Visualization
Prof. Bernhard Schmitzer, Uni Göttingen, summer term 2025

Problem sheet 8

- *Submission by **2025-07-09** 18:00 via StudIP **as a single PDF/ZIP**. Please combine all results into one PDF or archive. If you work in another format (markdown, jupyter notebooks), add a PDF converted version to your submission.*

- *Use Python 3 for the programming tasks as shown in the lecture. If you cannot install Python on your system, the GWDG jupyter server at `https://jupyter-cloud.gwdg.de/` might help. Your submission should contain the final images as well as the code that was used to generate them.*

- *Work in groups of up to three. Clearly indicate names and enrollment numbers of all group members at the beginning of the submission.*

Exercise 8.1: Flow matching.
Let $\mu : [0,1]^2 \to \mathbb{R}$ be the probability density function on the unit square $[0,1]^2$ that describes a uniform measure concentrated on the rectangle $[0.05, 0.65] \times [0.05, 0.95]$. Let ν be the analogous density for the rectangle $[0.35, 0.95] \times [0.05, 0.95]$. In this problem we investigate the flow matching technique introduced in:

- Lipman et al., Flow Matching for Generative Modeling, ICLR 2023
- Liu et al., Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow, ICLR 2023

It is not necessary to understand these methods for the present problem sheet (but it will help to have a rough idea). The file `flows.npz` contains samples from three stochastic interpolations between μ and ν related to this method. More concretely, it contains the array `t=np.linspace(0,1,nT)` for `nT=51`, which describes the list of times at which the stochastic interpolations are evaluated. Further, it contains the arrays `xA`, `xB`, and `xC`, each of which contain samples from one particular stochastic interpolation. These samples contain `nX=5000` points, at `nT` times (corresponding to the entries of `t` above) in two dimensions. For `x` $\in$ {`xA`, `xB`, `xC`}, `x[i,j,k]` contains the `k`-th coordinate of the `j`-th particle at time `t[i]`. For all three arrays, `x[0,:,:]` contains (the same) samples from μ and `x[-1,:,:]` contains (approximately) samples from ν. `xA` contains samples from the independent coupling of μ and ν that travel on straight lines. `xB` contains samples of the flow in the velocity field, obtained from `xA` via flow matching. In `xC` the start and end points of `xB` are connected again by straight lines, instead of the curved flow.

1. Visualize the trajectories of (some of) the particles in all three interpolations {`xA`, `xB`, `xC`}. One should clearly see the distinction between 'mixed' straight lines, curved flows, and 'aligned' straight lines.

2. Use dynamic visualization to show the movement of the individual particles. Again, the different behaviour of the three interpolations should be clearly visible.

3. Use dynamic visualization to show the probability distributions of the particles `x` $\in$ {`xA`, `xB`, `xC`}. Concretely, for some index `nt` $\in \{0, \ldots,$ `nT` $- 1\}$, estimate the probability density of the particles `x[nt,:,:]` on $[0, 1]^2$, for instance via 2d histograms, and show the evolution of these three densities over time. Which two densities are approximately equal?

Hint: This is (partially) interactive and will not work in a PDF. Therefore, for this problem sheet please submit a jupyter notebook.