

UTRECHT UNIVERSITY  
Department of Experimental Psychology

---

**Artificial Intelligence Master Thesis**

**Modelling Neural Organization in the Human Visual System**

**First examiner:**

Ben Harvey

**Candidate:**

Stan Bergey

**Second examiner:**

Surya Gayet

July 23, 2024

## **Abstract**

# Contents

|                               |   |           |
|-------------------------------|---|-----------|
| <b>1</b>                      | <b>Introduction</b>   | <b>3</b>  |
| <b>2</b>                      | <b>Method</b>   | <b>14</b> |
| 2.1                           | Dataset . . . . .   | 14        |
| 2.2                           | Masking the Visual System . . . . .                                       | 14        |
| 2.3                           | Selecting and Preprocessing the betas . . . . .                           | 15        |
| 2.4                           | Creating the Representational Sampling Spaces . . . . .                   | 17        |
| 2.5                           | Rotating the MDS spaces . . . . .   | 19        |
| 2.6                           | Fit the Gaussian models . . . . .   | 20        |
| 2.7                           | Creating a model of the Visual System . . . . .                           | 24        |
| 2.8                           | Voxel Performance . . . . .   | 25        |
| 2.9                           | Correlations between MDS spaces and Cortical Surface topography . . . . . | 26        |
| 2.10                          | Full Brain Model . . . . .  | 28        |
| <b>3</b>                      | <b>Results</b>  | <b>29</b> |
| 3.1                           | Roated MDS . . . . .  | 29        |
| 3.2                           | Best ROI and model creation . . . . .                                     | 30        |
| 3.3                           | Goodness of fit . . . . .   | 31        |
| 3.4                           | Voxel Performance . . . . .   | 32        |
| 3.5                           | $\sigma$ . . . . .  | 32        |
| 3.6                           | correlation . . . . .   | 32        |
| 3.7                           | Full brain . . . . .  | 32        |
| <b>4</b>                      | <b>Conclusion</b>   | <b>37</b> |
| 4.1                           | Discussion . . . . .  | 37        |
| <b>Appendix</b>               |   |           |
| <b>A</b>                      | <b>Appendix title</b>   | <b>38</b> |
| <b>A</b>                      | <b>Appendix A</b>   | <b>39</b> |
| A.1                           | First appendix section . . . . .  | 39        |
| <b>Bibliography</b> <b>42</b> |   |           |

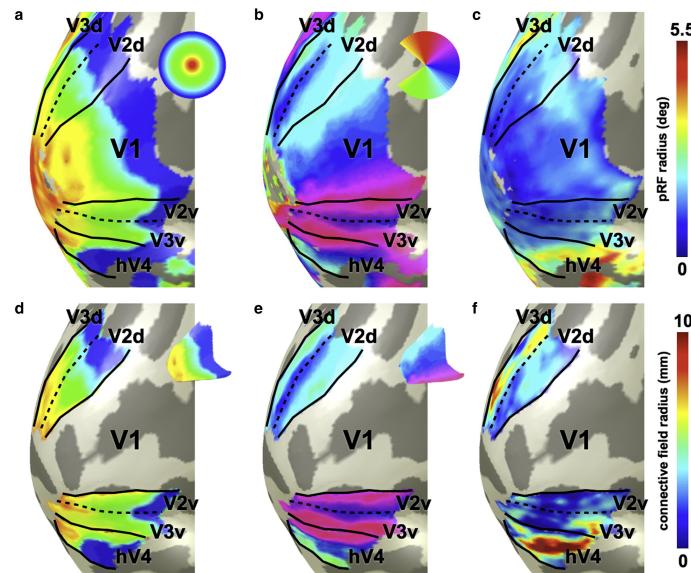
# 1. Introduction

One of the main objectives of neuroscience is to understand the representations of patterns of brain activity, how these patterns are correlated to different stimuli, and how to relate them to behavior (Kriegeskorte and Wei, 2021). The study of the relationship between the brain organs and behavior dates back to ancient times, with the first mention of the cerebral cortex being found in an Egyptian papyrus, dating back to at least 1700 BCE (Gross, 1999). Ancient Greek philosophers and scholars took an interest in the inside of the human skull as well, with writings about lesions and brain physiology being dated to pre-Hippocratic Greece (Gross, 1999; Rose, 1993). The early 20th century brought significant discoveries in neurophysiology Grant (2006), which combined with the improvement of brain recording technologies, led us to a better understanding of the brain's physiology and its functional organization (Marzi et al., 2009). Recent publications presented functional frameworks of the cortex and subcortex and identified cortical networks linked to cognitive functions (Ji et al., 2019; Power et al., 2011; Yeo et al., 2011). For example, different body sections have been associated with specific neural populations in the parietal area 5, indicating that the brain stores a model of our different body parts (Graziano and Botvinick, 2002). Motor functions have been identified in the posterior wall of the precentral gyrus (Amunts et al., 1997; Macdonell et al., 1999), and a combination of the prefrontal cortex (referred to as Broca's area) and the middle temporal, inferior temporal, fusiform, and angular gyri have been linked to the use of language (Binder et al., 1997; Friederici et al., 2017). Lastly, neural circuitry responsible for vision in humans and primates has been identified in the occipital lobe, with supportive networks in the temporal and parietal lobes( Wandell et al., 2009; Schall et al., 1995).

The study of brain activity saw a revolution with the development of functional Magnetic Imaging in the 1990s. This development kickstarted a significant increase in interest in neuroimaging research (Engel et al., 1994; Biswal et al., 1995; Bullmore and Sporns, 2009; Dosenbach et al., 2010). fMRI has been popular due to its ability to detect and define brain activity non-invasively. fMRI can detect changes in blood oxygenation (BOLD signal), which is correlated with neural activity (Kwong et al., 1992; Logothetis et al., 2001), although this has been debated (Heeger and Ress, 2002). Nevertheless, it is assumed that local neural activity leads to a change in local hemodynamics, which is captured by fMRI scanners. This change, and thus the neural activity of a neural population, is then translated to voxel activity. A voxel represents a local group of neurons, and the higher the magnetic field, the higher the voxel's resolution (Matthews and Jezzard, 2004). Each voxel

is associated with a time series of  $t$  times points, allowing us to detect changes in activity in each voxel over time (Smith, 2004). To determine which voxels are activated by a stimulus, we can use either correlation analysis or modeling of the data. A popular method is General Linear Modelling (GLM), where a linear model is initiated based on expected voxel activity after stimulation. If there's a good fit between the voxel data and the model, we can then conclude with some confidence that the changes in voxel activity were caused by the stimuli (Smith, 2004). These analyses have allowed us to study the functional structure of the brain non-invasively (Power et al., 2011), as well as the neural organization and activation patterns of specific regions, such as the motor cortex (Biswal et al., 1995) or the visual system (Engel et al., 1994).

Neural activations in the visual system originate from stimulation of the retinal neural population. Their neural activity then travels down the optic nerves toward the central ganglia, ending their course in the occipital poles (Catani et al., 2003). Visual information is managed contralaterally, meaning the left visual field for both eyes is processed by the right hemisphere and vice versa. Each position of the visual field can be represented by its eccentricity, which is the distance of that position to the focal point (Daniel and Whitteridge, 1961). The neural response to different visual field positions can then be modelled and used to represent neural activity on the cortical surface. Population Receptive Field (pRF) modelling is a method to estimate the activity of a neural population. Each neural population's pRF describes the location of the visual field to which it responds best based on a wide range of visual stimuli.



**Figure 1.1:** Stimuli and Neural-referred maps on the posterior medial surface of the occipital lobe of the left cerebral hemisphere. a and b show voxels' eccentricity and polar angle tunings using pRF modelling. d and e show the same tunings in V2, V3 and hV4 but derived from V1's best-fitting connective field models. From Haak et al. (2013).

---

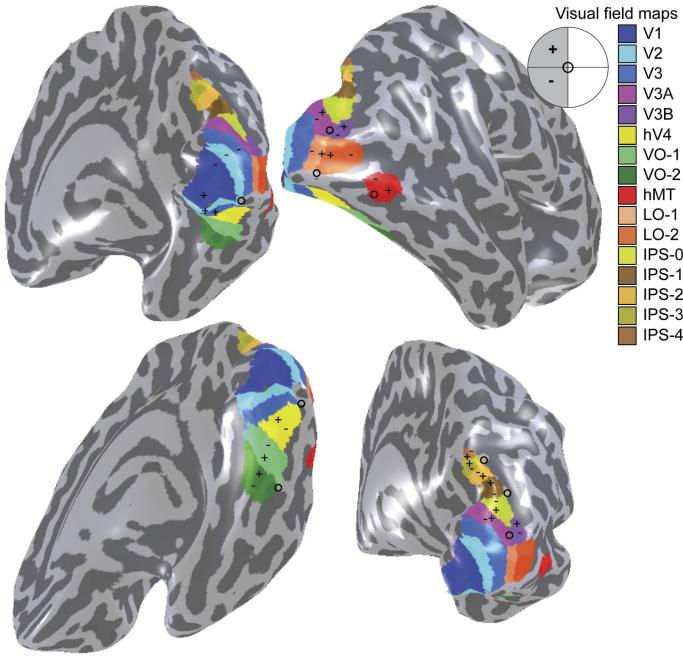
To achieve this, a pRF model estimates the position and the size of the area a voxel responds to. These responses are modelled through a Gaussian function, and the estimates of neural activity are obtained by finding the parameters of this Gaussian model that best predict the observation. pRF modelling aims to optimise the parameters  $x_0$ ,  $y_0$ , and  $s$ , where  $x_0$  and  $y_0$  are the coordinates of the centre of the pRF in the visual space, and  $s$  the spread, or size, of the receptive field. This is done by minimising the residual sum of squares between the prediction obtained through our models and the fMRI data (Dumoulin and Wandell, 2008). Using pRF modelling, we can display which voxel responds best to each visual field position. Furthermore, we can eccentricity and map these responses onto the cortical surface, as shown in Figure 1.1. Doing so reveals that different groups of neurons are tuned to various eccentricities and polar angles and indicates that populations with similar tunings are physically close to each other. In essence, a pRF determines the preferred spatial location of a neural population, which is described by the parameters of a Gaussian function. pRF models can estimate these parameters by generating predictions of neural activity and comparing these to the response in the fMRI data.

Similar modelling methodologies of using voxel's responses to stimuli to find the optimal parameters of Gaussian functions have been used before to model voxels' responses to timing and numerosity (Harvey et al., 2013; Harvey et al., 2020). Harvey et al. (2013) investigated voxel's responses to numerosity stimuli. Similar to the pRF modelling method, they fitted a Gaussian function for each voxel in the numerosity space: numerosity is one parameter of the Gaussian function, and tuning range is the other parameter. With these models, they found that some neural populations were tuned to small numerosities and that numerosity is organised topographically, forming ordered numerosity maps on the cortical surface. Harvey et al. (2020) utilised a similar methodology and investigated the neural responses to event duration and frequency. They fitted several models and found an anisotropic Gaussian function as the best predictor of observed neural activity. This Gaussian function takes the event's preferred duration and period as parameters, and the other parameters represent the responses to a range of timings. These models reveal timing preference in neural populations organised in maps containing timing-selective responses. These two studies reveal much about how timing and numerosity responses are organised in the cortical surface and show that these two sensory phenomena are well integrated into the brain's architecture. However, these studies also highlight a limitation of pRF-like modelling methods: the stimuli utilised were continuous and easily parameterised. Timing can be represented as a function of time, and numerosity as a discrete scale of consecutive numbers. But not all stimuli can be so easily represented as a function; much of the information we process in our everyday life cannot easily be functionalised and studied through pRF models. For example, quantifying the difference between natural scenes is non-trivial. We do not have an easy way to represent as a function the differences between scenes showing a face, a full human body, or some inanimate object, and such differences cannot be repre-

sented as a continuous variable. Therefore, in this project, we will propose an alternative to the pRF model that utilises decoding modelling to cover this limitation and allow us to model the neural responses to non-continuous high-dimension, hard-to-parameterised stimuli.

