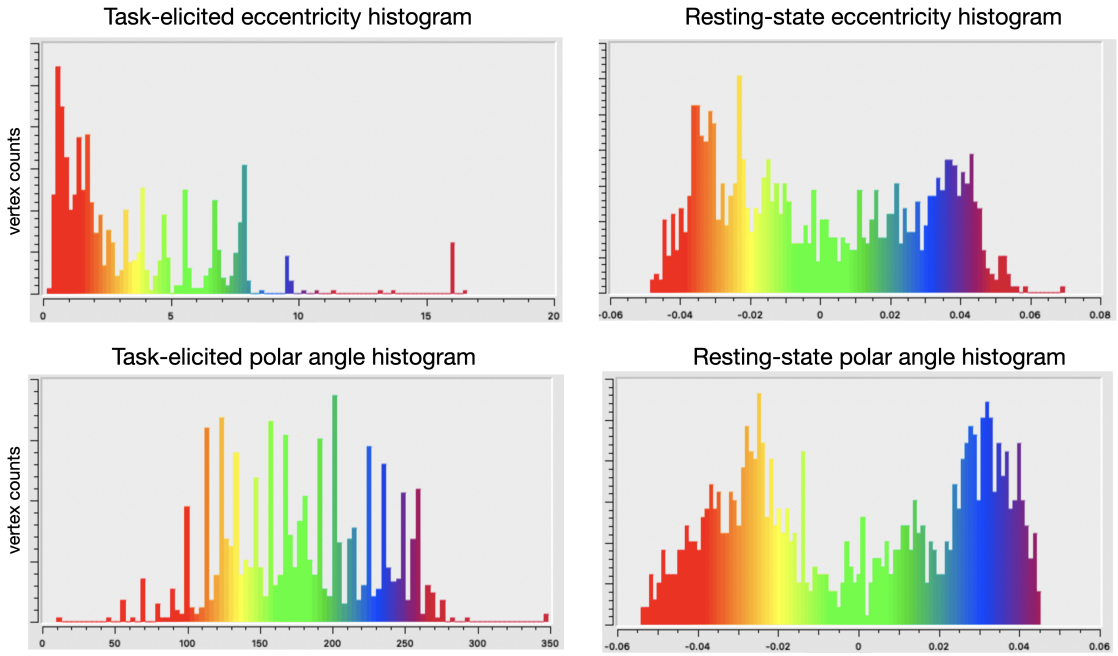

Figure C8: Histograms of task-elicited retinotopy (ground truth) and resting-state retinotopy (predictions) of an example subject, right hemisphere. The histograms of resting-state predicted retinotopy do not match well with the histograms of the actual retinotopy, complicating direct comparisons of the two. For example, the actual polar angle exhibits a roughly uniform distribution between $\pi/2$ and $3\pi/2$, while the predicted polar angle exhibits a bimodal pattern. This phenomenon might be explained by the inherent properties of the Laplacian Eigenmap technique. Specifically, this technique has a focus on preserving the local structure rather than the global structure. In other words, this technique prioritises keeping data points that were close in the high-dimensional space remain close in the lower-dimensional space, and is less concerned with the distances between points that were relatively far apart. This focus on local structure may lead to the suboptimal bimodal distribution that fails to preserve the distance relationships between data points that were more distantly separated.

References

[1] Ryota Kanai and Geraint Rees. "The structural basis of inter-individual differences in human behaviour and cognition". In: *Nature Reviews Neuroscience* 12.4 (2011), pp. 231–242.

[2] Martijn P van den Heuvel et al. "Efficiency of functional brain networks and intellectual performance". In: *Journal of Neuroscience* 29.23 (2009), pp. 7619–7624.

[3] Sophia Mueller et al. "Individual variability in functional connectivity architecture of the human brain". In: *Neuron* 77.3 (2013), pp. 586–595.

[4] Emre Bora et al. "Gray matter abnormalities in major depressive disorder: A meta-analysis of voxel-based morphometry studies". In: *Journal of Affective Disorders* 138.1-2 (2012), pp. 9–18.

[5] Madeleine Goodkind et al. "Identification of a common neurobiological substrate for mental illness". In: *JAMA Psychiatry* 72.4 (2015), pp. 305–315.

[6] John DE Gabrieli, Satrajit S Ghosh, and Susan Whitfield-Gabrieli. "Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience". In: *Neuron* 85.1 (2015), pp. 11–26.

[7] Michael D Fox et al. "Efficacy of transcranial magnetic stimulation targets for depression is related to intrinsic functional connectivity with the subgenual cingulate". In: *Biological psychiatry* 72.7 (2012), pp. 595–603.

[8] Kevin Murphy et al. "The impact of global signal regression on resting state correlations: are anti-correlated networks introduced?" In: *Neuroimage* 44.3 (2009), pp. 893–905.

[9] Karl J Friston, Andrew P Holmes, and Keith J Worsley. *How many subjects constitute a study?* 1999.

