# Mode-based morphometry: A multiscale approach to mapping human neuroanatomy

Trang Cao[a], James C. Pang[a], Ashlea Segal[a], Yu-Chi Chen[a], Kevin M. Aquino[b], Michael Breakspear[c], Alex Fornito[a]

[a]*The Turner Institute for Brain and Mental Health, School of Psychological Sciences, and Monash Biomedical Imaging, Monash University, 762-772 Blackburn Rd, Clayton VIC 3168, Australia*
[b]*School of Physics, University of Sydney, Physics Rd, Camperdown NSW 2006, Australia*
[c]*School of Psychological Sciences, University of Newcastle, University Dr, Callaghan NSW 2308, Australia*

## Abstract

Voxel-based morphometry (VBM) and surface-based morphometry (SBM) are two widely used neuroimaging techniques for investigating brain anatomy. These techniques rely on statistical inferences at individual points (voxels or vertices), clusters of points, or a priori regions-of-interest. They are powerful tools for describing brain anatomy, but offer little insights into the generative processes that shape a particular set of findings. Moreover, they are restricted to a single spatial resolution scale, precluding the opportunity to distinguish anatomical variations that are expressed across multiple scales. Drawing on concepts from classical physics, here we develop an approach, called mode-based morphometry (MBM), that can describe any empirical map of anatomical variations in terms of the fundamental, resonant modes—eigenmodes—of brain anatomy, each tied to a specific spatial scale. Hence, MBM naturally yields a multiscale characterization of the empirical map, affording new opportunities for investigating the spatial frequency content of neuroanatomical variability. Using simulated and empirical data, we show that the validity and reliability of MBM are either comparable or superior to classical vertex-based SBM for capturing differences in cortical thickness maps between two experimental groups. Our approach thus offers a robust, accurate, and informative method for characterizing empirical maps of neuroanatomical variability that can be directly linked to a generative physical process.

## 1. Introduction

Voxel-based morphometry (VBM) [1] and surface-based morphometry (SBM) [2] are the most commonly used techniques for studying neuroanatomical variations with magnetic resonance imaging (MRI). They are used to find associations between morphometric quantities (e.g., cortical thickness, surface area, grey matter volume) and diverse sociodemographic [3, 4, 5], behavioral [6, 7, 8], and clinical [9] variables. Throughout this article, we will focus principally on applications that examine neuroanatomical differences between two experimental groups (e.g., males and females, patients and controls), but our arguments and techniques easily generalize to cross-sectional analyses of inter-individual variability (e.g., brain-wide association analyses [10]).

The typical procedure in VBM- and SBM-based analyses of group differences involves fitting a statistical model at each individual voxel or vertex, respectively, and then applying some statistical threshold for inferring the spatial location(s) of significant mean differences [11, 12, 13]. In some cases, voxels/vertices are aggregated into anatomical regions-of-interest (ROIs) defined using an a priori atlas [14]. These approaches have been successful in describing where the differences are located in the brain, but offer no insights into the fundamental constraints that may have shaped those differences. Moreover, these approaches rest on the assumption that anatomical differences (1) are highly localized; (2) arise independently of each other; and (3) are confined to the spatial scale defined by the measurement resolution (i.e., local collections of voxels/vertices or atlas-based ROIs). These assumptions interact with the statistical thresholding procedure in a way that can obscure spatially extended patterns underlying the data (for an example, see Fig. S15 in ref. [15]). Moreover, the application of different statistical thresholds, derived using different methods, contributes to inconsistencies in findings reported in the literature [13, 16, 17, 18, 19]. Hence, a robust framework is needed that does not necessarily rely on a predefined threshold and can examine anatomical differences at multiple scales.

In diverse areas of physics and engineering, analysis at multiple scales can be obtained as the structure of a system can be comprehensively described by its structural eigenmodes, which are also referred to as modes, eigenfunctions, eigenvectors, or harmonics [20, 21, 22]; we will use these terms interchangeably in the text. Eigenmodes correspond to the natural, resonant modes of the system and represent an orthogonal basis set that can

describe any spatial pattern expressed by the system, much like the basis set of sines and cosines used to understand the frequency content of signals in Fourier analysis [20, 22]. Recent work has shown that eigenmodes derived either from a model of brain geometry, termed *geometric eigenmodes*, or from a graph-based model of the structural connectome based on diffusion MRI, termed *connectome eigenmodes*, can be used as a basis set for reconstructing diverse aspects of brain activity [23, 24, 25, 26, 27, 28, 15, 29, 30, 31, 32], for quantifying structure-function coupling in the brain [33, 34, 35], and for understanding atrophy patterns in neurodegeneration [36, 37, 38] and other conditions [39, 40]. In each of these cases, empirical spatial brain maps can be viewed as resulting from the preferential involvement, or excitation, of specific resonant modes of brain structure, thus offering insights into the generative physical mechanisms that shape the observed spatial pattern.

A further advantage of a mode-based approach is that, as in Fourier analyses, the modes are ordered by their topological (connectome eigenmodes) or spatial (geometric eigenmodes) frequency, offering a spectral decomposition of the data that provides insights into its multiscale organization. Indeed, recent work using geometric eigenmodes suggests that virtually all task-evoked activation maps obtained with functional MRI result from excitations of low-frequency modes spanning wavelengths >60 mm [15]. Other work indicates that individual differences in brain anatomy are most salient at relatively coarse scales, spanning wavelengths $\geqslant$37 mm [41]. Such findings challenge classical approaches to brain mapping that focus on identifying focal effects in isolated brain regions and underscore the benefits of a spatially-informed, multiscale approach.

Here, we leverage the advantages offered by a mode-based view of the brain to develop a new approach, called mode-based morphometry (MBM), for mapping anatomical differences between groups. This approach involves modeling such differences as a linear combination of eigenmodes and performing statistical inference on the modes, rather than on individual voxels or vertices. In this work, we use the geometric eigenmodes of the cortex, given their superior performance in explaining functional data [15], but note that our method can be adapted for use with other eigenmodes (e.g., connectome eigenmodes) as required. We focus on a surface-based analysis of cortical thickness (CT) differences, but other morphometric measures (e.g., surface area, volume, etc.) can easily be analyzed with the same method. We develop a framework for simulating CT maps with known ground truth and use them to compare the accuracy and reliability (under resampling) of MBM

with classical vertex-based SBM. We then evaluate the consistency of the findings obtained by the two techniques as applied to empirical data through the analyses of (i) CT differences between sexes and (ii) multi-site measures of CT differences between healthy controls and patients with Alzheimer's disease. In both simulated and empirical scenarios, our results indicate that MBM offers comparable, or superior, validity and consistency compared to SBM, while offering a direct insight into the generative mechanisms and multiscale characteristics of the data.

## 2. Methods

We begin by describing the details of the SBM and MBM approaches. We then describe our methods for evaluating the accuracy and consistency of the two approaches using simulated and empirical data.

### 2.1. Surface-based morphometry (SBM)

Traditional SBM compares neuroanatomical variations between two groups of interest using a general linear model (GLM) (Fig. 1a) given by

$$\mathbf{Y_i} = \mathbf{G}\alpha_\mathbf{i} + \mathbf{e}, \tag{1}$$

where $\mathbf{Y_i}$ is the data vector $(L \times 1)$ of $L$ measurements at vertex $i$ of the cortical surface, $\mathbf{G}$ is the design matrix $(L \times P)$ of effect variables, $\alpha_i$ is the parameter vector $(P \times 1)$, and $\mathbf{e}$ is an error vector $(L \times 1)$. Elements of $\alpha_i$ control the contribution of the corresponding effect columns of $\mathbf{G}$ to each vertex, $\mathbf{Y_i}$. Here, the GLM parameterizes the group differences as $t$ statistics by defining a design matrix $\mathbf{G}$ and a contrast vector $\mathbf{c}$ as follows. Each element in a column of $\mathbf{G}$ has value 1 or 0 to indicate a measurement, i.e., a subject, belonging to a group or not. Each column of $\mathbf{G}$ controls the contribution of each group. The estimated values of $\alpha_i$ representing the mean of each group are multiplied by the contrast vector $\mathbf{c}$ defined by $[1, -1]$ for calculating the $t$-statistic for each vertex (see [42, Appendix A] for more details).

In this work, we performed non-parametric inference on the $t$ statistics via permutation testing with 5000 iterations. At each iteration, we shuffled the group labels of all participants to create a $t$-map observed under random group assignment, resulting in an empirical null distribution at each vertex. We then used tail estimation on each null distribution [43] to calculate $p$
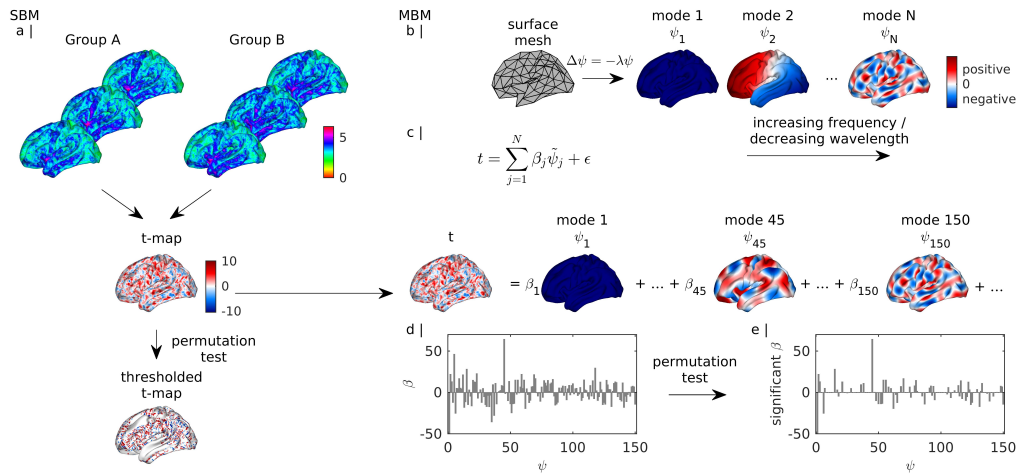
4

Figure 1: SBM and MBM analysis pipelines. (a) In SBM, a $t$-statistic is calculated independently at each vertex, quantifying point-wise group differences in CT. A thresholded $t$-map is derived by comparing the observed $t$-map and the distribution of null $t$-maps after permutation testing. (b) In MBM, eigenmodes are derived from a cortical surface mesh (solving Eq. (2)). The modes are ordered in increasing spatial frequency or decreasing spatial wavelength. Values in each mode are arbitrarily defined, with negative–zero–positive values in colored as blue–white–red. (c) An empirical $t$-map can be decomposed as a weighted sum of eigenmodes and errors using a GLM (Eq. 6), with weights given by $\beta_j$. The set of $\beta_j$ is called the $\beta$ spectrum. (d) An example $\beta$ spectrum with large $\beta_{45}$ indicating a dominant contribution from mode 45. (e) An example of statistically significant $\beta$s derived by comparing the observed $\beta$ spectrum and the distribution of null $\beta$ spectra after permutation testing.

values with arbitrarily high precision. In our analyses, we evaluated results with respect to (i) unthresholded differences; (ii) differences thresholded at $p < 0.05$ but uncorrected ($p_{\mathrm{unc}}$); and (iii) differences thresholded at $p < 0.05$ but corrected for false-discovery rate (FDR) ($p_{\mathrm{FDR}}$).