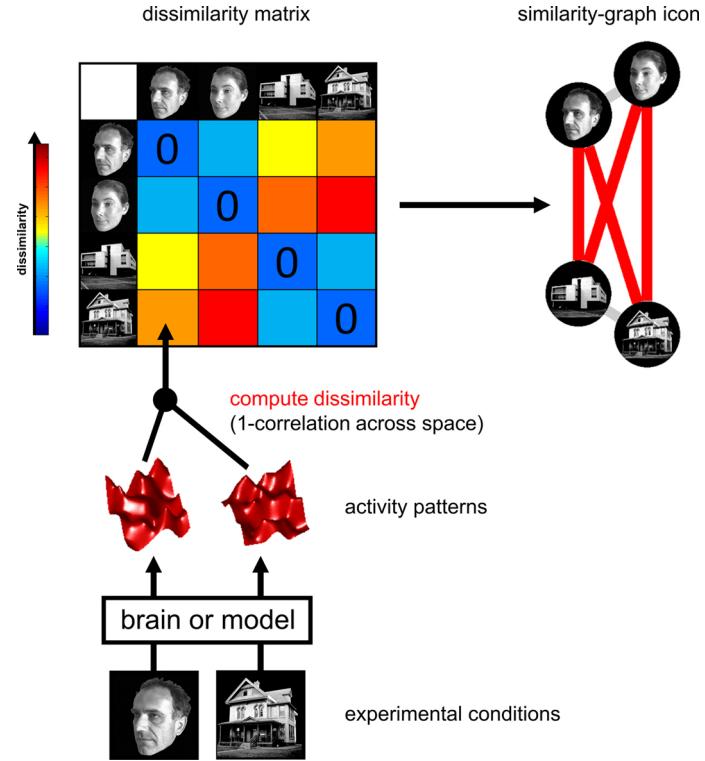
Nevertheless, the pRF method provides us with tools to set the boundaries between the different visual field maps of the visual system. As mentioned earlier, the human visual cortex fully spans over the occipital lobe and expands onto the temporal and parietal lobes. Each visual field map (also referred to as a Regions of Interest (ROI) in this study since they correspond to the areas we want to analyse) contains a complete representation of the visual field and transforms of the representation of its input, being either from previous visual field maps or the retinogeniculate pathway (Wandell et al., 2007). Using pRF modelling to draw boundaries, Wandell et al. (2007) propose a cortex subdivision in different visual field maps, illustrated in Figure 1.2. The first of these field maps is V1, which receives direct input from the retinal pathways and contains representations of the whole visual space. V1 is encircled by V2 and V3, which are subdivided along the horizontal meridian (Wandell et al., 2007). As illustrated in Figure 1.1 a, all these field maps represent visual field position in an organised way, with the central foveal responses near the occipital pole and peripheral visual field responses expanding onto the medial surface. Cells in these early visual field maps have been identified to respond to stereoscopic surfaces, edges, and orientation (von der Heydt et al., 2000). Then, the visual system spreads onto dorsal pathways and ventral pathways. Extending from dorsal V3, the Lateral Occipital (LO) maps, LO-1 and LO-2, have been identified to represent the contralateral visual hemifield with LO-1 mirroring V3's angle representation, LO-1 mirroring LO-1's. Additionally, these two maps have parallel eccentricity tuning, meaning their visual field representation differs slightly from those of V1, V2, and V3 (Wandell et al., 2007). Visual field maps have been identified in the temporal-occipital (TO) cortex, lateral to LO-2, and are seemingly tuned to visual motion (Amano et al., 2009). The ventral pathways consist of hV4, which shares previous ROIs' eccentricity representations and leads to the two maps in the ventral occipital (VO) cortex. VO-1 and VO-2 have been identified to respond to central visual position and are strongly tuned to colour and object stimuli (Wandell et al., 2007; Arcaro et al., 2009). Extending from VO-2 and spreading across the visual sulcus, we find the posterior parahippocampal cortex (PHC) maps, which are tuned to peripheral visual field locations (Arcaro et al., 2009). The study of fMRI data in the visual cortex allowed researchers to subdivide the visual systems into ROIs, each with different tuning preferences. Furthermore, each field map represents the whole visual image and its transformation, making these maps a good choice for our study since our goal is to model these transformations.

We previously established that pRF was limited to easily parameterised paradigms, such as numerosity or timing. As we argued, this limitation prevents us from analysing re-



**Figure 1.2:** Visual Field Maps in Human Visual Cortex. From Wandell et al. (2007).

sponses to stimuli that take place in a non-continuous, hard-to-parameterised space, such as natural scene images. To solve this issue, we can take advantage of decoding models. Decoding models can characterise an ROI's overall responses without considering the nature of the relationship between stimuli and neural activity, meaning that decoding models are not limited to continuous and easily parameterised stimuli. This is due to decoding models taking information from brain activity, unlike encoding models, which look at the characteristics of the stimuli. In other words, decoding models can derive the stimuli parameters by looking at the responses themselves, allowing us to study responses from non-continuous and hard-to-parameterised stimuli. Through decoding models, we can compare activity patterns and illustrate how different stimuli lead to different activity profiles, allowing us to make hypotheses about the distribution of these activity profiles (Diedrichsen and Kriegeskorte, 2017). For example, we can conjecture that the activity patterns caused by two similar stimuli will also be similar. By defining how the activity patterns caused by different conditions relate to each other, representational models can fully define the representational geometry of an ROI. From this, we can attempt to determine the neural code of a population thanks to the representation of neural activity given by our decoding model, although this relationship is quite ambiguous (Kriegeskorte and Wei, 2021; Diedrichsen and Kriegeskorte, 2017). Our goal is to not only study the representational geometry of neural activity in a non-continuous space but also to disambiguate the relationship between the representational geometry and the neural code of a neural population. In essence, decoding models allow us to study the neural representation of non-continuous stimuli, making it a valuable approach to cover pRF's modelling main limitation.



**Figure 1.3:** RSA FigureDiedrichsen and Kriegeskorte (2017).

In a multivariate scenario where fMRI data was obtained by measuring responses to different stimuli, we can go beyond activation-based analysis and study the representational content of our neural responses. An ROI might show the same involvement in different tasks when looking only at its overall activity, not allowing us to discriminate between conditions. However, multivariate pattern information in an ROI indicates the representational content of that neural population. Moreover, this approach does not consider the parameters of the stimuli, which allows us to study neural activity obtained in high-dimension, hard-to-parameterised experimental contexts, effectively covering pRF's limitations. Multivoxel Pattern Analysis (MVPA) provides us with methods to analyse these activity patterns by examining the neural representation of a neural population (or ROI) (Mur et al., 2009). In MVPA, the information in the activity patterns is assumed to contain the representational content of an ROI. (Mur et al., 2009). One way to study the representational content is by comparing activity patterns, which allows us to discriminate between conditions in a way that activity-based analysis cannot do. This can be done with Representational Similarity Analysis (RSA), an MVPA technique that quantitatively relates patterns of neural activity. Representing voxels in a high-dimension space, RSA portrays the activity patterns as positions in these dimensions, allowing for easy comparisons between these dimensions, as illustrated in Figure 1.3. (Kriegeskorte et al., 2008; Diedrichsen and Kriegeskorte, 2017). From this, we can measure the dissimilarities between each pattern representing each condition and draw condition-related hypotheses, as illustrated in

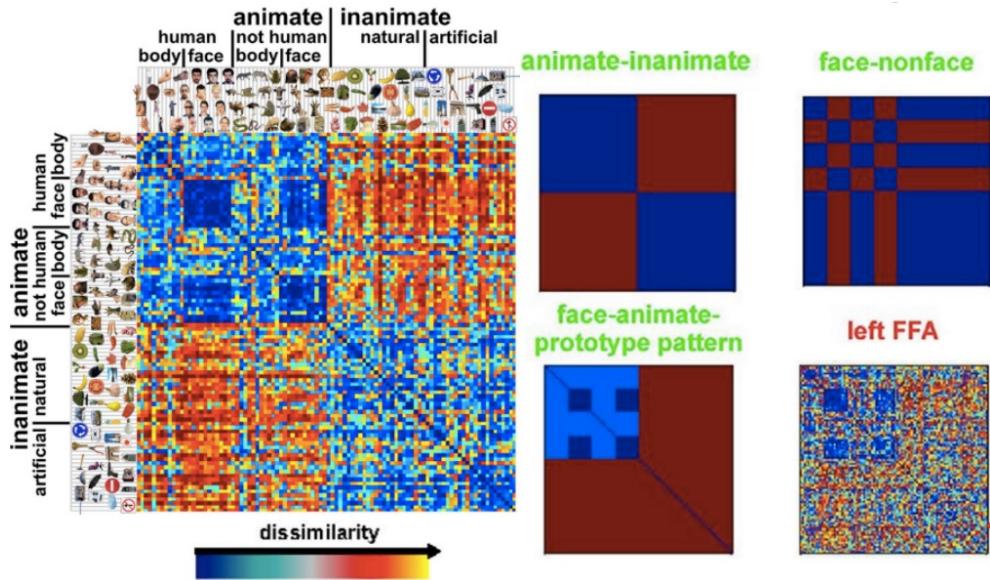
---

Figure 1.4 (Kriegeskorte et al., 2008). RSA has been used to reveal semantic categories in different regions of the brain (Devereux et al., 2013; Chen et al., 2016), examine cognitive control during a colour-word Stroop task (Freund et al., 2021) and discriminate object categories in EEG recording (Kaneshiro et al., 2015). In short, RSA allows us to investigate the representational content of an ROI by comparing the activity distributions between different conditions.

By comparing the dissimilarities between the activity patterns of each pair of conditions, we can obtain a Representational Dissimilarity Matrix (RDM) (see Figures 1.3 and 1.4). An RDM stores all pairwise dissimilarities obtained by taking 1 minus the correlation between the activity patterns associated with two conditions. These matrices can store an ROI’s geometrical representation, allowing for easy comparison between conditions. If a set of similar conditions leads to similar geometries and thus low dissimilarities, we can infer that the brain’s activity reflects these similarities (Kriegeskorte et al., 2008; Diedrichsen and Kriegeskorte, 2017). As seen in Figure 1.4, the responses to animate objects are quite similar to those to other animate objects and also for inanimate-inanimate pairs, while the responses between animate and inanimate stimuli are primarily different. RDMs have been used to identify biological categories in V1 (Connolly et al., 2012) or can be used to compare the neural representational between humans, primates, and artificial vision models (Khaligh-Razavi and Kriegeskorte, 2014). Furthermore, we can use Multidimensional Scaling (MDS) to simplify the pattern of response similarities for all stimuli. MDS is a general mathematical tool to arrange high-dimensional data into a low, 2D-dimensional space. Since each condition is represented as one dimension, RDMs are, by nature, high-dimension matrices. MDS allows us to project these dissimilarities into a low dimensions (usually 2) Cartesian Coordinates space. A point on this plane represents each condition, and the distance between them represents their dissimilarities (Douglas Carroll and Arabie, 1998). This allows us to project all conditions in a two-dimensional space, where conditions with similar activity patterns will be grouped together. Therefore, RSA allows us to extract the representational geometry of an ROI, resulting in an RDM, comparing each condition against all others. Then, thanks to MDS, we can simplify this result in a lower-dimension, more interpretable space.

We established that RSA allowed us to analyse neural activity gathered using high-dimension, non-parametric stimuli and gather the representational geometry of a neural population. However, unlike encoding modelling methods such as pRF modelling, defining the neural code through such representation is non-trivial. As explained previously, pRF modelling describes a neural population’s tuning to the visual space by finding the optimal parameters of a Gaussian function. Thus, pRF modelling straightforwardly expresses the neural code of a neural population through this Gaussian fitting. The non-

triviality of expressing the neural code through geometrical representation stems from the ambiguous relationship between them. While the neural code can determine the representational geometry, the inverse is not true (Kriegeskorte and Wei, 2021). Kriegeskorte and Wei (2021) illustrate this by stating that tuning functions after rotations will be a linear combination of the original tuning functions and, therefore, will lead to the same geometrical representation. Thus, we cannot infer that the tuning functions are based only on geometrical representations since different tuning functions can define the same representation. In other words, the representational geometry obtained through RSA cannot tell us how the information of a specific stimulus is distributed across neurons (Kriegeskorte and Wei, 2021). RSA only allow us to decode the distinction between several stimuli. Therefore, our goal for this project is to bridge the encoding and decoding model by attempting to define the neural code of the neural population in the visual system's ROIs while using RSA to parameterise the fitted functions. To achieve this, we will take inspiration from pRF modelling and use the MDS space's representation of our non-continuous conditions to find each voxel's optimal fitting parameters. The main difference with pRF modelling is that by tuning on the MDS space, we will be able to conclude if a voxel has a preferred position in the condition space instead of the visual field. From this, we will be able to infer if a voxel responds better to a specific condition or set of conditions. In essence, our goal is to combine both encoding and decoding modelling methods to cover each's limitations and determine the neural code of voxel in a high-dimension, hard-to-parameterize experimental context.



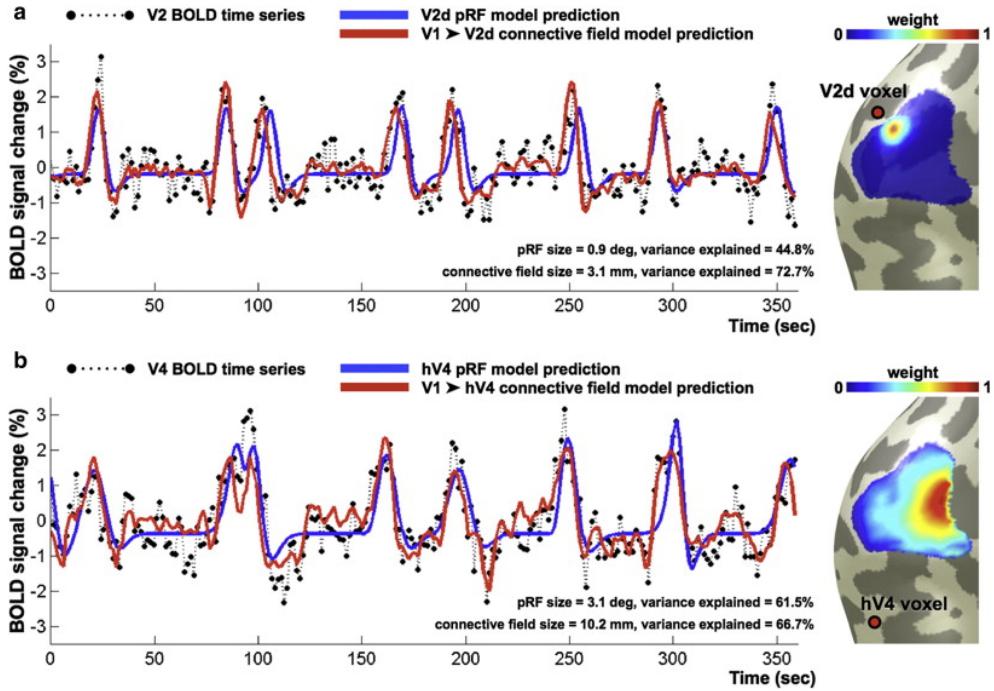
**Figure 1.4:** On the left, an RDM showing the dissimilarities between different conditions, separated in categories. The right shows hypotheses about the distribution of dissimilarities that can be tested against an RDM. Adapted from Kriegeskorte et al. (2008).

