A Stratification Matrix Viewer for Analysis of Neural Network Data

Philipp Harth1, Sumit Vohra1, Daniel Udvary2, Marcel Oberlaender2, Hans-Christian Hege1 and Daniel Baum1

1Zuse Institute Berlin (ZIB), Department of Visual and Data-Centric Computing, Germany
2Max Planck Institute for Neurobiology of Behavior – caesar, Bonn, Germany

Abstract
The analysis of brain networks is central to neurobiological research. In this context the following tasks often arise: (1) understand the cellular composition of a reconstructed neural tissue volume to determine the nodes of the brain network; (2) quantify connectivity features statistically; and (3) compare these to predictions of mathematical models. We present a framework for interactive, visually supported accomplishment of these tasks. Its central component, the stratification matrix viewer, allows users to visualize the distribution of cellular and/or connectional properties of neurons at different levels of aggregation. We demonstrate its use in four case studies analyzing neural network data from the rat barrel cortex and human temporal cortex.

CCS Concepts
• Human-centered computing → Visual analytics; • Computing methodologies → Model verification and validation; • Applied computing → Biological networks;

1. Introduction
Brain networks or connectomes [STK05] are graphs whose nodes represent functional units (individual neurons, neuron assemblies, brain regions) and whose links represent structural or functional connections between these units [FZB16]. We focus on neuroanatomy and structural connectivity on the micro- to mesoscale. Previously, we have created a detailed anatomical model of the rat barrel cortex [UHM∗22]. A key motivation for the tool presented here, is to make these data easily accessible online [cis] for other neuroscientists. This requires support for quick and systematic exploration and for identification of specific subsets of the data. Concurrently, we are working on mathematical models for predicting synaptic connectivity with the aim of elucidating principles of neural network architecture. These models are also applied to neural network datasets shared by other researchers. In this context, we needed a tool that supports us in filtering neurons of interest and comparing model-predicted with empirically observed connectivity at different levels of aggregation. We meet these requirements with a versatile tool that supports us in filtering neurons of interest and comparing model-predicted with empirically observed connectivity at different levels of aggregation. We meet these requirements with a versatile tool that supports us in filtering neurons of interest and comparing model-predicted with empirically observed connectivity at different levels of aggregation. We meet these requirements with a versatile tool that supports us in filtering neurons of interest and comparing model-predicted with empirically observed connectivity at different levels of aggregation.

2. Related work
Visualization techniques in connectomics have been reviewed by Pfister et al. [PKB∗14], by Margulies et al. [MBWG13], and recently by Beyer et al. [BTB∗22] for high resolution connectomics,
which is most relevant to our application. We also point to the survey by Behrisch et al. [BBH +16] on matrix reordering techniques and the work by Alper et al. [ABH +13] on weighted graph comparison techniques for brain networks. Early works on zoomable matrix views include Matrix Zoom [AV04] and ZAME [EDG +08].

**Matrix views in connectomics.** Several tools used in the connectomics community integrate matrix views to display connectivity information. Examples are NeuroVISAAS [SE12], Neuro-MAP [SBS +13], and newPrint [PCD +22]. Yang et al. [YSD +17] propose a hybrid visualization of matrices and node-link diagrams for brain network analysis. BrainCove [VMVB12] uses raycasting to render fMRI correlation matrices. The Graffinity tool addresses the analysis of connectivity in very large networks [KLS +17]. A tool specifically designed for the rat barrel cortex features a semi-abstract representation of the barrel field in combination with a heatmap showing synapse counts [DEHO12].

**Complex heatmaps and tabular layouts.** Complex heatmaps extend regular heatmaps by providing complex row and column annotations [GES16]. They gained increased attention in the bioinformatics community and are used in tools like Clustergrammer [FGR +17], which offers zoom, pan, and reordering with animated transitions. Responsive matrix cells [HBS +21] feature interactive charts that are embedded into a heatmap to support focus+context exploration of multivariate graphs. Neo: [GHM +22] presents a flexible nested tabular layout for the visualization of confusion matrices, which are used to evaluate machine learning models. Blocks [WK5 +22] allows generating rich tables via drag-and-drop, supporting data aggregation at user-specified levels of detail.