## 2.2. Mode-based morphometry (MBM)

MBM takes as input the unthresholded spatial map of $t$ statistics quantifying group mean differences in CT and reconstructs that map using a basis set of structural eigenmodes of the brain. These eigenmodes can be defined using relevant anatomical properties of interest (e.g., geometry or connectome). The assumptions in MBM are that (1) the anatomical group differences reflect the superposition of different resonant modes of neuroanatomy

that are preferentially expressed in one group over the other; and (2) these modes are orthogonal and intrinsic to the brain structure (see third paragraph of the Introduction for justification). Here, we focus on eigenmodes derived from the geometry of the cortex (geometric eigenmodes) because recent data [15] indicate that they offer a more accurate and parsimonious basis set for brain function and their derivation relies on a simpler processing pipeline. We provide further justification for this choice in the Discussion section.

The geometric eigenmodes are obtained by solving the eigendecomposition of the Laplace-Beltrami Operator (LBO), also known as the Helmholtz equation, on the cortical surface,

$$\Delta \psi = -\lambda \psi, \tag{2}$$

where $\Delta$ is the LBO and the solution $\psi = \{\psi_1, \psi_2, \dots\}$ is the family of geometric eigenmodes with corresponding family of eigenvalues $\lambda = \{\lambda_1, \lambda_2, \dots\}$. The LBO captures the intrinsic geometry of the cortical manifold, which includes the curvature of the cortical surface in this case, and is generally defined as [44, 45, 15]

$$\Delta := \frac{1}{W} \sum_{i,j} \frac{\partial}{\partial x_i} \left( g^{ij} W \frac{\partial}{\partial x_j} \right), \tag{3}$$

where $x_i, x_j$ are the local coordinates, $W := \sqrt{\det G}$, $G := (g_{ij})$, $g_{ij} := \langle \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \rangle$ is an inner product, $g^{ij} = (g_{ij})^{-1}$ is the inverse of $g_{ij}$, and det denotes the determinant. $\psi$ represents the set of spatially extended resonant modes, or eigenmodes, of the geometry. $\lambda$ is related to the intrinsic spatial wavelengths of the modes, such that

$$wavelength = \frac{2\pi}{\sqrt{\lambda}}. \tag{4}$$

Eq. (4) is derived from the relationship between eigenvalues $\lambda$ and resonant or natural frequency $k$ [45], with $k = \sqrt{\lambda} = 2\pi/wavelength$. Furthermore, based on their wavelengths, the modes can be grouped into spatial scales by considering the special case of a sphere, which is known to be topologically similar to the cortical surface [46]. By solving Eq. (2) on the sphere, degenerate solutions exist, such that certain eigenmodes will have the same number of nodal lines and wavelengths, and they can be aggregated into an

eigengroup $l$, similar to the angular momentum number in quantum physics. Each eigengroup comprises $2l + 1$ eigenmodes, and its wavelength is [46, 41]

$$wavelength = \frac{2\pi R_s}{\sqrt{l(l+1)}}, \tag{5}$$

where $R_s$ is the radius of the sphere (see Table S1 for an explicit list of wavelengths and eigenmode membership for each eigengroup on a sphere of $R_s = 67$ mm [41, 15]). Therefore, the eigenmodes of the cortical surface belonging to a given eigengroup have approximately the same spatial scale or wavelength given by Eq. (5).

The solutions of Eq. (2)—i.e., the geometric eigenmodes—correspond to the spatial part of the solutions of a wave equation describing the dynamics on a surface [47]. When applied to neuroanatomy, the eigenmodes represent an orthogonal basis set describing spatial patterns of the variations of cortical geometry at different spatial scales or wavelengths. The increasing order of the eigenvalues $0 \leqslant \lambda_1 < \lambda_2 < \ldots$ corresponds to a decreasing order of the wavelengths. The first eigenvalue $\lambda_1$ is approximately equal to zero with a wavelength that is very large compared to the size of the brain. Its corresponding eigenmode $\psi_1$ is a constant function with no nodal lines and can be used to represent a mean across the brain.

We solved the eigenmodes of the cortical surface represented by a triangular mesh using the LaPy Python library [47, 48]. Here, we used a population-averaged template of the cortical midthickness surface in fsLR-32k CIFTI surface [49], comprising $32,492$ vertices in each hemisphere (Fig. 1b). We used a population-averaged surface to obtain a common set of eigenmodes for all participants, enabling easier comparisons. Note that individual-specific eigenmodes can also be derived from individual surfaces. However, small differences in the geometry of the cortical surface can alter the spatial patterns of the eigenmodes, which makes comparison between individual-specific modes challenging, particularly for short wavelengths or more fine-grained modes [41]. Moreover, prior work indicates that relying on a single, canonical basis set of modes derived from a group template leads to a negligible difference in reconstructions of empirical data when compared to using individual-specific modes [15], suggesting that the group-derived modes provide a reasonable approximation for present purposes. Here, we performed analyses on the left hemisphere and thus used the modes derived from the left hemisphere.

7

We then used the eigenmodes to decompose the $t$-map of each hemisphere via a GLM given by (Fig. 1c)

$$t = \sum_{j=1}^{N} \beta_j \tilde{\psi}_j + \epsilon, \tag{6}$$

where $t = [t_1, t_2, \ldots, t_i, \ldots]$ is the $t$-map, $i$ represents a vertex, $\tilde{\psi}_j$ is the normalized $j$th eigenmode, $\beta_j$ is the coefficient quantifying the contribution of $\tilde{\psi}_j$ to the $t$-map, $N$ is the number of modes used, and $\epsilon$ is a constant vector (error vector). Note that $\beta_j$ is distinct from the eigenvalue $\lambda_j$, with the former corresponding to the contribution of eigenmode $j$ to the $t$-map and the latter corresponding to the wavelength of eigenmode $j$. In this work, we considered $N = 150$ modes spanning wavelengths $\geqslant 34$ mm, given recent work showing that the most salient aspects of brain function reside at spatial wavelengths longer than 34 mm [15, 41]. However, our results do not change with other choices of $N$ (see Fig. S1). We normalized the eigenfunction $\psi_j$ by dividing it with its Euclidean norm to ensure orthonormality of the basis set.

We call the set of $\beta_j$ coefficients the $\beta$ spectrum, which encodes the contribution of each eigenmode at a specific spatial scale, i.e., each frequency component, to the neuroanatomical property being investigated. In the present application, the investigated property is a $t$-statistic map quantifying group differences in CT. Thus, for example, if $\beta_{45}$ is large, the spatial pattern defined by mode 45 dominates group differences in CT (Figs. 1c and d).

We performed statistical inference on the $\beta$ estimates using permutation testing. To this end, we reconstructed null $t$-maps obtained after shuffling the group labels, as described in Section 2.1 for the SBM analysis, yielding an empirical distribution of null $\beta$ spectra. We then used tail approximation to estimate $p$-values for each observed $\beta$ value in the spectrum [43]. This procedure thus allows us to determine which specific modes make a significant contribution to the observed CT differences between groups. In this way, MBM performs inference at the level of the modes of brain structure rather than individual voxels. The analysis thus determines the specific modes that make a statistically significant contribution to the empirical spatial map. As in the SBM analysis, we considered unthresholded spectra and those thresholded at $p_{\text{unc}} < 0.05$ and $p_{\text{FDR}} < 0.05$ (Fig. 1e).
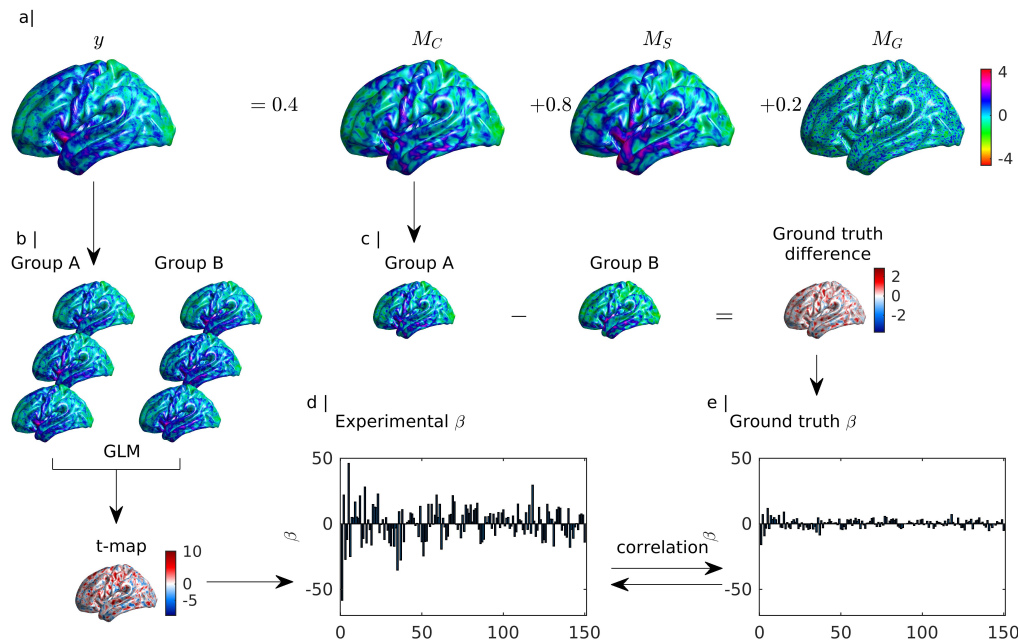
Figure 2: Framework for ground-truth simulations. (a) We generate a CT map using the model in Eq. (7). Here, we show an example with $\sigma_C = 0.4$ and $\sigma_S = 0.8$. (b) Using the simulated CT maps for groups $A$ and $B$, we estimate a $t$-map and its corresponding $\beta$ spectrum (d). (c) The ground-truth (GT) difference map is given by the subtraction of $\mathbf{M}_C^A$ and $\mathbf{M}_C^B$, from which the ground-truth $\beta$ spectrum (e) is obtained.

## 2.3. Validations using simulated data

One challenge in evaluating the validity of any new brain mapping technique is the lack of a ground truth for most applications. Here, we developed a framework for simulating group differences in CT with a known ground truth to compare the performance of SBM and MBM.

### 2.3.1. Simulation Framework

We simulated two experimental groups based on empirical CT maps using the simple model defined in Eq. (7). We simulated the CT map, $\mathbf{y}$, of a subject in each group by combining three elements: (1) a group-specific common phenotype map ($\mathbf{M}_C$), which represents a structured CT pattern, i.e., a ground truth, that is common to all members of a group, thus justifying their assignment to the same group; (2) an individual-specific structured noise map ($\mathbf{M}_S$), which represents structured CT variations that are specific

9

to each individual; and (3) Gaussian noise ($\mathbf{M}_G$), which represents measurement error. $\mathbf{M}_S$ reflects between-subject variance, which can be considered structured noise when compared with the group means. It is structured in the sense that it is autcorrelated. For simplicity, we assumed that the combination of these three elements was linear, such that the model for $\mathbf{y}$ was

$$\mathbf{y} = \sigma_C \mathbf{M}_C + \sigma_S \mathbf{M}_S + (1 - \sigma_S)\mathbf{M}_G, \tag{7}$$

where $\sigma_C$ $(0 \leqslant \sigma_C \leqslant 1)$ is a free parameter that controls the contribution of the group-specific CT phenotype, $\mathbf{M}_C$, and $\sigma_S$ $(0 \leqslant \sigma_S \leqslant 1)$ is a free parameter that controls the relative contributions of structured noise, $\mathbf{M}_S$, and Gaussian noise, $\mathbf{M}_G$ (Fig. 2a). When $\sigma_C = 0$, the generated maps only contain noise, and when $\sigma_C = 1$, the generated maps have equal weightings between the group-specific map, representing the ground-truth phenotype, and noise, including structured and Gaussian noise. When $\sigma_S = 0$, the generated maps are only affected by Gaussian noise, and when $\sigma_S = 1$, the generated maps are only affected by individual-specific structured noise. Note that the range of $\sigma_C$ and $\sigma_S$ were chosen for simplicity to capture the relation between $\mathbf{M}_C, \mathbf{M}_S$, and $\mathbf{M}_G$.

