

This is a README file with information of what each folder of this database contains and also, has information about the mice of the database.

=====
We have 5 mice:

Mouse 3 with id 24617

Mouse 4 with id 24705

Mouse 5 with id 25133

Mouse 6 with id 25341

Mouse 7 with id 25387

=====
Mice Information:

In the link below there are information on frames and more for mice 3 and 4 :
https://sideridis.netlify.app/#1_Datasets (the layer depths in here are not updated. The correct ones are in the section below)

IMPORTANT: The frame start and end of the Spontaneous period for each mouse is based on R indexing, which starts from 1 by default. If you want to reference the frames in python, you must subtract 1 (e.g. mouse 3 python index: spontaneous = (26919 - 1):(49988 - 1)), because in python the default index starts from 0.

Information about mouse 3,4,5,6,7 are below:

mouse 3 (24617):

spontaneous = 26919:49988

session = 3

scanIndex = 10

framesampling = 6.30072 Hz

NumberOfCells = 9314

mouse 4 (24705):

spontaneous = 26905:50052

session = 3

scanIndex = 24

framesampling = 6.30072 Hz

NumberOfCells = 7436

mouse 5 (25133):

spontaneous = 26955:50141

session = 6

scanIndex = 15

framesampling = 6.30072 Hz

NumberOfCells = 8689

mouse 6 (25341):

spontaneous = 27075:50131

session = 2
scanIndex = 20
framesampling = 6.30072 Hz
NumberOfCells = 7962

mouse 7 (25387):
spontaneous = 26976:49944
session = 3
scanIndex = 17
framesampling = 6.30072 Hz
NumberOfCells = 8237

=====
Layers:

Layer 2 neurons are neurons with z coordinate: $80 \leq z \leq 210$.
Layer 3 neurons are neurons with z coordinate: $285 \leq z \leq 330$.
Layer 2/3 neurons are neurons with z coordinate ≤ 340 .
Layer 4 neurons are neurons with z coordinate: $450 \geq z \geq 400$.

=====
The name of the datasets/files consist of: the mouse number(3-7) or mice id as shown above, what they represent (spiketrains, sttc etc.) and any exclusion criteria that was used to form this dataset (0.01Hz, V1 area, 15um, 0 or 2 dt, 1.5dc, L2/3 or L4). If a dataset has the layer in the name (for example: mouse3_L23) , then this dataset refers to layer 2/3 and the neurons are selected with exclusion 15um,V1,0.01Hz,1.5dc. If a dataset has the name mouse3_L23_L23 this means the dataset refers to L2/3 -> L2/3 or if the name is mouse3_L4_L23 refers to L4 -> L2/3 .

=====
In the link below are information about metrics we use and how we use them and also, definitions about variables we have created (example: STTC Groups):

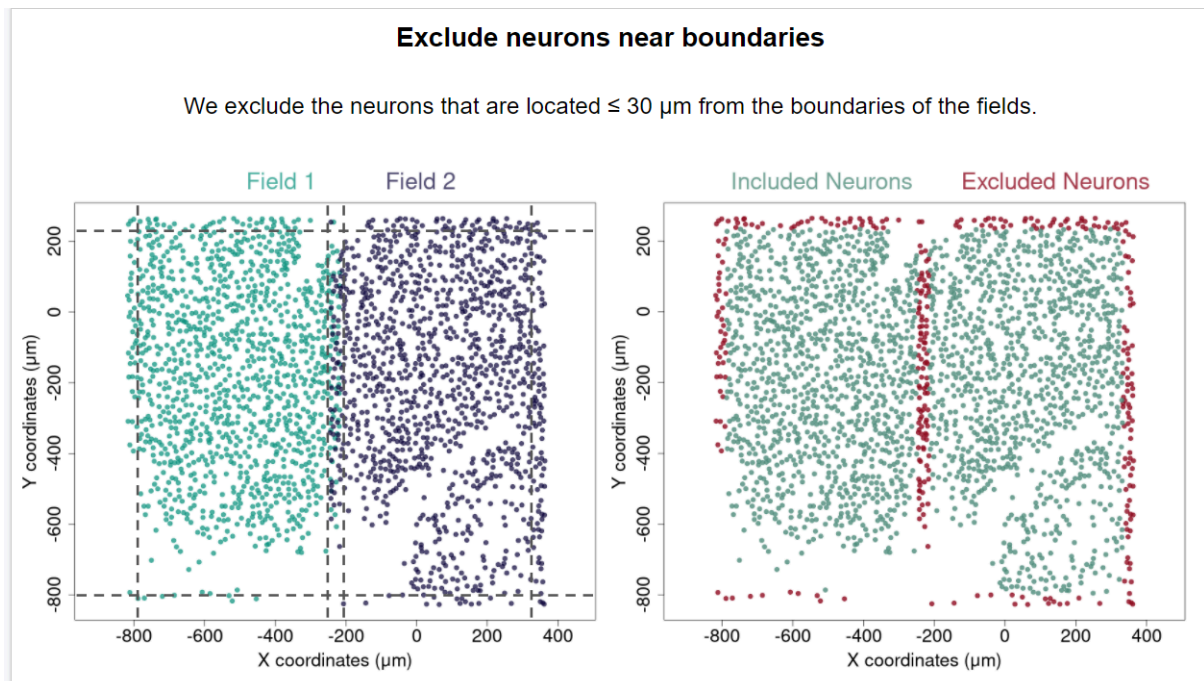
https://docs.google.com/document/d/1q8WKsf8mMhtV80IvTzju8RfCdIK6PzuCj4d_hUveQq0/edit?usp=sharing