We have previously established that the visual system can be divided into contiguous visual field maps. While it is clear that the neural activity from these regions influences

---

each other, the functional organisation between the different ROIs of the visual system is still a subject of debate (Heinze et al., 2011). Therefore, we aim to explore this functional organisation between the visual system’s field maps. To do so, we can take advantage of the low-dimensional representation of neural activity offered by MDS. Our goal is to model the functional organisation of brain activity between ROIs by modelling each voxel’s response to all stimuli, using each ROI’s MDS space as a sampling space. In other words, we can obtain the neural code thanks to this representational geometry. To achieve this, we aim to describe each voxel in the visual system’s cortical surface response as an area within each field map’s MDS space. We will do this by modelling neural responses as a symmetric Gaussian function, which will find each voxel’s preferred position on the MDS space. This methodology is quite similar to pRF but does not sample from the visual space. Instead, we sample from the MDS space, meaning we will define each voxel’s response based on activity from other voxels. This brings our methodology closer to Connective Field (CF) modelling (Haak et al., 2013). CF is similar to pRF modelling in that it estimates the location and size a voxel responds to, but it does so by taking the activity of other regions in the brain. Thus, unlike in pRF, where a voxel’s receptive field predicts its activity based on stimulus position, a voxel’s connective field does that through a function of activity elsewhere in the brain. For example, CF allows us to explain the activity in V2 thanks to the activity in V1. This is illustrated in Figure 1.1d, where the polar angle and eccentricity of V2 are defined using V1’s visual field mapping stimuli (Haak et al., 2013). Once done, the models of each voxel can be compared to the neuroimaging data to see how they perform, as illustrated in Figure 1.5. Thus, Similarly to Haak et al. (2013), we want to model the activity of voxels based on other voxels’ activity. Here, our sampling source for each voxel will be each other region’s MDS space. So, for a voxel in V1, we will use MDS spaces from V1 and all other regions to find the optimal parameters of the Gaussian function we want to fit. Then, our goal is to use these models to find, for each ROI, which source ROI leads to the best predictions. For example, if voxels in V1 that were sampling from V2’s MDS space led to the best-predicting performance, then we will be able to say that V2 best explains V1. In essence, we want to explain the functional organisation of the visual system’s cortical surface by finding the best sampling source for each ROI.

As established, we want to model each voxel’s responses by finding their preferred position on every ROI’s MDS space. Each MDS space represents an ROI’s responses to all stimuli, and we can combine this decoding approach with the encoding approach of fitting a Gaussian function to the data to find our optimal parameters. The parameters of this Gaussian function should be the ones that best predict a voxel’s response to all stimuli. Therefore, we can measure how well each model explains the variance in neural activity and, for each ROI, which target ROI explains this variance the best. Our research question is then as follows:



**Figure 1.5:** Fit of pRF models (Blue) and CF models (Red) on BOLD time-series in V2 (a) and V4 (b). From Haak et al. (2013)

*RQ1: Can the neural activity of voxels in an ROI be predicted using a model built by fitting a Gaussian response function that samples the activity from other ROIs, based on a 2D spatial representation derived from RSA of the activity produced by a set of stimuli in each ROIs? Can these models teach us something about the functional organisation of neural activity in the visual system's cortical surface?*

However, fitting our models with low-dimensional cortical surface (MDS) representations requires us to question whether these representations accurately represent this cortical topology. In our 2D MDS space, two different stimuli are expected to be represented with two different coordinates. This is also seen in higher-dimensional spaces before MDS analysis, where two different stimuli lead to different representations. We observe that neural populations in the cortical surface respond preferentially to specific categories. The activation profiles between preferred and non-preferred categories are clearly distinguishable, with an obvious step between the two (Mur et al., 2012). Therefore, category preferences should appear in spatially contiguous neural populations (Vul et al., 2012). Due to these discrete boundaries and apparent differences in activities, the representational space seems discontinuous on the cortical surface. This should be reflected in MDS since each class is represented by a range of nearby coordinates. Therefore, the functional clustering of similarly responding neural populations, as shown in Harvey et al. (2013) and Harvey et al. (2020), on the cortical surface should be reflected in the MDS space. This should be clearly seen for classes with two different centres of mass of response in the cortical surface: the difference in their activity patterns should be reflected in the RDM and, thus, in the MDS.

---

Hence, we ask ourselves if the MDS's organisation reflects the cortical surface's topological organisation. This will be done by answering the second research question, being:

*RQ2: Are the low dimensionality MDS space accurate geometrical representations of the visual system's cortical surface's topological organisation?*

To answer these questions, we will use an open-source data set containing the neural activity of eight subjects, recorded while exposed to 9000-10,000 colored natural scenes (Allen et al., 2022). Thanks to being given the condition of each set of activities, we can organize this data in MDS, allowing us to apply our decoding modeling approach to it. Hopefully, these models will shed some light on the neural organization of the visual system and will be able to predict neural activity to some extent.

## 2. Method

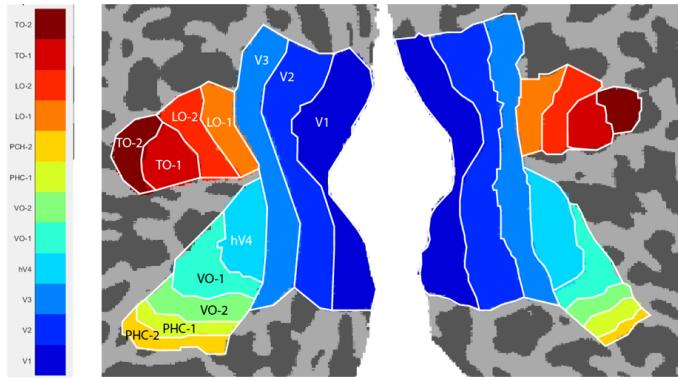
### 2.1 Dataset

The fMRI data we used was taken from the Natural Scenes Dataset (NSD), an open-source whole brain 7T fMRI dataset. This dataset was created by displaying annotated natural scenes to eight participants during thirty to forty-one-hour scanning sessions. Through all sessions, each participant viewed 9000 to 10,000 coloured natural scenes (referred to as conditions), with 1000 of these images being seen by all participants. The total number of scenes seen across all participants is around 73,000. All these images were taken from the COCO dataset (Lin et al., 2014) and were presented to the participants a maximum of three times. Then, single trial activations, which are referred to as **betas**, are obtained through General Linear Model analysis, reducing the noise and inaccuracies in the data. Each of these betas corresponds to estimates of the fMRI response amplitude of each voxel for each trial. Thus, for each subject, we obtained the betas of each voxel, and for each voxel, which contains a maximum of three beta values corresponding to each condition.

### 2.2 Masking the Visual System

While we can access the data for the whole brain, we are interested in the functional organization of the visual system. Thus, we need to determine which voxel belongs to the visual system and which ones do not. Furthermore, since we want to investigate the relationship between the different ROI of the visual system, we also need to indicate which voxel is part of which ROI. To achieve this, we create a system of masks, which allows us to associate an integer value for each voxel by drawing on the cortical surface, which corresponds to which ROI a voxel is a part of. Additionally, voxels outside of the visual system are associated with 0. This was done using a modified version of MATLAB software developed by the NSD team. Using these scripts, we can plot neuroimaging data on a representation of the cortical surface, which is also used later for visual representation of our results. Some of the data included in the NSD dataset allow us to draw the boundaries between different ROIs. Derived from pRF models, each voxel's polar angle and eccentricities illustrated clear boundaries between different groups of voxels, allowing us to draw and store these boundaries. Using that methodology, we identified and saved 12 ROI on each subject: V1, V2, V3, hV4, LO-1, LO-2, VO-1, VO-2, PHC-1, PHC-2, TO1 and TO-2. Thus, by using these masks alongside the dataset described above, we can know which voxels are included in each ROI, allowing us to associate all the betas values and our ROIs. The MATLAB script

that manages this is called *drawrois.m*.



**Figure 2.1:** Division of our 12 ROIs on the cortical surface. This illustration, as well as all other projections on the cortical surface presented in this document, are taken for subject 01

## 2.3 Selecting and Preprocessing the betas

### 2.3.1 Selection of the betas

With our goal being to model the response function of each voxel in the visual system, we needed to extract the responses of each voxel to each image the eight participants have been subjected to. To achieve this, our *load\_betas.py* script utilizes the *get\_betas* function provided by the publishers of the NSD dataset to extract the betas from the neuroimaging data (Allen et al., 2022). We load the nativesurface version of the betas, a cubic interpolation of the 1.0-mm betas averaged over three cortical surface depths. Since these betas contain the responses of every voxel of a subject’s cortical surface (about 200,000 voxels per subject), we only loaded the voxels situated in the visual system (around 30,000 voxels per subject) using our previously defined masks. Furthermore, the *get\_betas* function extracts the z-score of each response, which further reduces outliers and the variance between responses (Allen et al., 2022). At this step, our betas are contained in a matrix, each row containing the responses of a voxel to all images seen by a participant, as illustrated in Figure 2.2. Importantly, subject 06 and subject 08 contained missing (*NotANumber*, or *NaN*) values in their betas. To deal with this, we removed any voxel that contained missing values so it would not negatively affect the model fitting later in our analysis. The data we removed accounts for 2.9% of the betas of subject 06 and 0.22% for subject 08. We also adapted our masks to the missing data: mask indices of the voxels with missing values were also removed in all the masks to account for their removal.

### 2.3.2 Cross Validation Splitting

Next, we divided the betas into a training and test group via LeaveOne-Out Cross-Validation for each participant. Since each image from the CoCo dataset was seen at most three times by the participant, we chose to keep the activation of the last time each image was seen as

$$\text{betas} = \begin{bmatrix} \beta_{0,0} & \beta_{0,1} & \dots & \beta_{0,n\_conditions} \\ \beta_{1,0} & \beta_{1,1} & \dots & \beta_{1,\beta_{0,n\_conditions}} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{n\_voxels,0} & \beta_{n\_voxels,1} & \dots & \beta_{n\_voxels,n\_conditions} \end{bmatrix}$$

**Figure 2.2:** Matrix containing our nativesurface betas. Each row contains the responses of one voxel to all images seen by a participant ( 30,000 responses)

our test set. Then, our training set consists of the average value of the rest of the betas for each image, leaving us with two matrices of the same dimensionality for each participant. If a participant only saw an image once, we chose to discard that condition since the training example would not have a test counterpart. While this leads to the loss of some data, it is crucial to keep the dimensionality between both our RDM and MDS equal. Thus, we end up with two datasets for each of our 8 participants, containing the responses of most voxels to conditions seen more than once in the visual system.

### 2.3.3 Noise ceiling

We wanted to know the maximum variance we could explain in each voxel based on the training data for our analysis. To do this, we calculated the noise ceiling of every voxel in the visual system, which is the equivalent of the "theoretically best" model we could obtain with our data. We achieved this by comparing the distribution of values in the training set and the test set for each voxel individually. However, due to the distribution between voxels being often misaligned between training and test sets, we first rescaled the training betas of each voxel to bring it closer to the testing betas. This was done by finding the pseudoinverse of a matrix that contains the training betas of voxels and a row of 1s of the same size, as shown in equation 2.1. Then, this pseudoinverse was multiplied by the transposed matrix of the test betas (equation 2.2), which gives us a two-by-one matrix containing the slope and intercept of the linear transformation between the training and the test betas. Lastly, by multiplying the training betas with this scale matrix, we obtained new training values for each voxel that more closely align with the test betas, as shown in equation 2.3.

$$X = \text{pseudoinverse}\left(\begin{bmatrix} \beta_{train_{0,0}} & \beta_{train_{0,1}} & \dots & \beta_{train_{0,n\_conditions}} \\ 1 & 1 & \dots & 1 \end{bmatrix}\right) \quad (2.1)$$

$$scale = X * \text{transpose}\left(\begin{bmatrix} \beta_{test_{0,0}} & \beta_{test_{0,1}} & \dots & \beta_{test_{0,n\_conditions}} \end{bmatrix}\right) \quad (2.2)$$

$$rescaled\_train\_betas = \left[\beta_{train_{0,0}} \quad \beta_{train_{0,1}} \quad \dots \quad \beta_{train_{0,n\_conditions}}\right] * scale \quad (2.3)$$

Once we obtain our rescaled training betas for each voxel, we can calculate the per-voxel noise ceiling by measuring the variance in the error between the training betas and

the test betas. To do this, we divided the Residual Sum of Squares of these two vectors by the Variance in a test voxel, and our noise ceiling is one minus this result, as shown in equation 2.4. From this, we can know the maximum variance that can be explained in each voxel or its "best theoretical" model. The calculation of the per-voxel noise ceiling is found in the *noise\_ceiling.ipynb* notebook.

$$NC_{voxel} = 1 - RSS / Variance \quad (2.4)$$

$$RSS = \sum (row\_train - row\_test)^2$$

$$Variance = \sum (row\_test - mean(row\_test))^2$$

## 2.4 Creating the Representational Sampling Spaces

From both our training and test datasets, we created 2D representational spaces since we want to use the one obtained through the training set for fitting, and the one from the test data will be used to create predicted betas values. To achieve this, we need to know the relationships between the activity patterns of all pairs of conditions. Therefore, our first step was creating our RDMs, which store the dissimilarities between all conditions.