To specify two groups $A$ and $B$, we include the superscripts $A$ and $B$ to $\mathbf{M}_S$ and $\mathbf{M}_C$; for instance $\mathbf{M}_C^A$ denotes the common group CT phenotype map of a subject in group $A$. Under this model, classical inference on group differences is the same as asking whether there is a difference in $\mathbf{M}_C^A$ and $\mathbf{M}_C^B$ that can be reliably detected over and above the variability associated with $\mathbf{M}_S$ and $\mathbf{M}_G$.

We used 25 subjects for each group, unless otherwise stated, to mimic sample sizes historically used in the literature, although our general conclusions are not affected by this choice (see Fig. S2). To ensure that the generated maps have the same spatial structure as empirical data, we sampled maps $\mathbf{M}_C$ and $\mathbf{M}_S$ without replacement from 339 CT maps in the Human Connectome Project (HCP) data described in Sec. 2.4.1 [50]. Note that $\mathbf{M}_C$ corresponds to the same spatial map for all participants within a group, whereas a different $\mathbf{M}_S$ was selected for each individual within a group. We embedded actual empirical data within the simulations in this way to preserve the real structure and properties of CT maps as much as possible.

For each $\mathbf{y}$ map, we subtracted the minimum negative value across all $\mathbf{y}$ (both groups) to remove negative values caused by the addition of Gaussian noise to the CT maps. Given Eq. (1), this offset did not affect the $t$-map

10

comparing group mean CT differences, as the same constant value was subtracted to the two groups.

For each pair of parameters $\sigma_C$ and $\sigma_S$, we ran 100 experiments with fixed ground-truths $\mathbf{M}_C^A$ and $\mathbf{M}_C^B$. We chose 100 experiments to balance the computational cost and the reliability of the resulting performance metrics. Figure S3 confirms that our findings are not sensitive to the selection of different ground truths. From the generated CT maps of the two groups, the $t$-map and $\beta$ spectrum of each experiment were computed as described in Sections 2.1 and 2.2, respectively (Figs. 2b and d).

The parameters $\sigma_C$ and $\sigma_S$ allow us to evaluate the sensitivity of both SBM and MBM to variations in the level of real phenotype and structured and Gaussian noise. To determine the most plausible parameter regime that yields the most realistic estimates of empirical data, we compared the spatial variograms of the empirical and simulated CT maps under different parameter combinations (Fig. 3). A variogram provides a measure of spatial dependence, or spatial autocorrelation, of a random field, defined as the variance of the difference between field values at two locations across realizations of the field. We used the Python package BrainSMASH [51] to calculate the variograms. For the empirical data, we calculated the mean and variance of the variograms of the 339 CT maps from the HCP data (see Section 2.4.1). For each pair of parameters $\sigma_C$ and $\sigma_S$ in the simulations, we generated 500 CT maps and calculated the mean and variance of the corresponding variograms. The slope of the variogram encodes the spatial structure of the data [51]. For example, a positive slope means that there is high autocorrelation in the data; i.e., variations at small distances are smaller than variations at large distances. A flat variogram means that variations are independent of distances. Similar slopes refer to the similarity of the spatial autocorrelations. Thus, to compare the slope of the mean of variograms of empirical maps and that of simulated maps, we removed the offset between their minimum values, so that the two variograms start at the same point (Fig. 3a). In particular, the offset was calculated from the mean variogram that has a higher minimum value. Note that removing the offset does not affect the structure of the variograms. We then calculated the norm distance between the emirical and simulated mean variograms after subtracting the minimum offset.

Figure 3b shows the norm distances for combinations of $\sigma_C$ and $\sigma_S$. Figure 3a shows two examples of the empirical and simualted mean variograms (after subtracting the minimum offset) at parameter pairs ($\sigma_C = 1, \sigma_S = 0$)
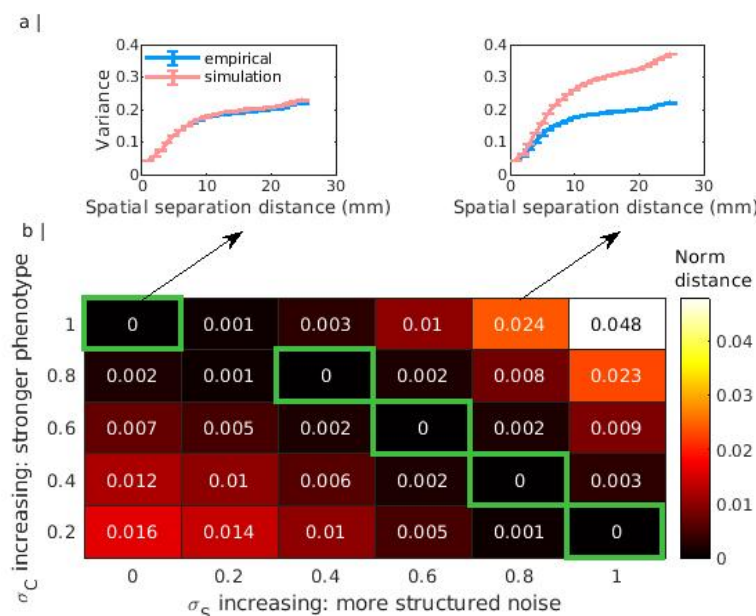
Figure 3: Comparing spatial variograms of empirical and simulated CT maps for different combinations of $\sigma_C$ and $\sigma_S$. (a) Two examples of mean variograms (after subtracting the minimum offset) of empirical and simulated maps for parameter pairs ($\sigma_C = 1, \sigma_S = 0$) and ($\sigma_C = 1, \sigma_S = 0.8$). The error bars (vertical bars) show the variance of the variograms. (b) Norm distances between the empirical and simulated mean variograms (after subtracting the minimum offset) for combinations of $\sigma_C$ and $\sigma_S$. The green boxes highlight the realistic regimes where the generated maps have a similar spatial structure as the empirical data (norm distance $\approx 0$).

and ($\sigma_C = 1, \sigma_S = 0.8$) (see Fig. S4 for the mean variograms of all $\sigma_C$ and $\sigma_S$ in the norm distance table). The green boxes in Fig. 3b highlight the realistic parameter regimes where the generated maps have a similar spatial structure to the empirical data; i.e., the norm distances are close to zero. We will highlight these realistic regimes in the Results section and examine the effects of parameter choices in these realistic regimes.

### 2.3.2. Performance evaluation

We compared the performance of SBM and MBM in terms of accuracy and consistency. We defined accuracy as the ability of each method to accurately capture the ground-truth group difference in the simulations. These ground-truth difference maps were obtained by subtracting the $\mathbf{M}_C^A$ and $\mathbf{M}_C^B$

maps (Fig. 2c). The accuracy of SBM was quantified as the product-moment correlation between this ground truth and the $t$-map obtained through classical analysis (see Section 2.1). Similarly, the accuracy of MBM was quantified as the product-moment correlation between the $\beta$ spectrum (see Section 2.2) of the ground-truth difference map (Fig. 2e) and the $\beta$ spectrum obtained for each experiment (Fig. 2d).

We defined consistency as the ability of each method to obtain consistent findings in the face of sampling variability across repeated experiments. Therefore, we compared the distributions of pairwise correlations between $t$-maps of 100 experiments for SBM and pairwise correlations between $\beta$ spectra of 100 experiments for MBM. When considering thresholded results, we used binary correlations [52, 53]. We present unthresholded and uncorrected thresholded results at $p_{\mathrm{unc}} \leqslant 0.05$ in the main text, and FDR-corrected results in Supplementary Fig. S5. The general conclusions are consistent for both uncorrected and corrected results. We also evaluated the performance of SBM and MBM on CT maps spatially smoothed with surface-based kernel sizes of 10 mm, 20 mm, or 30 mm full-width at half-maximum (FWHM) to match with the common smoothing practice of SBM. The Connectome Workbench software [54] was used to smooth the CT maps.

### 2.4. Validations using empirical data

The simulated data offer an opportunity to compare the accuracy of the proposed MBM method vs SBM with respect to a known ground truth. Ideally, we could also perform a similar analysis on empirical data; however, ground truth is difficult to ascertain for such applications. Hence, we can only assess the consistency of MBM and SBM in analyzing empirical data, provided that we have multiple runs of comparable experiments. Here, we evaluated consistency with respect to two different group comparisons and empirical datasets, as described below.

#### 2.4.1. Sex differences

The first empirical validation focused on sex differences in CT maps from HCP [50]. In particular, we analyzed FreeSurfer-derived CT maps [55] from 339 unrelated healthy young adults (ages 22 to 35; 182 females; no siblings). This subsample corresponds to unrelated subjects from the HCP S900 release. The CT maps were spatially normalized to the fsLR-32k CIFTI surface using FreeSurfer [56, 57] and Connectome Workbench [54].

For each experiment, we subsampled 25 subjects from each group of females and males. We then computed the $t$-map and $\beta$ spectrum for each experiment. We repeated this process 100 times, representing 100 experiments. We then compared the distributions of the correlations between $t$-maps and $\beta$ spectra obtained for each pair of experiments, as in the consistency analysis using simulated data.

### 2.4.2. Patient-control differences

The second empirical validation focused on patient vs control differences in CT maps from the Open Access Series of Imaging Studies (OASIS-3) [58]. We analyzed FreeSurfer-derived CT maps from 693 healthy individuals and individuals with Alzheimer's disease (ages 42 to 95; 407 females). For subjects with multiple sessions, we only analyzed data from their last scan. The OASIS-3 data were acquired from one imaging center with three different 3T scanners and one 1.5T scanner. Hence, we treated data from each 3T scanner as a separate site and analyzed each site independently (we do not use data from the 1.5T scanner). The numbers of cognitively normal healthy controls (HC) and individuals with Alzheimer's disease (AD) in each site are: site 1: 164 HC, 27 AD; site 2: 218 HC, 24 AD; and site 3: 142 HC, 118 AD. We excluded subjects with known history of active medicine-induced cognitive dysfunction, neurological diseases, seizure disorder, hydrocephalus, head injury, alcoholism, or Parkinson's disease to minimize confounds.

Each of the three sites thus represents an independent experimental replication. We therefore evaluated the consistency of the findings obtained at each site using correlations between $t$-maps and $\beta$ spectra between each pair of sites, as per the other analyses.

