Neural Data Science

Lecturer: Prof. Dr. Philipp Berens

Tutors: Jonas Beck, Ziwei Huang, Rita González Márquez

Summer term 2023

Name: Kathrin Root, Alexander Wendt, Patrick Weygoldt

Coding Lab 4

If needed, download the data files `nds_cl_4_*.csv` from ILIAS and save it in the subfolder `../data/`. Use a subset of the data for testing and debugging, ideally focus on a single cell (e.g. cell number x). The spike times and stimulus conditions are read in as pandas data frames. You can solve the exercise by making heavy use of that, allowing for many quite compact computationis. If you need help on that, there is lots of documentation and several good tutorials are available online. Of course, converting the data into classical numpy arrays is also valid.

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import matplotlib as mpl
import numpy as np
import scipy.optimize as opt

from scipy import signal as signal

import itertools
# %matplotlib qt6

%load_ext jupyter_black

%load_ext watermark
%watermark --time --date --timezone --updated --python --iversions --watermark -p sklearn
```

```
<IPython.core.display.HTML object>
```

```
Last updated: 2023-05-16 17:59:28CEST
```

```
Python implementation: CPython
```

```
Python version       : 3.11.3
IPython version      : 8.11.0

sklearn: 0.0.post1

pandas    : 1.5.3
numpy     : 1.24.3
matplotlib: 3.7.1
seaborn   : 0.12.2
scipy     : 1.10.1

Watermark: 2.3.1
```

```
plt.style.use("../matplotlib_style.txt")
```

Load data

```
spikes = pd.read_csv("../data/nds_cl_4_spiketimes.csv")  # neuron id, spike time
stims = pd.read_csv("../data/nds_cl_4_stimulus.csv")  # stimulus onset in ms, direction

stimDur = 2000.0  # in ms
nTrials = 11  # number of trials per condition
nDirs = 16  # number of conditions
deltaDir = 22.5  # difference between conditions

stims["StimOffset"] = stims["StimOnset"] + stimDur
```

We require some more information about the spikes for the plots and analyses we intend to make later. With a solution based on dataframes, it is natural to compute this information here and add it as additional columns to the `spikes` dataframe by combining it with the `stims` dataframe. We later need to know which condition (`Dir`) and trial (`Trial`) a spike was recorded in, the relative spike times compared to stimulus onset of the stimulus it was recorded in (`relTime`) and whether a spike was during the stimulation period (`stimPeriod`). But there are many options how to solve this exercise and you are free to choose any of them.

```
# you may add computations as specified above
spikes["Dir"] = np.nan
spikes["relTime"] = np.nan
spikes["Trial"] = np.nan
spikes["stimPeriod"] = np.nan
```

```
dirs = np.unique(stims["Dir"])
trialcounter = np.zeros_like(dirs)

for i, row in stims.iterrows():
    trialcounter[dirs == row["Dir"]] += 1

    i0 = spikes["SpikeTimes"] > row["StimOnset"]
    i1 = spikes["SpikeTimes"] < row["StimOffset"]

    select = i0.values & i1.values

    spikes.loc[select, "Dir"] = row["Dir"]
    spikes.loc[select, "Trial"] = trialcounter[dirs == row["Dir"]][0]
    spikes.loc[select, "relTime"] = spikes.loc[select, "SpikeTimes"] - row["StimOnset"]
    spikes.loc[select, "stimPeriod"] = True

spikes = spikes.dropna()
```

```
spikes.head()
```

	Neuron	SpikeTimes	Dir	relTime	Trial	stimPeriod
514	1	15739.000000	270.0	169.000000	1.0	True
515	1	15776.566667	270.0	206.566667	1.0	True
516	1	15808.466667	270.0	238.466667	1.0	True
517	1	15821.900000	270.0	251.900000	1.0	True
518	1	15842.966667	270.0	272.966667	1.0	True

Task 1: Plot spike rasters

In a raster plot, each spike is shown by a small tick at the time it occurs relative to stimulus onset. Implement a function `plotRaster()` that plots the spikes of one cell as one trial per row, sorted by conditions (similar to what you saw in the lecture). Why are there no spikes in some conditions and many in others?

If you opt for a solution without a dataframe, you need to change the interface of the function.

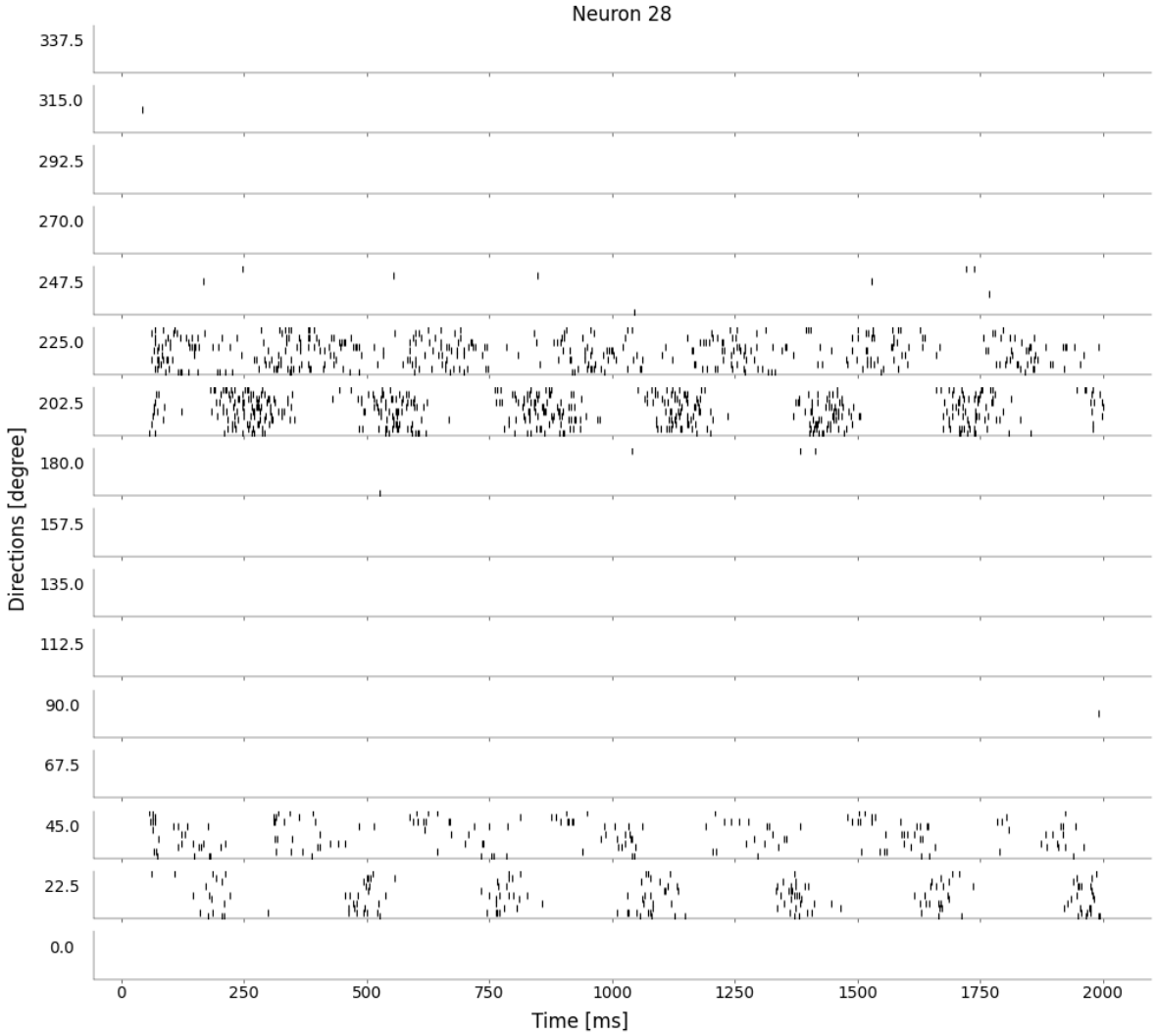
Grading: 2 pts

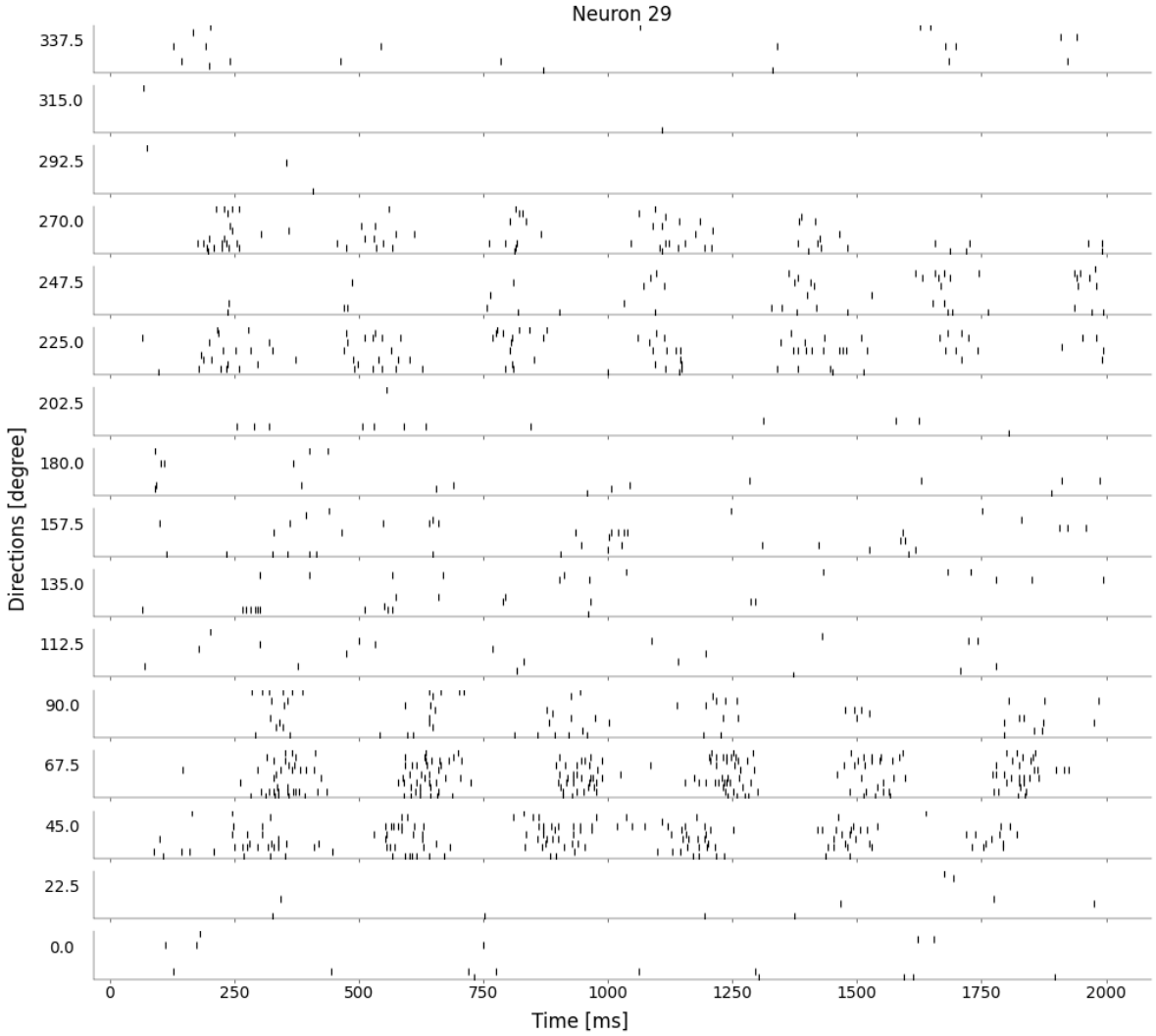
```
def plotRaster(spikes, neuron):
    """plot spike rasters for a single neuron sorted by condition

