Probabilistic Flow Circuits: Towards Unified Deep Models for Tractable Probabilistic Inference

Sahil Sidheekh Kristian Kersting Sriraam Natarajan

StARLinG LAB — The University of Texas at Dallas

Technische Universität Darmstadt

What ?

Motivation

The generative modeling dichotomy: tractability vs. expressivity

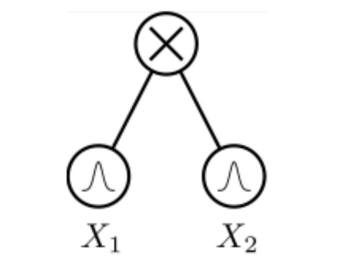
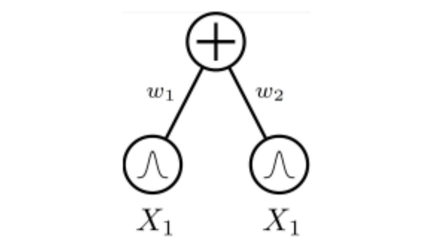
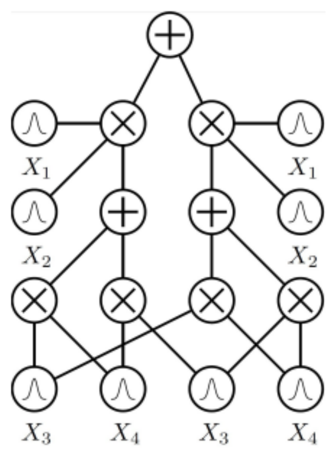
Tractable — **No Free Lunch** — **Expressive**

Naïve Bayes, Trees, And-Or-Graphs, SPN, Einsum Networks, Normalizing Flows, Autoregressive, VAEs, GANS

Can we borrow concepts from both ends of the spectrum to **bridge** this gap? Build **expressive** and **tractable** models?

Background

Tractability with Probabilistic Circuits

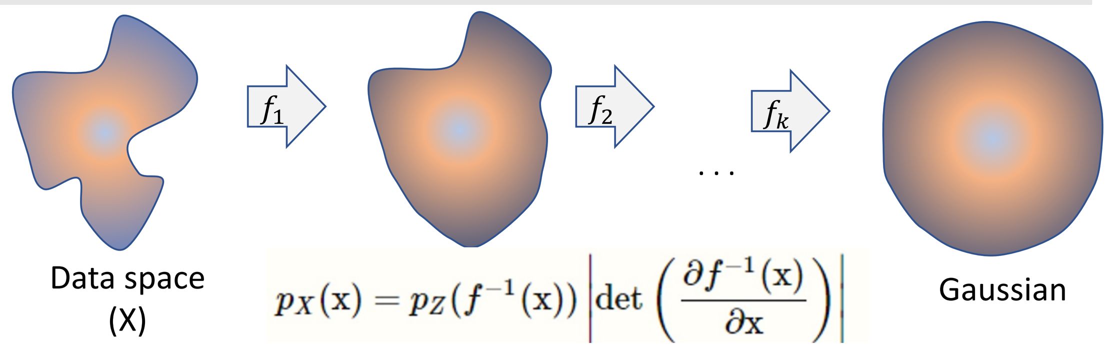
Leaf Nodes	Product Nodes	Sum Nodes	Stacked
	 $p(X_1,X_2)=p(X_1)\cdot p(X_2)$	 $p(X_1)=w_1p_1(X_1)+w_2p_2(X_1)$	
Simple univariate distribution	Represents **factorizations**	Represents **mixtures**	
	Enables tractability	Adds expressivity	

Expressivity with Normalizing Flows

Learn probability distributions using invertible transformations and the change of variables formula