### 2.5. Analyzing the frequency content of group different maps

We used MBM to examine the frequency content of group differences in CT, e.g., the dominant scale of the group differences, in two ways. First, we assigned eigenmodes to distinct eigengroups, as outlined in Section 2.2, and quantified the fraction of significant modes in each group. Second, we evaluated the accuracy of mode-based reconstructions of the empirical CT difference map following incremental, sequential removal of modes according to spatial wavelength. Specifically, we removed modes starting from long-wavelength to short-wavelength modes (from the first mode to mode $i$) or starting from short-wavelength to long-wavelength modes (from the last mode to mode $i$), and evaluated the correlation between the empirical CT difference
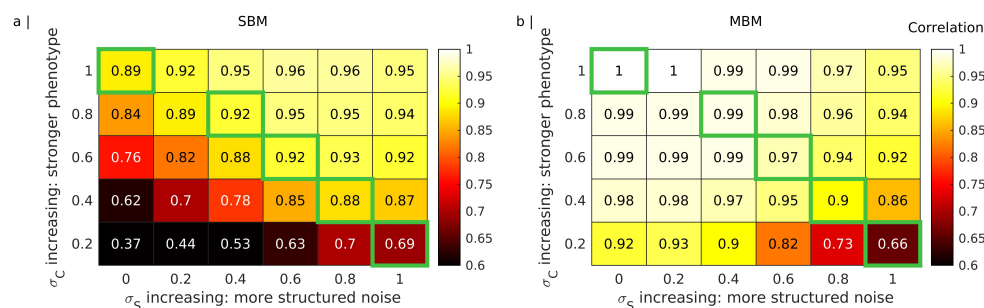
Figure 4: Accuracy of SBM and MBM with respect to ground-truth simulations for different combinations of $\sigma_C$ and $\sigma_S$. (a) Mean correlation between the $t$-map of an experiment and the ground-truth difference map. (b) Mean correlation between the $\beta$ spectra of the $t$-map of an experiment and the ground-truth difference map. The green boxes highlight the realistic parameter regimes where the generated maps have a similar spatial structure as the empirical data, as shown in Fig. 3.

map and the map reconstructed using the remaining set of modes obtained after each removal.

## 3. Results

### 3.1. Simulated data

Figure 4a shows the mean correlations between $t$-maps and the ground-truth CT difference map obtained with SBM, hereafter referred to as SBM correlations. As expected, the SBM correlation increases as $\sigma_C$ increases, which reflects a stronger contribution from the ground-truth CT pattern to the simulated CT map. Figure 4b shows a similar behavior for the mean correlations between the $\beta$ spectra of $t$-maps and the ground-truth difference map obtained with MBM, hereafter referred to as MBM correlations. Critically, MBM has better performance for nearly all parameter combinations, including in the realistic parameter regimes denoted by the green boxes. Differences between SBM and MBM were particularly salient in regimes characterized by a weak phenotype contribution and high levels of Gaussian noise (i.e., low $\sigma_C$ and low $\sigma_S$). This effect likely reflects the low-pass spatial filtering effect of the MBM reconstruction, which renders it more robust to high-frequency noise than SBM. Thus, across a wide range of signal-to-noise ratios, MBM more accurately captures the ground-truth difference than SBM.
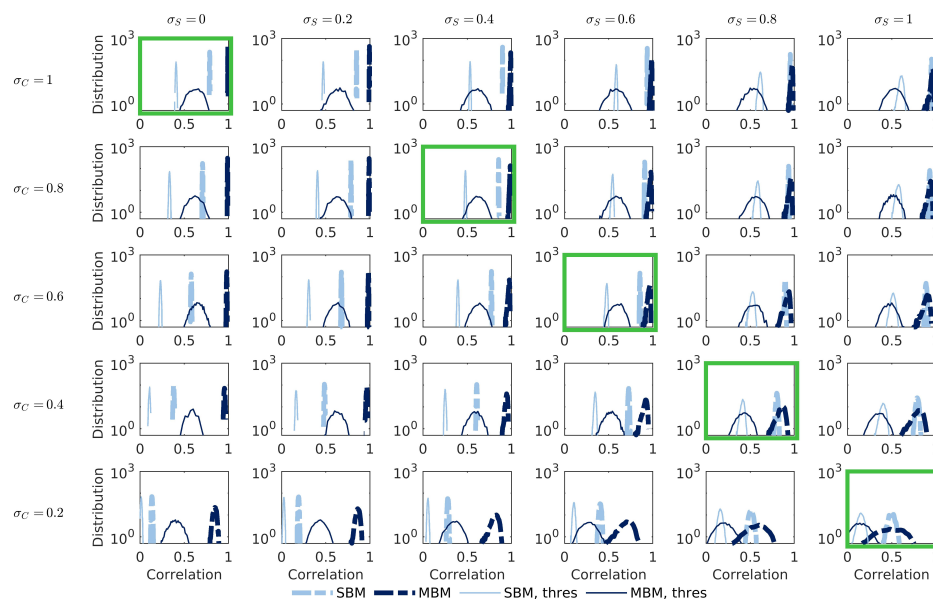
15

Figure 5: Distributions (in log scale) of correlations between pairs of 100 experiments for different combinations of $\sigma_C$ and $\sigma_S$. The panels show correlations between pairs of experimental $t$-maps (SBM), correlations between pairs of experimental $\beta$ spectra (MBM), binary correlations between thresholded $t$-maps (SBM, thres), and binary correlations between statistically significant $\beta$ spectra (MBM, thres). The green boxes highlight the realistic parameter regimes where the generated maps have a similar spatial structure as the empirical data, as shown in Fig. 3.

Figure 5 shows the results of our consistency analysis quantified in terms of the distributions of the SBM and MBM correlations, for both unthresholded and thresholded results, between each pair of 100 different experiments relying on the same ground truth. When $\sigma_C$ increases, the distributions shift to the right and become narrower. Thus, when the contribution of the ground-truth difference is more strongly expressed in individual CT maps, both SBM and MBM become more consistent. Across most parameter combinations, including those in the realistic regimes, the distribution of MBM correlations shifts to the right of the SBM correlations for both unthresholded and thresholded results, indicating that MBM generally yields more consistent findings. The exception to this rule is in cases of a weak phenotype contribution and high levels of structured noise (i.e., low $\sigma_C$ and high $\sigma_S$), where the peaks of the SBM and MBM distributions converge but the latter show a wider variation. This result arises because MBM is particularly sen-
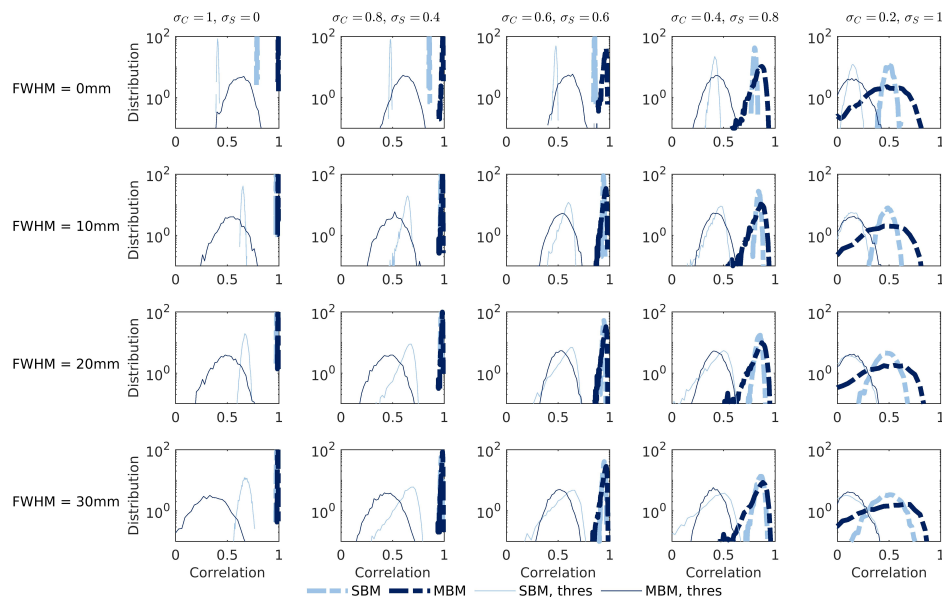
16

Figure 6: Distribution (in log scale) of pairwise correlations between experiments in the realistic parameter regimes and for different smoothing kernels. The panels show correlations for experimental $t$-maps (SBM), correlations for experimental $\beta$ spectra (MBM), binary correlations for thresholded $t$-maps (SBM, thres), and binary correlations for statistically significant $\beta$ spectra (MBM, thres).

sitive to structured spatial patterns extending over long wavelengths. MBM will thus have difficulty in reliably detecting a ground-truth difference in the presence of a high degree of structured noise such as, for instance, when individual variability swamps consistent patterns observed across different individuals belonging to the same group.

To further investigate this spatial filtering effect of MBM, we repeated the analyses after smoothing the CT data using surface-based smoothing kernels with FWHM of $10, 20$, and $30$ mm. Figure 6 shows the distributions of SBM and MBM correlations obtained for 5 pairs of $\sigma_C$ and $\sigma_S$ in realistic regimes, i.e., $(\sigma_C, \sigma_S) = \{(1, 0), (0.8, 0.4), (0.6, 0.6), (0.4, 0.8), (0.2, 1)\}$. When the smoothing kernel size increases, the SBM correlation distributions shift to the right and become broader, suggesting that smoothing improves SBM consistency on average, but large deviations from this average are also more common. On the other hand, the MBM correlation distributions are relatively unaffected by spatial smoothing. In general, smoothing only improves
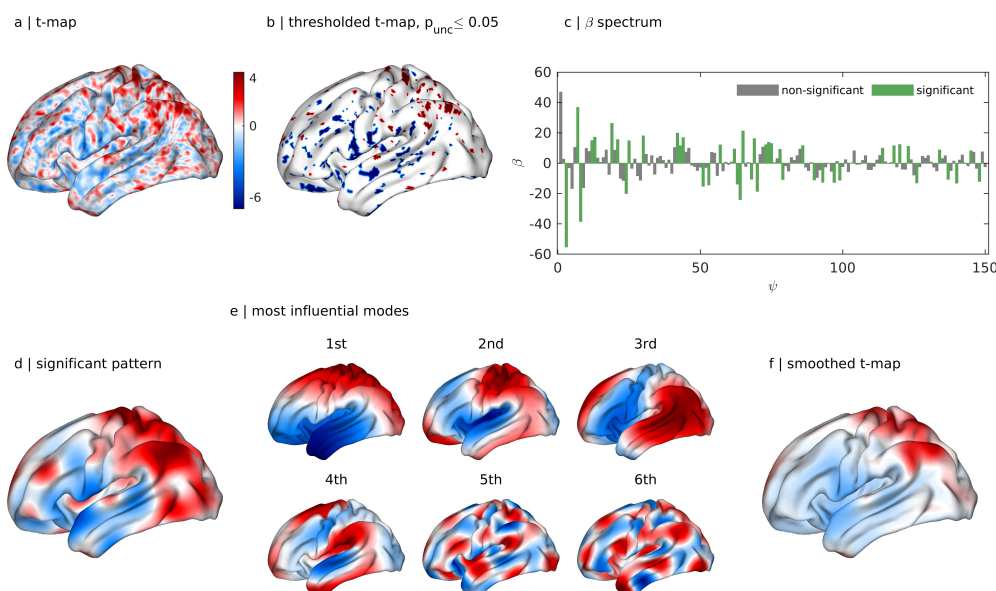
17

Figure 7: SBM and MBM analyses of CT differences between sexes. (a) Unthresholded $t$-map. (b) Thresholded $t$-map at $p_{unc} \leqslant 0.05$. Red and blue denote significantly thicker CT in females and males, respectively. (c) $\beta$ spectrum of the unthresholded $t$-map. The $\beta$'s of the significant modes, obtained via permutation testing at $p_{unc} \leqslant 0.05$, are colored green. (d) Significant pattern obtained by combining the significant modes weighted by their $\beta$'s. (e) Six most influential modes, i.e., significant modes with largest absolute $\beta$ values. The signs of modes with negative $\beta$'s were flipped to better visualize the similarity between the modes and the significant patterns. The number denotes the order of influence, not the mode index. (f) Smoothed $t$-map at FWHM = 30 mm.

the consistency of SBM to a level commensurate with MBM applied to un-smoothed data. This result implies that MBM offers similar advantages to smoothing without having to choose an arbitrary kernel size and shape.