    Parameters
    ----------

    spikes: pd.DataFrame
        Pandas DataFrame with columns
            Neuron | SpikeTimes | Dir | relTime | Trial | stimPeriod

    neuron: int
        Neuron ID


    Note
    ----

    this function does not return anything, it just creates a plot!
    """

    spikes_neuron = spikes[spikes["Neuron"] == neuron]
    dirs_all = np.sort(spikes["Dir"].unique())[::-1]
    # sort the spikes by direction
    spikes_neuron = spikes_neuron.sort_values(by="Dir")
    # plot the spikes for each trial in a raster plot
    dirs = spikes_neuron["Dir"].unique()[::-1]

    fig, ax = plt.subplots(len(dirs_all), figsize=(10, 9), sharex=True)

    for d, directions in enumerate(dirs_all):
        spikes_neuron_dir = spikes_neuron[spikes_neuron["Dir"] == directions]
        ax[d].scatter(
            spikes_neuron_dir["relTime"],
            spikes_neuron_dir["Trial"],
            marker="|",
            color="k",
            s=20,
            linewidths=0.8,
        )
        ax[d].set_ylabel(f"{directions}", rotation=0, labelpad=20)
        ax[d].set_yticks([])
```

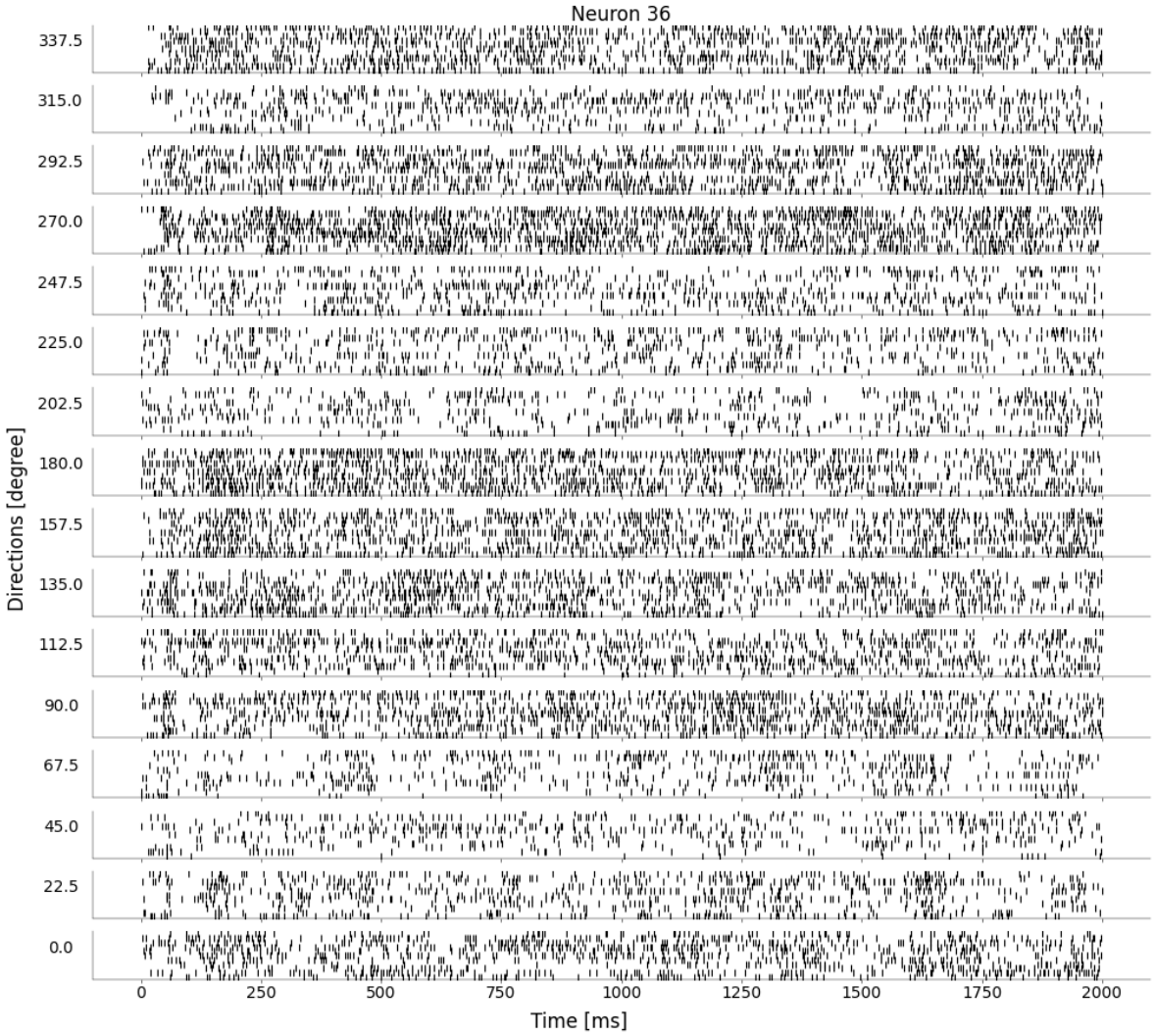
```
    ax[0].set_title(f"Neuron {neuron}")
    fig.supylabel("Directions [degree]")
    # create supylabel on the right side

    # fig.supylabel("Trials n")
    fig.supxlabel("Time [ms]")

    # insert your code here
    # stim direction should be on the y-axis and time on the x-axis
    # you can use plt.scatter or plt.plot to plot the responses to each stim
```

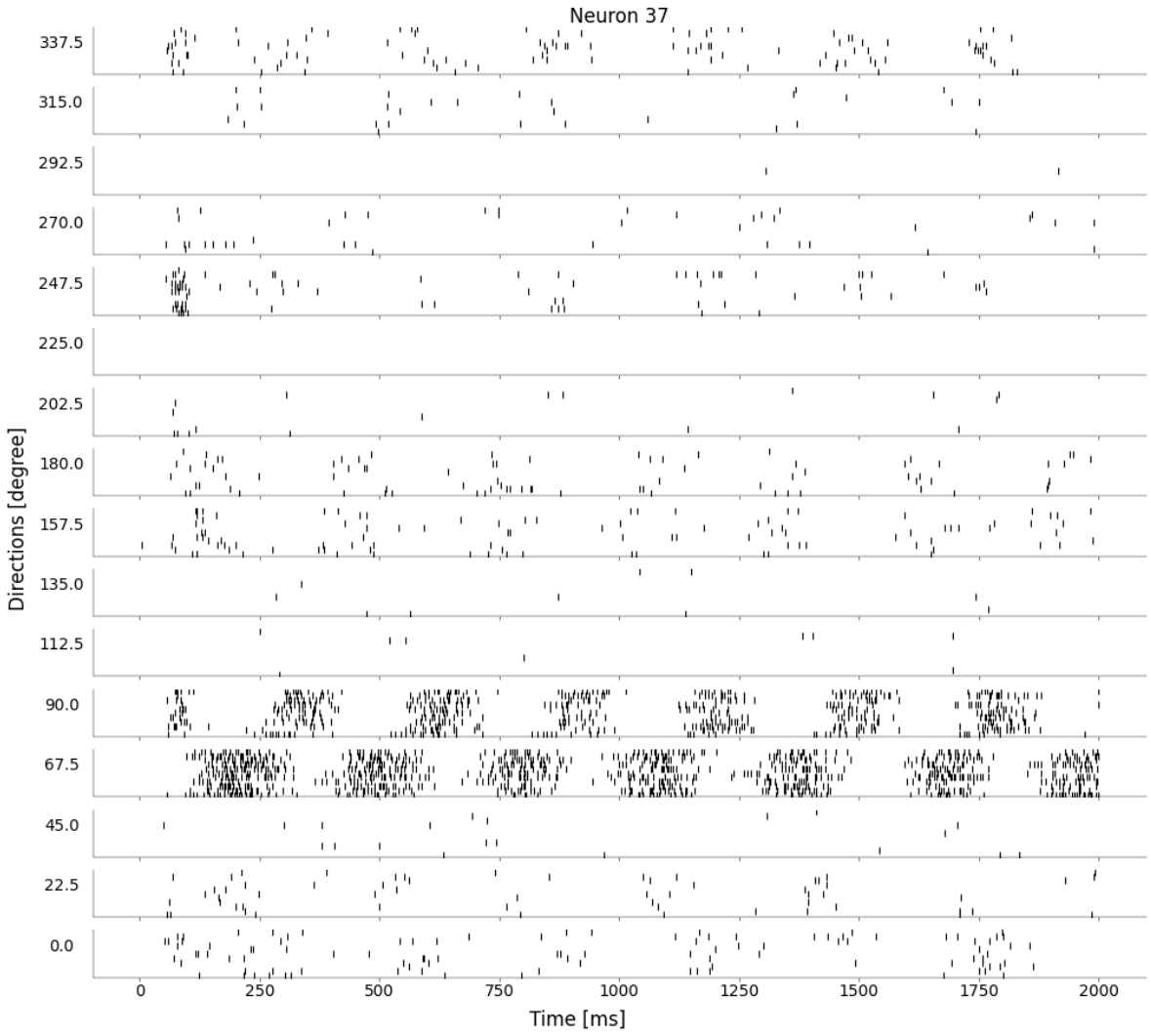
Show examples of different neurons. Good candidates to check are 28, 29, 36 or 37.