**Our contribution.** Unlike matrix views presented in previous works in connectomics, the stratification matrix is not a connectivity matrix but represents a more generic tabular layout that can encode both connectional and non-connectional information. This is illustrated in Fig. 1, where rows correspond to cell types and the columns to cortical depths, encoding the spatial distribution of neurons. In contrast, the stratification matrix shown in Fig. 2E represents a hierarchical connectivity matrix showing synapse counts.

### 3. Analysis tasks and requirements

Initial requirements for the tool arose from a request of a neuroscientist seeking detailed information about connectivity between various subpopulations of neurons in the rat barrel cortex [Kle]. We agreed that a flexible tabular interface was needed to provide access to such connectivity parameters at user-defined levels of aggregation. The interface should also support queries of non-connectional parameters, such as the spatial distribution of cells. Additional requirements arose from our ongoing research, testing mathematical models for synaptic connectivity on datasets shared by other researchers. These requirements are reflected in the following analysis tasks: T1: Query the cellular composition of reconstructed neural networks. This involves getting an overview of which types of cells are located where in the neural tissue volume. T2: Group neurons based on morphological features (e.g., axon length) and get an overview of the distribution of such features in the neuron population. T3: Examine connectivity features between pre- and postsynaptic neuron populations, such as the distribution of synaptic contacts between different cell types. T4: Compare connectivity features observed in the reconstructed neural network with those of mathematical models that aim to predict synaptic connectivity.

### 4. Design of visual components

The user interface of the application is shown in Fig. 1. After specifying a dataset in the **Project** panel, the **Selection** panel displays all available properties in two juxtaposed trees that control the row and column annotations of the matrix. The user selects property types and values of interest by toggling the respective check boxes in the tree. Currently, only categorical properties are supported and numeric features are mapped to a fixed number of bins in a preprocessing step. The hierarchical order of properties can be changed via drag-and-drop. If more than one property is selected per row or column, the **Stratification matrix** is hierarchical. When zooming out, lower level properties are collapsed and the corresponding values in the matrix are aggregated. Rows and columns can be zoomed simultaneously or separately. Matrix cells display the number of neurons, synapses, or other entities, depending on the dataset (see Supplementary Information). The visual encoding of matrix cells is controlled in the **Data channels** panel. A predefined set of color scales is provided, which can be normalized by the following maximal values: (1) global, (2) per row, and (3) per column. If the dataset provides more than one value per matrix cell (as in comparison tasks [ABH +13]), a second channel can be activated. In this case the matrix cells are split diagonally. The **Toolbox** provides utility functions to interact with the matrix and allows the inspection of neuron morphologies in an external anatomical viewer. Selected matrix cells are displayed in a bar chart, enabling direct comparison of values that do not belong to adjacent tiles. We support automatic matrix reordering [BBH +16] with an optimal leaf ordering algorithm from the reorder.js library [reo]. In two-channel mode, the user specifies which channel to use for reordering, including a mode in which the unsigned difference between both channels is used.

### 5. Implementation

The front-end is based on Meteor, a full-stack web development framework. The property selection panel uses the react-complex-tree. The stratification matrix viewer is based on the vector graphics library Paper.js and we use reorder.js for matrix reordering. All back-end routines were implemented in Python using Flask as web server. The neural network datasets were preprocessed to create precomputed NumPy arrays and boolean masks for indexing and intersecting property values, enabling fast query execution and retrieval of matrix data. The code is available at Github: [https://github.com/zibneuro/stratification-matrix](https://github.com/zibneuro/stratification-matrix)

### 6. Evaluation

Our evaluation is based on four case studies (Sect. 6.1) and a user survey (Sect. 6.2). The case studies use two recent neural network datasets to illustrate how the outlined analysis tasks T1-T4 can be solved with the tool. The first dataset is a detailed anatomical model of the **rat barrel cortex (RBC)** containing 550,000 neurons...
neurons and 5 billion synapses in a volume of 6.5 mm³ covering 24 cortical columns [UHM 22]. The second dataset is an electron microscopy-based reconstruction of 1 mm³ of the human temporal cortex (HTC) (Fig. 1B); it contains 57,000 cells of which 16,000 are neurons and 133 million synapses [SCJB 21].

6.1. Case studies

The case studies summarized here can be reproduced in the online version of the tool [cis]: timestamps for the accompanying video are provided at the end of each paragraph.