### 3.2. Empirical results

Having demonstrated that MBM yields more accurate and consistent results than SBM in simulated data, we now evaluate the performance of the two methods in actual empirical data. To reiterate, since there is no clear ground truth in such cases, we primarily focus on evaluating the consistency, under sampling variability, of results obtained with both methods.
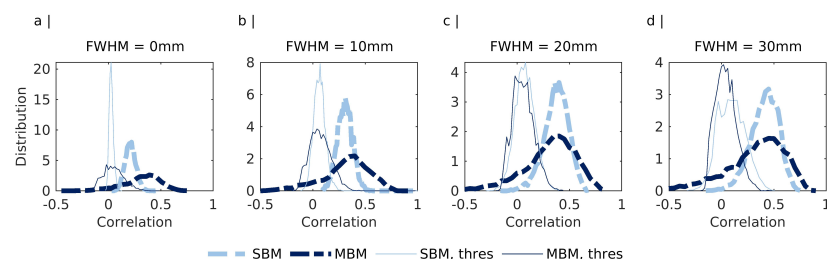
18

Figure 8: Distribution of pairwise correlations between experiments for different smoothing kernels in the empirical analysis of sex differences under resampling. The panels show correlations for experimental $t$-maps (SBM), correlations for experimental $\beta$ spectra (MBM), binary correlations for thresholded $t$-maps (SBM, thres), and binary correlations for statistically significant $\beta$ spectra (MBM, thres).

### 3.2.1. Sex differences

Figure 7 shows the SBM and MBM analyses of CT differences between sexes. Figures 7a and b show the $t$-map and thresholded $t$-map in a typical SBM. Figure 7c shows the $\beta$ spectrum obtained from MBM. The $\beta$ spectrum describes the influence of multiscale patterns, i.e., modes, on the CT differences. As shown in Fig. 7c, CT differences between sexes are influenced more by long-wavelength, i.e., low-frequency, modes than short-wavelength, i.e., high-frequency, modes. For example, $\beta$s of modes 1 to 50 have higher absolute amplitudes compared to those of modes 100 to 150. The significant modes with their $\beta$s (shown in green in Fig. 7c) are combined to show the significant pattern of the CT difference between sexes in Fig. 7d. The six most influential modes, i.e., significant modes with largest absolute $\beta$ values, are shown in Fig. 7e. Figure 7f shows the smoothed $t$-map at FWHM=30 mm, which is different from Fig. 7d, demonstrating that the complex pattern of CT difference extracted by MBM cannot be trivially obtained by simply smoothing the $t$-map.

Figure 8a shows that MBM yields more consistent results than SBM when considering sex differences in the HCP data under resampling. Figures 8b-d indicate that applying progressively larger smoothing kernels leads to a convergence in the consistency of both SBM and MBM, but SBM never displays clearly superior performance. These results align with our simulation findings. In fact, the wide distribution tails observed with MBM, particularly at higher smoothing kernels, suggest that the empirical sex difference maps studied here are characterized by a relatively weak true phenotype contribu-
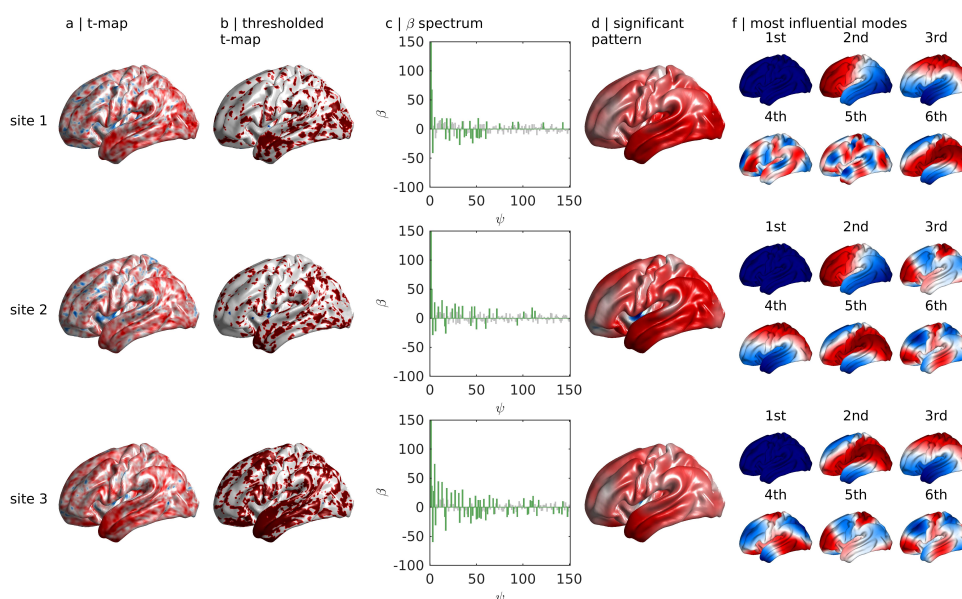
19

Figure 9: MBM analyses of CT differences between patients with Alzheimer's disease and healthy controls across three sites. (a) Unthresholded $t$-maps of the three sites. Negative–zero–positive values are colored as blue–white–red, with positive values indicating reduced thickness in patients. (b) Thresholded $t$-map at $p_{\mathrm{unc}} \leqslant 0.05$. Red and blue denote significantly thinner CT in patients and controls, respectively. (c) $\beta$ spectrum of the unthresholded $t$-map. The $\beta$'s of the significant modes, obtained via permutation testing at $p_{\mathrm{unc}} \leqslant 0.05$, are colored green. (d) Significant pattern obtained by combining the significant modes weighted by their $\beta$s. (e) Six most influential modes, i.e., significant modes with largest absolute $\beta$ values. The signs of modes with negative $\beta$'s were flipped to better visualize the similarity between the modes and the significant patterns. The number denotes the order of influence, not the mode index.

tion (i.e., a reliable, characteristic sex difference) and a strong contribution of structured noise; in other words, the data appear to correspond to a low $\sigma_C$, high $\sigma_S$ regime.

### 3.2.2. Alzheimer's disease

Figures 9a and b show the $t$-maps and thresholded $t$-maps of CT differences between patients with Alzheimer's disease and healthy controls across three sites. Figures 9c, d, and e show the results obtained from MBM, including the $\beta$ spectra, reconstructions using the significant patterns of the CT difference, and the six most influential modes. The reconstructions using significant modes show a consistent pattern of differences across the sites that is
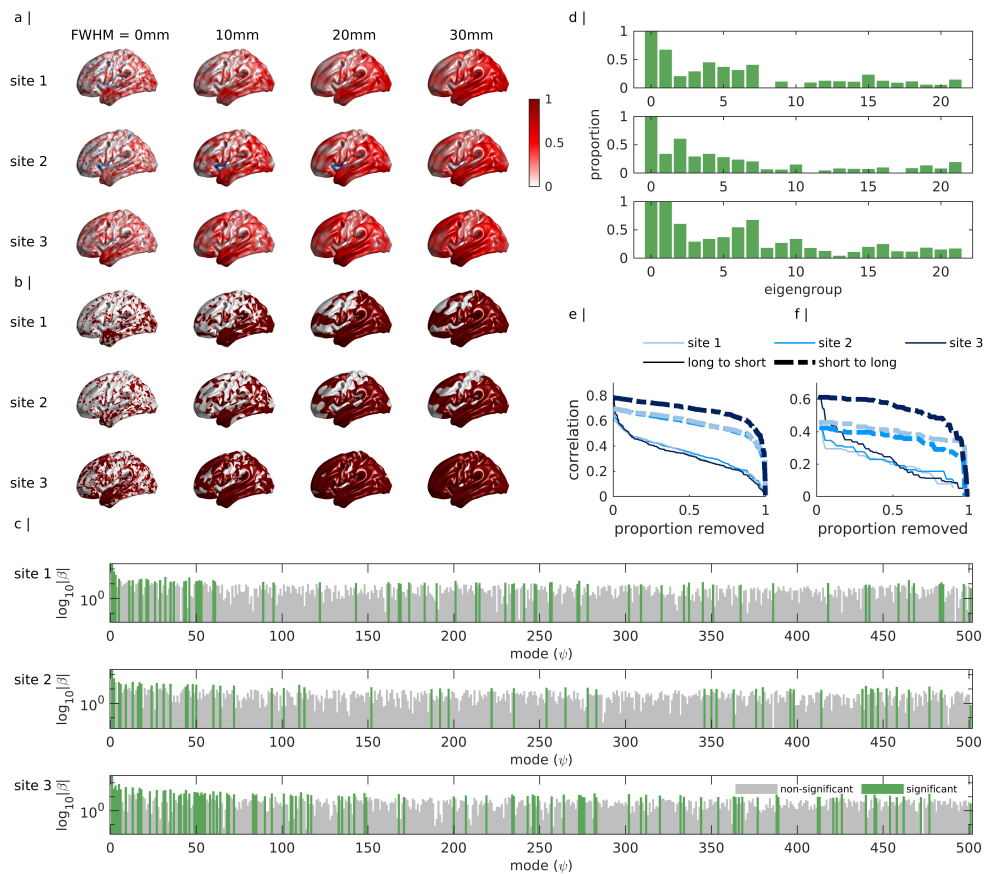
20

Figure 10: Comparing the performance of SBM and MBM analyses of CT differences between patients with Alzheimer's disease and healthy controls across three sites. (a) Unthresholded $t$-maps of the three sites in the Alzheimer's study with FWHM $= 0, 10, 20, 30$ mm. Negative–zero–positive values are colored as blue–white–red, with positive values indicating reduced thickness in patients. (b) Thresholded $t$-maps ($p_{\mathrm{unc}} \leqslant 0.05$) of the three sites with FWHM $= 0, 10, 20, 30$ mm. Red and blue denote significantly thinner CT in patients and controls, respectively. (c) Absolute values of $\beta$ spectra of the three sites without smoothing in log scale. Green and gray bars show significant and not significant $\beta$s, respectively. (d) The proportion of significant modes in each approximate eigengroup. (e) Correlation between the empirical $t$-map and its mode-derived reconstruction obtained using the full $\beta$ spectrum, after removing modes in order of decreasing or increasing spatial wavelength. (f) Correlation between the empirical $t$-map and its mode-derived reconstruction obtained using only significant modes from the full $\beta$ spectrum, after removing modes in order of decreasing or increasing spatial wavelength.

not directly apparent when inspecting the thresholded $t$-maps obtained with SBM. Accordingly, the six most influential modes show consistency across the datasets and all sites include the first mode, implying consistent global differences in CT between cases and controls.

Figures 10a and b show the effect of spatial smoothing on the unthresholded and thresholded $t$-maps obtained for each of the three sites with FWHM $= 0, 10, 20, 30$ mm, respectively (see Section 2.4.2). Figure 10c shows the absolute values of $\beta$ spectra of the unsmoothed data for each of the three sites obtained with MBM. Here, we consider $N = 500$ modes to observe the results across a broad range of spatial scales ranging from 18 mm and larger and to more comprehensively evaluate the spatial frequency content of the data, as outlined below. The statistically significant $\beta$'s are shown by the green bars. For SBM, smoothing generally increases the consistency of the spatial maps observed across sites, such that nearly the entire brain shows a CT reduction in patients compared to controls at $p_{\mathrm{unc}} < 0.05$ (Fig. 10b). The MBM $\beta$ spectra also offer qualitative evidence for consistency across sites, with similar modes being identified as significant, particularly at lower spatial frequencies. Figure 10d shows the proportion of significant modes in each eigengroup. These proportions are high in coarse-scale, long-wavelength eigengroups, particularly groups 0 to 7, corresponding to modes 1 to 64 and wavelengths $\geqslant 51$ mm. Moreover, the first, global mode is also always significant, which is not observed in the analysis of sex differences (Fig. 7). Thus, while Alzheimer's disease is associated with global reductions in CT, sex differences tend to be more focal.