```
plotRaster(spikes, 28)
plotRaster(spikes, 29)
plotRaster(spikes, 36)
plotRaster(spikes, 37)
```


Neuron 28
337.5
315.0
292.5
270.0
247.5
225.0
202.5
180.0
157.5
135.0
112.5
90.0
67.5
45.0
22.5
0.0
Directions [degree]
0
250
500
750
1000
1250
1500
1750
2000
Time [ms]


Neuron 29
337.5
315.0
292.5
270.0
247.5
225.0
202.5
180.0
157.5
135.0
112.5
90.0
67.5
45.0
22.5
0.0
Directions [degree]
0
250
500
750
1000
1250
1500
1750
2000
Time [ms]


Neuron 36
337.5
315.0
292.5
270.0
247.5
225.0
202.5
180.0
157.5
135.0
112.5
90.0
67.5
45.0
22.5
0.0
Directions [degree]
0
250
500
750
1000
1250
1500
1750
2000
Time [ms]



Task 2: Plot spike density functions

Compute an estimate of the spike rate against time relative to stimulus onset. There are two ways: * Discretize time: Decide on a bin size, count the spikes in each bin and average across trials. * Directly estimate the probability of spiking using a density estimator with specified kernel width.

Implement one of them in the function `plotPsth()`. If you use a dataframe you may need to change the interface of the function.

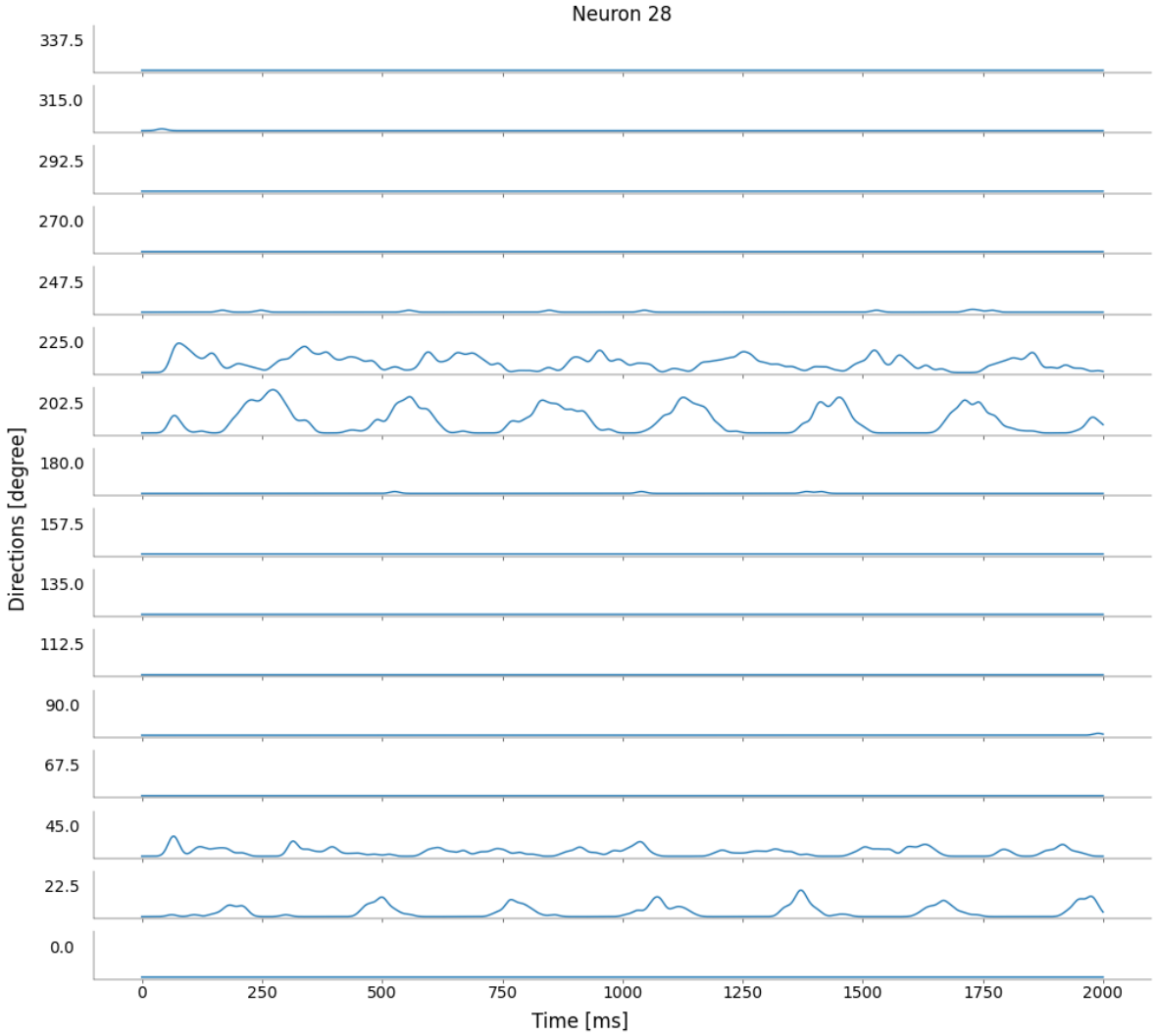
Grading: 2 pts

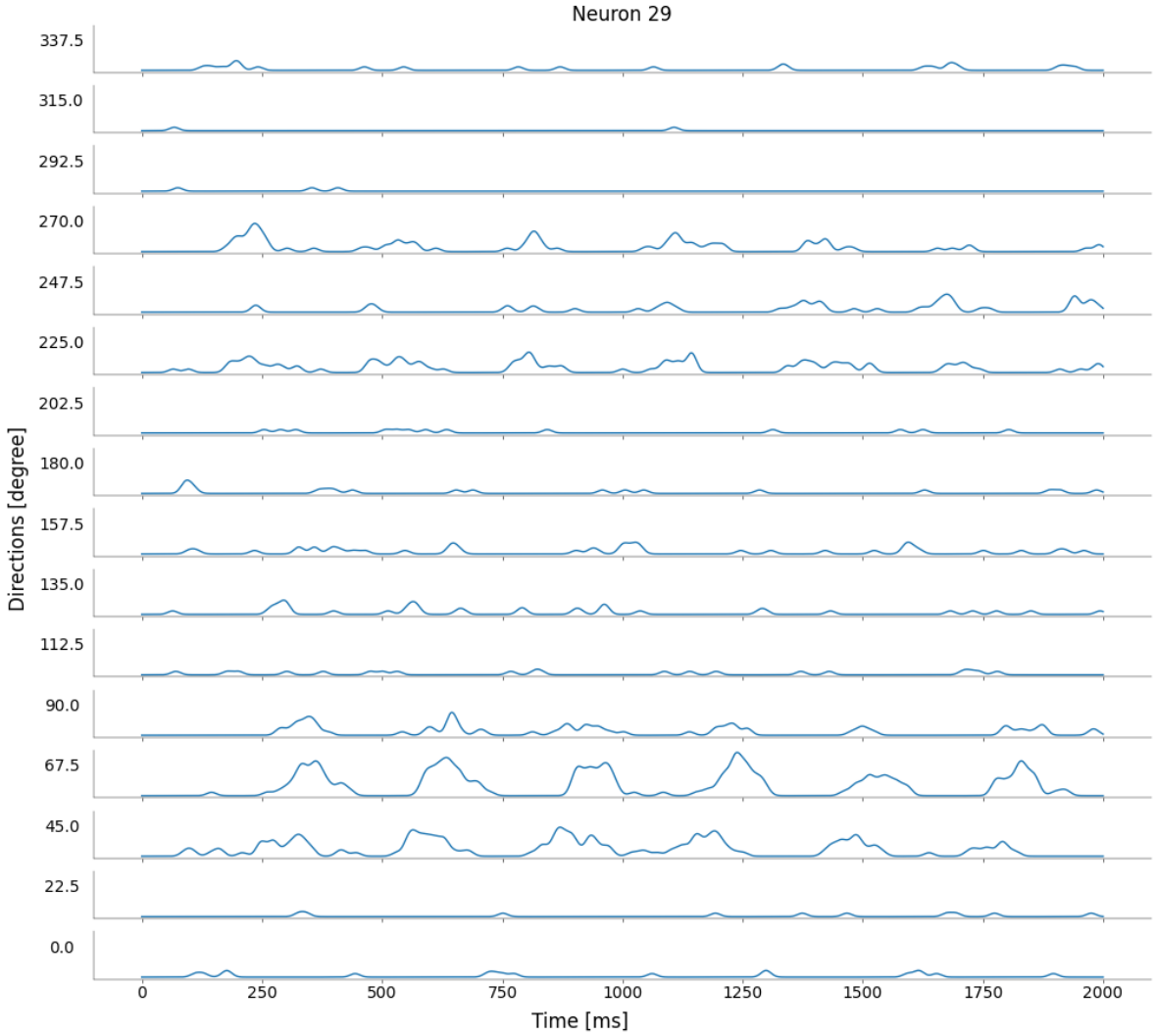
```
def plotPSTH(spikes, neuron):
    """Plot PSTH for a single neuron sorted by condition