Computing the distances between  $n$  data points, here our conditions, will result in a large  $n * n$  matrix. Each cell  $i, j$  of this matrix thus stores the difference between the activity, represented by our betas, elicited by condition  $i$  and the activity elicited by condition  $j$ . For each ROI, we calculate these differences by finding the correlation distance between all the betas corresponding to condition  $i$  against all the betas corresponding to  $j$ . This was done using the *spatial.distance.pdist* function from *scipy*, with "*correlation*" as a metric. Here, the response amplitude to a condition is represented by a vector of betas  $\beta_V$ , with  $V$  being the number of voxels in an ROI. Thus, for each pair of conditions, the dissimilarity between them is:

$$d(\beta_{V_i}, \beta_{V_j}) = 1 - \frac{(\beta_{V_i} - \bar{\beta}_{V_i}) \cdot (\beta_{V_j} - \bar{\beta}_{V_j})}{\|(\beta_{V_i} - \bar{\beta}_{V_i})\|^2 \|(\beta_{V_j} - \bar{\beta}_{V_j})\|^2} \quad (2.5)$$

Then, we can store all of these dissimilarities in our RDM. For each ROI, we will obtain the corresponding RDM that corresponds to the dissimilarity for each pair of voxels based on the activity in that ROI. Furthermore, for each ROI, two RDMs are computed, one using

the training betas and one using the test betas. Finally, our RDMs, of shape  $n * n$  with  $n$  being the number of images seen by a particular participant, can be constructed as follows:

$$RDM = \begin{bmatrix} d(\beta_{V_0}, \beta_{V_0}) & d(\beta_{V_0}, \beta_{V_1}) & \dots & d(\beta_{V_0}, \beta_{V_{n\_conditions}}) \\ \beta_d(\beta_{V_1}, \beta_{V_0}) & d(\beta_{V_1}, \beta_{V_1}) & \dots & d(\beta_{V_1}, \beta_{V_{n\_conditions}}) \\ \vdots & \vdots & \vdots & \vdots \\ d(\beta_{V_{n\_voxels}}, \beta_{V_0}) & d(\beta_{V_{n\_voxels}}, \beta_{V_1}) & \dots & d(\beta_{V_{n\_voxels}}, \beta_{V_{n\_conditions}}) \end{bmatrix}$$

For each participant, we obtain 24 matrices, all of the same dimensionality. These matrices' diagonals are always zero since we compare a condition to itself. Additionally, all of our RDMs are symmetrical alongside the diagonal since we compute the dissimilarities for all pairs of conditions. Therefore, taking only the upper (or lower) triangle of an RDM is enough to extract all the dissimilarities between different conditions. Then, the amount of dissimilarities will still amount to the number of pairs of distinct images  $p_s$ , being:

$$p_s = \frac{n * (n - 1)}{2} \quad (2.6)$$

As mentioned in Section I, these large matrixes of high dimensionality cannot be used for our analysis. By taking the  $p_s$  number of pairs of distinct images, we can create a 2D projection of the difference in response amplitude between each condition. Using an MDS algorithm, we can create a 2D Cartesian space, where each point corresponds to a condition, and the distance between each point will correspond to the distance between these two conditions in the RDM. This projection from the high-dimensionality of RDMs to 2D Cartesian spaces is done using the SMACOF algorithm, which iteratively minimizes a stress function between the original high-dimension representation of differences and its 2D projection (De Leeuw, 1977). To create our MDS spaces, we used an implementation of the *manifold.mds* function found in the *scikit-learn* library provided by the author of the NSD dataset. All the code for both RDM and MDS creation can be found in the *create\_rdm.py* script. This implementation function takes the upper triangle of an RDM and returns a  $2 * n$  matrix, which corresponds to the 2D coordinates of each condition seen by a participant for each ROI. Importantly, we create an MDS space for each ROI twice, once using the training RDM and once using the test MDS. Each of these MDS spaces is centered around the  $(0, 0)$  coordinates and has a diameter of 2, meaning that the maximum distance between a condition's position and the center is 1. Overall, each of these MDS spaces represents the geometrical organization of the responses amplitudes to each condition in ROI by organizing them

based on how different they are from each other.

## 2.5 Rotating the MDS spaces

As illustrated in Wandell et al. (2007), transformations in the visual system can affect the orientation of the original input. The transformations between each ROI can lead to rotated and/or inverted outputs, which means that the spatial integrity of the original retinal input will be lost the further information travels in the visual system. To attempt to correct these angular distortions, we aim to find the optimal rotation angle between each of our ROI independently for each subject. To do so, we use the Kabsch algorithm, which finds the optimal rotation matrix between two sets of coordinates. More precisely, the algorithm aims to find the rotation matrix that minimizes Root Mean Squared Deviation (RMSD) between these sets of coordinates (Kabsch, 1976). For two sets of points  $P$  and  $Q$  of size  $N$ , the algorithm goes as follows: We first find  $H$ , the cross-covariance matrix between  $P$  and  $Q$ .

$$H = P^T Q \quad (2.7)$$

Then, we calculate the Singular-Value Decomposition of this matrix  $H$ .

$$H = U \Sigma V^T \quad (2.8)$$

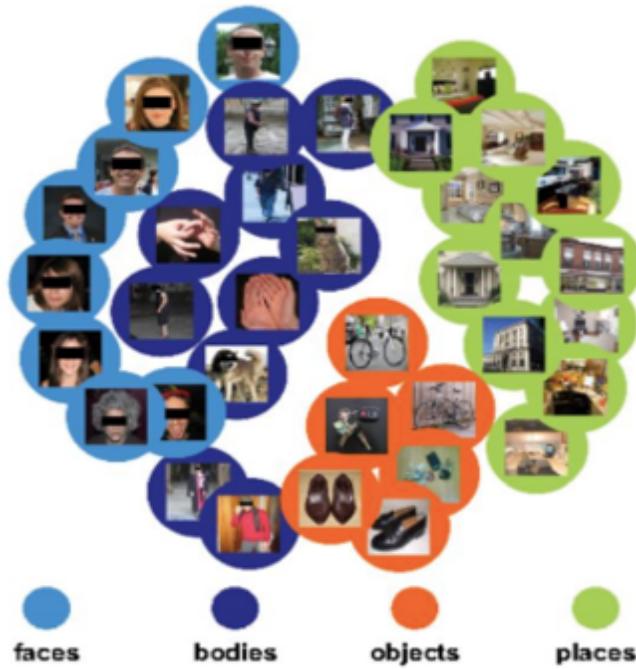
Finally, we can obtain the  $R$  optimal rotation matrix with the following formula:

$$R = U \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & d \end{pmatrix} V^T \quad (2.9)$$

Where  $d$  is:

$$d = \det(UV^T) = \det(U)\det(V) \quad (2.10)$$

Then, we calculated the optimal rotation matrix  $R$  for all ROI's MDS and their flipped counterpart against all others. Once done, we rotated all ROI's (source) MDS based on all other (target) MDSs. Then, we find the target MDS that led to the smallest error by computing the distances between all pairs of points between a target MDS and all MDSs rotated to it. The target MDS that led to the smaller overall distances, for all subjects, between itself and the rotated MDSs from other ROIs was then chosen as our main target to rotate to. Finally, we rotated all MDS based on the optimal rotation matrix for each participant's target MDS. If the flipped source MDS led to a smaller distance to the best target MDS, we chose to keep this flipped source MDS and rotated it to the best target. Thus, by aligning all the MDS to the best target MDS, we attempt to mitigate the lack of spatial integrity caused by the transformation of inputs between ROIs. Figure 2.3 illustrates a 2D MDS space that represents the organisation of our conditions. The *apply\_rotation.py* script in the source code automatically manages these rotations and saves the new rotated MDS spaces.



**Figure 2.3:** Illustration of an MDS Space. Assuming responses from similar conditions will elicit similar responses in the visual system, an ROI's MDS will have similar conditions grouped together and different conditions far apart.

## 2.6 Fit the Gaussian models

To answer our first research question and predict some of the neural activity or variance in a voxel, we want to fit a response function to the activity in each voxel. Now that we have our 2D MDS space, the next step is to fit such a response function and find its optimal parameters on the MDS space. These parameters tell us the preferred position of each voxel

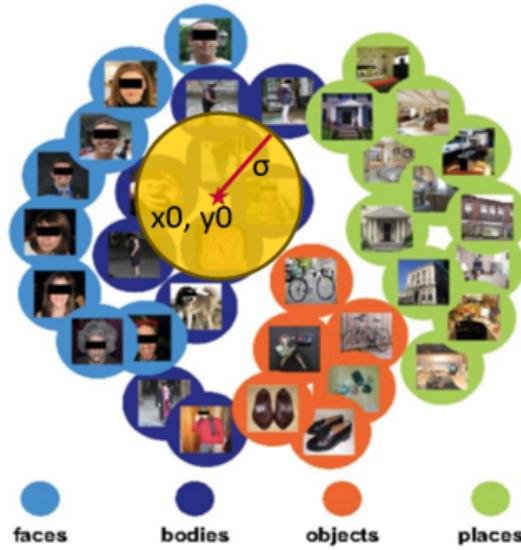
on the MDS space and the size of its preferred responses, i.e. its receptive field. From this, we can use the MDS spaces from our test dataset to obtain predicted betas values for each voxel. Finally, by comparing these predicted betas values to the test betas, we can assess how much variance there is in the activity of a voxel that our models can explain. By using this amount of variance explained in a voxel, we can investigate the relationships between our different ROI since we fitted our models using each ROI's MDS as a sampling space for each voxel.

Previous studies have shown that the receptive field of neurons in some areas of primates' visual system, mainly V4 and MT, can be modelled as a symmetrical, two-dimension circular Gaussian function (Motter, 2009; Kumano and Uka, 2010). Furthermore, pRF and Connective Field models utilize a circular symmetrical Gaussian function when estimating a voxel's response function (Dumoulin and Wandell, 2008; Haak et al., 2013). Thus, to stay consistent with these findings and methodologies, we also chose to find the optimal  $x$ ,  $y$  and  $\sigma$  parameters for each voxel using a circular symmetrical Gaussian function,  $g(v)$ , where  $v$  are the training betas of one voxel.

Our MDS representational spaces will be used as the "visual space" on which we will find each voxel's preferred position and the size of its receptive field by fitting our Gaussian function to the response amplitudes. We hypothesized that some neurons might respond preferably to some conditions. If these conditions are nearby in the MDS space, that should be reflected in the preferred position of that voxel after fitting and in the size of its receptive field. The preferred position  $x$ ,  $y$  of a voxel in the MDS space and the size of its receptive field found after fitting a Gaussian response function is illustrated in Figure 2.4.

Finally, our Gaussian function that estimates a response function for each of our voxels is defined in equation 2.11. The main goal of this function is to find the arrangement of parameters that best predict the response amplitudes to all our conditions using the training betas when sampling from each ROI's MDS space. This is done by minimizing the distance between each image position in the MDS space ( $x, y$ ) and the voxel's preferred position ( $x_0, y_0$ ). Due to the restrictive space of the MDS space, we applied several restrictions on our parameters, mainly to prevent the output preferred position  $x_0, y_0$  from being outside the MDS space and the  $\sigma$  receptive field from being too close to zero. These bounds are shown in Table 2.1

$$g(v) = V(x_0, y_0, \sigma, m, k) = m \cdot e^{-\frac{1}{2}((\frac{x-x_0}{\sigma})^2 + (\frac{y-y_0}{\sigma})^2)} + k \quad (2.11)$$



**Figure 2.4:** Illustration of an MDS Space and the preferred position  $x, y$  of one voxel, obtained after fitting using a circular symmetrical Gaussian function. This represents an ideal scenario in which a voxel responds preferably to one condition (bodies) and in which the images showing a body elicited similar responses, which is then translated into the MDS space.

|          | Lower Bound | Upper Bound |
|----------|-------------|-------------|
| $x_0$    | -1.05       | 1.05        |
| $y_0$    | -1.05       | 1.05        |
| $\sigma$ | 0.01        | inf         |
| $m$      | 0.1         | inf         |
| $k$      | - inf       | inf         |

**Table 2.1:** Lower and Upper bounds of our Gaussian function.

Where

$$x_0 = x_0 * \cos y_0 \quad (2.12)$$

and

$$y_0 = x_0 * \sin y_0 \quad (2.13)$$

Using this range allows a voxel's preferred position on an MDS space to be slightly outside that space, and not giving  $\sigma$  upper bounds allows it to model voxels that do not preferably respond to some specific conditions but respond equally to the entire MDS space

|          | Lower Bound | Upper Bound |
|----------|-------------|-------------|
| $x_0$    | -0.4        | 0.4         |
| $y_0$    | -0.4        | 0.4         |
| $\sigma$ | 0.01        | 2           |
| $m$      | 0.1         | 10          |
| $k$      | -2          | 2           |

**Table 2.2:** Lower and Upper bounds of the starting values of our parameters. Before fitting, each variable will be randomly initialized and thus be given a random value between its lower and upper bound.

if we allow for it. A voxel that responds indiscriminately from all conditions in an MDS might lead to a poor-performing model since it might mean that there is little to no relationship between its responses and the representational geometry of its conditions, which should result in low goodness of fit. Nevertheless, we decided not to bind the upper  $\sigma$  values to be able to discriminate between the average size of voxels' receptive fields between different ROIs later in our analysis. Lastly, we are only interested in models that describe a positive relationship between the response amplitude and the conditions, and thus, we only allow all positive values of  $m$ .

Furthermore, our Gaussian models are randomly initialized from a specified range for each variable, as shown in Table 2.2. The goal with these bounds was to guide the searching process by starting the fitting around the centre of the MDS space, which should allow the algorithm to find the optimal solution faster compared to starting from the edges of the MDS space. Lastly, the  $\sigma$  value is bounded to 2 to not start with a receptive field larger than the sampling space itself.

The fitting of the Gaussian function using the response amplitude as our data and the MDS space as our sampling space is done using the `curve_fit` function from the `scipy` library. `curve_fit` finds a function that minimizes the Residual Sum of Squares (RSS) between the newly found optimal parameters and the data. We initialized the function with our bounds, `trf` as an optimization method and  $ftol = 1e^{-6}$ . `trf` stands for Trust Region Reflective algorithm and is a robust method for sparse bounded search spaces, which is our case. Finally,  $ftol$ , similarly to a learning rate, indicates the change to the cost function for each iteration: we found this value to be a good trade-off between fast learning and finding good solutions. Our implementations of the Gaussian fitting can be found in two scripts: `fit_params.py` and `fit_params_inverse.py`. The former takes each ROI's MDS and fits each voxel on this sampling space. The latter selects a voxel first and then fits a Gaussian function to that voxel twelve times for each MDS.

