# Onlikorn: optimal transportation distances through recurrent sampling

**Anonymous Authors**[1]

## Abstract

We introduce a stochastic method to estimate optimal transport distances between two continuous probability distributions. Our algorithm handles streams of points sampled from these distributions, that it uses to iteratively enrich a non-parametric representation of the transportation plan. This plan is naturally represented as a mixture of simple (e.g. Gaussian) distributions, of increasing memory complexity. Compared to the classic Sinkhorn algorithm, our method handles new samples at each iteration, which permits a consistent estimation of the true regularized OT distance. We show how our algorithm can be casted as a block-convex mirror descent in the space of positive distributions; from this point of view, we analyse its convergence. Empirically, we demonstrate the efficiency of our estimation compared to concurrent methods.

## 1. Introduction

Optimal transport (OT) distances are fundamental in statistical learning, both as a tool for analyzing the convergence of various algorithms, and as a data-dependant term for estimating data density, e.g. using generative models. OT lifts a given distance over data points into a distance between distributions over the data space; as such, it allows to compare distributions with disjoint support. To alleviate the computational burden of optimal transport, it is common to regularize the linear problem that defines it, using an entropic barrier term. This approach, that has been rediscovered many times in the previous thirty years, allows to approximate OT distances using a matrix balancing algorithm, amenable to GPU computations.

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.

## References