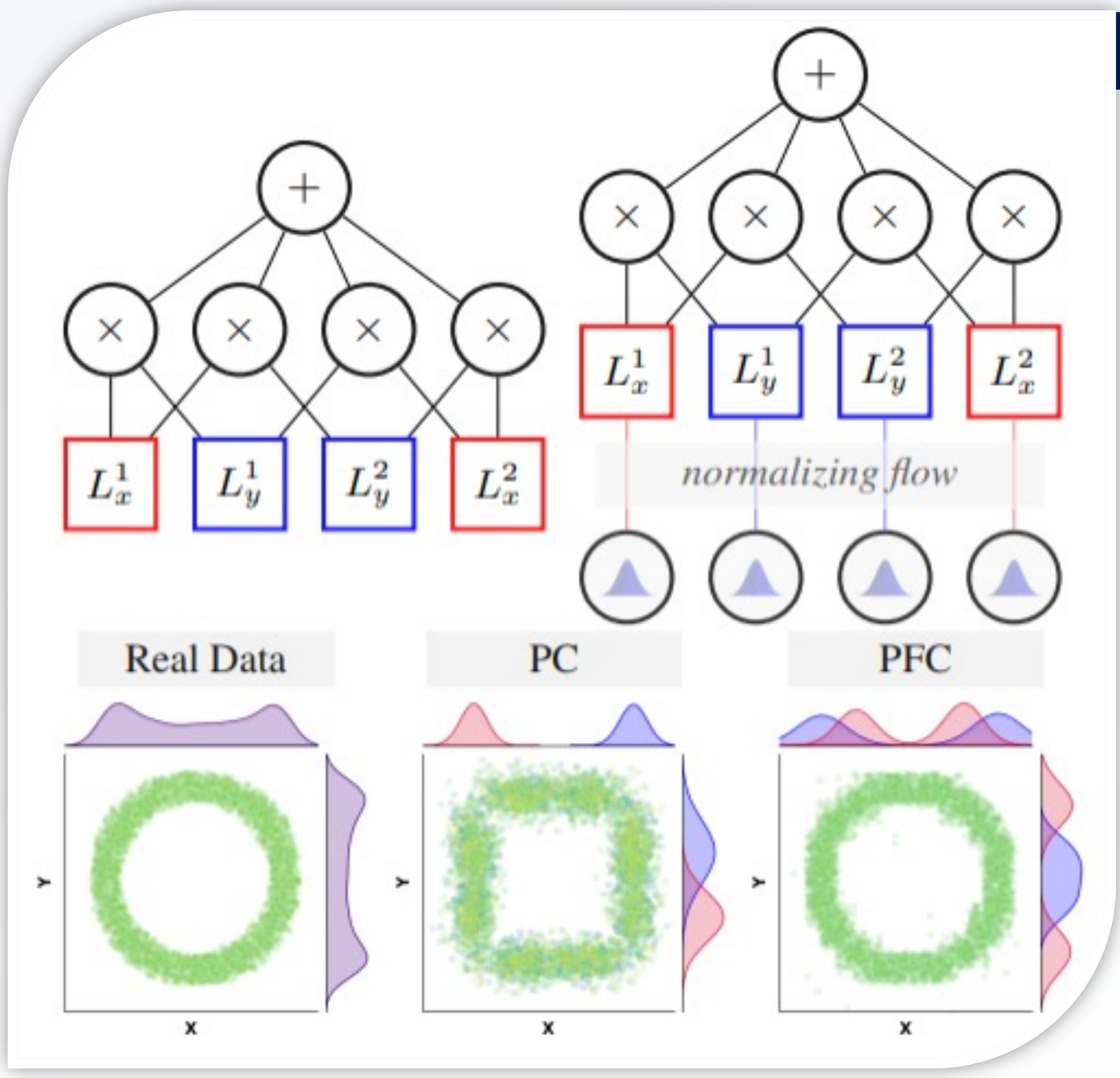
Neural Network parameterization

Supports **exact** density evaluation



$$p_X(x)=p_Z(f^{-1}(x))\left|\det\left(\frac{\partial f^{-1}(x)}{\partial x}\right)\right|$$

How ?