Case Study 1. The goal of the user is to understand the distribution of different cell types across different regions in the RBC (T1). The matrix cells encode the number of neurons (non-connectional information). The matrix in Fig. 2A shows a 12 × 24 grid, with rows corresponding to the 12 cell types and columns corresponding to the 24 cortical columns in the barrel cortex. The visual pattern shows that most neurons per cell type reside in the D and E columns. Activating the subregion property, which subdivides cortical columns into inside and septum, yields a 12 × [24/2] or 12 × (2/24) grid, depending on the property order. The resulting view shows that L4sp and L4ss neurons do not reside in the septum. Applying column-independent zoom collapses the matrix into a 12 × 2 grid (cell type × subregion) that gives an aggregated view, which quickly reveals that for all cell types except L3PY and L4PY, more neurons reside inside the cortical column than in the septum. Stratifing cell type by cortical depth of the soma produces the matrix shown in Fig. 1A. It shows that cell types occupy different depth ranges and overlap with each other to various degrees. For example, L4PY cells reside in a narrow depth range of 500-800 µm, whereas L6CT cells reside in a broad depth range of 1000-2200 µm. [video 0:55]

Case Study 2. The goal of the user is to obtain a breakdown of the cellular composition in the HTC and to identify individual neurons with long axon segments for further analysis (T1, T2). Matrix cells encode the number of neurons and glial cells in the human temporal cortex (non-connectional information). Stratifying the cells by cell type and layer shows that most of the cells are glial cells residing in the white matter. Adding misc annotations as property, reveals that most of these glial cells are oligodendrocytes and that this type outnumbers all other types, excluding astrocytes in the upper layers. Stratifying neuronal cells in a similar way shows that layer 1 is populated exclusively by inhibitory interneurons and that in all other layers excitatory pyramidal neurons outnumber the interneurons. Neurons classified as bipolar mainly reside in layers 5 and 6. Stratifying neurons by their morphological properties axon length (NAs) and dendrite length (NDe) shows that these quantities are correlated but also reveals that axons are much shorter than dendrites or completely missing. This is likely an artefact of the automated segmentation. To filter out neurons with large axons and a high number of outgoing synapses, the user selects all neurons grouped by layer for the rows of the matrix and the four highest bins of NAs (1745, 2171, 2664, 13177) combined with the number of outgoing synapses NSO ≥ 100 as columns in the matrix (Fig. 2B). Selecting matrix cells yields the corresponding neurons, which can then be displayed in an external anatomical viewer. [video 2:05]

Case Study 3. The goal of the user is to understand the characteristics of neuronal connections in the RBC (T3). Matrix cells encode the number of connected/unconnected neuron pairs in the rat barrel cortex (connectional information). Stratifying the neuron pairs by presynaptic cell type (rows) and postsynaptic cell type, connected/unconnected (columns) shows that most neuron pairs are in fact unconnected and that this is the case for all combinations of pre- and postsynaptic cell types. A known determinant for synaptic connectivity is the intersomatic distance of pre- and postsynaptic neurons. Stratifying the neuron pairs by synapse count and intersomatic distance results in the visual pattern shown in Fig. 2C. The first row shows that the number of unconnected neuron pairs (synapse count 0) increases with intersomatic distance. Conversely,
the number of connected pairs (one or more synapses) decreases with intersomatic distance. [video 5:26]

Case Study 4. The goal of the user is to inspect synaptic connectivity patterns in the HTC and compare them to connectivity patterns predicted by a mathematical model (T3, T4). The matrix cells encode the number of synapses (channel 1: empirically observed, channel 2: model-predicted). The user selects pre layer for the rows in the matrix and post layer for the columns. Most empirically observed synapses connect neurones residing in the same layer or in neighboring layers, which can be seen as a diagonal structure in the matrix (Fig. 2D). At this high level of aggregation, the model-predicted synapse counts are in good agreement with the empirically observed counts. Activating the cell type and compartment properties gives a hierarchical matrix that allows a more fine-grained comparison of empirical and model-predicted synapse counts (Fig. 2E). The compartment property describes at which specific location a synapse is formed at a postsynaptic neuron. At this subcellular resolution, limitations of the model become apparent, as recognizable by the many matrix cells with deviating values (Fig. 2E). To understand where the model fails, the matrix reordering functionality is applied, using the unsigned difference between channel 1 and 2 as target value. The resulting block structure of matrix cells provides a compact view of the synaptic connections with high deviations between empirical data and the model (Fig. 2E). Visual inspection reveals that the model systematically underestimates synaptic inputs from interneurons onto the axon initial segment and the soma of pyramidal neurons. Conversely, connections from pyramidal neurons onto these compartments are overestimated by the model. [video 6:45]