Finally, once we obtain our fitted Gaussian function for each voxel, we can use their parameters to predict the response amplitude of a voxel to each condition. To achieve this,

we use the MDS space obtained with our test betas and pass it through our fitted Gaussian functions. Doing this returns us a predicted betas value for each condition; thus, we obtained a predicted response amplitude for each voxel for each condition. Furthermore, we can compare these predicted values to the ground truth from the test data, giving us the amount of variance explained (VE) or goodness-of-fit by our model. This VE is calculated by taking the RSS and Variance between our predicted value,  $y_i$ , where  $i$  is a unique condition, the true activity for the same unique condition  $\hat{y}_i$  and its mean  $\bar{y}$ , as shown in equation 2.14. The reader should note that this is the same formula as for the noise ceiling calculation done in equation 2.4.

$$VE = 1 - \frac{RSS}{Variance} \quad (2.14)$$

$$RSS = \sum (y_i - \hat{y}_i)^2$$

$$Variance = \sum (y_i - \bar{y})^2$$

## 2.7 Creating a model of the Visual System

After fitting, each voxel was used to fit twelve different Gaussian models, which yielded twelve sets of optimal parameters. For the rest of our analysis and to create cortical surface representations, we selected a unique set of optimal parameters for each voxel. We designed two selection procedures, which led us to obtain two different models: the ‘only self’ (*oself*) model and the ‘*best roi*’ model.

The *oself* model simply only keeps the optimal parameters of a voxel that were found when fitting on that voxel’s own MDS space. Simply, if a voxel is in V1, we keep the  $x_0$ ,  $y_0$ ,  $\sigma$  and  $VE$  values obtained when fitting the Gaussian function using V1’s MDS space. The goal of this model is to study how voxels respond when using information from their own ROI as a sampling space. Furthermore, to obtain this model, we utilized the fitted parameters obtained from the *fit\_params.py*: in this case, we are interested in the fitting procedure that iterated over each ROI’s MDS.

The *best ROI* model is slightly trickier: We first took the average  $VE$  for all ROI and each sampling space. This was done by taking the fitted parameters from the other fitting procedure, found in *fit\_params\_inverse.py*: here, we want to keep the fitted variables when iterating over each voxel. We discarded the average  $VE$  for an ROI when sampling from itself, except for V1: we hypothesized that since V1 takes input from the retinogeniculate pathway, neurons from V1 might greatly influence each other. Then, we only kept the

sampling space for each ROI, which led to the highest average VE among all eight participants. For example, if sampling from V2's MDS led to the highest average VE in V1 over all subjects, we kept the  $x_0$ ,  $y_0$ ,  $\sigma$  and VE values for voxels in V1 obtained when V2 was our sampling space in all subjects. This *best ROI* model tells us the preferred sampling space of each other's ROI, allowing us to answer our first research question by investigating the relationship between each ROI based on their average VE. This is similar to how CF modeling uses the responses of a brain area that are different from the one it fits to, although here, we use MDS spaces instead of cortical surface data. This model, with its highest complexity than the *oself* model, should show higher goodness of fits since it allows for more source sampling spaces and indicates relationships between ROIs, unraveling the organization between them.

The code that we used to find the best sampling ROIs can be found in the *find\_best\_ROI.ipynb* notebook. The script that computes and stores the final models is called *create\_model\_bestroi.py*

## 2.8 Voxel Performance

After fitting, we found low VE values for many voxels. This can indicate that our model cannot explain much of the variance in a particular voxel, but it can also be because a voxel is very noisy, and not much variance can be explained. To verify this, we wanted to know the relationship between the noise ceiling we computed earlier and our new VE variable. This new "Voxel Performance" (VP) variable is simply the proportion of variance explained in a voxel compared to the theoretically best model. Thus, we divide the VE of a voxel by its noise ceiling to obtain this new VP variable, as shown in equation 2.15. We argue that this variable is more indicative than only VE since it tells us how well a voxel is really doing compared to the best theoretical result: a voxel with a low VE might be explaining a lot of the variance that can be explained, and thus we can consider that the response function obtained from this voxel's betas is performing well at explain the variance that can be explained, even if its overall goodness of fit appears to be low.

$$\text{Voxel\_performance} = \frac{\text{Variance\_Explained}}{\text{Noise\_Ceiling}} \quad (2.15)$$

## 2.9 Correlations between MDS spaces and Cortical Surface topography

### 2.9.1 Distances on the MDS Space

To answer our second research question, we wanted to find the correlation between the distances of each voxel's preferred position  $x_0, y_0$  on the MDS space and the distances between their position on the cortical surface for both the *oself* and *best roi* models. Firstly, we aligned our MDS spaces once more to manage further the fact that transformation between ROIs does not maintain spatial integrity. We found the rotation angle between subject 1's VO-1 and all others to achieve this. This is done by using the previously introduced Kab-sch algorithm to find the optimal rotation matrix between every subject's VO-1 and subject 1 VO-s, using the coordinates of conditions shared by each participant. Next, since the rotation matrix contains the cosine and the sinus of the angle  $n$  of the optimal rotation, we extracted these two values and computed new  $x_0$  and  $y_0$  values based on the optimal rotation angle. The equations to obtain these new coordinates are as follows:

$$\begin{aligned}x'_0 &= x_0 \cdot \cos(n) - y_0 \cdot \sin(n) \\y'_0 &= x_0 \cdot \sin(n) + y_0 \cdot \cos(n)\end{aligned}\tag{2.16}$$

Once done, we compute the pairwise distance for each voxel's new preferred position in the MDS space by measuring the Euclidean distances between each  $x_0, y_0$  position. We did this for each ROI, but this time separated them from each hemisphere as well: we did not want to calculate the distance between voxels in the left and right hemispheres since they are explicitly separated. Moreover, we separated V2 and V3 in **ventral** and **dorsal** regions to prevent arbitrarily large distances resulting from the anatomical structure of these regions. We also save the direction of their relationship, represented by an angle in the [-180, 180] range, and the difference between each voxel, which is simply  $x_{diff} = x_{01} - x_{02}$  and  $y_{diff} = y_{01} - y_{02}$ . Here, we only allow  $x_{diff}$  to be positive, meaning that we shift the  $x$  and  $y$  values accordingly to respect this constraint. The code that computes and saves these three variables can be found in *distances\_mds.py*.

### 2.9.2 Distances on the cortical surface

Next, we also computed and saved the distances of voxels on the cortical surface. Since voxels are organised in a 3D mesh on the brain's surface, we used an algorithm that iteratively finds the distances between points in such an environment. The **white** neuroimaging

file contains the positions of each voxel, represented as vertices, and the sets of edges making all the meshes in each subject. From there, finding the distance between two voxels is equivalent to finding the shortest path amongst these edges. The *meshes\_and\_distances.m* script achieves this by using an algorithm provided by the *GIBBON* package in MATLAB. Similarly to the distance on the MDS surface, we computed distances for each ROI in each hemisphere separately and split V2 and V3 into ventral and dorsal sub-regions. Unfortunately, this script does not account for the missing values but calculated the distances between all voxels, even the ones that are removed due to containing *Nan* values in their betas. However, this issue is later solved by removing the distances associated with removed voxels.

### 2.9.3 Spearman Correlation

With both distances from the preferred position of the MDS space and the cortical surface for each voxel, we calculated the Spearman Correlation for all distances in one ROI for each hemisphere separately. To select our sample size, we used a downscaling factor, which we selected based on how the data was initially processed. The original cubic voxels of size 1.8mm were downscaled to 1mm, meaning an upscaling of  $f = 1.8^3 * n_{voxels}$  voxels compared to the original data. Therefore, we set or scaling factor as the reverse of this upscaling:

$$scaling\_factor = \frac{1}{1.8^3} \quad (2.17)$$

Then, our sample size for our Spearman correlation becomes an approximation of the original number of recorded voxels, being:

$$sample\_size = \frac{n_{voxels}}{scaling\_factor} \quad (2.18)$$

Finally, for each ROI, we paired the distribution of distances between the MDS space and the cortical surface. Then, we computed the Spearman correlation 1000 times on different samples of size *sample\_size* of these distributions and kept the median  $\rho$  and  $p$  values for each ROI in both hemispheres. This was done using the *stats.spearmanr* function available in the *scipy* library. The code to compute the correlation is stored in the *compute\_correlation.ipynb* notebook.

### 2.9.4 Optimal Angle for x, y projections on the Cortical Surface

As an additional tool for visualization on the cortical surface, we decided to compute an "optimal" angle to rotate our  $x_0$  and  $y_0$  values when projecting them on the cortical surface. First, we take all of the pairwise differences between all distinct voxels and divide them by their distances on the cortical surface. Then, we take the average of all these differences over all participants and calculate the new optimal angle using the average difference between all  $x$ :  $\bar{x}$  and the average difference between all  $y$ :  $\bar{y}$  using the 2-argument arctangent, as illustrated in equation 2.19. We also stored the average angle for each ROI in each subject for statistical analysis. The code for the optimal angle calculation for each model is found in the *new\_angle.ipynb* notebook.

$$\text{angle} = \arctan 2(\bar{x}, \bar{y}) * \frac{180}{\pi} \quad (2.19)$$

## 2.10 Full Brain Model

As an additional visual analysis element, we fitted models using all voxels on the cortical surface. The methodology for loading and splitting the betas is the same as in section 2.3; the only difference is that we do not ignore the voxels outside the visual system. This resulted in large datasets (50GB per participant), which required splitting for lighter computations. Then, we fit a Gaussian model on each of the voxels and each of the 12 train MDS spaces and obtain VE values using the test MDS spaces, using the same procedure as for the *best roi* model. Our goal was to be able to illustrate the goodness of fit of all voxels on the cortical surface using different sampling spaces. While this is purely explorative, making these models can show how our models perform over the whole brain and observe if other brain regions have positive or negative goodness of fit when fitting on our MDS space. This would indicate that some regions can be stimulated or inhibited when a subject is faced with naturalistic visual stimuli, indicating that processing these visual stimuli requires more than activity from the visual system. The code for the full brain betas loading and fitting can be found in the *load\_betas\_full.py*, *clean\_betas.py*, *fit\_params\_fullbrain.py* and *variance\_full\_brain.py*

## 3. Results

### 3.1 Roated MDS

don't forget to include stats in methods

After rotating all our MDS spaces to every other, we found that rotating to VO-1 (that is, finding the optimal rotation matrix of every MDS when VO-1 is the target MDS to rotate to) led to the smallest overall distance for all subjects between each pair of images, when comparing the rotated and target MDS spaces. Additionally, when we compared the distribution of distances calculated when rotating to each ROI we found that rotating to VO-1 consistently led to the smallest overall distances compared to when rotating to other regions. The p-values of these sign rank tests are shown in Figure 3.1. Furthermore, the z values when VO-1 is our second set of measurements are always positive, as shown in the fifth column of Figure 3.2. This means that the distance elicited by rotating to VO-1 is always smaller compared to distances elicited by smaller regions, leading to positive outcomes when subtracting the latter from the former. Therefore, we chose to align the MDS spaces of each subject to its own VO-1's MDS space.

|       | V1     | V2     | V3     | hV4    | VO-1   | VO-2   | PHC-1  | PHC-2  | LO-1   | LO-2   | TO-1   | TO-2   |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| V1    | 1.0    | 0.4315 | 0.0052 | 0.9694 | 0.0001 | 0.0557 | 0.307  | 0.1202 | 0.6625 | 0.9084 | 0.4231 | 0.803  |
| V2    | 0.4315 | 1.0    | 0.0487 | 0.2721 | 0.0026 | 0.0039 | 0.666  | 0.2124 | 0.7407 | 0.7407 | 0.3827 | 0.7956 |
| V3    | 0.0052 | 0.0487 | 1.0    | 0.1435 | 0.0287 | 0.977  | 0.1383 | 0.8629 | 0.3185 | 0.1777 | 0.045  | 0.1672 |
| hV4   | 0.9694 | 0.2721 | 0.1435 | 1.0    | 0.0146 | 0.0152 | 0.6521 | 0.5142 | 0.8894 | 0.4958 | 0.2231 | 0.5584 |
| VO-1  | 0.0001 | 0.0026 | 0.0287 | 0.0146 | 1.0    | 0.3231 | 0.0706 | 0.666  | 0.2177 | 0.1071 | 0.0251 | 0.0274 |
| VO-2  | 0.0557 | 0.0039 | 0.977  | 0.0152 | 0.3231 | 1.0    | 0.3279 | 0.8105 | 0.2361 | 0.1113 | 0.0305 | 0.0569 |
| PHC-1 | 0.307  | 0.666  | 0.1383 | 0.6521 | 0.0706 | 0.3279 | 1.0    | 0.2935 | 0.5978 | 0.3932 | 0.1157 | 0.2658 |
| PHC-2 | 0.1202 | 0.2124 | 0.8629 | 0.5142 | 0.666  | 0.8105 | 0.2935 | 1.0    | 0.0071 | 0.0011 | 0.0    | 0.034  |
| LO-1  | 0.6625 | 0.7407 | 0.3185 | 0.8894 | 0.2177 | 0.2361 | 0.5978 | 0.0071 | 1.0    | 0.0313 | 0.0013 | 0.3521 |
| LO-2  | 0.9084 | 0.7407 | 0.1777 | 0.4958 | 0.1071 | 0.1113 | 0.3932 | 0.0011 | 0.0313 | 1.0    | 0.2361 | 0.9198 |
| TO-1  | 0.4231 | 0.3827 | 0.045  | 0.2231 | 0.0251 | 0.0305 | 0.1157 | 0.0    | 0.0013 | 0.2361 | 1.0    | 0.3374 |
| TO-2  | 0.803  | 0.7956 | 0.1672 | 0.5584 | 0.0274 | 0.0569 | 0.2658 | 0.034  | 0.3521 | 0.9198 | 0.3374 | 1.0    |

**Figure 3.1:** p values of the Wilcoxon signed-rank test between the distribution of distance caused by each ROI against all others. Each row represents the p-value of the sign rank test between the distribution of distances caused by taking that ROI as a target on that row against the distribution of distances elicited by all other ROI. For example, the second cell of the first row compares the distribution of distances when V1 is the target (so the distances of V3 rotated to V1, hV4 rotated to V1, etc...) against the distribution of distances when V2 is the target (so V3 rotated to V2, hV4 rotated to V2, etc...)