[10] Katherine S Button et al. "Power failure: why small sample size undermines the reliability of neuroscience". In: *Nature reviews neuroscience* 14.5 (2013), pp. 365–376.

[11] Michael B Miller and John Darrell Van Horn. "Individual variability in brain activations associated with episodic retrieval: a role for large-scale databases". In: *International journal of psychophysiology* 63.2 (2007), pp. 205–213.

[12] Anderson M Winkler et al. "Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies". In: *Neuroimage* 53.3 (2010), pp. 1135–1146.

[13] Kirk I Erickson, Regina L Leckie, and Andrea M Weinstein. "Physical activity, fitness, and gray matter volume". In: *Neurobiology of aging* 35 (2014), S20–S28.

[14] Nandita Vijayakumar et al. “Structural brain development: A review of methodological approaches and best practices”. In: *Developmental cognitive neuroscience* 33 (2018), pp. 129–148.

[15] Anderson M Winkler et al. “Joint analysis of cortical area and thickness as a replacement for the analysis of the volume of the cerebral cortex”. In: *Cerebral cortex* 28.2 (2018), pp. 738–749.

[16] Kathryn L Mills et al. “Inter-individual variability in structural brain development from late childhood to young adulthood”. In: *NeuroImage* 242 (2021), p. 118450.

[17] Emily S Finn et al. “Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity”. In: *Nature neuroscience* 18.11 (2015), pp. 1664–1671.

[18] Linda Geerligs, Mikail Rubinov, Richard N Henson, et al. “State and trait components of functional connectivity: individual differences vary with mental state”. In: *Journal of Neuroscience* 35.41 (2015), pp. 13949–13961.

[19] I Tavor et al. “Task-free MRI predicts individual differences in brain activity during task performance”. In: *Science* 352.6282 (2016), pp. 216–220.

[20] Ru Kong et al. “Spatial topography of individual-specific cortical networks predicts human cognition, personality, and emotion”. In: *Cerebral cortex* 29.6 (2019), pp. 2533–2551.

[21] Jeremy D Schmahmann et al. “Cerebral white matter: neuroanatomy, clinical neurology, and neurobehavioral correlates”. In: *Annals of the New York Academy of Sciences* 1142.1 (2008), pp. 266–309.

[22] Hui Zhang et al. “NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain”. In: *Neuroimage* 61.4 (2012), pp. 1000–1016.

[23] Peter Kochunov et al. “Multi-site study of additive genetic effects on fractional anisotropy of cerebral white matter: comparing meta and megaanalytical approaches for data pooling”. In: *Neuroimage* 95 (2014), pp. 136–150.

[24] Lotte GM Cremers et al. “Altered tract-specific white matter microstructure is related to poorer cognitive performance: The Rotterdam Study”. In: *Neurobiology of aging* 39 (2016), pp. 108–117.

[25] Sila Genc et al. “Development of white matter fibre density and morphology over childhood: A longitudinal fixel-based analysis”. In: *Neuroimage* 183 (2018), pp. 666–676.

[26] Christian K Tamnes et al. “Diffusion MRI of white matter microstructure development in childhood and adolescence: Methods, challenges and progress”. In: *Developmental cognitive neuroscience* 33 (2018), pp. 161–175.

[27] Derrek P Hibar et al. “Common genetic variants influence human subcortical brain structures”. In: *Nature* 520.7546 (2015), pp. 224–229.

[28] Lloyd T Elliott et al. “Genome-wide association studies of brain imaging phenotypes in UK Biobank”. In: *Nature* 562.7726 (2018), pp. 210–216.

[29] Weikang Gong, Christian F Beckmann, and Stephen M Smith. “Phenotype discovery from population brain imaging”. In: *Medical image analysis* 71 (2021), p. 102050.

[30] Weikang Gong et al. “Supervised phenotype discovery from multimodal brain imaging”. In: *IEEE Transactions on Medical Imaging* (2022).

[31] Catherine Lebel and Christian Beaulieu. “Longitudinal development of human brain wiring continues from childhood into adulthood”. In: *Journal of neuroscience* 31.30 (2011), pp. 10937–10947.

[32] Russell A Poldrack et al. “Long-term neural and physiological phenotyping of a single human”. In: *Nature communications* 6.1 (2015), p. 8885.

[33] Edward Ofori et al. “Longitudinal changes in free-water within the substantia nigra of Parkinson’s disease”. In: *Brain* 138.8 (2015), pp. 2322–2331.

[34] Harith Akram et al. “Connectivity derived thalamic segmentation in deep brain stimulation for tremor”. In: *NeuroImage: Clinical* 18 (2018), pp. 130–142.

[35] Francisca Ferreira et al. “Ventralis intermedius nucleus anatomical variability assessment by MRI structural connectivity”. In: *NeuroImage* 238 (2021), p. 118231.

