

COURANT INSTITUTE SCHOOL OF MATHEMATICS,
COMPUTING, AND DATA SCIENCE

Active Learning for Conditional Generative models via transport Neural Tangent Kernels (tNTK)

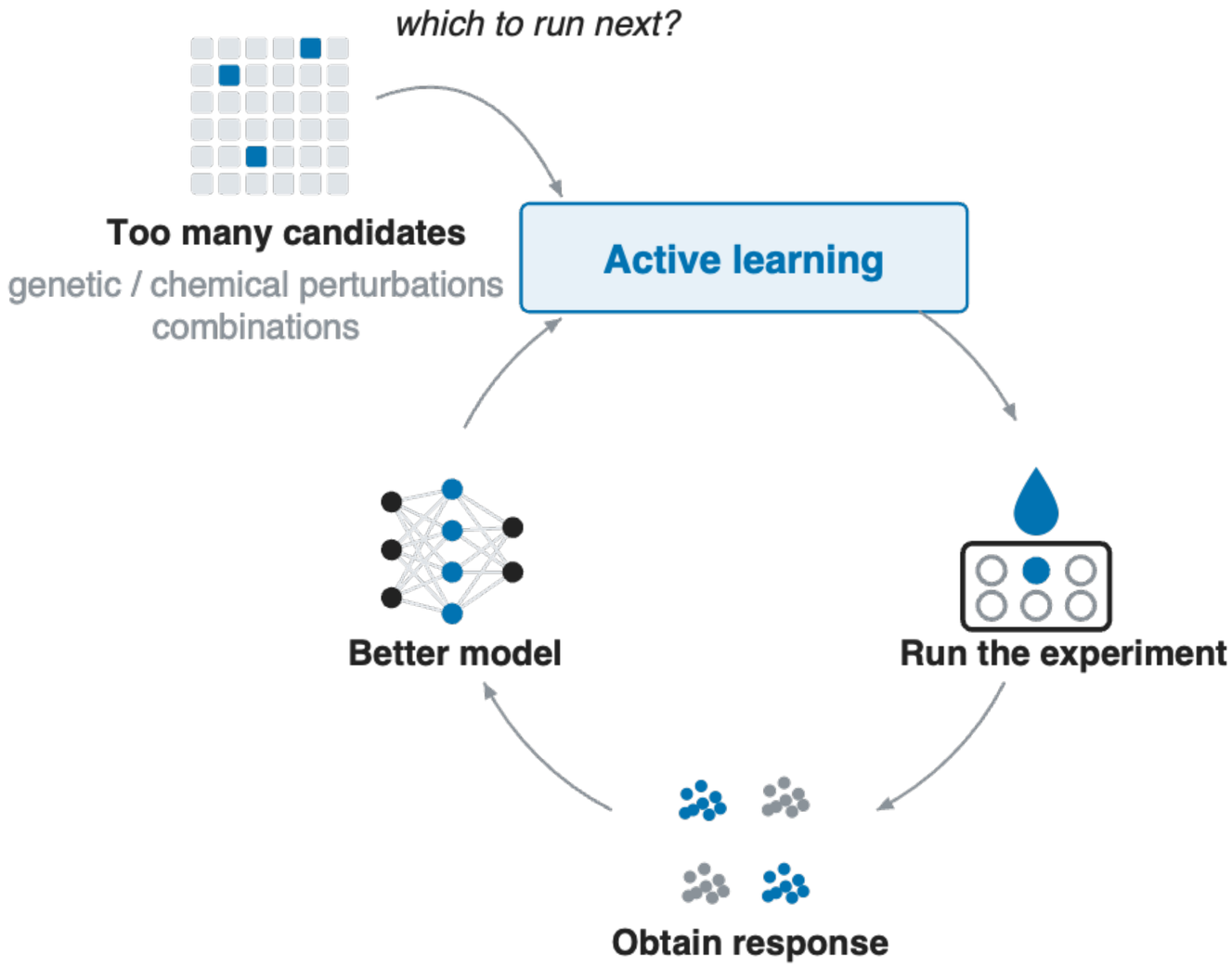
NeurIPS 2026

Jayoung Ryu, Nov 10th 2026
Work with Kyunghyun Cho & Romain Lopez

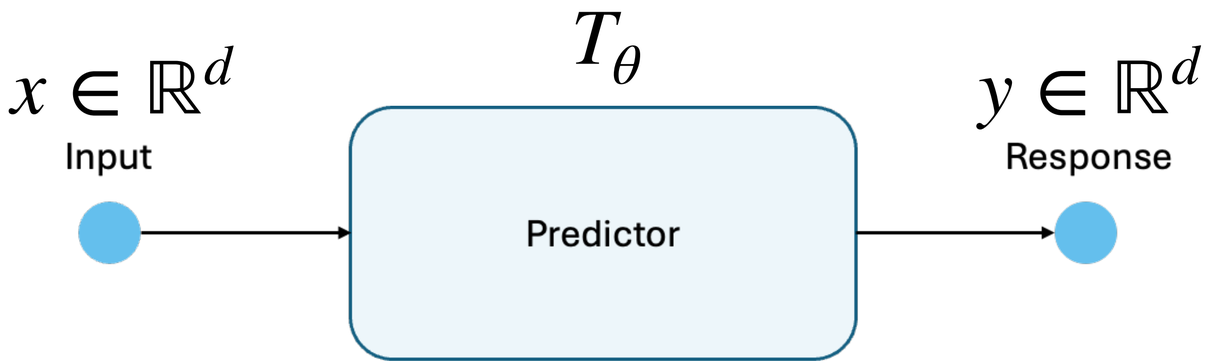
Prediction task in Science

- Response prediction in unseen samples
- With growing model capacity, training data is the bottleneck.
- For many tasks, the response is inherently random.

Active learning chooses which samples to label

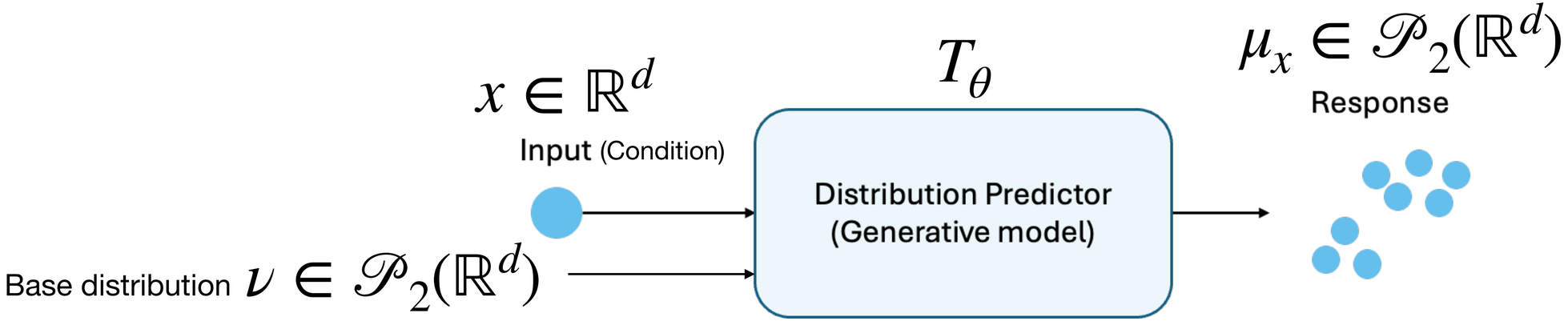


Active learning: problem setup



- Posterior risk $\mathscr{R}(x, \mathscr{D}) = \mathbb{E}_{\theta|\mathscr{D}}[\|T_\theta(x) - y_x\|^2]$
- Uncertainty of the output $Var_{\theta|\mathscr{D}}[T_\theta(x)] \leq \mathscr{R}(x, \mathscr{D})$
- Algorithm updates $\mathscr{D}$ that minimizes risk or the variance

Conditional generative model predicts conditional distributions



$$\min_\theta W_2^2(T_\theta(x)_\#\nu, \mu_x)$$

Pushforward

particles $y \sim \nu \to T_\theta(x)(y) \sim T_\theta(x)_\#\nu$

Transport NTK

Active learning algorithms use Gaussian process trick that uses NTK as the prior covariance

$$\Longrightarrow \text{We define } k_\theta^{\text{tNTK}}(x, x') = \text{Tr}\left[\int D_\theta T_\theta(x)(y)^\top D_\theta T_\theta(x')(y)\, d\nu(y)\right]$$

1. Upper bounds the variance

$$V_F(x; \theta \sim \mathcal{N}(\hat{\theta}, \Sigma)) \leq \text{Tr}\left(G_{\hat{\theta}}(x)\,\Sigma\right) + C(x)\|\Sigma\|_{\text{op}}^{3/2}$$

