This repository contains training data for an ML-based parameter optimization for modeling morphogen gradient formation by a source-diffusion-sink (SDS) mechanism, as described by the following one-dimensional (1D) reaction-diffusion partial differential equation:

$$\frac{\partial c(x,t)}{\partial t} = f_{\text{source}}(x) + D\frac{\partial^2}{\partial x^2}c(x,t) - k_{\text{sink}}c(x,t)\,. \tag{1}$$

Here, $c(x,t)$ is the scalar concentration field of the morphogen in space x at time t, D is the constant homogeneous diffusion coefficient, k_{sink} is the sink rate scaling with $c(x,t)$, describing morphogen degradation by the cells and proteases in the extracellular space, and f_{source} is the source term describing morphogen secretion by the source cells.

f_{source} depends on the location as only a group of cells produce the morphogen, i.e.,

$$f_{\text{source}}(x) = \begin{cases} k_{\text{source}} & \text{for} \quad 0 \leq x \leq w_{\text{source}} \\ 0 & \text{otherwise}\,, \end{cases} \tag{2}$$

with source width $w_{\text{source}} = 0.3L$ in a 1D diffusion domain of length L.

Solving Eq. (1) until steady-state results in a morphogen concentration profile that can flat, exponential, or step-wise depending on the parameters k_{source}, k_{sink}, and D. The goal is to generate a model that predicts a set of parameters given an input gradient. To train this model, we produce simulated training data by solving Eq. (1) until steady-state for different parameter sets.

The output folders contain a parameters.csv file, where the first row defines the parameter type of the respective column, and each row below contains a set of parameters with index i. The corresponding gradients folder contains gradient_i.csv files, each corresponding to one row i of the parameters.csv file.

Each gradient_i.csv file contains two columns: the first containing x, the second containing the steady-state concentration field $c(x, t_{\text{max}})$.