[36] Salvatore Bertino et al. “Ventral intermediate nucleus structural connectivity-derived segmentation: anatomical reliability and variability”. In: *Neuroimage* 243 (2021), p. 118519.

[37] Michael W Cole et al. “Activity flow over resting-state networks shapes cognitive task activations”. In: *Nature neuroscience* 19.12 (2016), pp. 1718–1726.

[38] Alexander D Cohen et al. “Regression-based machine-learning approaches to predict task activation using resting-state fMRI”. In: *Human brain mapping* 41.3 (2020), pp. 815–826.

[39] Elvis Dohmatob et al. “Brain topography beyond parcellations: local gradients of functional maps”. In: *NeuroImage* (2021), p. 117706.

[40] Gia Ngo et al. “Predicting Individual Task Contrasts From Resting-state Functional Connectivity using a Surface-based Convolutional Network”. In: *bioRxiv* (2021).

[41] Paul C Lauterbur. “Image formation by induced local interactions: examples employing nuclear magnetic resonance”. In: *nature* 242.5394 (1973), pp. 190–191.

[42] Peter Mansfield. “Multi-planar image formation using NMR spin echoes”. In: *Journal of Physics C: Solid State Physics* 10.3 (1977), p. L55.

[43] E Mark Haacke et al. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. Wiley-Liss, 1999.

[44] John Ashburner and Karl J Friston. “Voxel-based morphometry—the methods”. In: *Neuroimage* 11.6 (2000), pp. 805–821.

[45] Anders M Dale, Bruce Fischl, and Martin I Sereno. “Cortical surface-based analysis: I. Segmentation and surface reconstruction”. In: *Neuroimage* 9.2 (1999), pp. 179–194.

[46] R Cameron Craddock et al. “Imaging human connectomes at the macroscale”. In: *Nature methods* 10.6 (2013), pp. 524–539.

[47] David C Van Essen et al. “The Human Connectome Project: a data acquisition perspective”. In: *Neuroimage* 62.4 (2012), pp. 2222–2231.

[48] Matthew F Glasser and David C Van Essen. “Mapping human cortical areas in vivo based on myelin content as revealed by T1-and T2-weighted MRI”. In: *Journal of neuroscience* 31.32 (2011), pp. 11597–11616.

[49] Håkon Grydeland et al. “Intracortical myelin links with performance variability across the human lifespan: results from T1-and T2-weighted MRI myelin mapping and diffusion tensor imaging”. In: *Journal of Neuroscience* 33.47 (2013), pp. 18618–18630.

[50] G. Young and A. M. Halliday. “Neuroradiology”. In: ed. by G. Young and A. M. Halliday. New York, NY: Springer, 2007. Chap. Neuroradiology, pp. 32–57.

[51] Peter D Schellinger et al. “Evidence-based guideline: The role of diffusion and perfusion MRI for the diagnosis of acute ischemic stroke: Report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology”. In: *Neurology* 75.2 (2010), pp. 177–185.

[52] Alexey Surov et al. “Can apparent diffusion coefficient (ADC) distinguish breast cancer from benign breast findings? A meta-analysis based on 13 847 lesions”. In: *BMC Cancer* 16.1 (2016), pp. 1–12.

[53] Talia M Nir et al. “Effectiveness of regional DTI measures in distinguishing Alzheimer’s disease, MCI, and normal aging”. In: *NeuroImage: Clinical* 3 (2013), pp. 180–195.

[54] Catherine Lebel et al. “Diffusion tensor imaging of white matter tract evolution over the lifespan”. In: *Neuroimage* 60.1 (2012), pp. 340–352.

[55] Catherine Lebel and Sean Deoni. “The development of brain white matter microstructure”. In: *Neuroimage* 182 (2018), pp. 207–218.

[56] Timothy EJ Behrens et al. “Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?” In: *neuroimage* 34.1 (2007), pp. 144–155.

[57] Saad Jbabdi and Heidi Johansen-Berg. “Tractography: where do we go from here?” In: *Brain Connectivity* 1.3 (2011), pp. 169–183.

[58] Marco Catani and Michel Thiebaut de Schotten. “A diffusion tensor imaging tractography atlas for virtual in vivo dissections”. In: *Cortex* 44.8 (2005), pp. 1105–1132.

[59] Peter Kochunov et al. “Genetics of microstructure of cerebral white matter using diffusion tensor imaging”. In: *Neuroimage* 53.3 (2010), pp. 1109–1116.

[60] Seiji Ogawa et al. “Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging.” In: *Proceedings of the National Academy of Sciences* 89.13 (1992), pp. 5951–5955.

[61] PA Bandettini et al. “Time course EPI of human brain function during task activation”. In: *Magnetic Resonance in Medicine* 25.2 (1992), pp. 390–397.

[62] KK Kwong et al. “Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation”. In: *Proceedings of the National Academy of Sciences* 89.12 (1992), pp. 5675–5679.