To evaluate the spatial frequency content of the difference maps in more detail, Figs. 10e and f show how correlations between the empirical and reconstructed $t$-maps are affected by the removal of a proportion of eigenmodes from the reconstruction (see Section 2.5). The reconstruction uses modes from the full $\beta$ spectrum in Fig. 10e and from the $\beta$ spectrum of significant modes in Fig. 10f. The correlation declines more rapidly when removing long-wavelength modes first, indicating that they make a dominant contribution to CT differences. These findings indicate that CT differences in Alzheimer's disease are preferentially expressed over coarse scales spanning nearly the entire brain. Such broad patterns will be missed by classical analyses that focus only on point-wise inferences.

Cross-site consistency is more directly quantified in Fig. 11. Here, we considered $N = 150$ modes to be consistent with the other analyses in the paper and also because the group differences are primarily expressed over
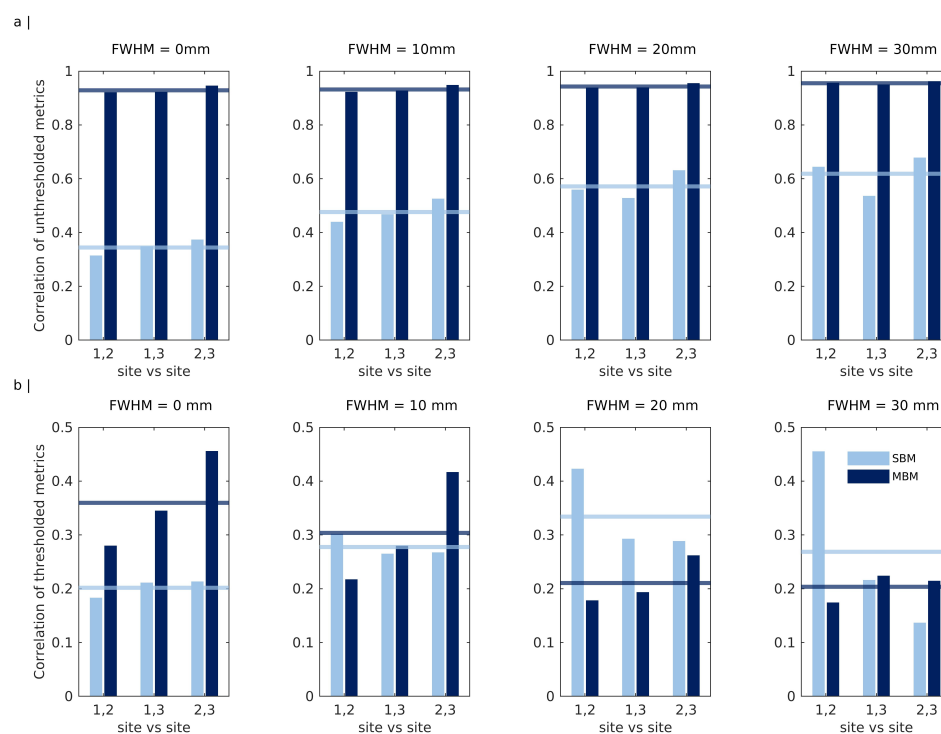
Figure 11: Consistency of SBM and MBM results in explaining multi-site CT differences between patients with Alzheimer's disease and healthy controls at different smoothing kernels. (a) Pairwise correlations between sites for unthresholded results. The light and dark blue horizontal lines represent the mean correlation of the unthresholded $t$-maps (for SBM) and of $\beta$ spectra (for MBM), respectively. (b) Pairwise binary correlations between sites for thresholded results. The light and dark blue horizontal lines represent the mean binary correlation of the thresholded $t$-maps (for SBM) and of $\beta$ spectra (for MBM), respectively.

coarse scales, as observed in Fig. 10d. Figure 11a shows the pairwise site correlations for unthresholded $t$-maps and $\beta$ spectra, while Fig. 11b shows the binary pairwise site correlations for the thresholded $t$-maps and $\beta$ spectra. Analysis of the unthresholded maps reveals that MBM provides highly consistent results, which are unaffected by spatial smoothing. Spatial smoothing improves the performance of SBM, but it never reaches the level of MBM.

Considering the thresholded results, smoothing improves the performance of SBM and appears to be optimal at FWHM = 20 mm, where the average pairwise site correlations of SBM are clearly superior to MBM. However, this average SBM correlation at FWHM = 20 mm does not surpass the

23

average MBM correlation on unsmoothed data (i.e., Fig. 11b, FWHM = 0), indicating that MBM generally yields more consistent conclusions about group CT differences in Alzheimer's disease than classical SBM. Together, these findings support the reliability, parsimony, and simplicity of MBM given that (i) it yields more consistent inferences; (ii) it can adequately summarize the effects of interest with $\sim$150 values (i.e., the $\beta$ spectrum) rather than the $>$30,000 vertices per hemisphere typically used in SBM; and (iii) it can be applied directly to unsmoothed data, obviating the need to select a particular smoothing kernel size and shape.

## 4. Discussion

In this paper, we developed a method for mapping neuroanatomical group differences at multiple spatial scales using a basis set of eigenmodes derived from cortical geometry, called MBM. Our analysis shows that, when compared with classical SBM, MBM shows comparable or superior performance in terms of both accuracy—i.e., capturing ground-truth differences in simulated data—and consistency—i.e., yielding consistent findings under sampling variability in both simulated and empirical scenarios. MBM also enables a spectral decomposition of the group differences, offering insights into the spatial scales at which those differences are most salient. In contrast, SBM is restricted to a single scale that is defined by the vertex mesh resolution and smoothing kernel applied. MBM does not require a choice of a specific smoothing kernel. Furthermore, MBM offers the added advantage of directly linking group differences to an underlying generative process, where the differences are modeled as resulting from the differential involvement of distinct resonant modes of brain structure.

### 4.1. MBM yields more accurate and consistent inferences about ground-truth simulations

We compared the accuracy and consistency of MBM and SBM with respect to simulated and empirical data. Our simulations offered insight into the relative performance of the two approaches under varying ratios of phenotype and (structured and unstructured) noise. This framework allowed us to evaluate the accuracy with which SBM and MBM can uncover ground-truth differences. Our analysis showed that MBM was more accurate than SBM for most simulation parameter values. The performance of MBM was

particularly strong in parameter regimes associated with high levels of unstructured (i.e., Gaussian) noise, indicating that it is likely to perform better in noisy data due to the low-pass spatial filtering effect of the eigenmode reconstructions.

When structured noise dominated the ground-truth difference, MBM and SBM showed comparable mean accuracy, but MBM correlations were more variable across experiments, indicating less consistency. To the extent that structured noise in our model emulates subject-specific neuroanatomical features, this finding indicates that MBM will perform inconsistently in cases where the contribution of such features (i.e., inter-individual variability) match or swamp the ground-truth phenotype that is common to all subjects in the same group. This is because MBM is particularly sensitive to detecting structured (i.e., autocorrelated) patterns in the spatial maps. However, group mean comparisons of CT (or other neuroanatomical properties) are not likely to be very meaningful in such scenarios, as the small contribution of the ground-truth difference indicates that there is little in common between different people assigned to the same group.

Spatially smoothing the data improved the performance of SBM to the point where it matched MBM for smoothing kernels with FWHM $\geqslant 20$ mm. However, there are no consistent heuristics for choosing a proper smoothing kernel. In addition, smoothing kernels should be chosen carefully since they can impose geometric effects on the data as the kernel replaces the value at each point with a weighted average of its spatial neighbors [32]. A particular advantage of MBM is that it does not require the selection of a specific smoothing kernel size.

### 4.2. MBM yields more consistent results in empirical data

It was not possible to evaluate the accuracy of MBM and SBM in empirical data due to the lack of ground truth, but we could evaluate the consistency of the findings across multiple repetitions of the same experiment. For both the analyses of differences between sexes and between healthy controls and patients with Alzheimer's disease, MBM was more consistent than SBM at low levels of smoothing, with the two converging at higher levels of smoothing.

In the case of sex differences, the distributions of MBM correlations were generally wider than those of SBM correlations. When interpreted with respect to our simulations, this result suggests that while there may be some consistent sex differences in CT that can be observed under resampling, they

25

are likely subtle relative to the effects of individual variability. This inter-
pretation aligns with ongoing debates over the consistency of sex differences
in neuroanatomy [3, 18, 19].

Alzheimer's disease is likely to be associated with a more salient and
robust CT phenotype, given the well-described stages of atrophy that are
known to occur in the condition [59, 60, 61]. Accordingly, both unthresh-
olded and thresholded analyses were more consistent for MBM compared to
SBM. Increasing smoothing kernel size improved the performance of SBM,
but the average consistency never surpassed that observed for MBM in the
unsmoothed data. These findings support the utility of applying MBM to
minimally smoothed data.

### 4.3. MBM offers insights into scale-dependent group differences

Structural eigenmodes are ordered by spatial wavelength, opening the
opportunity to analyze the spatial frequency content of CT differences, much
like a Fourier decomposition is routinely used to examine spectral properties
of EEG signals. We presented an analysis of the frequency content of CT
differences in Alzheimer's disease, showing that the proportion of significant
modes was higher in eigengroups with long wavelengths, with most of the
differences found for modes with wavelengths $\geqslant$51 mm. The first, global
mode was consistently significant across the three sites, indicating a robust
global difference in CT between cases and controls. Reconstruction accuracy
also declined more rapidly when removing long-wavelength modes, indicating
that they make a dominant contribution to CT differences between groups.
Together, these findings suggest that CT differences in Alzheimer's disease
are most salient at coarse spatial scales that are not adequately captured by
classical point-wise analysis, such as SBM, where a specific resolution scale
is imposed by the mesh resolution and smoothing kernel size. In contrast,
MBM provides a natural way of characterizing group differences across a
wide range of spatial scales.

A further consideration is that statistical inference in SBM often assumes
that each point-wise location (e.g., surface vertex) is independent. This as-
sumption is incorrect, since CT and many other neuroanatomical properties
are spatially autocorrelated; i.e., the value at one point depends on oth-
ers. The violation of this assumption is worsened when spatial smoothing
is applied to the data. Dependencies are sometimes later considered if some
form of cluster-based thresholding is applied [62], but this is an ad hoc char-
acterization. In MBM, distinct spatial locations are not considered to be

independent but instead form part of brain-wide modes with varying spatial wavelengths. Critically, since the eigenmodes themselves are orthogonal, mode-specific inference is entirely justified.

### 4.4. MBM ties neuroanatomical differences to a generative process

Classical approaches to mapping neuroanatomical differences are purely phenomenological, relying on statistical analyses to identify differences between groups without offering a direct explanation for the mechanisms through which those differences have emerged. An important advantage of MBM is that the results can be linked to a direct physical interpretation, in which group differences in neuroanatomy are explained as the involvement of different resonant modes of brain anatomy. An intuitive analogy can be drawn from plucking a violin string [31]. The eigenmodes correspond to the string's harmonics, each of which is associated with its own preferred vibration frequency. The musical note generated by plucking the string results from a superposition of these modes. This basic idea has been used to understand how the structure of a system constrains its dynamics in diverse areas of physics and engineering, including the electromagnetic response of different media [20], the vibrational patterns of different structures [21], and aeroelasticity [22].