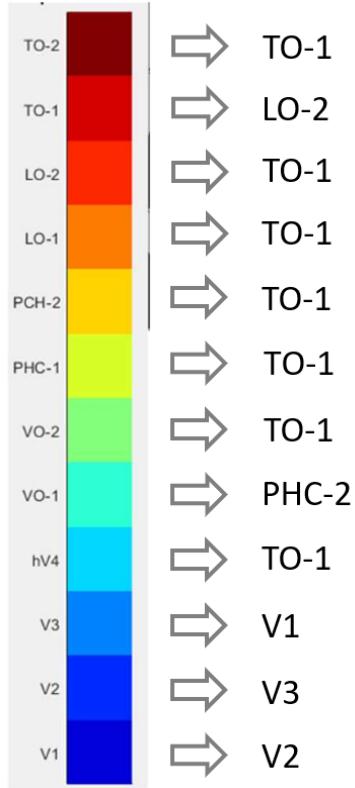
|       | V1      | V2      | V3      | hV4     | VO-1   | VO-2    | PHC-1   | PHC-2   | LO-1    | LO-2    | TO-1    | TO-2    |
|-------|---------|---------|---------|---------|--------|---------|---------|---------|---------|---------|---------|---------|
| V1    | 0.0     | 0.7866  | 2.7962  | -0.0384 | 4.0145 | 1.9137  | 1.0216  | 1.554   | -0.4365 | -0.1151 | -0.801  | 0.2494  |
| V2    | -0.7866 | 0.0     | 1.9713  | 1.0983  | 3.0169 | 2.8874  | 0.4317  | 1.247   | 0.3309  | -0.3309 | -0.8729 | -0.259  |
| V3    | -2.7962 | -1.9713 | 0.0     | -1.4629 | 2.1871 | -0.0288 | -1.4821 | 0.1727  | -0.9976 | -1.3478 | -2.0048 | -1.3813 |
| hV4   | 0.0384  | -1.0983 | 1.4629  | 0.0     | 2.4413 | 2.4269  | 0.4509  | 0.6523  | -0.1391 | -0.6811 | -1.2183 | -0.5851 |
| VO-1  | -4.0145 | -3.0169 | -2.1871 | -2.4413 | 0.0    | -0.988  | -1.8082 | -0.4317 | -1.2326 | -1.6116 | -2.2399 | -2.2063 |
| VO-2  | -1.9137 | -2.8874 | 0.0288  | -2.4269 | 0.988  | 0.0     | 0.9784  | -0.2398 | -1.1847 | -1.5924 | -2.1631 | -1.9041 |
| PHC-1 | -1.0216 | -0.4317 | 1.4821  | -0.4509 | 1.8082 | -0.9784 | 0.0     | 1.0504  | -0.5276 | -0.8537 | -1.5732 | -1.1127 |
| PHC-2 | -1.554  | -1.247  | -0.1727 | -0.6523 | 0.4317 | 0.2398  | -1.0504 | 0.0     | -2.6907 | -3.2759 | -4.849  | -2.12   |
| LO-1  | 0.4365  | -0.3309 | 0.9976  | 0.1391  | 1.2326 | 1.1847  | 0.5276  | 2.6907  | 0.0     | -2.1535 | -3.2087 | -0.9305 |
| LO-2  | 0.1151  | 0.3309  | 1.3478  | 0.6811  | 1.6116 | 1.5924  | 0.8537  | 3.2759  | 2.1535  | 0.0     | -1.1847 | 0.1007  |
| TO-1  | 0.801   | 0.8729  | 2.0048  | 1.2183  | 2.2399 | 2.1631  | 1.5732  | 4.849   | 3.2087  | 1.1847  | 0.0     | 0.9593  |
| TO-2  | -0.2494 | 0.259   | 1.3813  | 0.5851  | 2.2063 | 1.9041  | 1.1127  | 2.12    | 0.9305  | -0.1007 | -0.9593 | 0.0     |

**Figure 3.2:** z values of the same Wilcoxon signed-rank test. Positive values indicate that the second measurement contains smaller distances than the first one on average. Taking the same example as for Figure 3.1 above, the distances elicited by V2 as the target MDS to rotated two leads to smaller distances between V2 and the rotated source MDSs than when V1 is the target, on average.

## 3.2 Best ROI and model creation

To select the ROI that 'best explains' each other ROI, we looked at which other sampling space (as in, MDS space) led to the highest average VE over an ROI when fitting our Gaussian function. For V1, we also include the average VE obtained when V1 was the sampling space. The result of this selection is shown in Figure 3.3. Here, we see that TO-1 is the preferred sampling space for most regions in the mid-to-late ROIs. V1, V2, V3 show a circular relationship: V1 can be best explained by V2, V2 by V3 and V3 by V1. Although these ROIs were selected as the ones that "best explain" other regions since they obtained the numerical best average over participants, it should be noted that not all these are statically the best. In fact, Figure 3.4 shows the p-value of the Wilcoxon signed-rank test we ran to compare the distribution of VE across all subjects and hemispheres of an ROI when fitting on the "best" sampling space against all others. It shows that for some ROIs, the numerical best MDS space fitted on does not lead to a statistically better distribution than another MDS space: for example, V1 can also be "best explained" by itself since the distribution of VE when V1 voxels are fitted on V1 is not statistically different than when fitting on V2 ( $p = 0.795987$ ). Similarly, VO-1 can be best explained by TO-1 as well since its distribution of average VE is not statistically different than the one obtained with PHC-2, its numerical best, as a sampling space. Nevertheless, to obtain unambiguous models in our next step, we chose to keep the numerical best candidates as our sampling space due to them performing significantly better compared to the majority of other sampling spaces, as illustrated by the p values of our signed-rank test in Figure 3.4

From this, we create our models of the visual systems, which only store one of each fitted variable per voxel. The *oself* model stores the  $x_0$ ,  $y_0$ ,  $\sigma$  and  $VE$  that were obtained



**Figure 3.3:** Illustration of which ROI "best explains" each other ROI. For each ROI on the left side, we select the ROI that leads to the highest average VE when using its MDS when fitting the Gaussian function. For example, voxels in V1 had their highest average VE when fitted on V2: thus, we chose V2 as the region that 'best explains' activity in V1

when a voxel was fitting on its own MDS space, while the *best\_roi* model takes the fitted variable obtain when sampling from the MDS space that best explains its ROI. For example, for voxels in V1, we extract variables that were obtained when fitting on V2's MDS space. Figure 3.5 shows the projection of  $VE$ ,  $VP$ ,  $x_0$ ,  $y_0$ ,  $\sigma$  of both models on the cortical surface for subject 01.

### 3.3 Goodness of fit

Once we obtained our model, we can project the  $VE$  value of each voxel on the cortical surface, as done in Figure 3.5 b and c. For both models, we observe that voxels do rather well in early areas, in particular V1. These figures also indicate that our models performed the best in the ventral pathway, in particular in TO-1 and TO-2. This is further confirmed by Figure 3.6

|       | V1       | V2       | V3       | hV4      | VO-1     | VO-2     | PHC-1    | PHC-2    | LO-1     | LO-2     | TO-1     | TO-2     |
|-------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| V1    | 0.795987 | 1.0      | 0.108941 | 0.000935 | 0.022895 | 0.001123 | 0.008361 | 0.009726 | 0.000935 | 0.001609 | 0.015086 | 0.000776 |
| V2    | 0.679116 | 0.003205 | 1.0      | 0.001919 | 0.015086 | 0.001123 | 0.000531 | 0.001919 | 0.000776 | 0.000935 | 0.006133 | 0.000438 |
| V3    | 1.0      | 0.379375 | 0.958761 | 0.062671 | 0.569494 | 0.062671 | 0.070326 | 0.078731 | 0.019971 | 0.022895 | 0.301054 | 0.001609 |
| hV4   | 0.062671 | 0.001609 | 0.351979 | 0.001123 | 0.255291 | 0.002708 | 0.469113 | 0.876722 | 0.009726 | 0.004455 | 1.0      | 0.002282 |
| VO-1  | 0.019971 | 0.006133 | 0.019971 | 0.005234 | 0.017378 | 0.002282 | 0.002708 | 1.0      | 0.011286 | 0.070326 | 0.605095 | 0.005234 |
| VO-2  | 0.001123 | 0.000438 | 0.000935 | 0.000438 | 0.000643 | 0.000438 | 0.000438 | 0.196108 | 0.437967 | 0.162673 | 1.0      | 0.000438 |
| PHC-1 | 0.000531 | 0.000531 | 0.000531 | 0.000438 | 0.000531 | 0.000438 | 0.000438 | 0.005234 | 0.002708 | 0.162673 | 1.0      | 0.000438 |
| PHC-2 | 0.000531 | 0.000531 | 0.000531 | 0.000531 | 0.000531 | 0.000531 | 0.000438 | 0.000643 | 0.000776 | 0.029873 | 1.0      | 0.000531 |
| LO-1  | 0.000935 | 0.000531 | 0.000776 | 0.001346 | 0.000776 | 0.001919 | 0.001609 | 0.038607 | 0.000776 | 0.029873 | 1.0      | 0.000935 |
| LO-2  | 0.000438 | 0.000438 | 0.000438 | 0.000438 | 0.000438 | 0.000438 | 0.000438 | 0.162673 | 0.029873 | 0.087936 | 1.0      | 0.000438 |
| TO-1  | 0.000438 | 0.000438 | 0.000438 | 0.000438 | 0.000438 | 0.000438 | 0.000438 | 0.13373  | 0.836139 | 1.0      | 0.087936 | 0.000935 |
| TO-2  | 0.000438 | 0.000438 | 0.000438 | 0.000438 | 0.000438 | 0.000438 | 0.000438 | 0.002708 | 0.097989 | 0.108941 | 1.0      | 0.004455 |

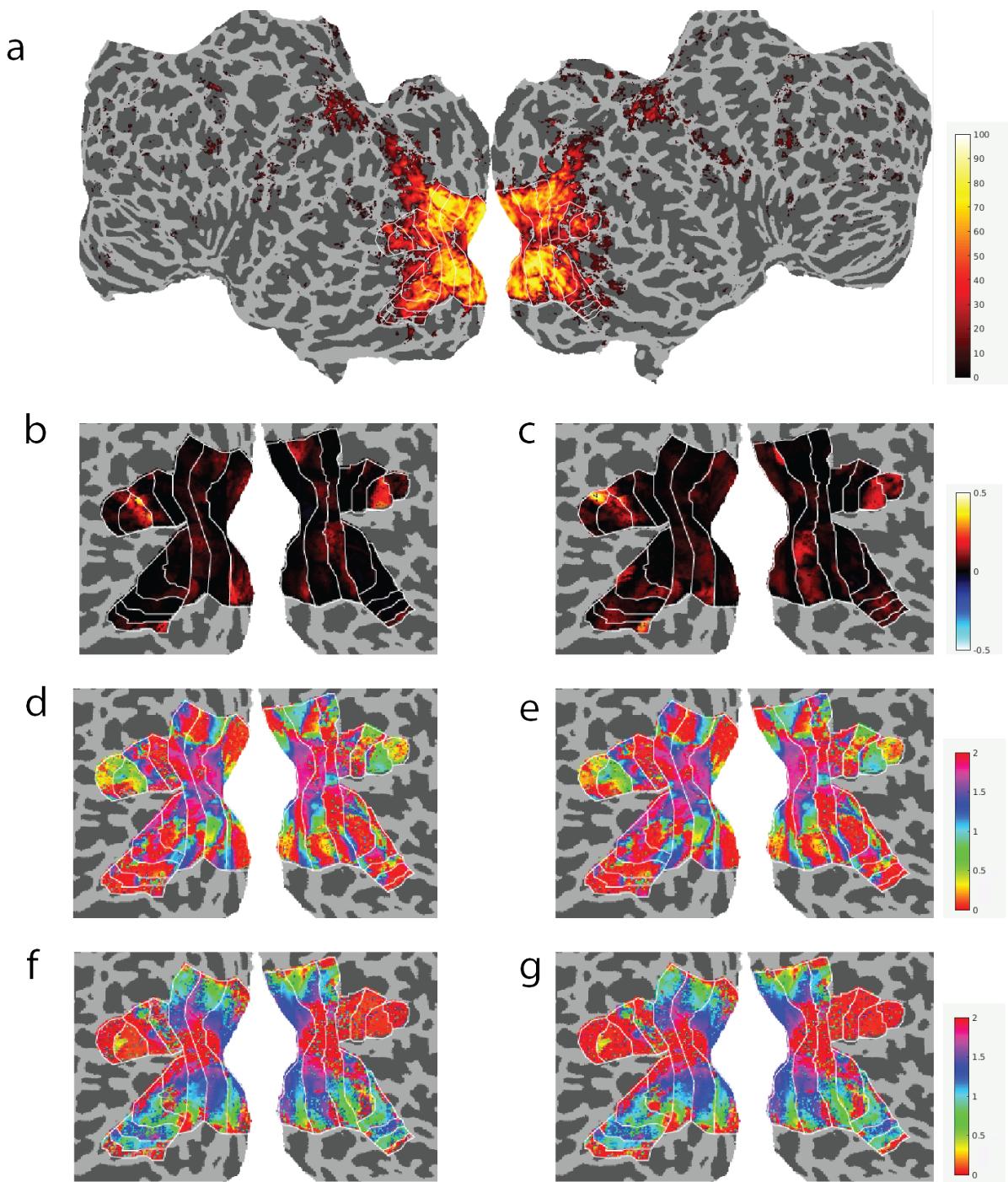
**Figure 3.4:** p-values of the Wilcoxon test that compares the distributions of VE between each ROI which voxels were fitted using each MDS space, against the distributions of VE for that ROI when fitted on the "best" MDS (which leads to the highest VE), over all subjects and both hemispheres. For example, the top left cell compared the distribution of average VE across both hemispheres and all subjects of voxels in V1 fitted using V1's MDS space against when voxels in V2 were fitted using the "best" MDS, here being V2. We observed that for some, these distributions are not statistically different. For example, V1 can also be "best explained" by itself.