Probabilistic Flow Circuit (PFC)

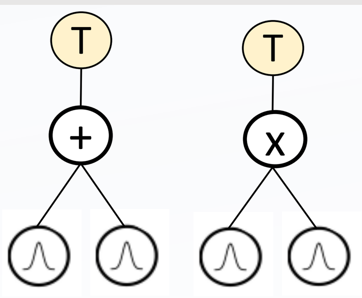
Has added Expressivity	Retains Tractability
Can model arbitrarily complex distributions at the leaves	As it encodes **factorizations** of the joint distribution

Integrating Flows with PCs

Introduce **invertible transform nodes** arbitrarily in a PC [Pevny et. Al 2020]

Represents **normalizing** flows

Adds expressivity



$$T(N(\boldsymbol{x}))=N(f(\boldsymbol{x}))|\det J_f|$$

$N=\{\oplus, \otimes, \bigcirc\}$

But this affects decomposability → Tractability violated → Need more structure

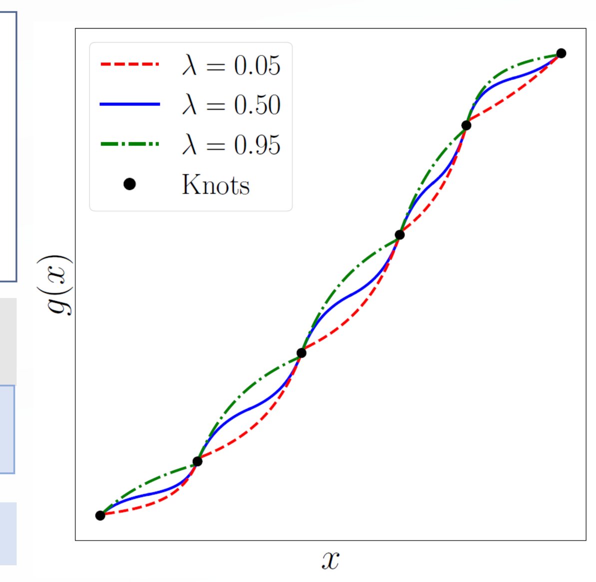
τ-Decomposability

- When defined over a product node, T needs to transform the variables involved in the scope of its children **independently**
- A **necessary condition** for tractability of marginal and conditional

Implications: Reduces to having flows at the **leaves**!

What transformation to use? Linear Rational Splines (LRS)

1. Divide the data space into K bins and fit a polynomial function f_k within each bin
2. Use monotone linear rational functions of the form: $f(x)=\frac{ax+b}{cx+d}$