    Parameters
    ----------

    spikes: pd.DataFrame
        Pandas DataFrame with columns
            Neuron | SpikeTimes | Dir | relTime | Trial | stimPeriod

    neuron: int
        Neuron ID


    Note
    ----

    this function does not return anything, it just creates a plot!
    """

    def gaussian_pdf(x, loc, scale):
        pdf = np.exp(-((x - loc) ** 2) / (2.0 * scale**2)) / (
            np.sqrt(2 * np.pi) * scale
        )
        return pdf

    def find_nearest(array, value):
        array = np.asarray(array)
        idx = (np.abs(array - value)).argmin()
        return idx

    dirs_all = np.sort(spikes["Dir"].unique())[::-1]
    max_trials = 11
    spikes_neuron = spikes[spikes["Neuron"] == neuron]
    # insert your code here
    sigma = 0.1
    tmax = 20 * sigma
    ktime = np.arange(-tmax, tmax, 0.001)
    kernel_o = gaussian_pdf(ktime, 0, sigma)
    spikes_neuron = spikes_neuron.sort_values(by="Dir")
    # plot the spikes for each trial in a raster plot
```

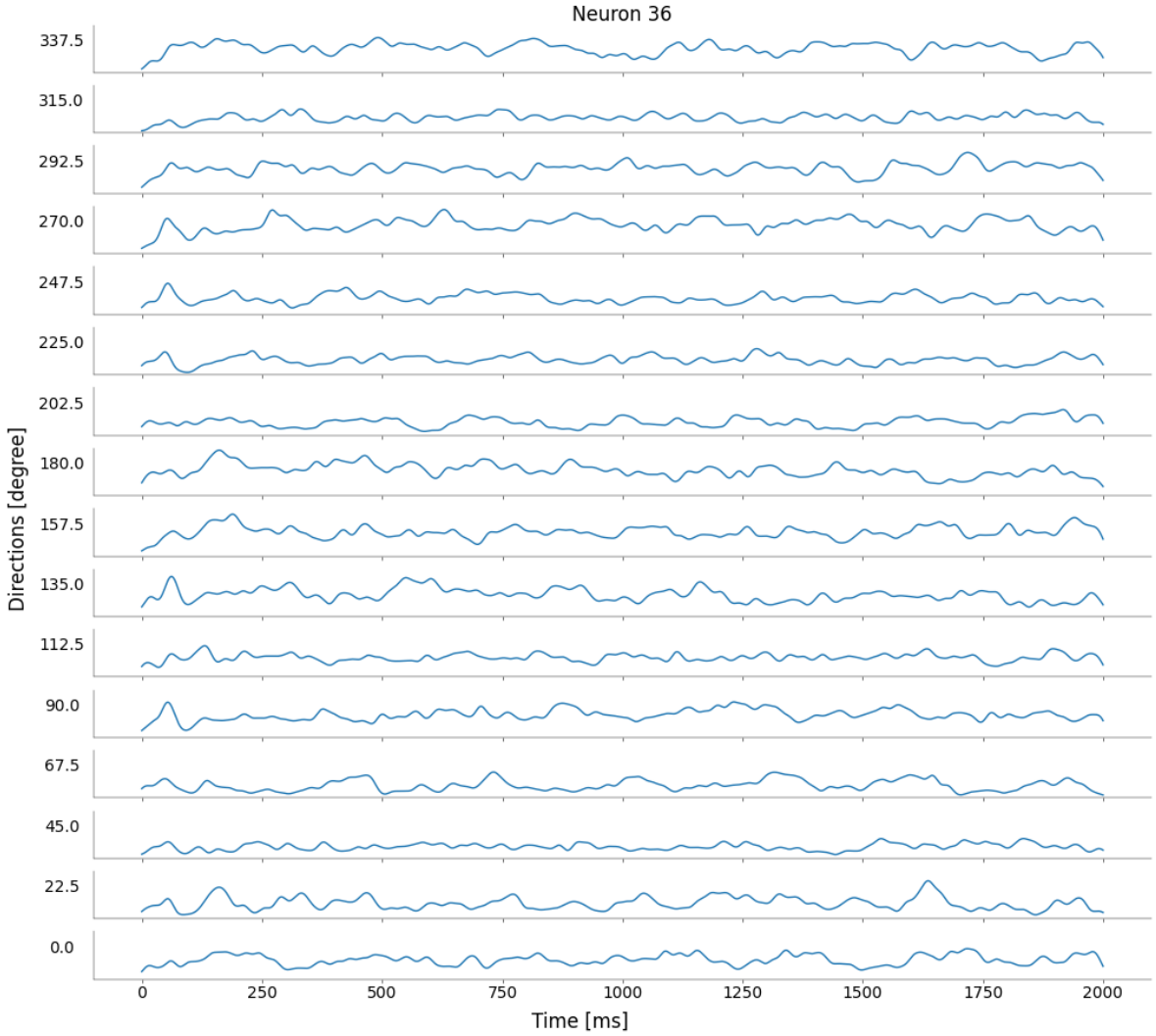
```
    dirs = spikes_neuron["Dir"].unique()
    time = np.arange(0, 2000.0, 0.1)
    fig, ax = plt.subplots(len(dirs_all), figsize=(10, 9), sharex=True, sharey=True)

    for d, directions in enumerate(dirs_all):
        spikes_neuron_dir = spikes_neuron[spikes_neuron["Dir"] == directions]
        # create empty array for the rate
        rates = np.zeros((len(spikes_neuron_dir["Trial"].unique()), len(time)))
        for t, trial in enumerate(np.sort(spikes_neuron_dir["Trial"].unique())):
            spikes_neuron_dir_trial = spikes_neuron_dir[
                spikes_neuron_dir["Trial"] == trial
            ]
            spikes = spikes_neuron_dir_trial["relTime"].to_numpy()
            index = []
            for spike in spikes:
                idx = find_nearest(time, spike)
                index.append(idx)
            binaryrate = np.zeros(len(time))
            binaryrate[index] = 1.0
            rate = np.convolve(binaryrate, kernel_o, mode="same")
            rates[t, :] = rate
        mean_rates = np.sum(rates, axis=0) / max_trials
        ax[d].plot(time, mean_rates)
        ax[d].set_ylabel(f"{directions}", rotation=0, labelpad=20)
        ax[d].set_yticks([])
    ax[0].set_title(f"Neuron {neuron}")
    fig.supylabel("Directions [degree]")
    # create supylabel on the right side

    # fig.supylabel("Trials n")
    fig.supxlabel("Time [ms]")
```

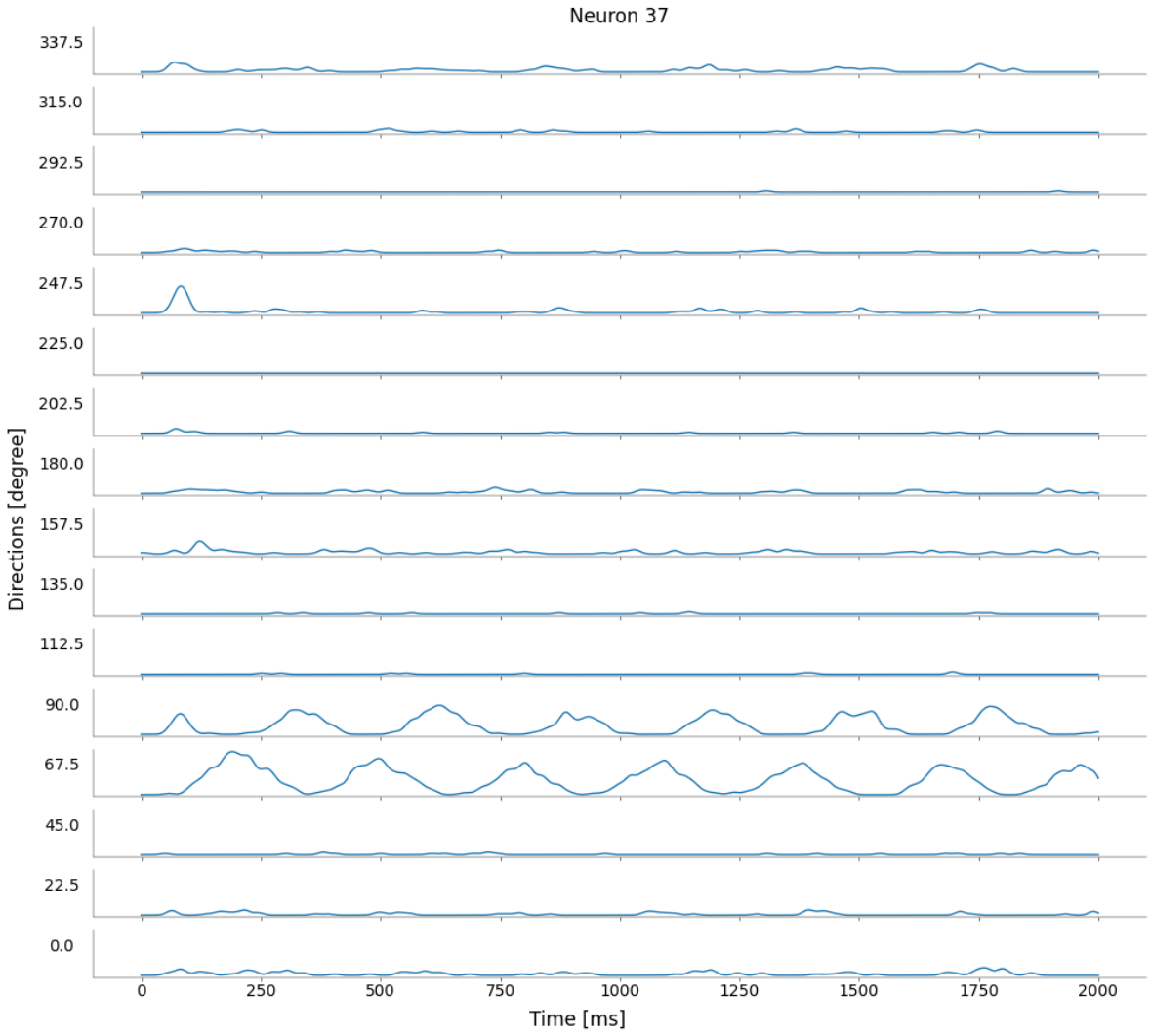
Show examples of different neurons. Good candidates to check are 28, 29, 36 or 37.