In the cortex, these resonant modes define the principal axes of structural variation and thus represent a fundamental basis set for understanding anatomical constraints on any spatially patterned process. For instance, the second, third, and fourth eigenmodes considered here correspond to spatial variations along the anterior-posterior, dorso-ventral, and medial-lateral axes, which are known to define many fundamental properties of cortical organization, such as regional variations in cell density [63], CT [64], genetic influences on neuroanatomy [65], and gene expression gradients that shape brain development [66, 67].

An important choice in such analyses concerns the neuroanatomical properties that should be used to define the eigenmodes. Two distinct approaches have emerged in the literature. One approach involves deriving eigenmodes from a discrete, graph-based model of the connectome, under the assumption that inter-regional connectivity represents the primary anatomical constraint on dynamical processes in the brain [23]. The other approach involves deriving eigenmodes from a model of the geometry of the cortex, as used here. This approach follows from a specific form of neural field theory, a well-validated

biophysical model of brain dynamics that characterizes how dynamical processes spatially propagate as waves through a continuous cortical medium [46]. Recent work indicates that geometric modes offer a more parsimonious account of diverse aspects of brain function than connectome eigenmodes [15]. Geometric eigenmodes also offer a practical advantage, as they can be extracted from T1-weighted MRI data alone using standard procedures, whereas connectome eigenmodes rely on the application of complex preprocessing pipelines to diffusion MRI data, which require many choices that can affect the final results [68, 69, 70]. For these reasons, we have focused on geometric modes in our analysis, but the approach developed here is sufficiently general that it can be used with any anatomical basis set.

## 5. Limitations and Conclusions

We derived the geometric eigenmodes using a population-average template surface, which does not completely account for individual differences in brain shape. While low-frequency modes tend to be consistent between people [41], the spatial patterns of high-frequency modes tend to diverge due to individual differences in neuroanatomy. This divergence makes it difficult to compare results across individuals. Past work suggests that modes derived from a population-average template can reconstruct brain function to a degree that is comparable to individual-specific modes [15], but further work is required to develop techniques that can better capture individual differences in eigenmode architecture. Future work could also incorporate volumetric analyses in subcortical regions to enable whole-brain inferences, given the strong coupling between geometry and function found in these areas [15].

In summary, we have introduced here a new multiscale approach, which we call mode-based morphometry (MBM), for mapping neuroanatomical differences between groups where the differences are modeled as arising from the involvement of distinct, resonant modes of brain anatomy. Using both simulated and empirical data, we show that MBM offers more accurate and consistent inferences than classical approaches (i.e., surface-based morphometry), while also providing insights into the spatial frequency content of the differences and allowing a direct link to putative generative physical processes.

28

## Data and code availability

Raw and preprocessed HCP data can be accessed at https://db.humanconnectome.org/. Raw and preprocessed OASIS-3 data can be accessed at https://www.oasis-brains.org/. An open-source toolbox implementing MBM will be available at https://github.com/BMHLab/MBM upon publication of the article. Code and sample data to reproduce the analysis results and figures of this study will be openly available at https://github.com/BMHLab/MBM_paper upon publication of the article.

## Acknowledgments

## References

[1] J. Ashburner, K. J. Friston, Voxel-based morphometry–the methods, NeuroImage 11 (2000) 805–821.

[2] B. Fischl, A. M. Dale, Measuring the thickness of the human cerebral cortex from magnetic resonance images, Proceedings of the National Academy of Sciences 97 (20) (2000) 11050–11055.

[3] A. R. DeCasien, E. Guma, S. Liu, A. Raznahan, Sex differences in the human brain: a roadmap for more careful analysis and interpretation of a biological reality, Biology of Sex Differences 13 (1) (Jul. 2022).

[4] S. Ducharme, M. D. Albaugh, T.-V. Nguyen, J. J. Hudziak, J. M. Mateos-Pérez, A. Labbe, A. C. Evans, S. Karama, Trajectories of cortical thickness maturation in normal brain development — The importance of quality control procedures, NeuroImage 125 (Jan. 2016).

[5] H. Kiesow, L. Q. Uddin, B. C. Bernhardt, J. Kable, D. Bzdok, Dissecting the midlife crisis: disentangling social, personality and demographic determinants in social brain anatomy, Communications Biology 4 (1) (Jun. 2021).

[6] E. Smith, A. Thurm, D. Greenstein, C. Farmer, S. Swedo, J. Giedd, A. Raznahan, Cortical thickness change in autism during early childhood, Human Brain Mapping 37 (7) (Jul. 2016).

[7] for the ENIGMA-Major Depressive Disorder Working Group, L. Schmaal, D. P. Hibar, P. G. Sämann, G. B. Hall, B. T. Baune, N. Jahanshad, J. W. Cheung, T. G. M. van Erp, D. Bos, M. A. Ikram, M. W. Vernooij, W. J. Niessen, H. Tiemeier, A. Hofman, K. Wittfeld, H. J. Grabe, D. Janowitz, R. Bülow, M. Selonke, H. Völzke, D. Grotegerd, U. Dannlowski, V. Arolt, N. Opel, W. Heindel, H. Kugel, D. Hoehn, M. Czisch, B. Couvy-Duchesne, M. E. Rentería, L. T. Strike, M. J. Wright, N. T. Mills, G. I. de Zubicaray, K. L. McMahon, S. E. Medland, N. G. Martin, N. A. Gillespie, R. Goya-Maldonado, O. Gruber, B. Krämer, S. N. Hatton, J. Lagopoulos, I. B. Hickie, T. Frodl, A. Carballedo, E. M. Frey, L. S. van Velzen, B. W. J. H. Penninx, M.-J. van Tol, N. J. van der Wee, C. G. Davey, B. J. Harrison, B. Mwangi, B. Cao, J. C. Soares, I. M. Veer, H. Walter, D. Schoepf, B. Zurowski, C. Konrad, E. Schramm, C. Normann, K. Schnell, M. D. Sacchet, I. H. Gotlib, G. M. MacQueen, B. R. Godlewska, T. Nickson, A. M. McIntosh, M. Papmeyer, H. C. Whalley, J. Hall, J. E. Sussmann, M. Li, M. Walter, L. Aftanas, I. Brack, N. A. Bokhan, P. M. Thompson, D. J. Veltman, Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group, Molecular Psychiatry 22 (6) (2017) 900–909.

[8] Y. Zhao, Q. Zhang, C. Shah, Q. Li, J. A. Sweeney, F. Li, Q. Gong, Cortical Thickness Abnormalities at Different Stages of the Illness Course

in Schizophrenia: A Systematic Review and Meta-analysis, JAMA Psychiatry 79 (6) (2022) 560.

[9] G. Douaud, S. Lee, F. Alfaro-Almagro, C. Arthofer, C. Wang, P. McCarthy, F. Lange, J. L. R. Andersson, L. Griffanti, E. Duff, S. Jbabdi, B. Taschler, P. Keating, A. M. Winkler, R. Collins, P. M. Matthews, N. Allen, K. L. Miller, T. E. Nichols, S. M. Smith, SARS-CoV-2 is associated with changes in brain structure in UK Biobank, Nature 604 (7907) (2022) 697–707.

[10] S. Marek, B. Tervo-Clemmens, F. J. Calabro, D. F. Montez, B. P. Kay, A. S. Hatoum, M. R. Donohue, W. Foran, R. L. Miller, T. J. Hendrickson, S. M. Malone, S. Kandala, E. Feczko, O. Miranda-Dominguez, A. M. Graham, E. A. Earl, A. J. Perrone, M. Cordova, O. Doyle, L. A. Moore, G. M. Conan, J. Uriarte, K. Snider, B. J. Lynch, J. C. Wilgenbusch, T. Pengo, A. Tam, J. Chen, D. J. Newbold, A. Zheng, N. A. Seider, A. N. Van, A. Metoki, R. J. Chauvin, T. O. Laumann, D. J. Greene, S. E. Petersen, H. Garavan, W. K. Thompson, T. E. Nichols, B. T. T. Yeo, D. M. Barch, B. Luna, D. A. Fair, N. U. F. Dosenbach, Reproducible brain-wide association studies require thousands of individuals, Nature 603 (7902) (2022) 654–660.

[11] J. Ashburner, K. J. Friston, Why Voxel-Based Morphometry Should Be Used, NeuroImage 14 (6) (2001) 1238–1243.

[12] M. Jenkinson, C. F. Beckmann, T. E. J. Behrens, M. W. Woolrich, S. M. Smith, FSL, NeuroImage 62 (2012) 782–790.

[13] B. Fischl, M. I. Sereno, R. B. Tootell, A. M. Dale, High-resolution intersubject averaging and a coordinate system for the cortical surface, Human Brain Mapping 8 (4) (1999) 272–284.

[14] R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, M. S. Albert, R. J. Killiany, An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest, NeuroImage 31 (3) (2006) 968–980.

[15] J. C. Pang, K. M. Aquino, M. Oldehinkel, P. A. Robinson, B. D. Fulcher, M. Breakspear, A. Fornito, Geometric constraints on human brain function, bioRxiv (Oct. 2022).

[16] B. C. Dickerson, A. Bakkour, D. H. Salat, E. Feczko, J. Pacheco, D. N. Greve, F. Grodstein, C. I. Wright, D. Blacker, H. D. Rosas, R. A. Sperling, A. Atri, J. H. Growdon, B. T. Hyman, J. C. Morris, B. Fischl, R. L. Buckner, The Cortical Signature of Alzheimer's Disease: Regionally Specific Cortical Thinning Relates to Symptom Severity in Very Mild to Mild AD Dementia and is Detectable in Asymptomatic Amyloid-Positive Individuals, Cerebral Cortex 19 (2009) 497–510.

[17] C. G. Schwarz, J. L. Gunter, H. J. Wiste, S. A. Przybelski, S. D. Weigand, C. P. Ward, M. L. Senjem, P. Vemuri, M. E. Murray, D. W. Dickson, J. E. Parisi, K. Kantarci, M. W. Weiner, R. C. Petersen, C. R. Jack, A large-scale comparison of cortical thickness and volume methods for measuring Alzheimer's disease severity, NeuroImage: Clinical 11 (2016) 802–812.

[18] C. M. Williams, H. Peyre, R. Toro, F. Ramus, Sex differences in the brain are not reduced to differences in body size, Neuroscience & Biobehavioral Reviews 130 (2021) 509–511.

[19] L. Eliot, A. Ahmed, H. Khan, J. Patel, Dump the "dimorphism": Comprehensive synthesis of human brain studies reveals few male-female differences beyond size, Neuroscience & Biobehavioral Reviews 125 (2021) 667–697.

[20] D. B. Melrose, R. C. McPhedran, Electromagnetic Processes in Dispersive Media, Cambridge University Press, USA, 1991.

[21] C. Beards, Structural Vibration: Analysis and Damping, Elsevier Science, 1996.

[22] C. A. Felippa, K. Park, C. Farhat, Partitioned analysis of coupled mechanical systems, Computer Methods in Applied Mechanics and Engineering 190 (24-25) (2001) 3247–3270.

[23] S. Atasoy, I. Donnelly, J. Pearson, Human brain networks function in connectome-specific harmonic waves, Nature Communications 7 (1) (2016) 10340.

[24] J. Rué-Queralt, K. Glomb, D. Pascucci, S. Tourbier, M. Carboni, S. Vulliémoz, G. Plomp, P. Hagmann, The connectome spectrum as

a canonical basis for a sparse representation of fast brain activity, NeuroImage 244 (Dec. 2021).

