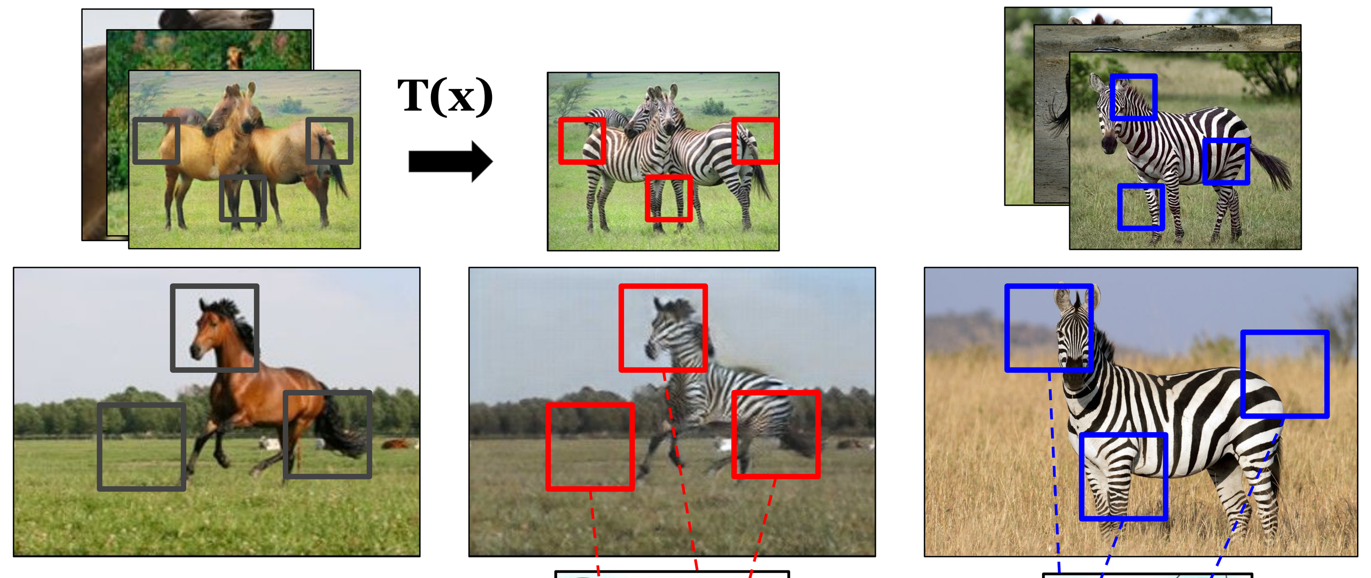
Log-Likelihood Ratio Minimizing Flows

Towards Robust and Quantifiable Neural Distribution Alignment

Visualizations, paper and code are available at:
ai.bu.edu/lrmf

Ben Usman usmn@bu.edu
Avneesh Sud asud@google.com
Nick Dufour ndufour@google.com
Kate Saenko saenko@bu.edu

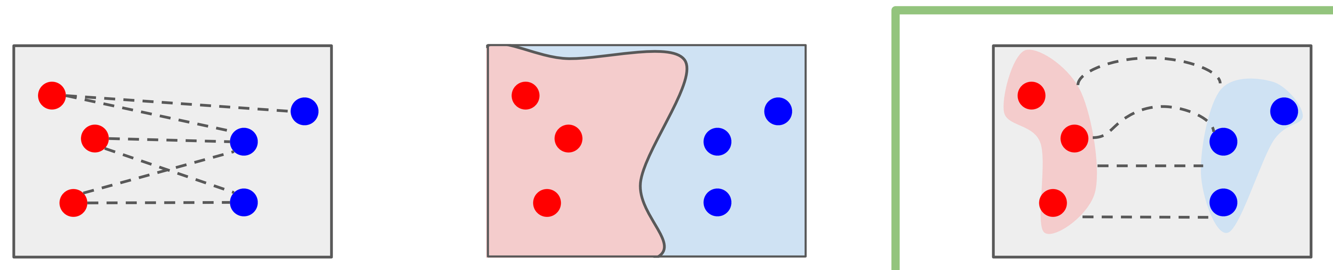
Task



A - source samples;
B - target samples;
T(A) - transformed source samples.