=====
README structure:

Folder name:
Information about what this folder contains

00_30um data (old -NOT IN USE):

30um is an important exclusion criterion of all our analysis. At some point the criterion changed from 30um to 15um, so every dataset that was in the database and it was computed on 30um was stored here , if we need to reference something in the future. The way the neurons 30um away from boundaries are computed is shown below:



00_old layer depths data (NOT IN USE):

Neurons are split into layers. In our analysis, we focus primarily on L2/3 and L4. The layers are classified depending on their depth. This classification changed (at some point), so every dataset that was computed with the old layers is now stored here for future reference.

Behavioural Data:

Here are the behavioural data of each mouse. The behavioural data are about pupil radius (pixels) and treadmill velocity (cm/s) of the mice. Pupil radius computed as 'The ellipse is fitted by a deep lab cut into the pupil, then a long radius of the ellipse is used'. Information about behavioural data can be found here:

<https://sideridisvasileios.gitbook.io/tnl-lab/dataset-format#behavioural-data>

Inside Behavioural Data, there is a folder named Behavioral Data Cleaned which contains only the spontaneous period frames for every mouse. Also, it contains one value of behavioural information of each frame and because we have sometimes more values for one frame - this value is the mean of all the values of the specific frame.

CellMembership: (Area Information)

Here is the information about the area of the visual cortex in which each neuron belongs, namely V1 RL AL LM. Information about cell membership can be found here:

<https://sideridisvasileios.gitbook.io/tnl-lab/dataset-format#cell-membership>

Clustering Coefficient:

Here is the computed clustering coefficient for each layer and each mouse. Information of how clustering coefficient is computed can be found here:

https://docs.google.com/document/d/1q8WKsf8mMhtV80IvTzju8RFCdIK6PzuCj4d_hUveQq0/edit#heading=h.kill3ujth7jx

	A	B	C
1	ref_neuron	cc_obs	cc_null
2	368	0	0
3	399	0.01666666667	0.05666666667
4	400	0	0.01904761905
5	401	0.0009250693802	0.007400555042
6	402	0	0.09090909091
7	403	0	0.01056338028
8	420	3.32E-05	0.008262817322
9	422	0	0.1666666667
10	423	0.003278688525	0.0262295082

10 rows example of the dataset. Columns:

ref_neuron: neuronID of the reference neuron.

cc_obs: observed clustering coefficient value computed as shown in the link above.

cc_null: null clustering coefficient value computed as shown in the link above.