[25] S. Naze, T. Proix, S. Atasoy, J. R. Kozloski, Robustness of connectome harmonics to local gray matter and long-range white matter connectivity changes, NeuroImage 224 (2021) 117364.

[26] Jennifer A. Cummings, Benjamin Sipes, Daniel H. Mathalon, Ashish Raj, Predicting Functional Connectivity From Observed and Latent Structural Connectivity via Eigenvalue Mapping, Frontiers in Neuroscience 16 (Mar. 2022).

[27] H. Behjat, A. Tarun, D. Abramian, M. Larsson, D. V. D. Ville, Voxel-Wise Brain Graphs from Diffusion MRI: Intrinsic Eigenspace Dimensionality and Application to fMRI (Sep. 2022).

[28] S. Ghosh, A. Raj, S. S. Nagarajan, Estimating Brain Functional Connectivity from Common Subspace Mapping Between Structural and Functional Connectomes (Apr. 2022).

[29] N. C. Gabay, T. Babaie-Janvier, P. A. Robinson, Dynamics of cortical activity eigenmodes including standing, traveling, and rotating waves, Physical Review E 98 (4) (2018) 042413.

[30] K. N. Mukta, P. A. Robinson, J. C. Pagès, N. C. Gabay, X. Gao, Evoked response activity eigenmode analysis in a convoluted cortex via neural field theory, Physical Review E 102 (6) (2020) 062303.

[31] P. A. Robinson, J. A. Henderson, N. C. Gabay, K. M. Aquino, T. Babaie-Janvier, X. Gao, Determination of Dynamic Brain Connectivity via Spectral Analysis, Frontiers in Human Neuroscience 15 (2021).

[32] J. A. Henderson, K. M. Aquino, P. A. Robinson, Empirical estimation of the eigenmodes of macroscale cortical dynamics: Reconciling neural field eigenmodes and resting-state networks, Neuroimage: Reports 2 (3) (2022) 100103.

[33] M. G. Preti, D. Van De Ville, Decoupling of brain function from structure reveals regional behavioral specialization in humans, Nature Communications 10 (1) (2019) 4747.

[34] A. Griffa, E. Amico, R. Liégeois, D. Van De Ville, M. G. Preti, Brain structure-function coupling provides signatures for task decoding and individual fingerprinting, NeuroImage 250 (2022) 118970.

[35] Z.-Q. Liu, B. Vázquez-Rodríguez, R. N. Spreng, R. N. Spreng, B. C. Bernhardt, R. F. Betzel, B. Misic, Time-resolved structure-function coupling in brain networks, Communications biology 5 (1) (Jun. 2022).

[36] F. Abdelnour, S. Mueller, A. Raj, Relating Cortical Atrophy in Temporal Lobe Epilepsy with Graph Diffusion-Based Network Models, PLOS Computational Biology 11 (10) (2015) e1004564.

[37] F. Abdelnour, A. Raj, O. Devinsky, T. Thesen, Network Analysis on Predicting Mean Diffusivity Change at Group Level in Temporal Lobe Epilepsy, Brain Connectivity 6 (8) (2016) 607–620.

[38] F. Abdelnour, M. Dayan, O. Devinsky, T. Thesen, A. Raj, Algebraic relationship between the structural network's Laplacian and functional network's adjacency matrix is preserved in temporal lobe epilepsy subjects, NeuroImage 228 (2021) 117705.

[39] G. Orrù, P. Maurel, J. Coloigner, Structural And Functional Interplay In Anxiety Related Classification: A Graph Signal Processing Approach, in: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), 2021, pp. 271–274.

[40] M. B. Wang, J. P. Owen, P. Mukherjee, A. Raj, Brain network eigenmodes provide a robust and compact representation of the structural connectome in health and disease, PLOS Computational Biology 13 (6) (2017) e1005550.

[41] Y.-C. Chen, A. Arnatkevičiūtė, E. McTavish, J. C. Pang, S. Chopra, C. Suo, A. Fornito, K. M. Aquino, The individuality of shape asymmetries of the human cerebral cortex, eLife 11 (Oct. 2022).

[42] R. A. Poldrack, A. Mumford, T. E. Nichols, Handbook of Functional MRI Data Analysis, Cambridge University Press, USA, 2011.

[43] A. M. Winkler, G. R. Ridgway, G. Douaud, T. E. Nichols, S. M. Smith, Faster permutation inference in brain imaging, NeuroImage 141 (2016) 502–516.

[44] I. Chavel, Eigenvalues in Riemannian Geometry, Academic Press, USA, 1984.

[45] M. Reuter, F.-E. Wolter, M. Shenton, M. Niethammer, Laplace–Beltrami eigenvalues and topological features of eigenfunctions for statistical shape analysis, Computer-Aided Design 41 (10) (2009) 739–755.

[46] P. Robinson, X. Zhao, K. Aquino, J. Griffiths, S. Sarkar, G. Mehta-Pandejee, Eigenmodes of brain activity: Neural field theory predictions and comparison with experiment, NeuroImage 142 (2016) 79–98.

[47] M. Reuter, F.-E. Wolter, N. Peinecke, Laplace–Beltrami spectra as 'Shape-DNA' of surfaces and solids, Computer-Aided Design 38 (4) (2006) 342–366.

[48] C. Wachinger, P. Golland, W. Kremen, B. Fischl, M. Reuter, Alzheimer's Disease Neuroimaging Initiative, BrainPrint: a discriminative characterization of brain morphology, NeuroImage 109 (2015) 232–248.

[49] ThomasYeoLab, fs_lr_32k template[Online]. Available: `https://github.com/ThomasYeoLab/CBIG/tree/master/data/templates/surface/fs_LR_32k`.

[50] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. J. Behrens, E. Yacoub, K. Ugurbil, The WU-Minn Human Connectome Project: An overview, NeuroImage 80 (2013) 62–79.

[51] J. B. Burt, M. Helmer, M. Shinn, A. Anticevic, J. D. Murray, Generative modeling of brain maps with spatial autocorrelation, NeuroImage 220 (2020) 117038.

[52] G. U. Yule, On the Methods of Measuring Association Between Two Attributes, Journal of the Royal Statistical Society 75 (6) (1912) 579–652.

[53] A. Arnatkeviciūtė, B. D. Fulcher, R. Pocock, A. Fornito, Hub connectivity, neuronal diversity, and gene expression in the Caenorhabditis elegans connectome, PLOS Computational Biology 14 (2) (2018) e1005989.

[54] D. S. Marcus, M. P. Harms, A. Z. Snyder, M. Jenkinson, J. A. Wilson, M. F. Glasser, D. M. Barch, K. A. Archie, G. C. Burgess, M. Ramaratnam, M. Hodge, W. Horton, R. Herrick, T. Olsen, M. McKay, M. House, M. Hileman, E. Reid, J. Harwell, T. Coalson, J. Schindler, J. S. Elam, S. W. Curtiss, D. C. Van Essen, Human Connectome Project informatics: Quality control, database services, and data visualization, NeuroImage 80 (2013) 202–219.

[55] M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, D. C. Van Essen, M. Jenkinson, The Minimal Preprocessing Pipelines for the Human Connectome Project, NeuroImage 80 (2013) 105–124.

[56] A. M. Dale, B. Fischl, M. I. Sereno, Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction, NeuroImage 9 (2) (1999) 179–194.

[57] B. Fischl, M. I. Sereno, A. M. Dale, Cortical Surface-Based Analysis: II: Inflation, Flattening, and a Surface-Based Coordinate System, NeuroImage 9 (2) (Feb. 1999).

[58] P. J. LaMontagne, T. L. Benzinger, J. C. Morris, S. Keefe, R. Hornbeck, C. Xiong, E. Grant, J. Hassenstab, K. Moulder, A. G. Vlassenko, M. E. Raichle, C. Cruchaga, D. Marcus, Oasis-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and alzheimer disease (2019). doi:10.1101/2019.12.13.19014902.

[59] L. Pini, M. Pievani, M. Bocchetta, D. Altomare, P. Bosco, E. Cavedo, S. Galluzzi, M. Marizzoni, G. B. Frisoni, Brain atrophy in Alzheimer's Disease and aging, Ageing Research Reviews 30 (2016) 25–48.

[60] G. B. Frisoni, N. C. Fox, C. R. Jack, P. Scheltens, P. M. Thompson, The clinical use of structural MRI in Alzheimer disease, Nature Reviews. Neurology 6 (2) (2010) 67–77.

[61] P. M. Thompson, K. M. Hayashi, G. de Zubicaray, A. L. Janke, S. E. Rose, J. Semple, D. Herman, M. S. Hong, S. S. Dittmer, D. M. Doddrell, A. W. Toga, Dynamics of gray matter loss in Alzheimer's disease, The Journal of Neuroscience: The Official Journal of the Society for Neuroscience 23 (3) (2003) 994–1005.

[62] S. M. Smith, T. E. Nichols, Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference, NeuroImage 44 (1) (2009) 83–98.

[63] D. J. Cahalane, C. J. Charvet, B. L. Finlay, Systematic, balancing gradients in neuron density and number across the primate isocortex, Frontiers in Neuroanatomy 6 (2012) 28.

[64] P. Rakic, Evolution of the neocortex: a perspective from developmental biology, Nature Reviews Neuroscience 10 (10) (2009) 724–735.

[65] C.-H. Chen, M. Fiecas, E. D. Gutiérrez, M. S. Panizzon, L. T. Eyler, E. Vuoksimaa, W. K. Thompson, C. Fennema-Notestine, D. J. Hagler, T. L. Jernigan, M. C. Neale, C. E. Franz, M. J. Lyons, B. Fischl, M. T. Tsuang, A. M. Dale, W. S. Kremen, Genetic topography of brain morphology, Proceedings of the National Academy of Sciences 110 (42) (2013) 17089–17094.

[66] B. K. Ip, I. Wappler, H. Peters, S. Lindsay, G. J. Clowry, N. Bayatti, Investigating gradients of gene expression involved in early human cortical development, Journal of Anatomy 217 (4) (2010) 300–311.

[67] A. Fornito, A. Arnatkevičiūtė, B. D. Fulcher, Bridging the Gap between Connectome and Transcriptome, Trends in Cognitive Sciences 23 (1) (2019) 34–50.

[68] S. Oldham, A. Arnatkeviciūtė, R. E. Smith, J. Tiego, M. A. Bellgrove, A. Fornito, The efficacy of different preprocessing steps in reducing motion-related confounds in diffusion MRI connectomics, NeuroImage 222 (2020) 117252.

[69] M. Bastiani, M. Cottaar, S. P. Fitzgibbon, S. Suri, F. Alfaro-Almagro, S. N. Sotiropoulos, S. Jbabdi, J. L. R. Andersson, Automated quality control for within and between studies diffusion MRI data using a non-parametric framework for movement and distortion correction, NeuroImage 184 (2019) 801–812.

[70] M. Gajwani, S. J. Oldham, J. C. Pang, A. Arnatkevičiūtė, J. Tiego, M. A. Bellgrove, A. Fornito, Can hubs of the human connectome be identified consistently with diffusion MRI?, Tech. rep. (Dec. 2022).

[71] W. J. Goscinski, P. McIntosh, U. Felzmann, A. Maksimenko, C. Hall, T. Gureyev, D. Thompson, A. Janke, G. Galloway, N. Killeen, P. Raniga, O. Kaluza, A. Ng, G. Poudel, D. Barnes, T. Nguyen, P. Bonnington, G. Egan, The multi-modal Australian ScienceS Imaging and Visualization Environment (MASSIVE) high performance computing infrastructure: applications in neuroscience and neuroinformatics research, Frontiers in Neuroinformatics 8 (2014).