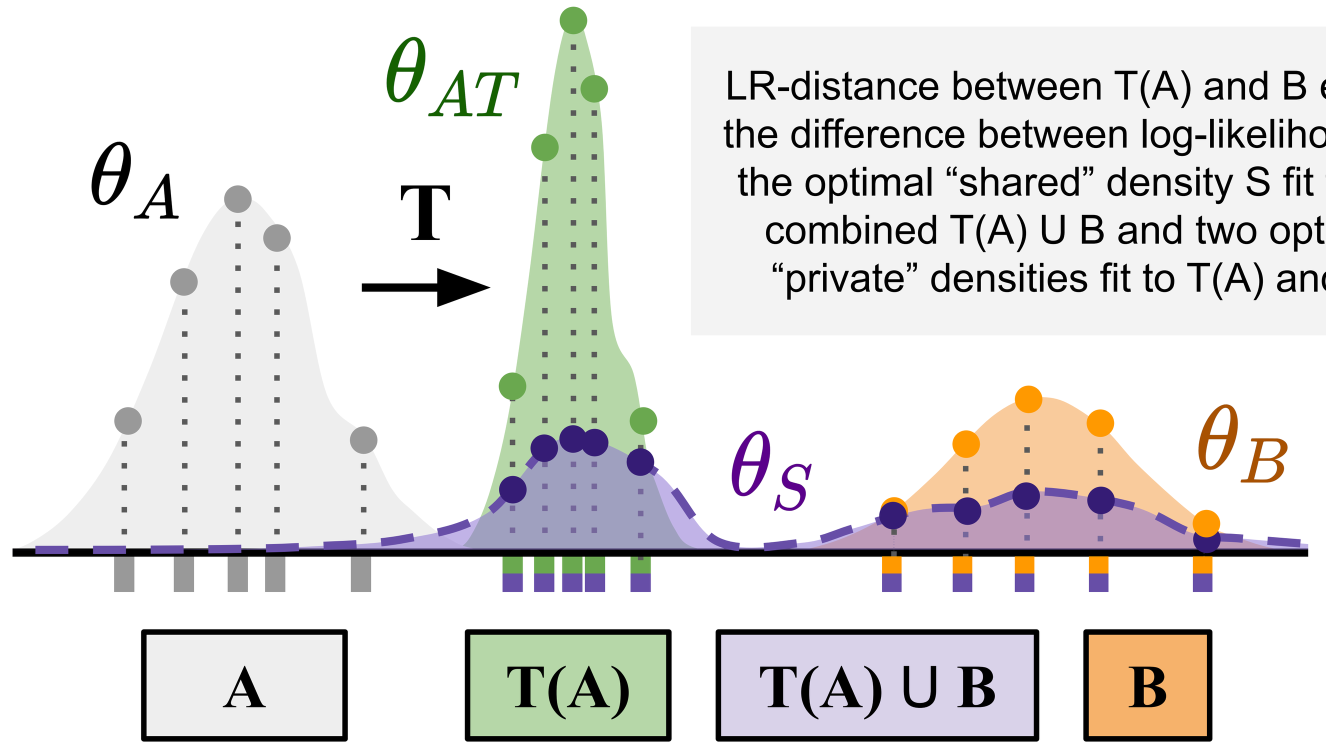
We want to know how to efficiently and reliably:
a) **measure** and b) **minimize**
the statistical distance between T(A) and B?