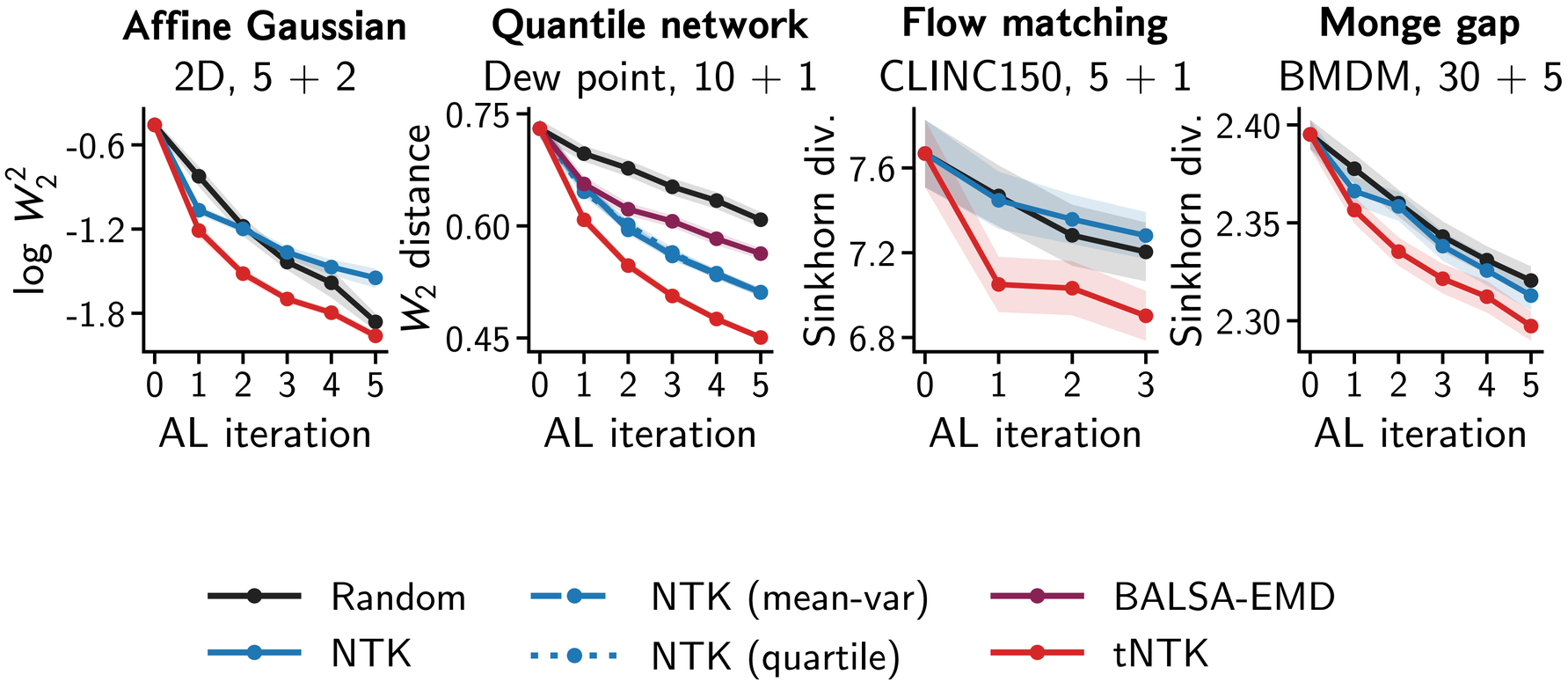
$$Var_\theta(f_\theta(x)) = \text{Tr}(\Sigma\, \nabla_\theta f_\theta(x)\, \nabla_\theta f_\theta(x')^\top)\,.$$

$$G_\theta(x) = \int D_\theta T_\theta(x)(y)^\top D_\theta T_\theta(x)(y) d\nu(y) \in \mathbb{R}^{p\times p}, \quad \theta \in \mathbb{R}^p; \qquad k(x,x) = \text{Tr}(G_\theta(x))\,.$$

2. Upper bounds the metric derivative of parameter perturbation curve

tNTK outperforms in 27/31 configurations

- Configurations: model class / initial batch size / batch size combinations



Conclusion

- Acquisition criteria that is aware of output distributions is needed for active learning of conditional generative models.
- tNTK is the drop-in replacement of kernel-based batched active learning approach
- tNTK achieves better active learning than baselines in multiple generative model classes

Thank you!