



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Modeling Cellular Communication

A System of Communication Channels

Your Name

June 17, 2025

Case

- Dataset: Mouse 1 - slice 153
- Cell type source: *Astro*
- Cell type target: *L2/3 IT*
- View of Anndata object: 1715 cells x 18450 genes
- Sample L LR interactions from *mouse consensus*:
5: {(Tac2 - Tacr3), (Angpt2 - Tek), (Actr2 - Adrb2), (Tnf - Tnfrsf1a),
(Angpt1 - Tek)}

Sparse COO tensors: these tensors are the Coordinate format, which is one of the storage formats for implementing sparse tensors.

In COO format, the specified element indices are stored as tuples of element indices and the corresponding values. In particular,

- **indices**: collected in tensor of size $(\text{ndim}, \text{nse})$ with type `torch.int64`
- **values**: collected in tensor of size $(\text{nse},)$ with `int` or `float` type

where `ndim` is the dimensionality of the tensor and `nse` is the number of specified elements.

Memory usage

The fully connected tensor is of dimension $N \times N \times L$. If we do not consider sparsity, the memory usage would be:

$$\text{Memory usage} = N^2 \cdot L \cdot \text{size of float} = 1715^2 \cdot 5 \cdot 32 \approx 0.055 \text{ GB}$$

If we need to consider sparsity. Definitely we need it if we want to display networks. Sparsity can be defined by thresholding the interaction scores.

- If we consider the Xth percentile of the interaction scores for each LR pair, we would get the same number of non-zero elements for each LR pair.
- If we consider the Xth percentile of the interaction scores for all LR pairs, we would need to normalize the scores to make them comparable.
- If we consider less nodes as an aggregate in space for each cellular type, we would consider less edges and therefore have less memory usage.

Let's consider the first sparsity definition, where we threshold the interaction scores for each LR pair independently.

Number of non-zero elements in the whole tensor: $147'064 \approx 0.56GB$

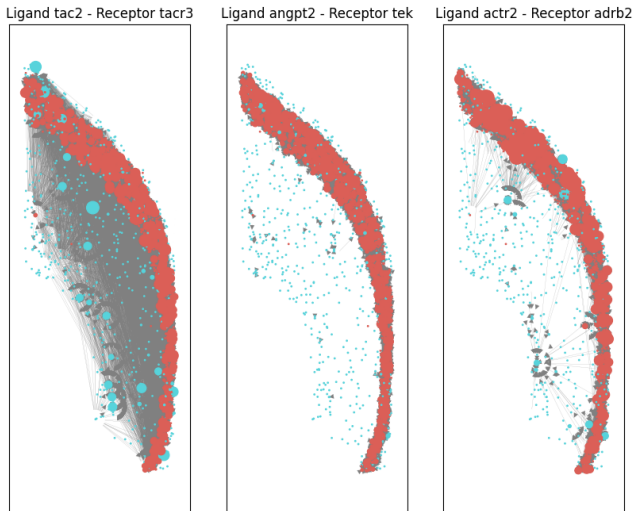


Figure: Example of the 99th percentile of pairwise interaction scores for 3 different LR pairs.

While if we consider the second sparsity definition, we would need before to zscore each LR pair.

Let's say that the intracellular communication is modeled naively as the correlation between the receptor of one layer and the ligand of another layer.

Definition

Let a receptor R be part of some layers of my model. For each R i will have a set of ligands L that are ligands of other layers.

Let's start considering the case where we evaluate the correlations between all R of one layers and L of all other layers

More than the correlations, let's think of the product of the expression...Since we do have only a value for each cell.