Prior Work



Non-parametric (MMD, EMD)	Adversarial (GAN, Monge–Kantorovich)	Log-Likelihood Ratio Minimizing Flows (LRMF) - ours
closed-form estimators exists ✓	min-max objective	closed-form upper bound ✓
can be minimized via gradient descent ✓	adversarial training	can be minimized via gradient descent ✓
model-free (metric-based)	powerful implicit data model (NN) ✓	any tractable density + normalizing flow ✓

Proposed Method: LRMF



LR-distance between T(A) and B equals the difference between log-likelihoods of the optimal "shared" density S fit to the combined T(A) U B and two optimal "private" densities fit to T(A) and B.

Observation 1 (⇒ Lemma 2.1):
The likelihood of the "shared" model (**S**) trained on the "combined" dataset is always **lower** than likelihoods of "private" models trained on each dataset alone (**AT**, **B**), **unless** both datasets are from the **same** distribution.

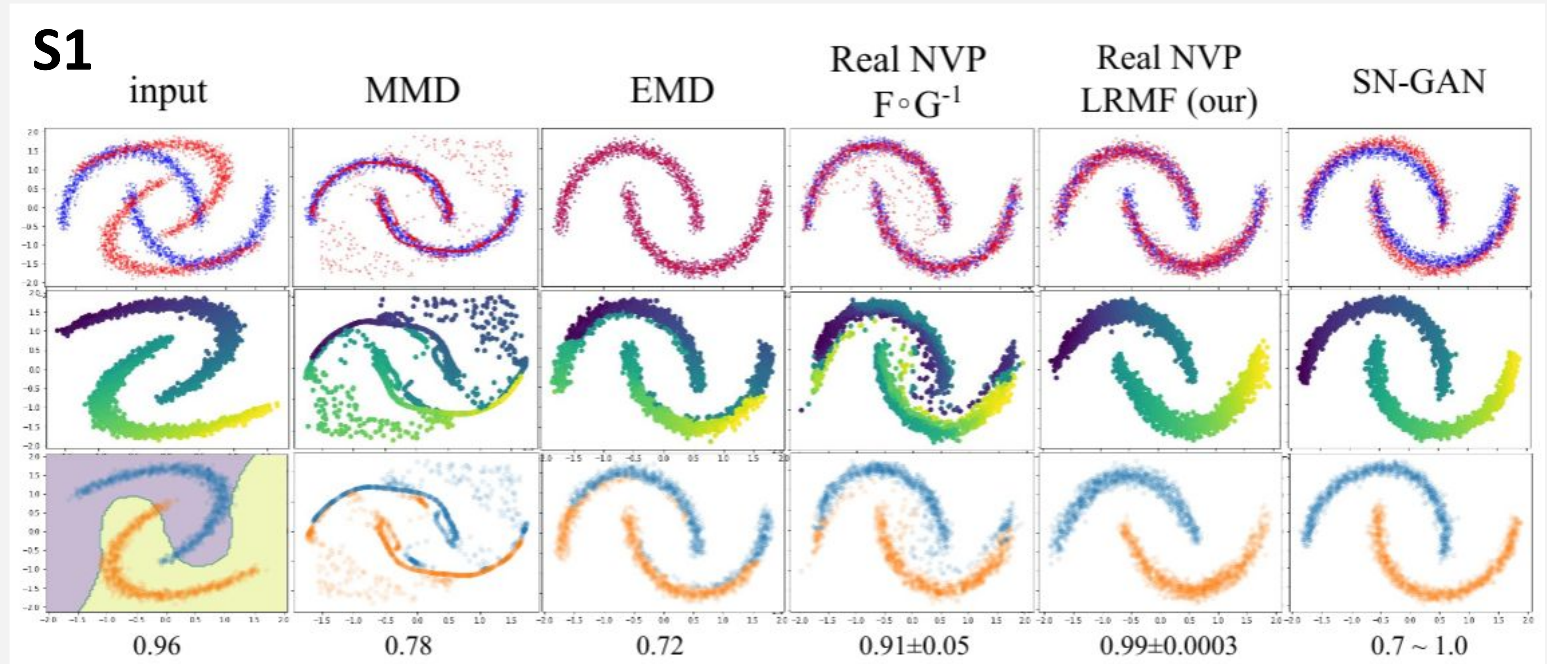
Observation 2 (⇒ Lemma 2.2):
The maximum likelihood of the transformed dataset T(A) can be **approximated in closed-form** if the maximum likelihood of A is known and T(x) is a normalizing flow.