[63] Marcus E Raichle et al. “A default mode of brain function”. In: *Proceedings of the National Academy of Sciences* 98.2 (2001), pp. 676–682.

[64] Michael D Greicius et al. “Default-mode network activity distinguishes Alzheimer’s disease from healthy aging: evidence from functional MRI”. In: *Proceedings of the National Academy of Sciences* 101.13 (2004), pp. 4637–4642.

[65] Stephen M Smith et al. “Correspondence of the brain’s functional architecture during activation and rest”. In: *Proceedings of the national academy of sciences* 106.31 (2009), pp. 13040–13045.

[66] Lucina Q Uddin, Kaustubh Supekar, and Vinod Menon. “Reconceptualizing functional brain connectivity in autism from a developmental perspective”. In: *Frontiers in human neuroscience* 7 (2013), p. 458.

[67] Karl J Friston et al. “Event-related fMRI: characterizing differential responses”. In: *Neuroimage* 7.1 (1998), pp. 30–40.

[68] Keith J Worsley et al. “A unified statistical approach for determining significant signals in images of cerebral activation”. In: *Human brain mapping* 4.1 (1996), pp. 58–73.

[69] Karl J Friston et al. “Statistical parametric maps in functional imaging: a general linear approach”. In: *Human brain mapping* 2.4 (1994), pp. 189–210.

[70] Deanna M Barch et al. “Function in the human connectome: task-fMRI and individual differences in behavior”. In: *Neuroimage* 80 (2013), pp. 169–189.

[71] Simon B Eickhoff et al. “Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty”. In: *Human brain mapping* 30.9 (2009), pp. 2907–2926.

[72] W Jeffrey Elias et al. “A pilot study of focused ultrasound thalamotomy for essential tremor”. In: *New England Journal of Medicine* 369.7 (2013), pp. 640–648.

[73] Rubens Gisbert Cury et al. “Thalamic deep brain stimulation for tremor in Parkinson disease, essential tremor, and dystonia”. In: *Neurology* 89.13 (2017), pp. 1416–1423.

[74] Percival Bailey. *Introduction to Stereotaxis with an Atlas of Human Brain: In Three Volumes*. Georg Thieme Verlag, 1959.

[75] T Hirai and EG Jones. “A new parcellation of the human thalamus on the basis of histochemical staining”. In: *Brain Research Reviews* 14.1 (1989), pp. 1–34.

[76] Marc N Gallay et al. “Human pallidothalamic and cerebellothalamic tracts: anatomical basis for functional stereotactic neurosurgery”. In: *Brain Structure and Function* 212 (2008), pp. 443–463.

[77] Edward G. Jones. *The Thalamus*. Cambridge: Cambridge University Press, 2007.

[78] Sharleen T Sakai. “Cerebellar thalamic and thalamocortical projections”. In: *Handbook of the Cerebellum and Cerebellar Disorders*. Springer, 2021, pp. 661–680.

[79] Richard P Dum and Peter L Strick. “An unfolded map of the cerebellar dentate nucleus and its projections to the cerebral cortex”. In: *Journal of neurophysiology* 89.1 (2003), pp. 634–639.

[80] Roberta Calzavara et al. “Neurochemical characterization of the cerebellar-recipient motor thalamic territory in the macaque monkey”. In: *European Journal of Neuroscience* 21.7 (2005), pp. 1869–1894.

[81] Kenneth B Baker et al. “Deep brain stimulation of the lateral cerebellar nucleus produces frequency-specific alterations in motor evoked potentials in the rat in vivo”. In: *Experimental neurology* 226.2 (2010), pp. 259–264.

[82] Cameron C McIntyre and Philip J Hahn. “Network perspectives on the mechanisms of deep brain stimulation”. In: *Neurobiology of disease* 38.3 (2010), pp. 329–337.

[83] Rick C Helmich et al. “Cerebral causes and consequences of parkinsonian resting tremor: a tale of two circuits?” In: *Brain* 135.11 (2012), pp. 3206–3226.

[84] Nicholas Gravbrot et al. “Advanced imaging and direct targeting of the motor thalamus and dentato-rubro-thalamic tract for tremor: a systematic review”. In: *Stereotactic and Functional Neurosurgery* 98.4 (2020), pp. 220–240.

[85] Anne Morel, Michel Magnin, and Daniel Jeanmonod. “Multiarchitectonic and stereotactic atlas of the human thalamus”. In: *Journal of Comparative Neurology* 387.4 (1997), pp. 588–630.

[86] Timothy EJ Behrens et al. “Characterization and propagation of uncertainty in diffusion-weighted MR imaging”. In: *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* 50.5 (2003), pp. 1077–1088.

[87] Heidi Johansen-Berg et al. “Functional–anatomical validation and individual variation of diffusion tractography-based segmentation of the human thalamus”. In: *Cerebral cortex* 15.1 (2005), pp. 31–39.