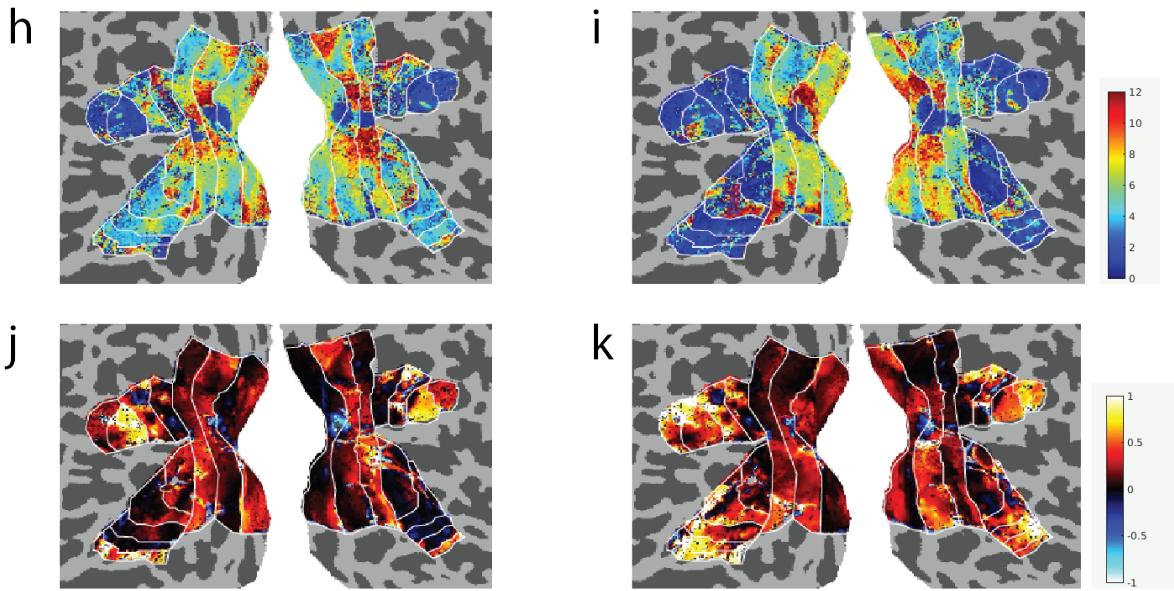
## 3.4 Voxel Performance

### 3.5 sigma

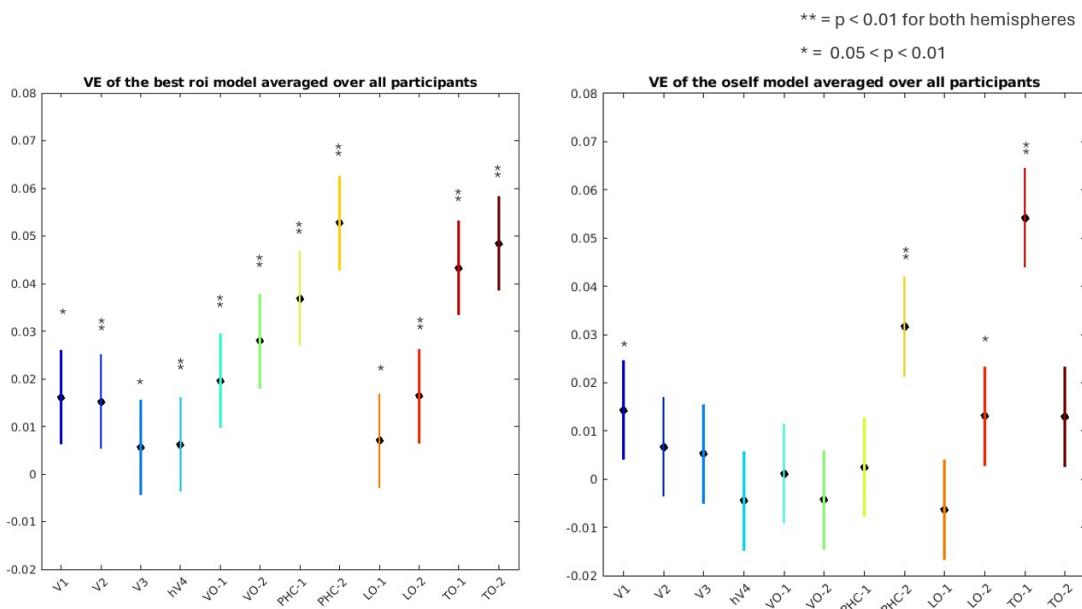
### 3.6 correlation

### 3.7 Full brain





**Figure 3.5:** Projections of our model's fitted variables on the cortical surface, taking subject's *01\_oself* and *best\_roi* models, except **a**, showing the goodness-of-fit of the pRF models ( $R^2$ ). These values were obtained directly from the NSD Dataset (Allen et al., 2022). **b** and **c** show the goodness-of-fit (VE) of both models. **d** and **e** is the  $x_0$  of the preferred position of each voxel on the MDS space, rotated an extra time using the optimal angle we calculated using directions of change in the MDS space. **f** and **g** show the  $y_0$  values of the voxels' preferred position. **h** and **i** are the  $\sigma$  size of the response function of each voxel, and lastly, **j** and **k** are the Voxel Performance of each voxel.



**Figure 3.6:** Average Variance Explained across both hemispheres and all participants. The star symbols indicate if the distribution of VE for each ROI is significantly different than 0 or not. The *best\_roi* model can consistently explain some of the variance in all regions, unlike the *oself* model. Both models show the same trend of difference existing between regions, with later regions tending to perform better.

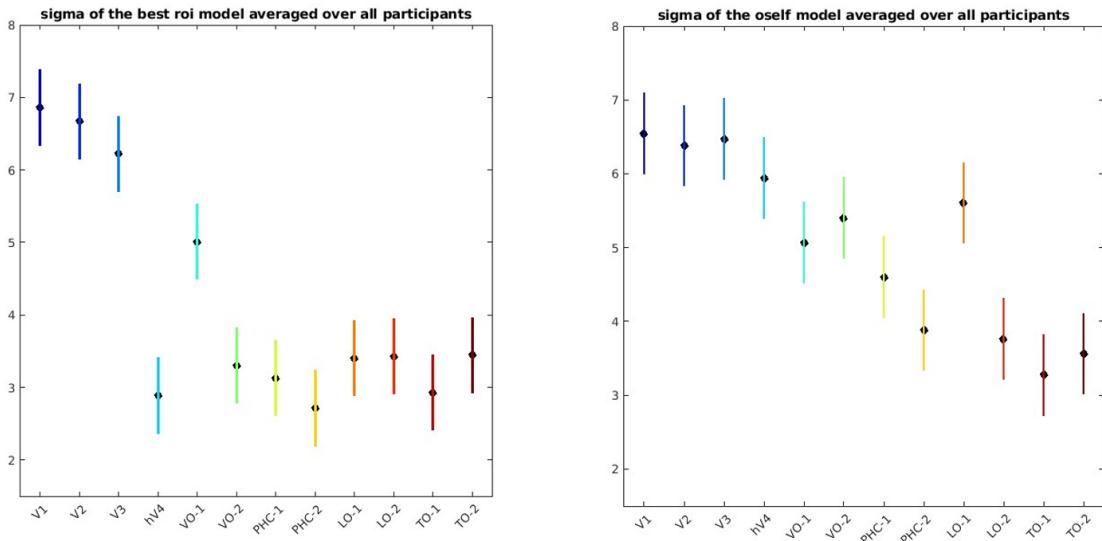
Median Voxel Performance of the oself model, per ROI for each subject

|        | V1        | V2        | V3        | hV4       | VO-1      | VO-2      | PHC-1     | PHC-2     | LO-1      | LO-2      | TO-1     | TO-2      |
|--------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|----------|-----------|
| subj01 | 0.040025  | 0.12344   | 0.170305  | 0.042499  | 0.019779  | -0.006485 | -0.012187 | 0.551065  | 0.00959   | 0.196468  | 0.654974 | 0.256429  |
| subj02 | 0.169498  | 0.024915  | 0.100674  | -0.015865 | 0.018734  | -0.039091 | 0.006008  | 0.519051  | 0.175942  | 0.175634  | 0.560543 | -0.009759 |
| subj03 | -0.156835 | -0.044242 | -0.305162 | -0.183178 | 0.044949  | -0.128011 | 0.197814  | 0.387031  | -0.321619 | 0.006857  | 0.477796 | 0.110832  |
| subj04 | 0.118806  | 0.021032  | -0.008325 | -0.255032 | -0.153751 | -0.289939 | -0.232878 | -0.035614 | -0.188762 | -0.165723 | 0.541446 | -0.036305 |
| subj05 | 0.008787  | 0.046391  | 0.016487  | -0.047193 | -0.046799 | 0.061238  | 0.371224  | 0.580817  | -0.081235 | 0.095299  | 0.341767 | 0.31094   |
| subj06 | 0.049803  | 0.069936  | 0.039297  | -0.087255 | -0.01695  | 0.010565  | -0.175242 | 0.055851  | 0.048446  | 0.436096  | 0.752979 | -0.002285 |
| subj07 | -0.064616 | -0.32504  | -0.689098 | -1.093173 | 0.017325  | -0.364204 | -0.078304 | -0.164701 | -3.85799  | -0.020043 | 0.026926 | -0.018189 |
| subj08 | -0.023141 | -0.041292 | -0.038437 | -0.034179 | -0.096871 | -0.166312 | -0.027234 | 0.295557  | -1.019041 | -0.042964 | 0.25848  | -0.241625 |

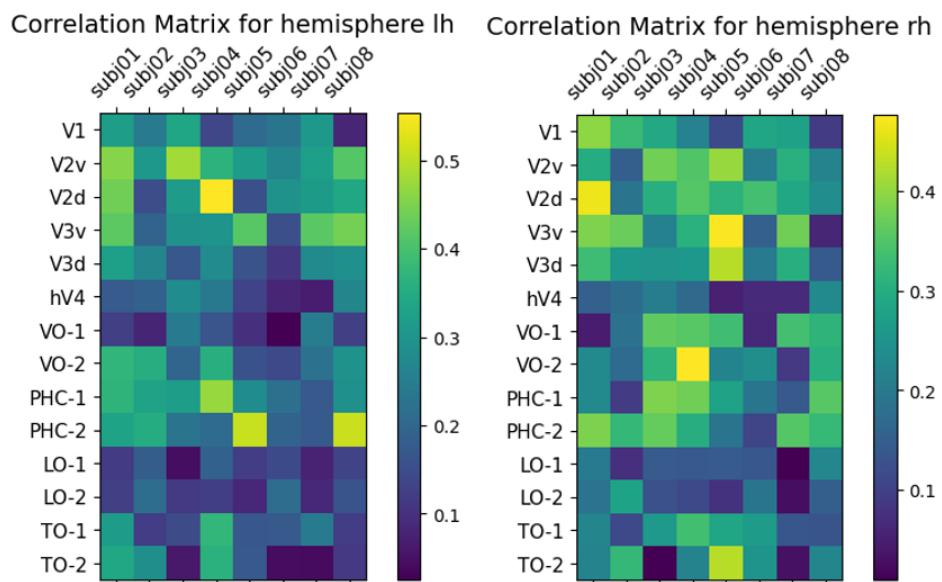
Median Voxel Performance of the best ROI model, per ROI for each subject

|        | V1        | V2        | V3        | hV4       | VO-1      | VO-2      | PHC-1     | PHC-2    | LO-1      | LO-2      | TO-1     | TO-2     |
|--------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|----------|-----------|-----------|----------|----------|
| subj01 | 0.151741  | 0.150579  | 0.101542  | 0.030396  | 0.322279  | 0.366542  | 0.616868  | 0.648627 | 0.059659  | 0.221925  | 0.435081 | 0.652332 |
| subj02 | 0.137578  | 0.110969  | 0.107965  | 0.068051  | 0.460134  | 0.588035  | 0.724519  | 0.82972  | 0.264172  | 0.289377  | 0.397471 | 0.743818 |
| subj03 | -0.084067 | -0.103247 | -0.117294 | -0.002923 | 0.09979   | 0.468392  | 0.56771   | 0.755124 | -0.061765 | 0.137805  | 0.286574 | 0.891751 |
| subj04 | 0.098773  | 0.145301  | -0.078031 | -0.083545 | 0.334309  | 0.596198  | 0.285461  | 0.045964 | -0.197897 | -0.147622 | 0.143574 | 0.816652 |
| subj05 | 0.098821  | 0.110138  | -0.013705 | 0.049634  | 0.190296  | 0.381423  | 0.620916  | 0.733828 | 0.074531  | 0.303765  | 0.317482 | 0.300054 |
| subj06 | 0.079139  | 0.052306  | -0.020764 | 0.188738  | 0.08082   | 0.417105  | 0.584806  | 0.732001 | 0.40386   | 0.492922  | 0.696161 | 0.554036 |
| subj07 | -0.218451 | -0.397111 | -0.056539 | -0.054193 | -0.578494 | -0.025509 | -0.015429 | 0.060564 | -0.168707 | 0.005847  | 0.219962 | 0.039993 |
| subj08 | -0.07221  | 0.026404  | -0.070775 | -0.055738 | -0.060925 | 0.138313  | 0.264977  | 0.376521 | -0.513227 | -0.272586 | 0.2691   | 0.093515 |

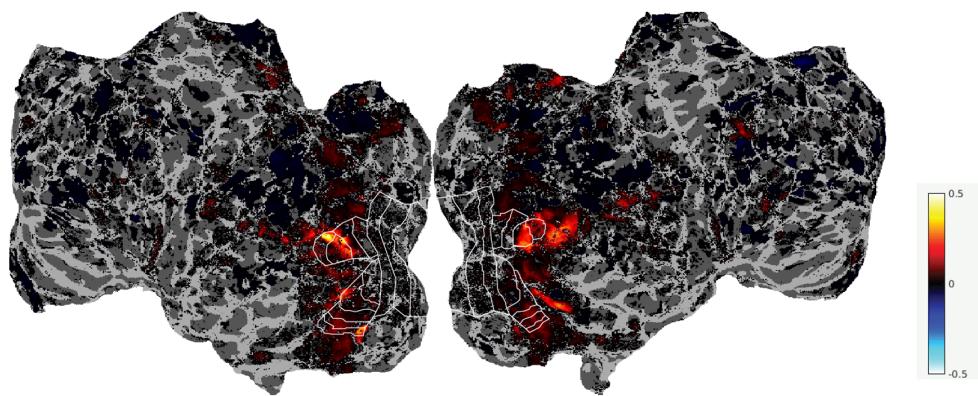
**Figure 3.7:** Median Voxel Performance of each ROI for each subject, with *oself* on top and *best\_roi* below. Both models' voxel performance increases in later areas for both ventral and dorsal pathways: for some ROI, most of the variance that can be explained is explained.