```
plotPSTH(spikes, 28)
plotPSTH(spikes, 29)
plotPSTH(spikes, 36)
plotPSTH(spikes, 37)
```

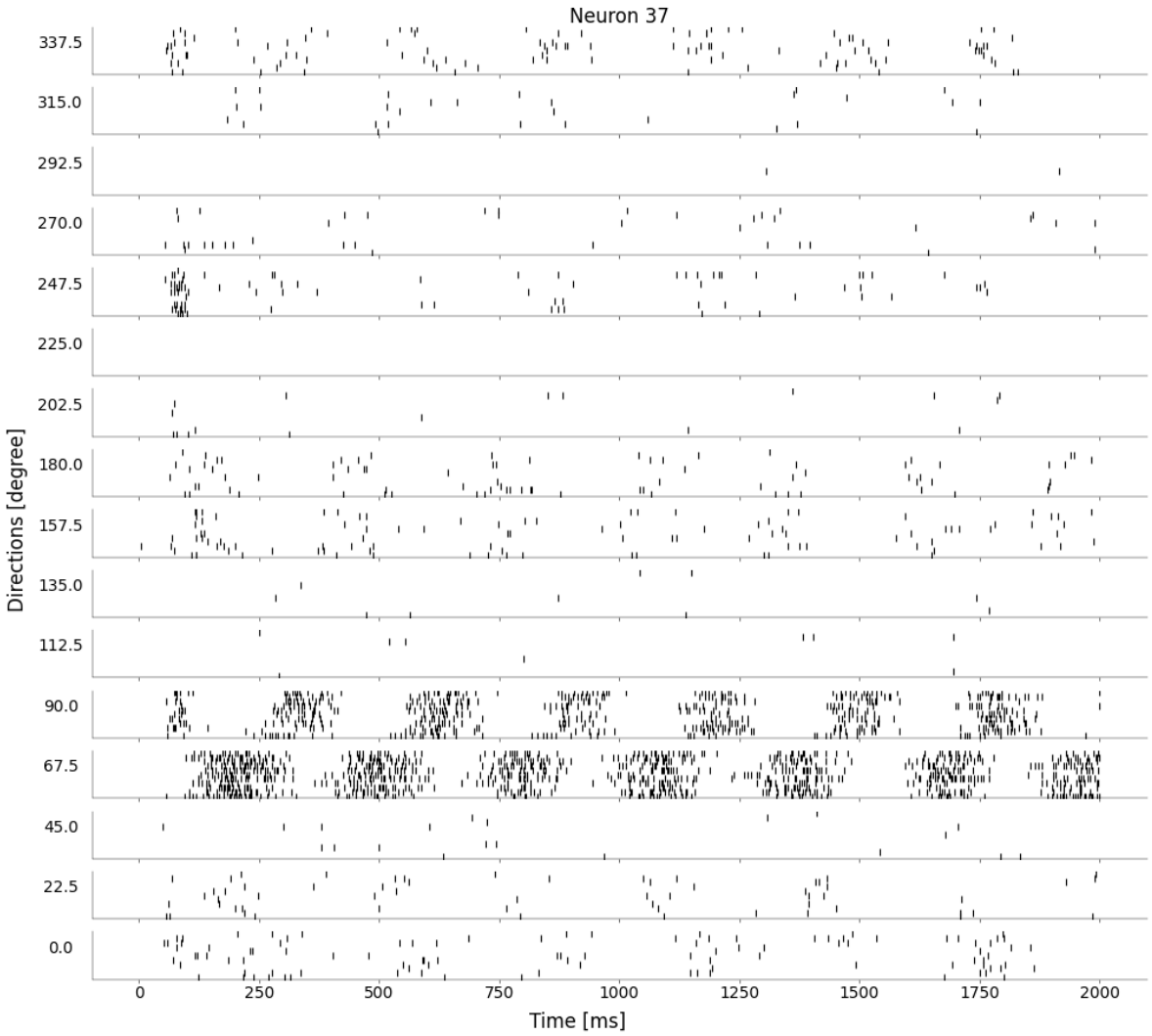

Neuron 28
337.5
315.0
292.5
270.0
247.5
225.0
202.5
180.0
157.5
135.0
112.5
90.0
67.5
45.0
22.5
0.0
Directions [degree]
0
250
500
750
1000
1250
1500
1750
2000
Time [ms]

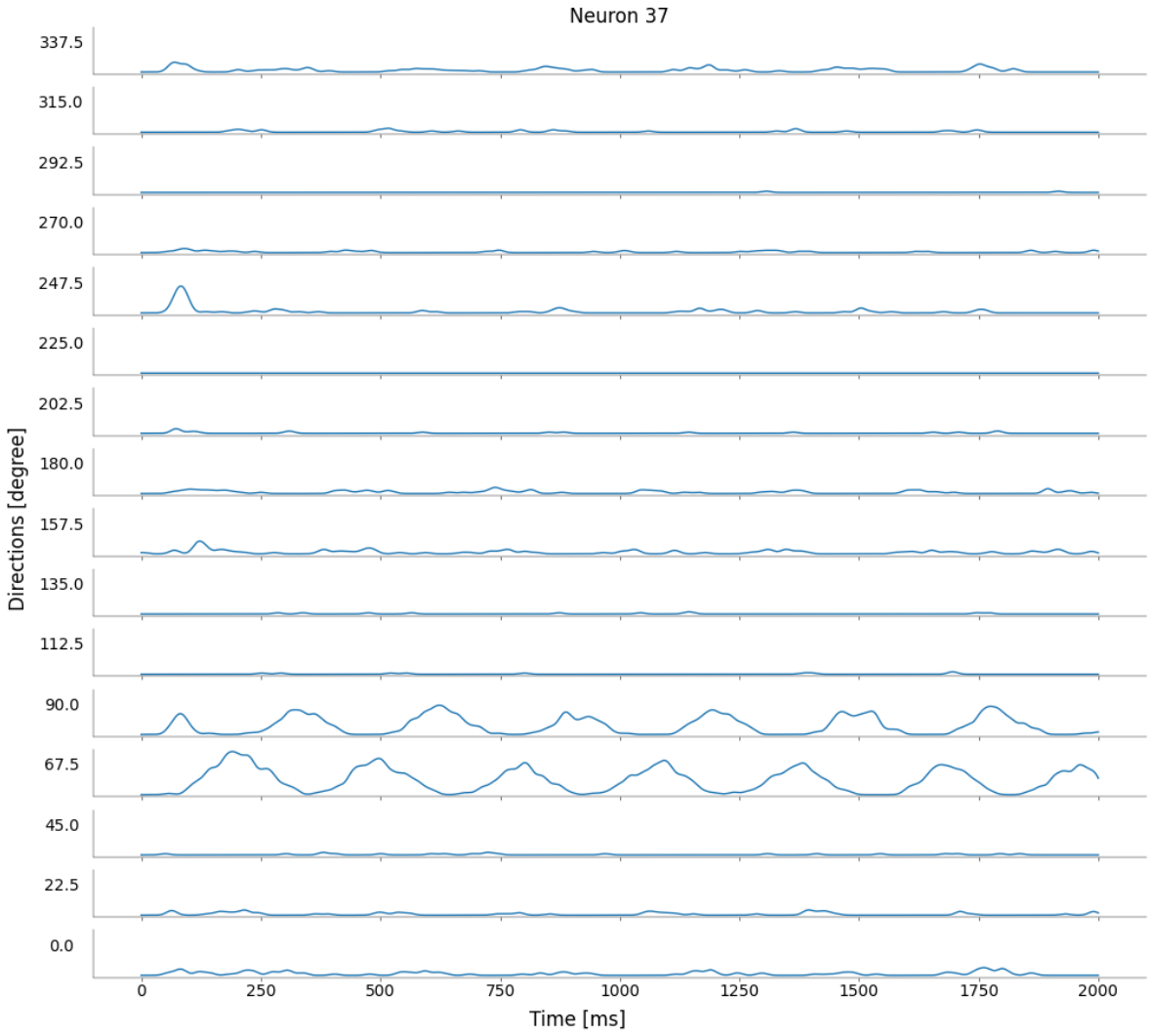

Neuron 29
337.5
315.0
292.5
270.0
247.5
225.0
202.5
180.0
157.5
135.0
112.5
90.0
67.5
45.0
22.5
0.0
Directions [degree]
0
250
500
750
1000
1250
1500
1750
2000
Time [ms]


Neuron 36
337.5
315.0
292.5
270.0
247.5
225.0
202.5
180.0
157.5
135.0
112.5
90.0
67.5
45.0
22.5
0.0
Directions [degree]
0
250
500
750
1000
1250
1500
1750
2000
Time [ms]




```
neuron = 37
plotRaster(spikes, neuron)
plotPSTH(spikes, neuron)
```


Neuron 37
337.5
315.0
292.5
270.0
247.5
225.0
202.5
180.0
157.5
135.0
112.5
90.0
67.5
45.0
22.5
0.0
Directions [degree]
0
250
500
750
1000
1250
1500
1750
2000
Time [ms]



Task 3: Fit and plot tuning functions

The goal is to visualize the activity of each neuron as a function of stimulus direction. First, compute the spike counts of each neuron for each direction of motion and trial. The result should be a matrix `x`, where x_{jk} represents the spike count of the j-th response to the k-th direction of motion (i.e. each column contains the spike counts for all trials with one direction of motion). If you used dataframes above, the `groupby()` function allows to implement this very compactely. Make sure you don't loose trials with zero spikes though. Again, other implementations are completely fine.

Fit the tuning curve, i.e. the average spike count per direction, using a von Mises model. To capture the non-linearity and direction selectivity of the neurons, we will fit a modified von

Mises function:

$$f(\theta) = \exp(\alpha + \kappa(\cos(2 * (\theta - \phi)) - 1) + \nu(\cos(\theta - \phi) - 1))$$

Here, θ is the stimulus direction. Implement the von Mises function in `vonMises()` and plot it to understand how to interpret its parameters ϕ, κ, ν, α. Perform a non-linear least squares fit using a package/function of your choice. Implement the fitting in `tuningCurve()`.

Plot the average number of spikes per direction, the spike counts from individual trials as well as your optimal fit.

Select two cells that show nice tuning to test you code.

Grading: 3 pts

```
def vonMises(theta, alpha, kappa, nu, phi):
    """Evaluate the parametric von Mises tuning curve with parameters p at locations theta

    Parameters
    ----------

     : np.array, shape=(N, )
        Locations. The input unit is degree.

    theta, alpha, kappa, nu, phi: float
        Function parameters

    Return
    ------
    f: np.array, shape=(N, )
        Tuning curve.
    """
    theta = np.radians(theta)
    phi = np.radians(phi)

    # insert your code here
    f = np.exp(
        alpha + kappa * (np.cos(2 * (theta - phi)) - 1) + nu * (np.cos(theta - phi) - 1)
    )

    # ----------------------------------
    # Implement the Mises model (0.5 pts)
    # ----------------------------------
```

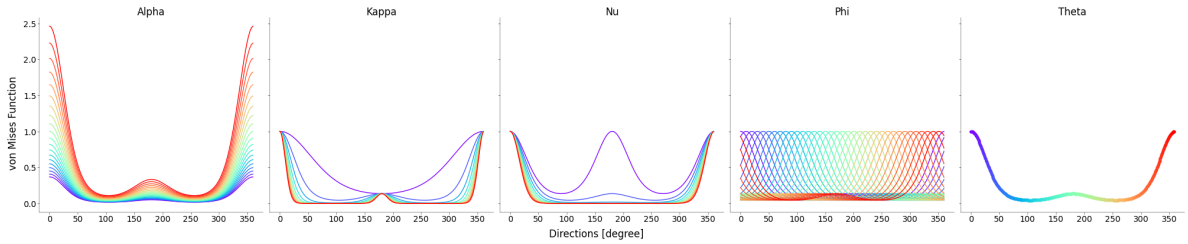
```
    return f
```

Plot the von Mises function while varying the parameters systematically.