[88] A Jakab et al. “Generation of individualized thalamus target maps by using statistical shape models and thalamocortical tractography”. In: *American Journal of Neuroradiology* 33.11 (2012), pp. 2110–2116.

[89] Ludvic Zrinzo et al. “Reducing hemorrhagic complications in functional neurosurgery: a large case series and systematic literature review”. In: *Journal of neurosurgery* 116.1 (2012), pp. 84–94.

[90] Daniel C Alexander et al. “Image quality transfer via random forest regression: applications in diffusion MRI”. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014: 17th International Conference, Boston, MA, USA, September 14-18, 2014, Proceedings, Part III 17*. Springer. 2014, pp. 225–232.

[91] Daniel C Alexander et al. “Image quality transfer and applications in diffusion MRI”. In: *NeuroImage* 152 (2017), pp. 283–298.

[92] Ryutaro Tanno et al. “Bayesian image quality transfer with CNNs: exploring uncertainty in dMRI super-resolution”. In: *Medical Image Computing and Computer Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part I 20*. Springer. 2017, pp. 611–619.

[93] Stefano B Blumberg et al. “Deeper image quality transfer: Training low-memory neural networks for 3d images”. In: *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I*. Springer. 2018, pp. 118–125.

[94] Mojgan Hodaie et al. “Chronic anterior thalamus stimulation for intractable epilepsy”. In: *Epilepsia* 43.6 (2002), pp. 603–608.

[95] Robert Fisher et al. “Electrical stimulation of the anterior nucleus of thalamus for treatment of refractory epilepsy”. In: *Epilepsia* 51.5 (2010), pp. 899–908.

[96] Kyung Jin Lee, Young Min Shon, and Chul Bum Cho. “Long-term outcome of anterior thalamic nucleus stimulation for intractable epilepsy”. In: *Stereotactic and functional neurosurgery* 90.6 (2012), pp. 379–385.

[97] Jukka Peltola et al. “Deep brain stimulation of the anterior nucleus of the thalamus in drug resistant epilepsy in the MORE multicenter patient registry”. In: *Neurology* (2023).

[98] Ihtsham U Haq et al. “A case of mania following deep brain stimulation for obsessive compulsive disorder”. In: *Stereotactic and functional neurosurgery* 88.5 (2010), pp. 322–328.

[99] Paul Krack et al. “Five-year follow-up of bilateral stimulation of the subthalamic nucleus in advanced Parkinson’s disease”. In: *New England Journal of Medicine* 349.20 (2003), pp. 1925–1934.

[100] Maria C Rodriguez-Oroz et al. “Bilateral deep brain stimulation in Parkinson’s disease: a multicentre study with 4 years follow-up”. In: *Brain* 128.10 (2005), pp. 2240–2249.

[101] Alim-Louis Benabid et al. “Deep brain stimulation of the subthalamic nucleus for the treatment of Parkinson’s disease”. In: *The Lancet Neurology* 8.1 (2009), pp. 67–81.

[102] Jens Volkmann et al. “Safety and efficacy of pallidal or subthalamic nucleus stimulation in advanced PD”. In: *Neurology* 63.5 (2004), pp. 915–920.

[103] Valerie C Anderson et al. “Pallidal vs subthalamic nucleus deep brain stimulation in Parkinson disease”. In: *Archives of Neurology* 62.4 (2005), pp. 554–560.

[104] Vincent JJ Odekerken et al. “Subthalamic nucleus versus globus pallidus bilateral deep brain stimulation for advanced Parkinson’s disease (NSTAPS study): a randomised controlled trial”. In: *The Lancet Neurology* 12.1 (2013), pp. 37–44.

[105] David C Van Essen et al. “The WU-Minn human connectome project: an overview”. In: *Neuroimage* 80 (2013), pp. 62–79.

[106] Stamatios N Sotiropoulos et al. “Advances in diffusion MRI acquisition and processing in the Human Connectome Project”. In: *Neuroimage* 80 (2013), pp. 125–143.

[107] Matthew F Glasser et al. “The minimal preprocessing pipelines for the Human Connectome Project”. In: *Neuroimage* 80 (2013), pp. 105–124.

[108] Jesper LR Andersson, Stefan Skare, and John Ashburner. “How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging”. In: *Neuroimage* 20.2 (2003), pp. 870–888.

[109] Jesper LR Andersson and Stamatios N Sotiropoulos. “An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging”. In: *Neuroimage* 125 (2016), pp. 1063–1078.

[110] Mark Jenkinson et al. “Improved optimization for the robust and accurate linear registration and motion correction of brain images”. In: *Neuroimage* 17.2 (2002), pp. 825–841.

[111] Mark Jenkinson et al. “Fsl”. In: *Neuroimage* 62.2 (2012), pp. 782–790.

[112] Jesper LR Andersson, Mark Jenkinson, and Stephen Smith. “Non-linear optimisation FMRIB technical report TR07JA1”. In: *University of Oxford FMRIB Centre: Oxford UK* (2007).