**Figure 3.8:** Average sigma across both hemispheres and all participants. Both models show the same trend: the size of the response functions in early areas is much larger than the diameters of the MDS space (2). This indicates that voxels in these areas do not, for the vast majority, respond preferentially to a type of image. The average  $\sigma$  is much lower in later areas, especially for the *best\_roi* model, indicating that some voxels respond preferably to some categories.



**Figure 3.9:** Correlation both hemis.



**Figure 3.10:** Full brain.

## **4. Conclusion**

### **4.1 Discussion**

### **Appendices**

## **A. Appendix title**

## **A. Appendix A**

### **A.1 First appendix section**

# Bibliography

- Kriegeskorte, N., & Wei, X.-X. (2021). Neural tuning and representational geometry. *Nature Reviews Neuroscience*, 22(11), 703–718.
- Gross, C. G. (1999). *Brain, vision, memory: Tales in the history of neuroscience*. MIT Press.
- Rose, F. C. (1993). European neurology from its beginnings until the 15th century: An overview. *Journal of the History of the Neurosciences*, 2(1), 21–44.
- Grant, G. (2006). The 1932 and 1944 nobel prizes in physiology or medicine: Rewards for ground-breaking studies in neurophysiology. *Journal of the History of the Neurosciences*, 15(4), 341–357.
- Marzi, C., Paulesu, E., & Bottini, G. (2009). The physiology of mind.
- Ji, J. L., Spronk, M., Kulkarni, K., Repovš, G., Anticevic, A., & Cole, M. W. (2019). Mapping the human brain's cortical-subcortical functional network organization. *Neuroimage*, 185, 35–57.
- Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., Vogel, A. C., Laumann, T. O., Miezin, F. M., Schlaggar, B. L., & Petersen, S. E. (2011). Functional network organization of the human brain. *Neuron*, 72(4), 665–678. <https://doi.org/10.1016/j.neuron.2011.09.006>
- Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology*.
- Graziano, M. S., & Botvinick, M. M. (2002). How the brain represents the body: Insights from neurophysiology and psychology. *Common mechanisms in perception and action: Attention and performance XIX*, 19, 136–157.
- Amunts, K., Schlaug, G., Jäncke, L., Steinmetz, H., Schleicher, A., Dabringhaus, A., & Zilles, K. (1997). Motor cortex and hand motor skills: Structural compliance in the human brain. *Human brain mapping*, 5(3), 206–215.
- Macdonell, R. A., Jackson, G. D., Curatolo, J. M., Abbott, D. F., Berkovic, S. F., Carey, L. M., Syngeniotin, A., Fabinyi, G., & Scheffer, I. E. (1999). Motor cortex localization using functional mri and transcranial magnetic stimulation. *Neurology*, 53(7), 1462–1462.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience*, 17(1), 353–362.
- Friederici, A. D., Chomsky, N., Berwick, R. C., Moro, A., & Bolhuis, J. J. (2017). Language, mind and brain. *Nature human behaviour*, 1(10), 713–722.
- Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2009). Visual cortex in humans. *Encyclopedia of neuroscience*, 10, 251–257.
- Schall, J. D., Morel, A., King, D. J., & Bullier, J. (1995). Topography of visual cortex connections with frontal eye field in macaque: Convergence and segregation of processing streams. *Journal of Neuroscience*, 15(6), 4464–4487.
- Engel, S. A., Rumelhart, D. E., Wandell, B. A., Lee, A. T., Glover, G. H., Chichilnisky, E.-J., & Shadlen, M. N. (1994). fMRI of human visual cortex. *Nature*, 369(6481), 525–525. <https://doi.org/10.1038/369525a0>
- Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine*, 34(4), 537–541. <https://doi.org/10.1002/mrm.1910340409>
- Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3), 186–198. <https://doi.org/10.1038/nrn2575>

- Dosenbach, N. U. F., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., Nelson, S. M., Wig, G. S., Vogel, A. C., Lessov-Schlaggar, C. N., Barnes, K. A., Dubis, J. W., Feczko, E., Coalson, R. S., Pruett, J. R., Barch, D. M., Petersen, S. E., & Schlaggar, B. L. (2010). Prediction of individual brain maturity using fMRI. *Science (New York, N.Y.)*, 329(5997), 1358–1361. <https://doi.org/10.1126/science.1194144>
- Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., Kennedy, D. N., Hoppel, B. E., Cohen, M. S., & Turner, R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences*, 89(12), 5675–5679. <https://doi.org/10.1073/pnas.89.12.5675>
- Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, 412(6843), 150–157. <https://doi.org/10.1038/35084005>
- Heeger, D. J., & Ress, D. (2002). What does fMRI tell us about neuronal activity? *Nature Reviews Neuroscience*, 3(2), 142–151. <https://doi.org/10.1038/nrn730>
- Matthews, P. M., & Jezzard, P. (2004). Functional magnetic resonance imaging. *Journal of Neurology, Neurosurgery & Psychiatry*, 75(1), 6–12.
- Smith, S. M. (2004). Overview of fMRI analysis. *The British Journal of Radiology*, 77(suppl\_2), S167–S175. <https://doi.org/10.1259/bjr/33553595>
- Catani, M., Jones, D. K., Donato, R., & ffytche, D. H. (2003). Occipito-temporal connections in the human brain. *Brain*, 126(9), 2093–2107. <https://doi.org/10.1093/brain/awg203>
- Daniel, P. M., & Whitteridge, D. (1961). The representation of the visual field on the cerebral cortex in monkeys. *The Journal of Physiology*, 159(2), 203–221.
- Haak, K. V., Winawer, J., Harvey, B. M., Renken, R., Dumoulin, S. O., Wandell, B. A., & Cornelissen, F. W. (2013). Connective field modeling. *NeuroImage*, 66, 376–384. <https://doi.org/10.1016/j.neuroimage.2012.10.037>
- Dumoulin, S. O., & Wandell, B. A. (2008). Population receptive field estimates in human visual cortex. *NeuroImage*, 39(2), 647–660.
- Harvey, B. M., Klein, B. P., Petridou, N., & Dumoulin, S. O. (2013). Topographic representation of numerosity in the human parietal cortex. *Science*, 341(6150), 1123–1126.
- Harvey, B. M., Dumoulin, S. O., Fracasso, A., & Paul, J. M. (2020). A network of topographic maps in human association cortex hierarchically transforms visual timing-selective responses. *Current Biology*, 30(8), 1424–1434.
- Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron*, 56(2), 366–383.
- von der Heydt, R., Zhou, H., & Friedman, H. S. (2000). Representation of stereoscopic edges in monkey visual cortex. *Vision Research*, 40(15), 1955–1967. [https://doi.org/10.1016/S0042-6989\(00\)00044-4](https://doi.org/10.1016/S0042-6989(00)00044-4)
- Amano, K., Wandell, B. A., & Dumoulin, S. O. (2009). Visual Field Maps, Population Receptive Field Sizes, and Visual Field Coverage in the Human MT+ Complex. *Journal of Neurophysiology*, 102(5), 2704–2718. <https://doi.org/10.1152/jn.00102.2009>
- Arcaro, M. J., McMains, S. A., Singer, B. D., & Kastner, S. (2009). Retinotopic organization of human ventral visual cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 29(34), 10638–10652. <https://doi.org/10.1523/JNEUROSCI.12807-09.2009>
- Diedrichsen, J., & Kriegeskorte, N. (2017). Representational models: A common framework for understanding encoding, pattern-component, and representational-similarity analysis. *PLOS Computational Biology*, 13(4), e1005508. <https://doi.org/10.1371/journal.pcbi.1005508>
- Mur, M., Bandettini, P. A., & Kriegeskorte, N. (2009). Revealing representational content with pattern-information fmri—an introductory guide. *Social cognitive and affective neuroscience*, 4(1), 101–109.
- Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis—connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, 4.

## Bibliography

---

- Devereux, B. J., Clarke, A., Marouchos, A., & Tyler, L. K. (2013). Representational similarity analysis reveals commonalities and differences in the semantic processing of words and objects. *Journal of Neuroscience*, 33(48), 18906–18916.
- Chen, Y., Shimotake, A., Matsumoto, R., Kunieda, T., Kikuchi, T., Miyamoto, S., Fukuyama, H., Takahashi, R., Ikeda, A., & Ralph, M. L. (2016). The ‘when’ and ‘where’ of semantic coding in the anterior temporal lobe: Temporal representational similarity analysis of electrocorticogram data. *Cortex*, 79, 1–13.
- Freund, M. C., Bugg, J. M., & Braver, T. S. (2021). A representational similarity analysis of cognitive control during color-word stroop. *Journal of Neuroscience*, 41(35), 7388–7402.
- Kaneshiro, B., Perreau Guimaraes, M., Kim, H.-S., Norcia, A. M., & Suppes, P. (2015). A representational similarity analysis of the dynamics of object processing using single-trial eeg classification. *Plos one*, 10(8), e0135697.
- Connolly, A. C., Guntupalli, J. S., Gors, J., Hanke, M., Halchenko, Y. O., Wu, Y.-C., Abdi, H., & Haxby, J. V. (2012). The Representation of Biological Classes in the Human Brain. *Journal of Neuroscience*, 32(8), 2608–2618. <https://doi.org/10.1523/JNEUROSCI.5547-11.2012>
- Khaligh-Razavi, S.-M., & Kriegeskorte, N. (2014). Deep Supervised, but Not Unsupervised, Models May Explain IT Cortical Representation. *PLOS Computational Biology*, 10(11), e1003915. <https://doi.org/10.1371/journal.pcbi.1003915>
- Douglas Carroll, J., & Arabie, P. (1998, January). Chapter 3 - Multidimensional Scaling. In M. H. Birnbaum (Ed.), *Measurement, Judgment and Decision Making* (pp. 179–250). Academic Press. <https://doi.org/10.1016/B978-012099975-0.50005-1>
- Heinze, J., Kahnt, T., & Haynes, J.-D. (2011). Topographically specific functional connectivity between visual field maps in the human brain. *NeuroImage*, 56(3), 1426–1436. <https://doi.org/10.1016/j.neuroimage.2011.02.077>
- Mur, M., Ruff, D. A., Bodurka, J., Weerd, P. D., Bandettini, P. A., & Kriegeskorte, N. (2012). Categorical, Yet Graded – Single-Image Activation Profiles of Human Category-Selective Cortical Regions. *Journal of Neuroscience*, 32(25), 8649–8662. <https://doi.org/10.1523/JNEUROSCI.2334-11.2012>
- Vul, E., Lashkari, D., Hsieh, P.-J., Golland, P., & Kanwisher, N. (2012). Data-driven functional clustering reveals dominance of face, place, and body selectivity in the ventral visual pathway. *Journal of Neurophysiology*, 108(8), 2306–2322. <https://doi.org/10.1152/jn.00354.2011>
- Allen, E. J., St-Yves, G., Wu, Y., Breedlove, J. L., Prince, J. S., Dowdle, L. T., Nau, M., Caron, B., Pestilli, F., Charest, I., et al. (2022). A massive 7t fmri dataset to bridge cognitive neuroscience and artificial intelligence. *Nature neuroscience*, 25(1), 116–126.
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014* (pp. 740–755). Springer International Publishing. [https://doi.org/10.1007/978-3-319-10602-1\\_48](https://doi.org/10.1007/978-3-319-10602-1_48)
- De Leeuw, J. (1977). Application of convex analysis to multidimensional scaling. *Recent developments in statistics*, 133–145.
- Kabsch, W. (1976). A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, 32(5), 922–923.
- Motter, B. C. (2009). Central V4 Receptive Fields Are Scaled by the V1 Cortical Magnification and Correspond to a Constant-Sized Sampling of the V1 Surface. *The Journal of Neuroscience*, 29(18), 5749–5757. <https://doi.org/10.1523/JNEUROSCI.4496-08.2009>
- Kumano, H., & Uka, T. (2010). The Spatial Profile of Macaque MT Neurons Is Consistent With Gaussian Sampling of Logarithmically Coordinated Visual Representation. *Journal of Neurophysiology*, 104(1), 61–75. <https://doi.org/10.1152/jn.00040.2010>