```
# ------------------------------------------------------------------------------
# plot von Mises curves with varying parameters and explain what they do (0.5 pts)
# ------------------------------------------------------------------------------
alpha = np.arange(-1, 1, 0.1)
kappa = np.arange(0, 10, 1)
nu = np.arange(0, 10, 1)
phi = np.arange(0, 360, 10)
theta = np.arange(0, 360, 1)

x = np.linspace(
    0,
    360,
    len(theta),
)
# color
fig, ax = plt.subplots(1, 5, figsize=(20, 4), sharey=True)
color = plt.cm.rainbow(np.linspace(0, 1, len(alpha)))

for i, a in enumerate(alpha):
    f = vonMises(theta, a, 1, 1, 0)
    # color space
    ax[0].plot(x, f, label=f"alpha = {a}", color=color[i])

    # ax[0].legend()
color = plt.cm.rainbow(np.linspace(0, 1, len(kappa)))
for i, k in enumerate(kappa):
    f = vonMises(theta, 0, k, 1, 0)
    ax[1].plot(x, f, label=f"kappa = {k}", color=color[i])
    # ax[1].legend()

color = plt.cm.rainbow(np.linspace(0, 1, len(nu)))
for i, n in enumerate(nu):
    f = vonMises(theta, 0, 1, n, 0)
    ax[2].plot(x, f, label=f"nu = {n}", color=color[i])
    # ax[2].legend()
color = plt.cm.rainbow(np.linspace(0, 1, len(phi)))
for i, p in enumerate(phi):
```

```
    f = vonMises(theta, 0, 1, 1, p)
    ax[3].plot(x, f, label=f"phi = {p}", color=color[i])

    # ax[3].legend()
color = plt.cm.rainbow(np.linspace(0, 1, len(theta)))
for i, t in enumerate(theta):
    f = vonMises(t, 0, 1, 1, 0)
    ax[4].scatter(t, f, label=f"theta = {t}", color=color[i])
ax[0].set_title("Alpha")
ax[1].set_title("Kappa")
ax[2].set_title("Nu")
ax[3].set_title("Phi")
ax[4].set_title("Theta")
fig.supxlabel("Directions [degree]")
fig.supylabel("von Mises Function")
```

```
Text(0.02, 0.5, 'von Mises Function')
```



- alpha scales up the peak of the function
- kappa is the width of the peak
- nu scales the peaks relative to each other
- phi shifts the function along the x-axis
- theta is the stimulus direction

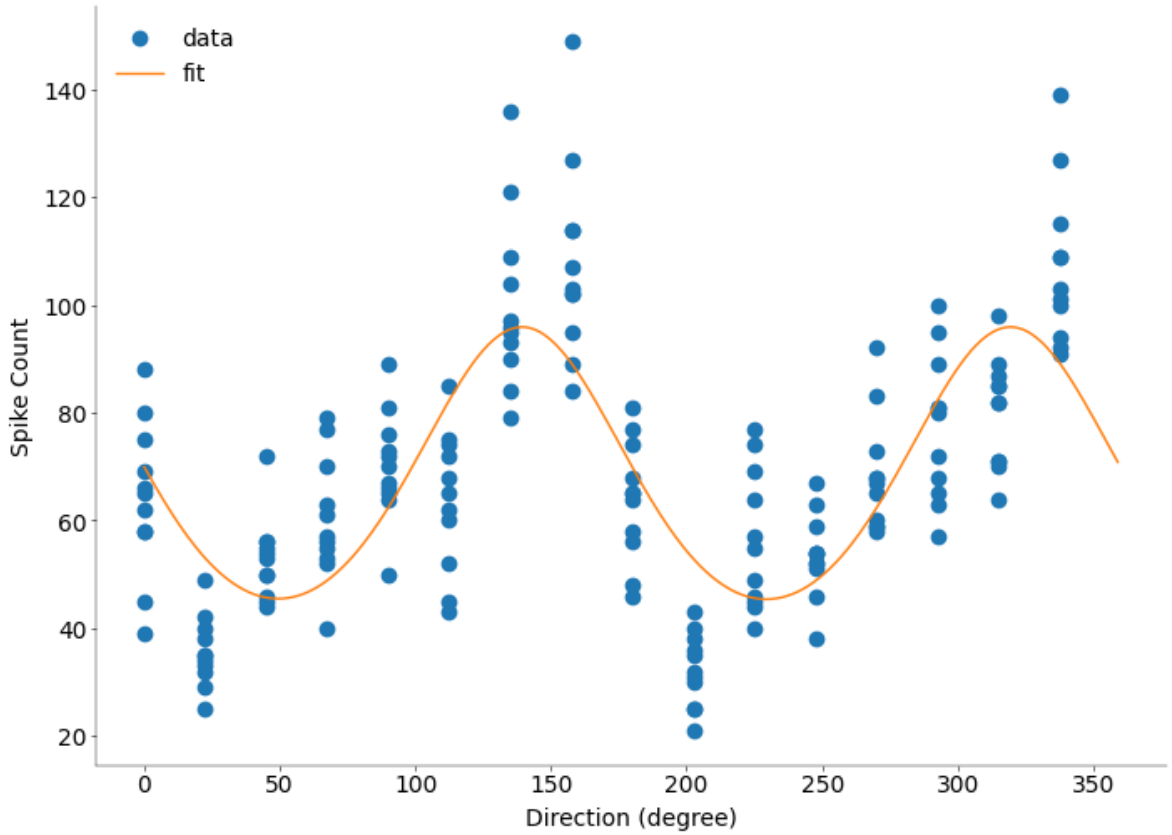
```
def tuningCurve(counts, dirs, show=True):
    """Fit a von Mises tuning curve to the spike counts in count with direction dir using

    Parameters
    ----------

    counts: np.array, shape=(total_n_trials, )
        the spike count during the stimulation period
```

```
    dirs: np.array, shape=(total_n_trials, )
        the stimulus direction in degrees

    show: bool, default=True
        Plot or not.


    Return
    ------
    popt: np.array or list, (4,)
        parameter vector of tuning curve function
    """

    # insert your code here
    upper_bounds = (np.inf, np.inf, np.inf, 360)
    lower_bounds = (0, 0, 0, 0)
    bounds = (lower_bounds, upper_bounds)

    try:
        popt, pcov = opt.curve_fit(vonMises, dirs, counts, maxfev=1000)
    except RuntimeError:
        popt, pcov = opt.curve_fit(
            vonMises, dirs, counts, maxfev=1000000, bounds=bounds, method="trf"
        )

    x = np.arange(0, 360, 1)

    y = vonMises(x, *popt)

    if show == True:
        fig, ax = plt.subplots(figsize=(7, 5))
        ax.plot(dirs, counts, "o", label="data")
        ax.plot(x, y, label="fit")
        ax.set_xlabel("Direction (degree)")
        ax.set_ylabel("Spike Count")
        plt.legend()

        return
    else:
        return popt
```

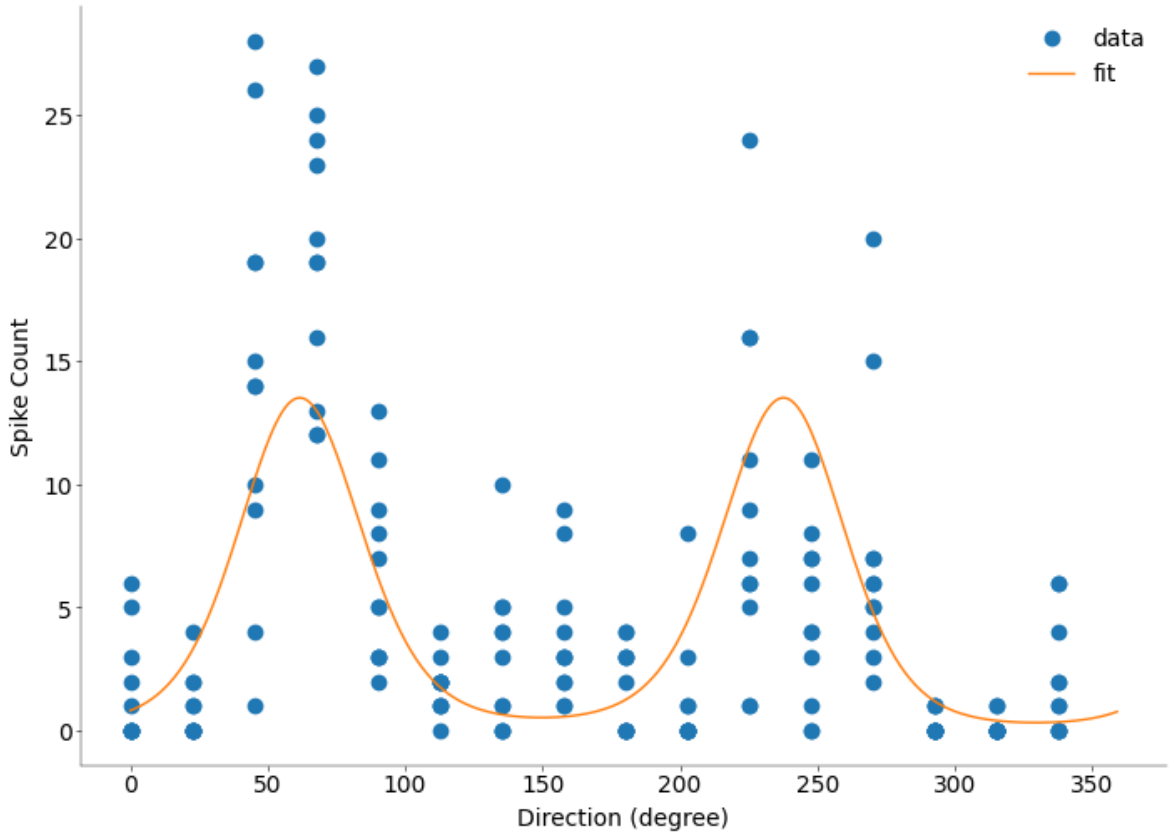
Plot tuning curve and fit for different neurons. Good candidates to check are 28, 29 or 37.