[113] Karla L Miller et al. “Multimodal population brain imaging in the UK Biobank prospective epidemiological study”. In: *Nature neuroscience* 19.11 (2016), pp. 1523–1536.

[114] Fidel Alfaro-Almagro et al. “Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank”. In: *Neuroimage* 166 (2018), pp. 400–424.

[115] Eelke Visser et al. “Automatic segmentation of the striatum and globus pallidus using MIST: Multimodal Image Segmentation Tool”. In: *NeuroImage* 125 (2016), pp. 479–497.

[116] Peter J Basser, James Mattiello, and Denis LeBihan. “Estimation of the effective self-diffusion tensor from the NMR spin echo”. In: *Journal of Magnetic Resonance, Series B* 103.3 (1994), pp. 247–254.

[117] Bruce Fischl et al. “Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain”. In: *Neuron* 33.3 (2002), pp. 341–355.

[118] Bruce Fischl et al. “Automatically parcellating the human cerebral cortex”. In: *Cerebral cortex* 14.1 (2004), pp. 11–22.

[119] Rahul S Desikan et al. “An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest”. In: *Neuroimage* 31.3 (2006), pp. 968–980.

[120] Christophe Destrieux et al. “Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature”. In: *Neuroimage* 53.1 (2010), pp. 1–15.

[121] Yuchun Tang et al. “A probabilistic atlas of human brainstem pathways based on connectome imaging data”. In: *Neuroimage* 169 (2018), pp. 227–239.

[122] Shaun Warrington et al. “XTRACT-Standardised protocols for automated tractography in the human and macaque brain”. In: *Neuroimage* 217 (2020), p. 116923.

[123] Saad Jbabdi et al. “Model-based analysis of multishell diffusion MR data for tractography: How to get over fitting problems”. In: *Magnetic resonance in medicine* 68.6 (2012), pp. 1846–1855.

[124] Stamatios N Sotiropoulos et al. “Fusion in diffusion MRI for improved fibre orientation estimation: An application to the 3T and 7T data of the Human Connectome Project”. In: *Neuroimage* 134 (2016), pp. 396–409.

[125] Moises Hernandez-Fernandez et al. “Using GPUs to accelerate computational diffusion MRI: From microstructure estimation to tractography and connectomes”. In: *Neuroimage* 188 (2019), pp. 598–615.

[126] GŘnther Grabner et al. “Symmetric atlasing and model based segmentation: an application to the hippocampus in older adults”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2006, pp. 58–66.

[127] Francesco Sammartino et al. “Tractography-based ventral intermediate nucleus targeting: Novel methodology and intraoperative validation”. In: *Movement Disorders* 31.8 (2016), pp. 1217–1225.

[128] Shuai Zheng et al. “Conditional random fields as recurrent neural networks”. In: *Proceedings of the IEEE international conference on computer vision*. 2015, pp. 1529–1537.

[129] Yongyue Zhang, Michael Brady, and Stephen Smith. “Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm”. In: *IEEE transactions on medical imaging* 20.1 (2001), pp. 45–57.

[130] Diederik P Kingma and Jimmy Ba. “Adam: A method for stochastic optimization”. In: *arXiv preprint arXiv:1412.6980* (2014).

[131] Gavin JB Elias et al. “Probabilistic mapping of deep brain stimulation: insights from 15 years of therapy”. In: *Annals of Neurology* 89.3 (2021), pp. 426–443.

[132] Matteo Figini et al. “Image quality transfer enhances contrast and resolution of low-field brain mri in african paediatric epilepsy patients”. In: *arXiv preprint arXiv:2003.07216* (2020).

[133] Cathie Sudlow et al. “UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age”. In: *Plos med* 12.3 (2015), e1001779.

[134] O Parker Jones et al. “Resting connectivity predicts task activation in pre-surgical populations”. In: *NeuroImage: Clinical* 13 (2017), pp. 378–385.

[135] David G Ellis and Michele R Aizenberg. “Structural brain imaging predicts individual-level task activation maps using deep learning”. In: *bioRxiv* (2020).

[136] Michael W Cole et al. “Intrinsic and task-evoked network architectures of the human brain”. In: *Neuron* 83.1 (2014), pp. 238–251.

[137] Fenna M Krienen, BT Thomas Yeo, and Randy L Buckner. “Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 369.1653 (2014), p. 20130526.

[138] Maxwell L Elliott et al. “General functional connectivity: Shared features of resting-state and task fMRI drive reliable and heritable individual differences in functional brain networks”. In: *Neuroimage* 189 (2019), pp. 516–532.

[139] Christian F Beckmann and Stephen M Smith. “Probabilistic independent component analysis for functional magnetic resonance imaging”. In: *IEEE transactions on medical imaging* 23.2 (2004), pp. 137–152.

[140] Vince D Calhoun, Kent A Kiehl, and Godfrey D Pearlson. “Modulation of temporally coherent brain networks estimated using ICA at rest and during cognitive tasks”. In: *Human brain mapping* 29.7 (2008), pp. 828–838.