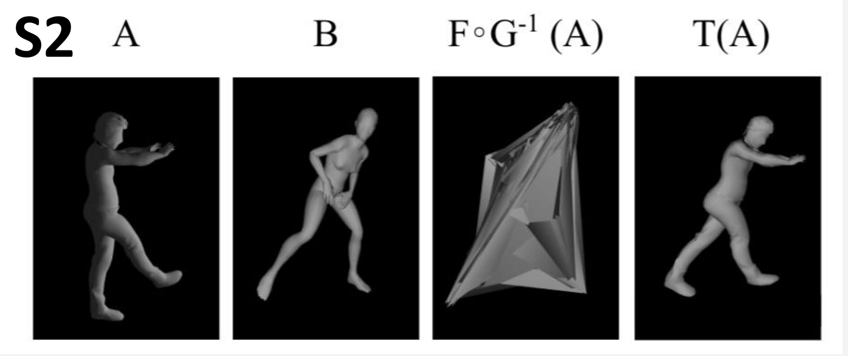
Conclusion (⇒ Theorem 2.3):
We can find the optimal flow T* that minimizes the adversarial LR-distance (the "gap" between shared private likelihoods) by minimizing a **non-adversarial LRMF**:

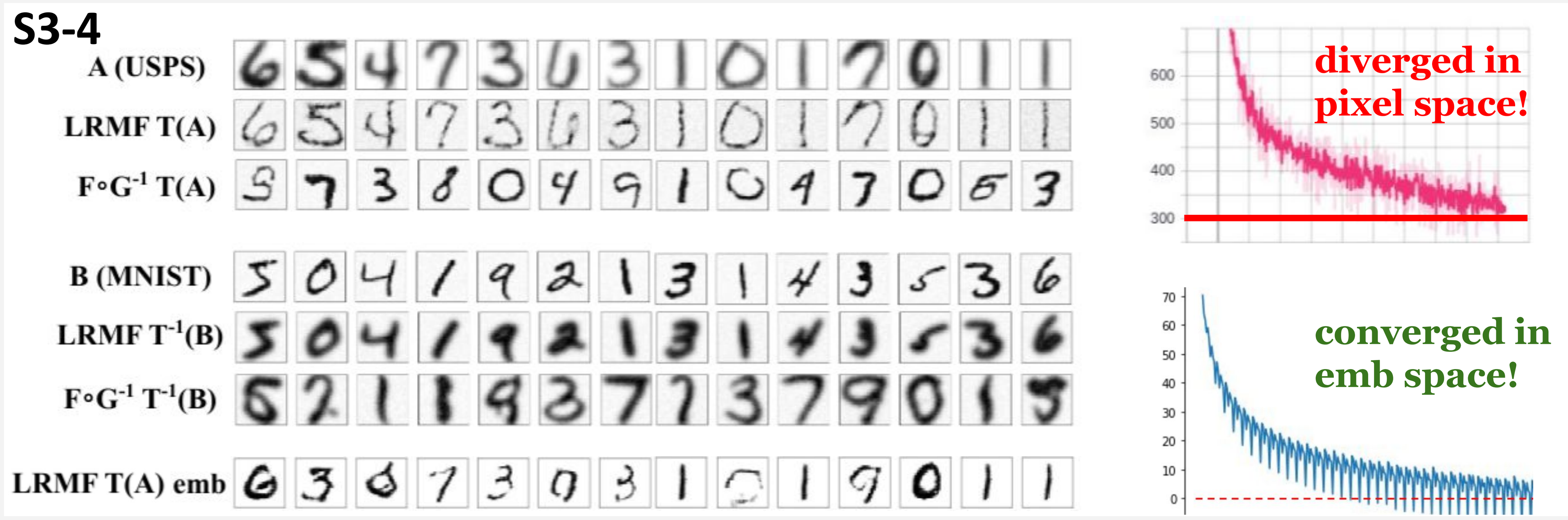
$$\mathcal{L}_{\text{LRMF}}(A, B, \phi, \theta_S) = -\log\det|\nabla_x T(A;\phi)| - \\ -\log P_M(T(A;\phi);\theta_S) - \log P_M(B;\theta_S) + c(A,B)$$