Coordinates:

Here are the x,y,z coordinates of each neuron of each mouse. More information can be found here:

<https://sideridisvasileios.gitbook.io/tnl-lab/dataset-format#coordinates>

Datasets for Hyperactive Epochs (Spike trains and STTC):

Here are the spike trains and STTC datasets for the hyperactive epochs (burst and no-burst). Examples of STTC and Spike train datasets are shown in the [STTC](#) and [Spike train](#) folder documentation. Information about the definition of hyperactive epochs can be found here:

https://docs.google.com/document/d/1q8WKsf8mMhtV80IvTzju8RFCdIK6PzuCj4d_hUveQq0/edit#heading=h.bulgfikfibe

Deconvolved:

Here is the deconvolved signal for each mouse. More information can be found here:

<https://sideridisvasileios.gitbook.io/tnl-lab/dataset-format#deconvolved>

DF over F:

Here is the df over f signal of each mouse. More information can be found here:

<https://sideridisvasileios.gitbook.io/tnl-lab/dataset-format#df-f>

Direction Tuning:

Here is the direction tuning (direction and orientation preference) for each mouse. (19 January 2022 - only mice 3,4,5 have orientation preference data). Inside this folder, there is a README file that explains the datasets of each mouse.

Fluorescence:

Here is the fluorescence signal of each mouse. More information can be found here:

<https://sideridisvasileios.gitbook.io/tnl-lab/dataset-format#fluorescence>

HighMembershipNeurons:

Here are the high membership neurons for each layer of each mouse.

High membership neurons definition can be found here:

https://docs.google.com/document/d/1q8WKsf8mMhtV80IvTzju8RFCdIK6PzuCj4d_hUveQq0/edit#heading=h.cvetowvcomvd

	A
1	NeuronID
2	402
3	423
4	447
5	448
6	451
7	475
8	478
9	479
10	481

10 rows example of the dataset. Columns:

NeuronID: neuronID of high membership neurons

Neurons per field:

The most important and basic data of this dataset are coming to us in fields. Each mouse has 8 fields, so in this folder are the neurons that correspond to each field. In the preprocessing, the fields are combined in one dataset, so all datasets in the database are whole. (f1 in the name of the dataset is field 1, f2 is field 2, etc)

	A
1	1273
2	1274
3	1275
4	1276
5	1277
6	1278
7	1279
8	1280
9	1281
10	1282

10 rows example of the dataset. Columns:

first column: neuronID of the field we are looking

Neurons per layer:

Here are the neurons that belong to each layer of each mouse. Exclusion criteria for the computation of the neurons per layer are: V1 area, 15um, 0.01Hz, 1.5dc.

	A
1	Neurons
2	133
3	135
4	136
5	137
6	138
7	139
8	140
9	141
10	142

10 rows example of the dataset. Columns:

Neurons: neuronID of neurons of the mouse and layer we are looking

Normalized Degree of Connectivity:

Here is the normalized degree of connectivity , computed for all mice and layers, also computed for correlated (edges with z-score > 4) and anticorrelated (edges with z-score < -4). Information about the computation of degree of connectivity can be found here:

https://docs.google.com/document/d/1q8WKsf8mMhtV80IvTzju8RFCdIK6PzuCj4d_hUveQq0/edit#heading=h.63yiut3mxvig

	A	B
1	neuron	ndoc
2	368	0.08280254777
3	399	0.1809666542
4	400	0.1577369801
5	401	0.3109778943
6	402	0.2150618209
7	403	0.2322967404
8	420	0.3184713376
9	422	0.009741476208
10	423	0.2896215811

10 rows example of the dataset. Columns:

neuron: neuronID of reference neuron

ndoc: normalized degree of connectivity of that reference neuron computed as the link above shows.

Quantile Spiketrains and STTC:

Here are the population and pupil quantile spiketrains and STTC datasets. These datasets are based on the original spiketrains, but split into 4 categories based on their population cofiring for population or pupil radius for pupil. The categories are (Q = Quantile) Q0-Q25, Q25-Q50, Q50-Q75, Q75-Q100. Examples of STTC and Spiketrain datasets are shown in the [STTC](#) and [Spiketrain](#) folder documentation.

Information about the quantiles are here:

https://docs.google.com/document/d/1q8WKsf8mMhtV80IvTzju8RFCdIK6PzuCj4d_hUveQq0/edit#heading=h.9nygk4ksj2q8

RawFluorescence:

Here is the raw fluorescence signal for each mouse. More information can be found here:

<https://sideridisvasileios.gitbook.io/tnl-lab/dataset-format#raw-fluorescence>

Robustness of stat. sign. edges (STTC and Spiketrains):

Here are the spiketrains and STTC of robustness. Each mouse has 12 spiketrain and 12 STTC files. Each file corresponds to the activity of the mouse in the first 5,10,15,20,25....55,60 minutes of the spontaneous recording. Examples of STTC and Spiketrain datasets are shown in the [STTC](#) and [Spiketrain](#) folder documentation.