6.2. User survey

To evaluate and further improve the tool, we carried out a user survey. Of the 25 people invited, which include collaborators in the field of computational neuroscience (including data science & machine learning) and members of our own visualization research group, 14 participated in the survey. We excluded one neuroscientist from the respondents due to his later co-authorship. The breakdown of the remaining participants by their self-ascribed field of work is: neuroscience (4), visualization (5), data science/machine learning (3), computer science (1). Eight participants were male, four female, and one preferred not to say.

We first asked the participants to watch a 6:17 minutes video explaining the intended usage of the tool (see additional material) and invited them to explore it online. Participants then had to rate the following statements on a 5-item Likert scale [Lik32]: Q1: Overall, I thought the tool was easy to use. Q2: I thought the matrix interaction functions (e.g., zoom) were easy to use and I found the utility functions provided in the toolbox were appropriate. Q3: The tool would make it easier to explore and download structural parameters of neural networks (e.g., the cellular composition of the barrel cortex). Q4: The tool would make it easier to communicate quantitative insights about neural networks. Q5: The tool itself and/or the concept of a stratification matrix could be of interest in my own ongoing research (including fields other than neuroscience), if the necessary adaptions were made.

Q1, Q2 were only answered by those who explored the tool online; participants could give optional free text comments. The results of the numeric survey items are shown in Fig. 3. The usability and feature scope of the tool (Q1, Q2) as well as its expected utility for the intended application (Q3, Q4) were rated positively. The potential utility of the tool for their own ongoing research, including fields other than neuroscience, received mixed responses by the participants (Q5). One neuroscientist wrote: “The tool is already applicable to my own research as I am working on the rat barrel cortex connectivity data and the tool makes it very easy to get connectivity information.” One computer scientist reasoned about the potential of stratification matrices to communicate research findings in a standardized form: “... when I see the tool I think that it – more precisely, specific well-expressive selections with well-accepted colour maps – could become some kind of standard for communicating what was researched in a paper, ...”

7. Conclusion

We have presented a matrix viewer that supports the analysis of both connectional (e.g., synaptic contacts) and non-connectional (e.g., cellular composition) information in neural network data. Depending on the nature of the data and the visual encoding chosen by the user, the stratification matrix takes the form of a hierarchical connectivity matrix, a nested table, or a heatmap. This flexibility will render the tool versatile for quite different applications, not only in the context of neurobiological research. The stratification matrix viewer was positively received by participants of a survey, which encouraged us to make the tool available through BarrelCorTextInSilico [cis], where it now facilitates access to structural parameters of the rat barrel cortex [UHM22]. The tool helps us to identify neurons of interest in a dense reconstruction of the human temporal cortex [SCB21], revealed interesting spatial distributions of neuronal and glial cells, and allowed us to compare a mathematical model of synaptic connectivity with empirical observations. We believe that the tool could be of use also in other domains, as illustrated in Fig. 2F, were we apply it to an epidemiological dataset (see Supplementary Information). Future work will include automatic zoom-dependent binning of numerical properties, a tighter integration with external anatomical viewers, and various usabil-
ity and design improvements suggested by survey participants and reviewers.

Acknowledgements
Deutsche Forschungsgemeinschaft (SPP 2041 Computational Connectomics) under award number 347210657 and National Institute of Neurological Disorders and Stroke (NINDS) of the National Institutes of Health under award number 1U19NS104653.

References


[cis] Barre`CortexinSilico. URL: https://cortexinsilico.zib.de/matrix. 1.3.4


[LiK32] LIKERT R.: A technique for the measurement of attitudes. Arch. Psychol. 22 (1932), 5–55. 4


[re] reorder.js. URL: https://github.com/jdfekete/reorder.js. 2


