S1: hi everyone, my name is Huy. Today I will present our paper Unbalanced co optimal transport.
This is the joint work with Hicham janati, nicolas courty, remi flamary, ievgen redko, pinar demetci, ritambhara singh

S2: Let us start with the unbalanced optimal transport. It is a generalization of balanced optimal transport and can be used to compare two positive measures whose supports live in the same underlying space
It does not require the hard constraints on the marginal distributions, so is more robust to outliers and has shown great interest in many real l applications.
However, it also has two major limitations. The first one is that it cannot compare measures whose supports are in different underlying spaces, for example one in R3 and one in R5, and
the second one is that, only alignments between samples are considered and there is no feature alignment, which can also be interpretable and allow to recover relations between the features of two different datasets.

S3: To overcome these two drawbacks, first we define the notion of sample-feature space. Here, we are given two compact measure spaces, Xs equipped with the measure mu_s and Xf equipped with the measure mu_f.
The superscript “s’’ indicates the sample space and the superscript “f’’ indicates the feature space.
We also define a scalar integrable function xi defined on the product space and we call the triplet comprising of these two spaces and the function xi a sample-feature space.
Xi is also called an interaction between the sample and feature spaces.
As an illustration, in the discrete setting, one can see the interaction between two points in the sample and feature spaces as the coordinate of the input matrix.

S4: Now, given two sample-feature spaces, we can define the unbalanced CO-Optimal Transport as shown in the formula. The first part of the objective function represents the transport cost using two interactions, and the second part containing the KL divergence roughly says that it is possible that a point can’t find a good match, so is not aligned to any other points.
Here, we are learning not only sample alignments pie_s, but also feature alignments pie_f, which will show great interest in many practical situations.
Note that, the purpose of using the product measure in the KL divergence are two-fold.
First, it allows to show the existence of minimizer and
second, it will be useful to show the robustness of Unbalanced CO Optimal transport.

S5: One of the main interest of UCOOT is its robustness to outliers. Before stating the main theorem, we first introduce some related notions, which might look complicated at first but in fact pretty intuitive.
Given two clean s.f. spaces X1 and X2, we introduce the noisy version of X1, which called X1 tilde, where in each of feature and sample space, we inject the outliers Os and Of, equipped with the noisy distribution eps_s and eps_f, respectively. Here, in X1 tilde, we consider a convex combination of noisy and clean distributions on its sample and feature spaces.
Once we introduce the outliers, we can define the two following costs: the first one delta zero represents the minimal transportation cost between the outliers and the clean sample-feature space X2.
The second one, delta infinity calculates the maximal transportation cost between the noisy X1 tilde and the clean X2.
Intuitively, the outliers will cause these two costs abnormally larger than the transportation cost between the clean X1 and X2.

S6: Now, we can show that COOT is very sensitive to the outliers because the lower bound on the right hand side is unbounded, so COOT can be made arbitrarily large. By contrast, UCOOT is very robust to outliers. The impact of outliers, which is the term delta infinity is rapidly saturated in the exponential term, so is always controlled and the UCOOT will never explode. In particular, if the outlier is too impactful, it will not receive any mass, so will not be aligned to any point.

S7: As an illustration, here we use the parameter tau to control the impact of outliers, the larger it is, the more impactful. If it is too much, then the UCOOT will be eventually constant, indicating that no mass is transported to the outliers. Clearly This is not the case for CO-Optimal Transport, where every point, even outlier must be aligned and this makes COOT easily explodes.

S8: Now, we illustrate the advantage of using UCOOT over COOT on the MNIST dataset.

In this experiment, we artificially add outliers to both feature (or, the pixels) and samples (which are the images). We see that very small pertubation is enough to degrade significantly the performance of COOT, whereas that of UCOOT remains stable, regardless of level of noise.

S9: Next, we apply UCOOT on heterogeneous domain adaptation. We use the Caltech-Office dataset and compare UCOOT with other 3 competing methods, in terms of accuracy. The source and targe domains contain the CaffeNet and GoogleNet pretrained embeddings of the images, respectively. We can see that in most tasks, UCOOT outperforms by large margin and on average, UCOOT is the most performing method.

S10: in this third experiment, we apply UCOOT to perform single cell multi omics alignement. We use the dataset generated by the CITE-seq experiment which simultaneously measures gene expression and antibody abundance in single cells. Here, we know the ground-truth correspondences on both the samples (i.e., cells) and the features (which are genes and their encoded antibodies), thus allowing us to quantify and compare the alignment performance of UCOOT and COOT. As you can see, overall, UCOOT gives much better alignments.

More details can be found in the paper. Thank you for your attention.