```
def get_data(spikes, neuron):
    spk_by_dir = (
        spikes[spikes["Neuron"] == neuron]
        .groupby(["Dir", "Trial"])["stimPeriod"]
        .sum()
        .astype(int)
        .reset_index()
    )

    dirs = spk_by_dir["Dir"].values
    counts = spk_by_dir["stimPeriod"].values

    # because we count spikes only when they are present, some zero entries in the count v
    for i, Dir in enumerate(np.unique(spikes["Dir"])):
        m = nTrials - np.sum(dirs == Dir)
        if m > 0:
            dirs = np.concatenate((dirs, np.ones(m) * Dir))
            counts = np.concatenate((counts, np.zeros(m)))

    idx = np.argsort(dirs)
    dirs_sorted = dirs[idx]  # sorted dirs
    counts_sorted = counts[idx]

    return dirs_sorted, counts_sorted
```

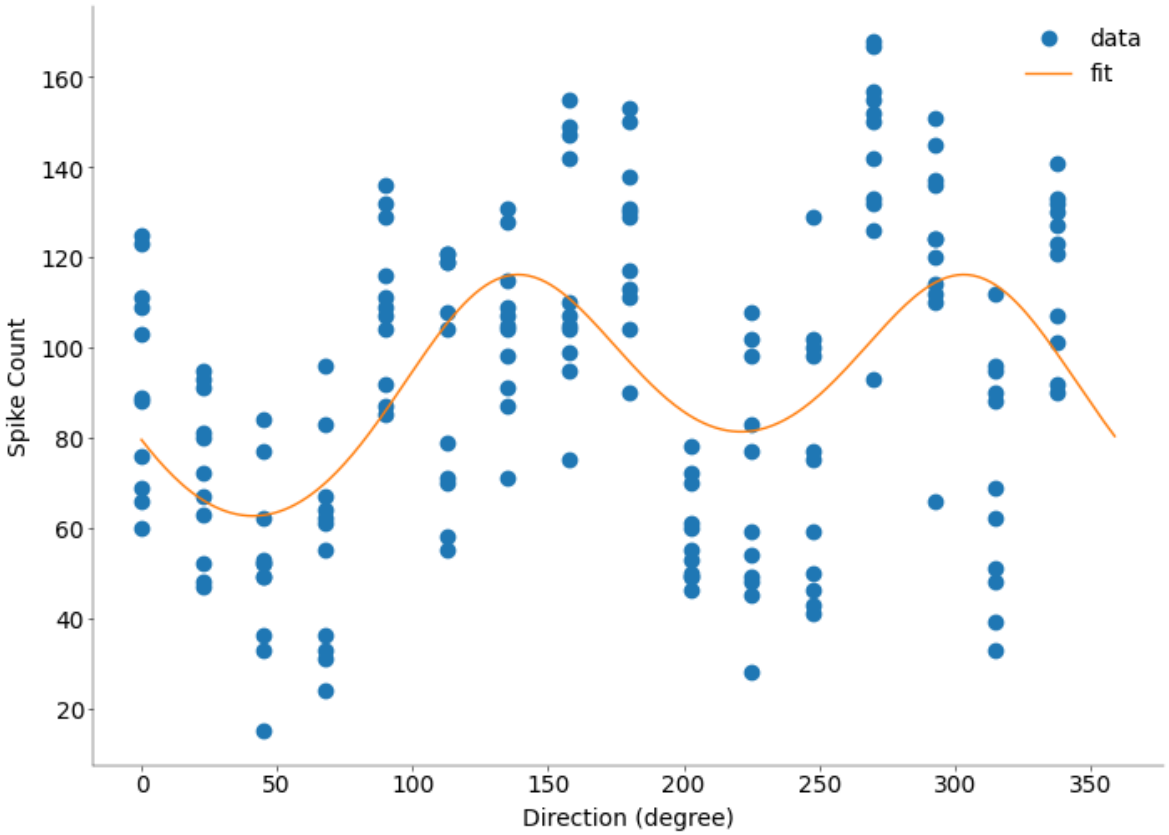
```
dirs, counts = get_data(spikes, 30)
tuningCurve(counts, dirs, show=True)
```



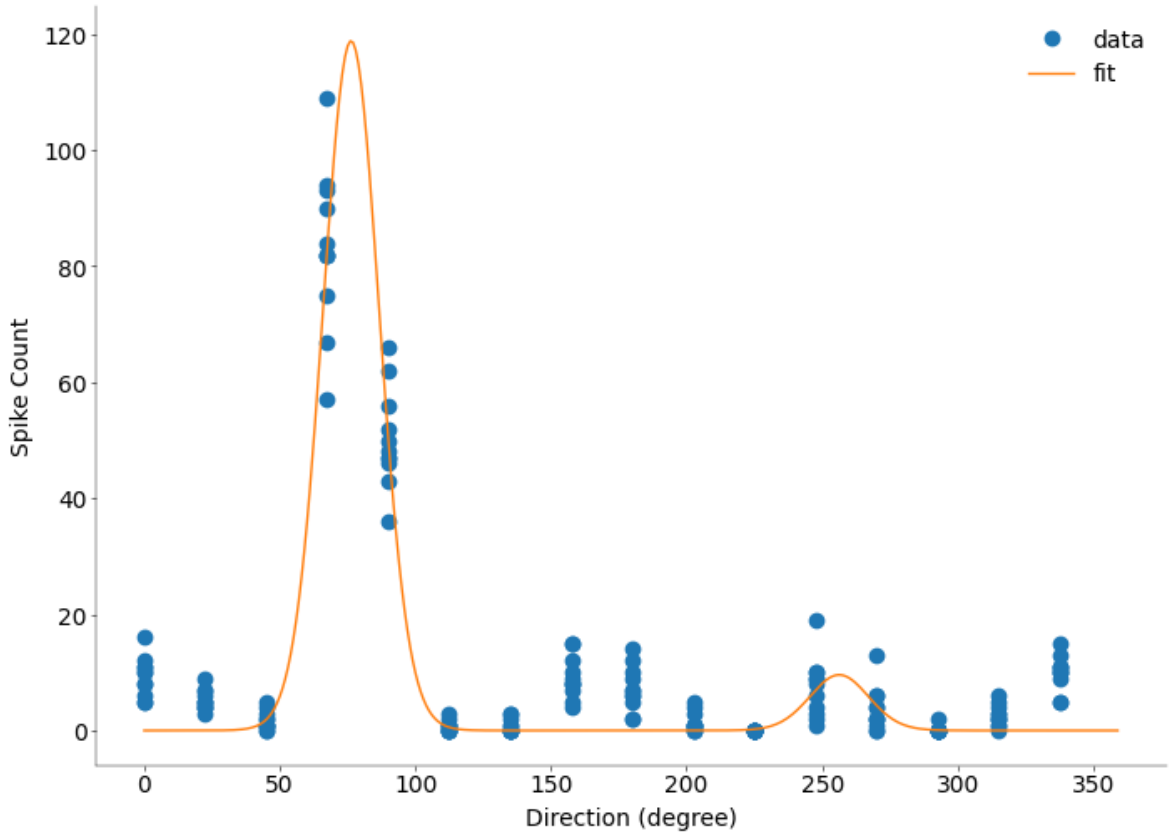

```
dirs, counts = get_data(spikes, 29)
tuningCurve(counts, dirs, show=True)
# add plot
```



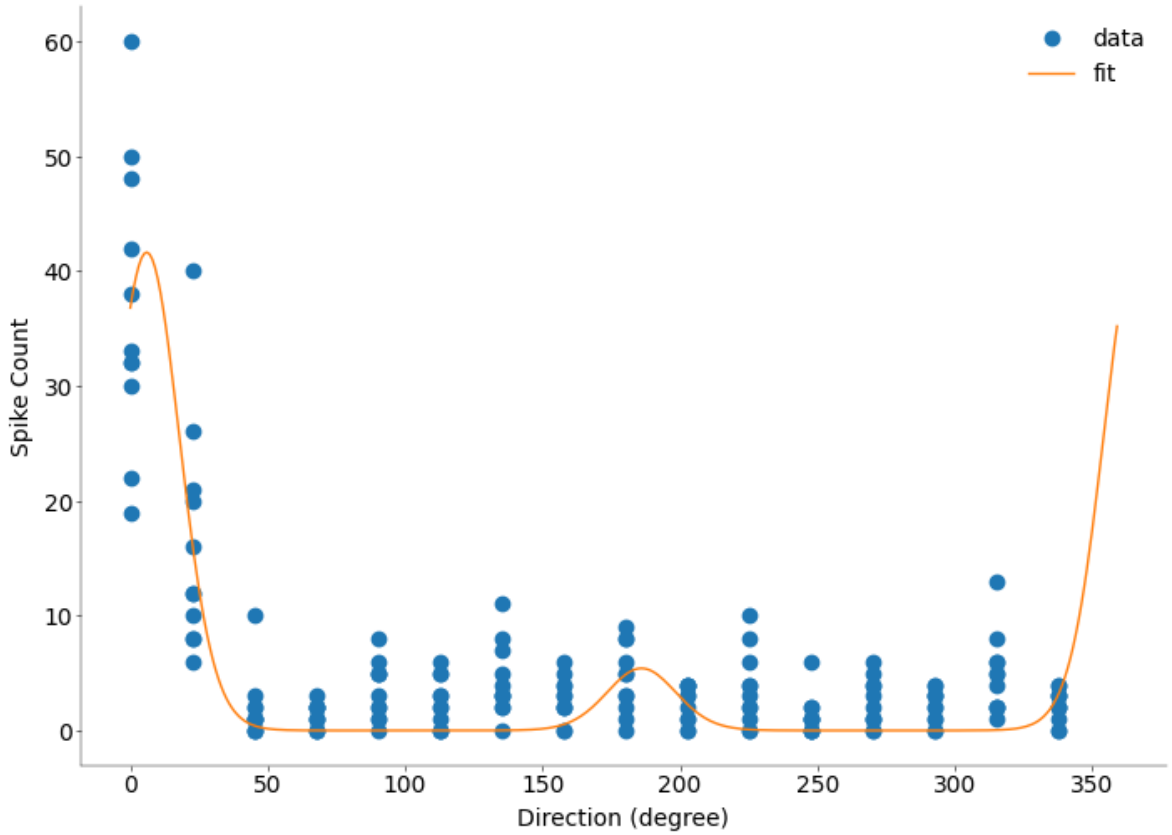

```
dirs, counts = get_data(spikes, 36)
tuningCurve(counts, dirs, show=True)
# add plot
```