RQA Periods (Spiketrains and STTC):

Here are the spiketrains and STTC datasets for the RQA periods. Recurrence quantification analysis (RQA) is a method of nonlinear data analysis (cf. chaos theory) for the investigation of dynamical systems. It quantifies the number and duration of recurrences of a dynamical system presented by its phase space trajectory. Examples of STTC and Spiketrain datasets are shown in the [STTC](#) and [Spiketrain](#) folder documentation.

Spiketrains (Eventograms):

Here are the spiketrains of all mice. More information can be found here: <https://sideridisvasileios.gitbook.io/tnl-lab/dataset-format#spiketrains-eventograms>

STTC:

Here are the STTC datasets of all mice. More information about the STTC can be found here:

<https://sideridisvmsasileios.gitbook.io/tnl-lab/dataset-format#sttc-weight-s-matrix>

Also, sttc was computed using this method and steps here:

<https://sideridisvasileios.gitbook.io/tnl-lab/extract-spiketrains#set-up-and-run-sttc-algorithm>

STTC Groups:

Here are the STTC group (relaxed, strict and unidirectional) datasets for each mouse and each layer, for 0dt,2dt - anticorrelated and correlated. There is also a README file inside this folder that explains the way the definitions of the sttc groups. Information about sttc groups' definition can be found here:

https://docs.google.com/document/d/1q8Wksf8mMhtV80IvTzju8RFCdIK6PzuCj4d_hUveQq0/edit#heading=h.uazffznk4ns0

	Reference_neuron	STTC_members	Group_size	Group_size_normalized	STTC_mean	STTC_median	mean_size	median_size	max_size	seventy_percentile_size	ninty_percentile_size	nintyfive_percentile_size	ref_neuron_with_max_size
1	368	4233_4239_4263_4278_4279_4281_4306_4317_4338_4397_4...	110	0.092281879	0.03487388	0.03344505	0.19361695	0.11818182	0.9090909	0.26363636	0.5363636	0.6545455	732
2	399	4210_4218_4224_4233_4234_4236_4237_4239_4246_4247_4...	222	0.186241611	0.03839842	0.03594380	0.16079123	0.08558559	0.8603604	0.20720721	0.4459459	0.5675676	6300
3	400	4226_4237_4239_4247_4249_4262_4263_4264_4281_4293_4...	273	0.229026846	0.04220270	0.04104180	0.20361167	0.10256410	0.9047619	0.28571429	0.5831502	0.6959707	3625
4	401	4226_4239_4240_4247_4257_4258_4262_4263_4264_4265_4...	473	0.396812081	0.04652113	0.04366570	0.18630633	0.09936575	0.9112051	0.25919662	0.5137421	0.6279070	3625
5	402	4264_4265_4268_4293_4294_4299_4300_4305_4318_4321_4...	386	0.323825503	0.04594607	0.04186660	0.19093623	0.08548223	0.9430052	0.24352332	0.5756477	0.7046632	1069
6	403	4226_4249_4262_4264_4265_4263_4293_4294_4296_4299_4...	397	0.333053691	0.04573966	0.04260150	0.18288338	0.09068010	0.9369521	0.25440806	0.5768262	0.6992443	557
7	420	4215_4226_4240_4262_4264_4265_4268_4266_4289_4294_4...	564	0.473154362	0.07297866	0.06662840	0.17028385	0.07801418	0.8847518	0.21099291	0.5212766	0.6294526	446
8	422	4520_4879	2	0.001677852	0.03242350	0.03242350	0.10621956	0.00000000	1.0000000	0.00000000	0.50000000	0.50000000	420
9	423	4226_4233_4239_4249_4257_4262_4263_4266_4268_4281_4...	439	0.368288591	0.04435920	0.04157810	0.18755115	0.10478360	0.8678815	0.26651481	0.5079727	0.6127563	3625
10	424	4241_4265_4268_4293_4295_4296_4298_4299_4300_4302_4...	291	0.244127517	0.04691694	0.04427010	0.19517845	0.06872852	0.9656357	0.23367698	0.6391753	0.7594502	420

10 rows example of the dataset. Columns:

Reference_neuron: neuronID of reference neuron

STTC_members: the sttc members of the group as a long string with ` _ ` in between neuronIDs

Group_size: the size of the sttc group

Group_size_normalized: the size sttc group normalized. if the sttc group is in the same layer as the reference neuron then the normalized size is group_size / (number of neurons of that layer - 1), else the normalized size is group_size / number of neurons of that layer

STTC_mean: mean STTC value of the edges from the sttc group members to the reference neuron

STTC_median: median STTC value of the edges from the sttc group members to the reference neuron

!Every size from here on references the overlap size of each group with all the others, which is calculated as the the intersection of the sttc group of the reference neuron with all the others / reference's neuron group size!

mean_size: mean of the overlap size of the reference neuron

median_size: median of the overlap size of the reference neuron

max_size: max of the overlap size of the reference neuron

seventy_percentile_size: seventy percentile of the overlap size of the reference neuron

ninty_percentile_size: ninety percentile of the overlap size of the reference neuron

nintyfive_percentile_size: ninety five percentile of the overlap size of the reference neuron

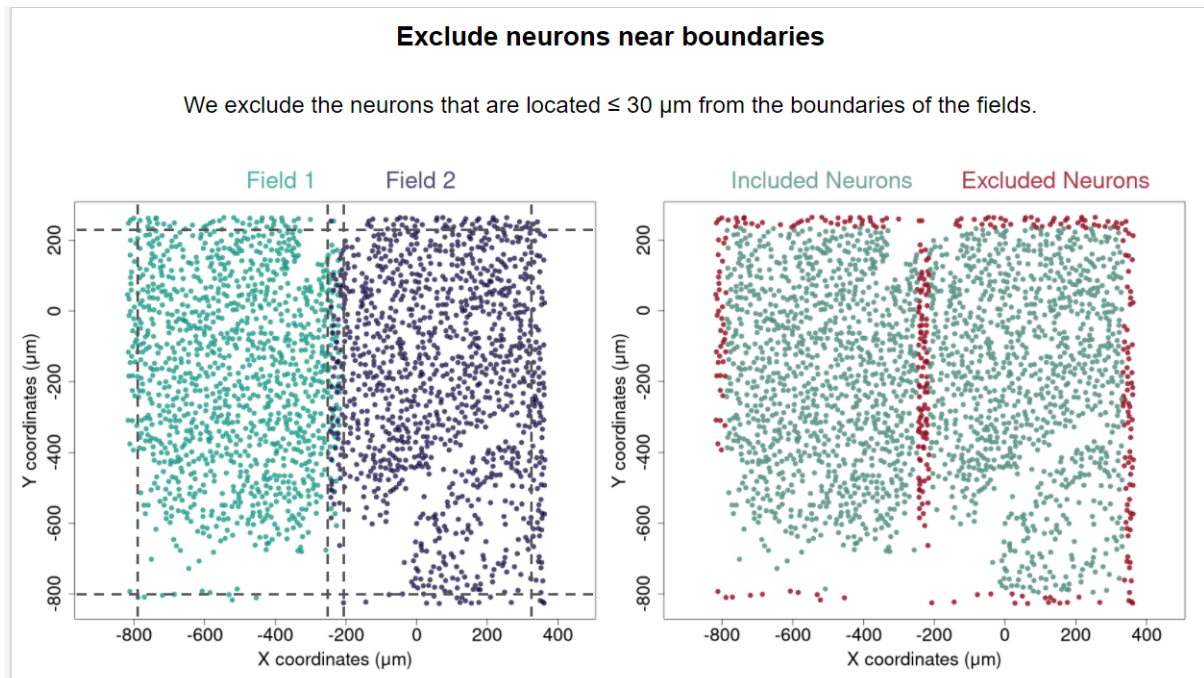
ref_neuron_with_max_size: the neuronID of the neuron that has the maximum overlap size with the reference neuron

WithinBoundaries:

Here are the neurons for each mouse that are within the 15um boundaries that we are using as an exclusion criteria for our analysis.

More information can be found here:

<https://sideridisvasileios.gitbook.io/tnl-lab/dataset-format#within-boundaries-of-field-of-view-fov> (this link was made for 30um datasets. Tthe structure of the dataset is the same, the only thing changes is 30um to 15um).



The picture above shows how the 30um,15um or any um away from the boundaries is calculated and which neurons are included.