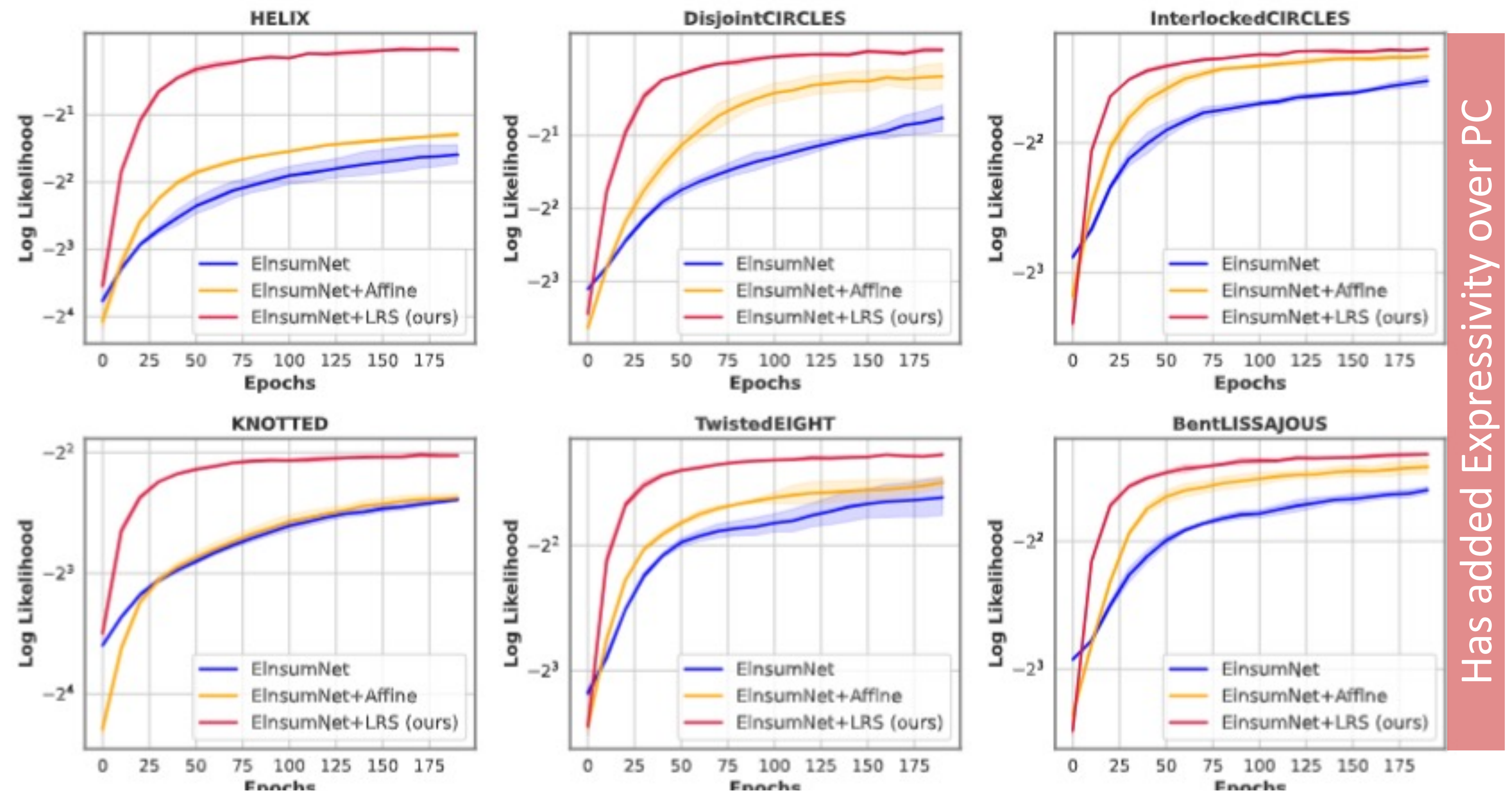
A PFC with LRS transformations is tractable for

Evidential	Marginal	Conditional	MAP

If you use a Student's-t base distribution and PC is deterministic

What do we get ?