[141] Martijn Van Den Heuvel, Rene Mandl, and Hilleke Hulshoff Pol. "Normalized cut group clustering of resting-state FMRI data". In: *PloS one* 3.4 (2008), e2001.

[142] Pierre Bellec et al. "Multi-level bootstrap analysis of stable clusters in resting-state fMRI". In: *Neuroimage* 51.3 (2010), pp. 1126–1139.

[143] BT Thomas Yeo et al. "The organization of the human cerebral cortex estimated by intrinsic functional connectivity". In: *Journal of neurophysiology* (2011).

[144] R Cameron Craddock et al. "A whole brain fMRI atlas generated via spatially constrained spectral clustering". In: *Human brain mapping* 33.8 (2012), pp. 1914–1928.

[145] Christian F Beckmann et al. "Group comparison of resting-state FMRI data using multi-subject ICA and dual regression". In: *Neuroimage* 47.Suppl 1 (2009), S148.

[146] Nicola Filippini et al. "Distinct patterns of brain activity in young carriers of the APOE-ε4 allele". In: *Proceedings of the National Academy of Sciences* 106.17 (2009), pp. 7209–7214.

[147] Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[148] Samuel J Harrison et al. "Large-scale probabilistic functional modes from resting state fMRI". In: *NeuroImage* 109 (2015), pp. 217–231.

[149] Samuel J Harrison et al. "Modelling subject variability in the spatial and temporal characteristics of functional modes". In: *NeuroImage* 222 (2020), p. 117226.

[150] Seyedeh-Rezvan Farahibozorg et al. "Hierarchical modelling of functional brain networks in population and individuals from big fMRI data". In: *bioRxiv* (2021).

[151] Gholamreza Salimi-Khorshidi et al. "Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers". In: *NeuroImage* 90 (2014), pp. 449–468.

[152] Ahmad R Hariri et al. "The amygdala response to emotional stimuli: a comparison of faces and scenes". In: *Neuroimage* 17.1 (2002), pp. 317–323.

[153] Mark W Woolrich et al. "Temporal autocorrelation in univariate linear modeling of FMRI data". In: *Neuroimage* 14.6 (2001), pp. 1370–1386.

[154] Mark W Woolrich et al. "Multilevel linear modelling for FMRI group analysis using Bayesian inference". In: *Neuroimage* 21.4 (2004), pp. 1732–1747.

[155] Emma C Robinson et al. "MSM: a new flexible framework for Multimodal Surface Matching". In: *Neuroimage* 100 (2014), pp. 414–426.

[156] Emma C Robinson et al. "Multimodal surface matching with higher-order smoothness constraints". In: *Neuroimage* 167 (2018), pp. 453–465.

[157] Lisa D Nickerson et al. "Using dual regression to investigate network shape and amplitude in functional connectivity analyses". In: *Frontiers in neuroscience* 11 (2017), p. 115.

[158] Stephen M Smith et al. "Group-PCA for very large fMRI datasets". In: *Neuroimage* 101 (2014), pp. 738–749.

[159] Janine Diane Bijsterbosch et al. "The relationship between spatial configuration and functional connectivity of brain regions". In: *Elife* 7 (2018), e32992.

[160] Janine Diane Bijsterbosch et al. “The relationship between spatial configuration and functional connectivity of brain regions revisited”. In: *Elife* 8 (2019), e44890.

[161] Agustin Lage-Castellanos et al. “Methods for computing the maximum performance of computational models of fMRI responses”. In: *PLoS computational biology* 15.3 (2019), e1006397.

[162] Timothy S. Coalson, David C. Van Essen, and Matthew F. Glasser. “The impact of traditional neuroimaging methods on the spatial localization of cortical areas”. In: *Proceedings of the National Academy of Sciences* 115.27 (2018), E6356–E6365. eprint: `https://www.pnas.org/content/115/27/E6356.full.pdf`. URL: `https://www.pnas.org/content/115/27/E6356`.

[163] Michael W Cole et al. “The functional relevance of task-state functional connectivity”. In: *Journal of Neuroscience* 41.12 (2021), pp. 2684–2702.

[164] Maxwell L Elliott et al. “What Is the Test-Retest Reliability of Common Task-Functional MRI Measures? New Empirical Evidence and a Meta-Analysis”. In: *Psychological Science* (2020), p. 0956797620916786.

[165] Martin I Sereno et al. “Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging”. In: *Science* 268.5212 (1995), pp. 889–893.

[166] Brian A Wandell, Serge O Dumoulin, and Alyssa A Brewer. “Visual field maps in human cortex”. In: *Neuron* 56.2 (2007), pp. 366–383.

[167] Yong Liu et al. “Impaired long distance functional connectivity and weighted network architecture in Alzheimer’s disease”. In: *Cerebral Cortex* 24.6 (2014), pp. 1422–1435.