```
dirs, counts = get_data(spikes, 37)
tuningCurve(counts, dirs, show=True)
# this makes no sense the optimization is going to infinity
```




```
dirs, counts = get_data(spikes, 32)
tuningCurve(counts, dirs, show=True)
# add plot
```



Task 4: Permutation test for direction tuning

Implement a permutation test to quantitatively assess whether a neuron is direction/orientation selective. To do so, project the vector of average spike counts, $m_k = \frac{1}{N}\sum_j x_{jk}$ on a complex exponential with two cycles, $v_k = \exp(\psi i \theta_k)$, where θ_k is the k-th direction of motion in radians and $\psi \in 1, 2$ is the fourier component to test (1: direction, 2: orientation). Denote the projection by $q = m^T v$. The magnitude $|q|$ tells you how much power there is in the ψ-th fourier component.

Estimate the distribution of |q| under the null hypothesis that the neuron fires randomly across directions by running 1000 iterations where you repeat the same calculation as above but on a random permutation of the trials (that is, randomly shuffle the entries in the spike count matrix x). The fraction of iterations for which you obtain a value more extreme than what you observed in the data is your p-value. Implement this procedure in the function `testTuning()`.

Illustrate the test procedure for one of the cells from above. Plot the sampling distribution of |q| and indicate the value observed in the real data in your plot.

How many cells are tuned at $p < 0.01$?

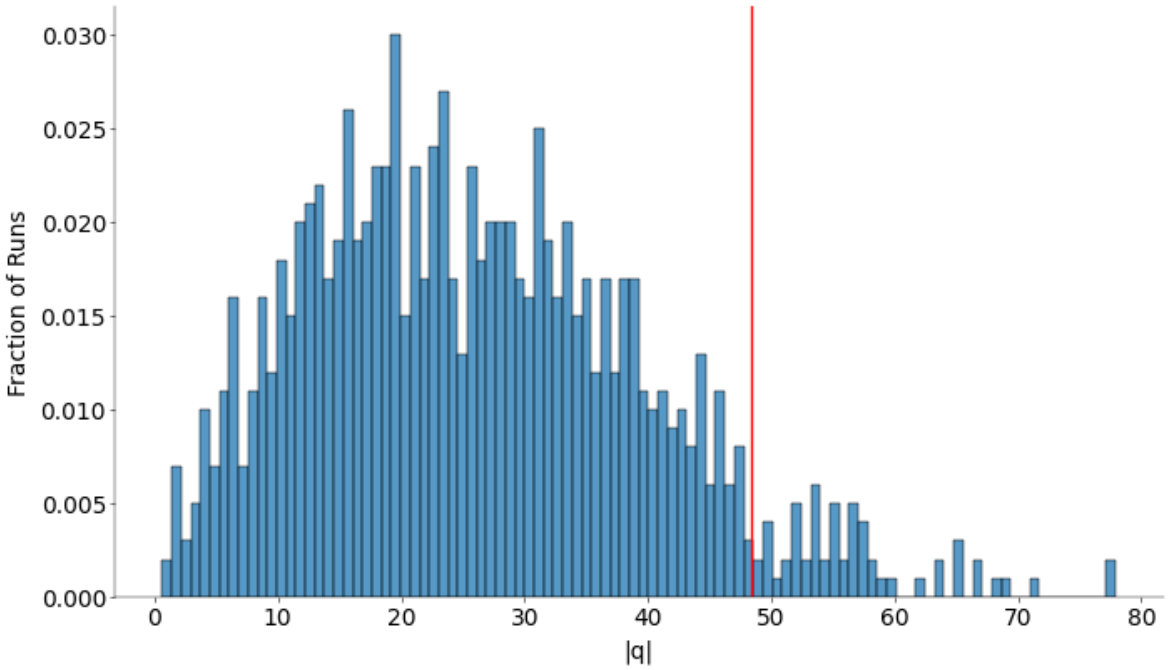
Grading: 3 pts

```
def testTuning(counts, dirs, psi=1, niters=1000, show=False):
    """Plot the data if show is True, otherwise just return the fit.

    Parameters
    ----------

    counts: np.array, shape=(total_n_trials, )
        the spike count during the stimulation period

    dirs: np.array, shape=(total_n_trials, )
        the stimulus direction in degrees

    psi: int
        fourier component to test (1 = direction, 2 = orientation)

    niters: int
        Number of iterations / permutation

    show: bool
        Plot or not.

    Returns
    -------
    p: float
        p-value
    q: float
        magnitude of second Fourier component

    qdistr: np.array
        sampling distribution of |q| under the null hypothesis

    """
    np.random.seed(42)
    dirs_unique = np.unique(dirs)
    means = np.zeros(len(dirs_unique))
    for d, directions in enumerate(np.unique(dirs)):
        means[d] = np.mean(counts[dirs == directions])
```

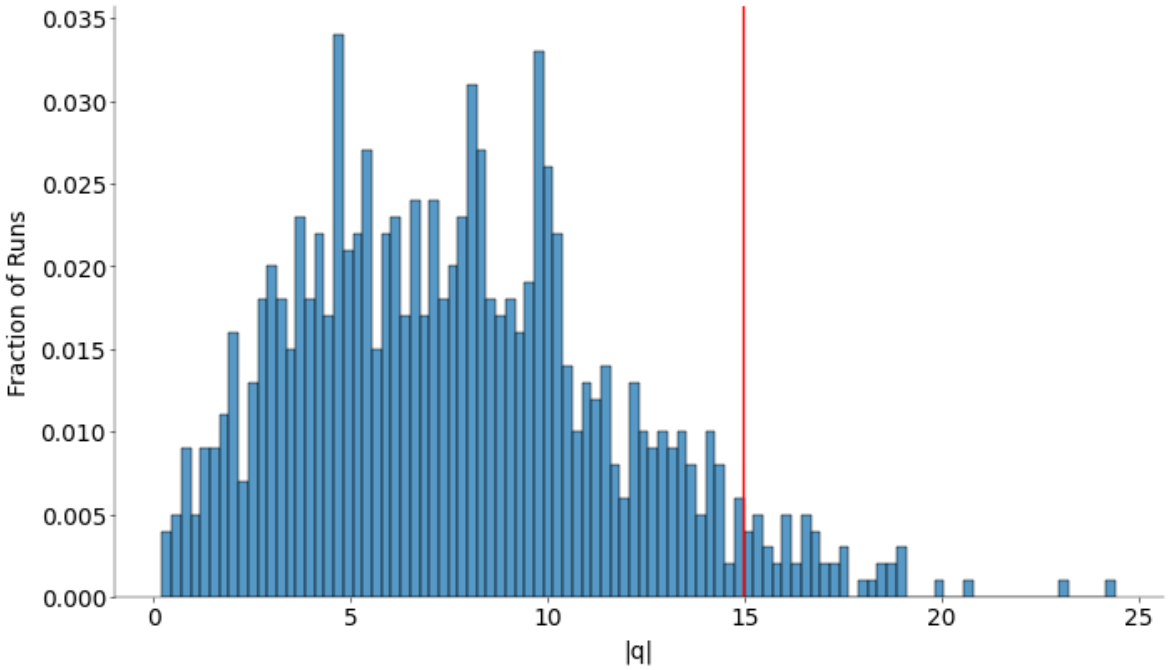
```
    # imaginary exponential function mu
    dirs_unique_rad = np.deg2rad(dirs_unique)

    nu = np.exp((1j * psi * dirs_unique_rad) * (2 * np.pi))

    q = means @ nu
    abs_q = np.absolute(q)

    counts_shuffle = np.array(counts)
    qs_shuffle = np.zeros(niters)
    valid_counter = 0

    for i in range(niters):
        np.random.shuffle(counts_shuffle)
        means = np.zeros(len(dirs_unique))
        for d, directions in enumerate(np.unique(dirs)):
            means[d] = np.mean(counts_shuffle[dirs == directions])

        q_shuffle = means @ nu
        abs_q_shuffle = np.absolute(q_shuffle)
        qs_shuffle[i] = abs_q_shuffle

        if abs_q_shuffle > abs_q:
            valid_counter += 1

    p = valid_counter / niters
    # print(p)

    if show == True:
        fig, ax = plt.subplots(figsize=(7, 4))

        sns.histplot(qs_shuffle, bins=100, stat="proportion", ax=ax, label="null")
        ax.axvline(abs_q, color="red", label="observed")
        ax.set_xlabel("|q|")
        ax.set_ylabel("Fraction of Runs")

        # you can use sns.histplot for the histogram
    else:
        return p, abs_q, qs_shuffle
```

Show null distribution for the example cell:

```
# ------------------------------------------------------------
# Plot null distributions for example cells 28 & 29. (0.5 pts)
# ------------------------------------------------------------

dirs, counts = get_data(spikes, 37)
testTuning(counts, dirs, psi=1, niters=1000, show=True)
# add plot
```




```
dirs, counts = get_data(spikes, 29)
testTuning(counts, dirs, psi=1, niters=1000, show=True)
# add plot
```



Test all cells for orientation and direction tuning

```
# ------------------------------------------------------
# Test cells for orientation / direction tuning (0.5 pts)
# ------------------------------------------------------
p_dirs = np.zeros(len(np.unique(spikes["Neuron"])))

p_oris = np.zeros(len(np.unique(spikes["Neuron"])))

for n, neuron in enumerate(np.unique(spikes["Neuron"])):
    dirs, counts = get_data(spikes, neuron)
    p_dir, abs_q, qs_shuffle = testTuning(counts, dirs, psi=1, niters=1000, show=False)
    p_dirs[n] = p_dir
    p_ori, abs_q, qs_shuffle = testTuning(counts, dirs, psi=2, niters=1000, show=False)
    p_oris[n] = p_ori


# print(p_ori)
# collect p values for orientation / direction selectivity
# collect p values for orientation / direction selectivity
```

Number of direction tuned neurons:

```
# count cells with p > 0.01 (which ones are they?)
print(
    np.sum(p_dirs < 0.01),
)
idx = np.where(p_dirs < 0.01)[0]
print(np.unique(spikes["Neuron"])[idx])
```

```
10
[ 3  7 12 13 16 20 30 34 36 38]
```

Number of orientation tuned neurons:

```
# count cells with p > 0.01 (which ones are they?)
print(np.sum(p_oris < 0.01))
idx = np.where(p_oris < 0.01)[0]
print(np.unique(spikes["Neuron"])[idx])
```

```
16
[ 2  3  6  7 13 18 20 26 27 28 29 31 32 34 37 38]
```