Empirical Evaluations



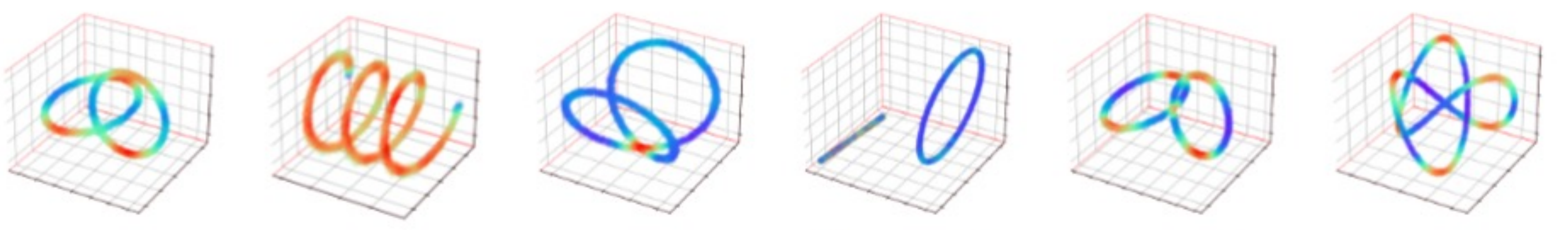
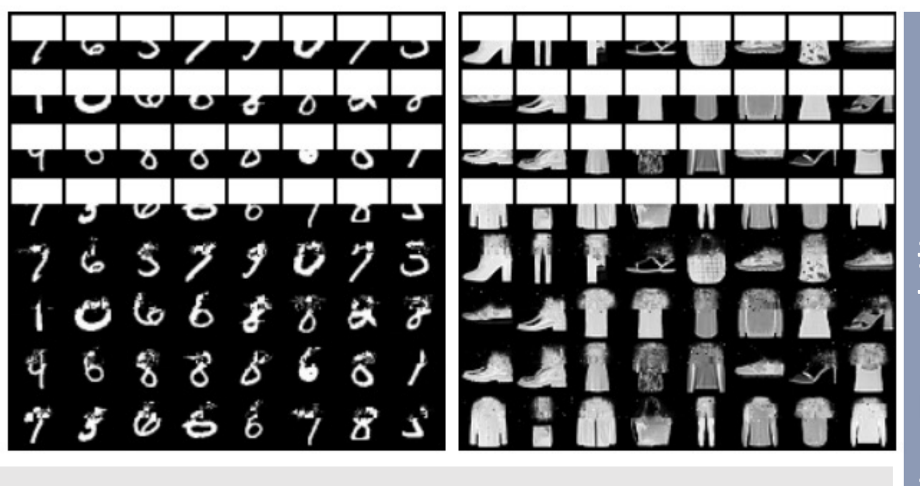
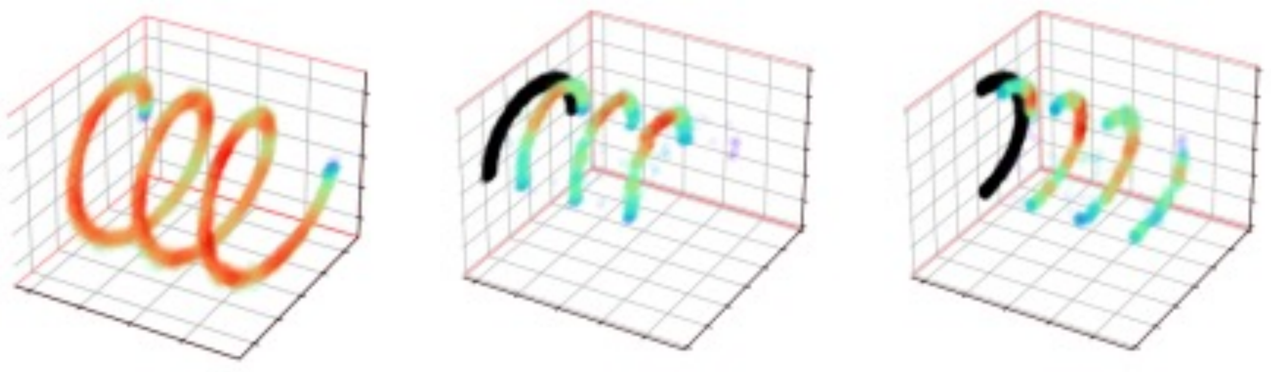


Image Inpainting

Can fill in missing data by sampling from the conditional distribution over occluded pixels



Controlled generation



Retains Tractability

Acknowledgements

SN acknowledges the support by the U. S. Army Research Laboratory and the U. S. Army Research Office (ARO) under grant number W911NF2010224.

References

[1.] Choi, Y., Antonio Vergari, and Guy Van den Broeck. "Probabilistic circuits: A unifying framework for tractable probabilistic models." 2020

[2.] Pevný, Tomáš, et al. "Sum-product-transform networks: Exploiting symmetries using invertible transformations." PGM 2020.

[3.] Peharz, Robert, et al. "Einsum networks: Fast and scalable learning of tractable probabilistic circuits." ICML 2020.

https://starling.utdallas.edu/ https://sahilsid.github.io/ @STARLing_lab