[168] Amalia Papanikolaou et al. “Population receptive field analysis of the primary visual cortex complements perimetry in patients with homonymous visual field defects”. In: *Proceedings of the National Academy of Sciences* 111.16 (2014), E1656–E1665.

[169] Ione Fine and Geoffrey M Boynton. “Pulse trains to percepts: the challenge of creating a perceptually intelligible world with sight recovery technologies”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 370.1677 (2015), p. 20140208.

[170] Oliver P Hinds et al. “Accurate prediction of V1 location from cortical folds in a surface coordinate system”. In: *Neuroimage* 39.4 (2009), pp. 1585–1599.

[171] Noah C Benson et al. “The retinotopic organization of striate cortex is well predicted by surface topology”. In: *Current Biology* 22.21 (2012), pp. 2081–2085.

[172] Mark M Schira and Alex R Wade. “Understanding the functional consequences of an individual retinotopic map”. In: *Vision research* 62 (2012), pp. 1–6.

[173] Noah C Benson et al. “Intrinsic fluctuations in the fMRI signal: A novel target for manipulation by cortical stimulation”. In: *NeuroImage* 84 (2014), pp. 1067–1079.

[174] Zheng Wang et al. “The relationship of anatomical and functional connectivity to resting-state connectivity in primate somatosensory cortex”. In: *Neuron* 78.6 (2015), pp. 1116–1126.

[175] Fernanda L Ribeiro, Steffen Bollmann, and Alexander M Puckett. “Predicting the retinotopic organization of human visual cortex from anatomy using geometric deep learning”. In: *Neuroimage* 244 (2021), p. 118624.

[176] Bruce Fischl et al. “Cortical folding patterns and predicting cytoarchitecture”. In: *Cerebral cortex* 18.8 (2008), pp. 1973–1980.

[177] Oliver P Hinds et al. “Accurate prediction of V1 location from cortical folds in a surface coordinate system”. In: *Neuroimage* 39.4 (2008), pp. 1585–1599.

[178] Jakob Heinzle, Thorsten Kahnt, and John-Dylan Haynes. “Topographically specific functional connectivity between visual field maps in the human brain”. In: *Neuroimage* 56.3 (2011), pp. 1426–1436.

[179] Timothy O Laumann et al. “Functional system and areal organization of a highly sampled individual human brain”. In: *Neuron* 87.3 (2015), pp. 657–670.

[180] Andrew S Bock et al. “Resting-state retinotopic organization in the absence of retinal input and visual experience”. In: *Journal of Neuroscience* 35.36 (2015), pp. 12366–12382.

[181] Ella Striem-Amit et al. “Functional connectivity of visual cortex in the blind follows retinotopic organization principles”. In: *Brain* 138.6 (2015), pp. 1679–1695.

[182] Joseph C Griffis et al. “Retinotopic patterns of functional connectivity between V1 and large-scale brain networks during resting fixation”. In: *Neuroimage* 146 (2017), pp. 1071–1083.

[183] Koen V Haak, Andre F Marquand, and Christian F Beckmann. “Connectopic mapping with resting-state fMRI”. In: *Neuroimage* 170 (2018), pp. 83–94.

[184] Koen V Haak and Christian F Beckmann. “Objective analysis of the topological organization of the human cortical visual connectome suggests three visual pathways”. In: *Cortex* 98 (2018), pp. 73–83.

[185] David M Watson and Timothy J Andrews. “An evaluation of how connectopic mapping reveals visual field maps in V1”. In: *Scientific reports* 12.1 (2022), p. 16249.

[186] Matthew F Glasser et al. “A multi-modal parcellation of human cerebral cortex”. In: *Nature* 536.7615 (2016), pp. 171–178.

[187] Ludovica Griffanti et al. “ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging”. In: *NeuroImage* 95 (2014), pp. 232–247.

[188] Matthew F Glasser et al. “Using temporal ICA to selectively remove global noise while preserving global signal in functional MRI data”. In: *NeuroImage* 181 (2018), pp. 692–717.

[189] An T Vu et al. “Tradeoffs in pushing the spatial resolution of fMRI for the 7T Human Connectome Project”. In: *NeuroImage* (2016).

[190] Serge O Dumoulin and Brian A Wandell. “Population receptive field estimates in human visual cortex”. In: *NeuroImage* 39.2 (2008), pp. 647–660.

[191] Noah C Benson et al. “The Human Connectome Project 7 Tesla retinotopy dataset: Description and population receptive field analysis”. In: *Journal of vision* 18.13 (2018), pp. 23–23.

[192] Kendrick N Kay et al. “Compressive spatial summation in human visual cortex”. In: *Journal of neurophysiology* 110.2 (2013), pp. 481–494.

[193] Fernanda L Ribeiro et al. “Variability of visual field maps in human early extrastriate cortex challenges the canonical model of organization of V2 and V3”. In: *bioRxiv* (2022), pp. 2022–10.