Results

1. **Setup 1:** A RealNVP LRMF: 1) produced good alignment, 2) converged to zero loss, 3) preserved the manifold structure better.
2. **Setup 2:** It preserved the manifold structure of aligned mesh vertex distributions better then our AlignFlow-inspired F ∘ G⁻¹ baseline.
3. **Setup 3:** A RealNVP LRMF successfully aligned digit images in the emb space of a VAE and the LRMF loss converged to zero.
4. **Setup 4:** A GLOW LRMF failed to converge, and the non-zero final loss value explicitly indicates this.

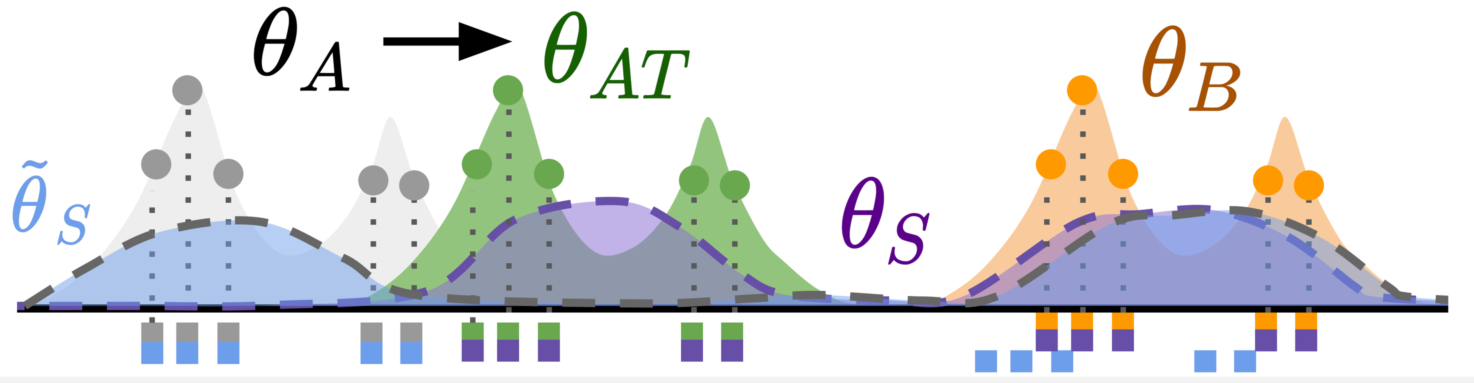






Special Cases and Limitations

1. **Special cases: 1)** Gaussian LRMF ⇔ matching mean and variance; **2)** minimizing an infinite capacity LRMF loss ⇔ training a GAN with a closed-form D(x) ⇔ minimizing Jensen-Shannon divergence.



If A and B are far apart, any planar *T(x; b)* does not change the likelihood of *T(A)* or *S*, so the LRMF objective is locally constant w.r.t. the transformation parameter *b*.

2. **Limitations**: transformation gradients of LRMF between two gaussian mixtures vanish as distribution means become further away from each other:

$$||[\partial\mathcal{L}_{\text{LRMF}}(A+\mu, B, \phi, \theta)/\partial\phi]|| \propto \exp(-\mu^2)$$