

On the Hardness of Approximating Distributions with Tractable Probabilistic Models

John Leland & YooJung Choi

Tractable Probabilistic Models

For a probabilistic model to be **tractable** it must meet the following requirements:

Tractable Probabilistic Models

For a probabilistic model to be **tractable** it must meet the following requirements:

- The model can perform exact inference queries.
- The queries can be performed in polynomial time in the size of the model.

Tractable Probabilistic Models

For a probabilistic model to be **tractable** it must meet the following requirements:

- The model can perform exact inference queries.
- The queries can be performed in polynomial time in the size of the model.

Examples of TPMs: probabilistic circuits [3], low tree-width graphical models [5], determinantal point processes [1], etc.

Setup

- We restrict to probabilistic circuits as a unifying framework for TPMs.

Setup

- We restrict to probabilistic circuits as a unifying framework for TPMs.
- Many TPMs can be modeled as classes of PCs satisfying structural properties.

Setup

- We restrict to probabilistic circuits as a unifying framework for TPMs.
- Many TPMs can be modeled as classes of PCs satisfying structural properties.
 - Structural properties to ensure tractable inference:

Setup

- We restrict to probabilistic circuits as a unifying framework for TPMs.
- Many TPMs can be modeled as classes of PCs satisfying structural properties.
 - Structural properties to ensure tractable inference:

Setup

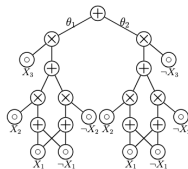
- We restrict to probabilistic circuits as a unifying framework for TPMs.
- Many TPMs can be modeled as classes of PCs satisfying structural properties.
 - Structural properties to ensure tractable inference:

Condition	Inference
Decomposability	Marginals
Determinism	MAP

Setup

- We restrict to probabilistic circuits as a unifying framework for TPMs.
- Many TPMs can be modeled as classes of PCs satisfying structural properties.
 - Structural properties to ensure tractable inference:

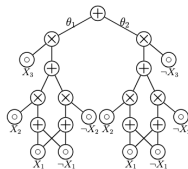
Condition	Inference
Decomposability	Marginals
Determinism	MAP



Setup

- We restrict to probabilistic circuits as a unifying framework for TPMs.
- Many TPMs can be modeled as classes of PCs satisfying structural properties.
 - Structural properties to ensure tractable inference:

Condition	Inference
Decomposability	Marginals
Determinism	MAP

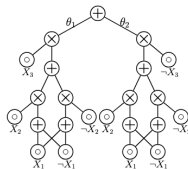


- NP-hard to represent distributions exactly with TPMs

Setup

- We restrict to probabilistic circuits as a unifying framework for TPMs.
- Many TPMs can be modeled as classes of PCs satisfying structural properties.
 - Structural properties to ensure tractable inference:

Condition	Inference
Decomposability	Marginals
Determinism	MAP

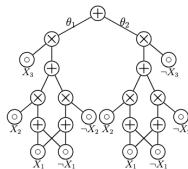


- NP-hard to represent distributions exactly with TPMs
- Exponential gap in expressive efficiency between decomposable circuits and deterministic, decomposable circuits in the exact case! [2, 4]

Setup

- We restrict to probabilistic circuits as a unifying framework for TPMs.
- Many TPMs can be modeled as classes of PCs satisfying structural properties.
 - Structural properties to ensure tractable inference:

Condition	Inference
Decomposability	Marginals
Determinism	MAP



- NP-hard to represent distributions exactly with TPMs
- Exponential gap in expressive efficiency between decomposable circuits and deterministic, decomposable circuits in the exact case! [2, 4]
- **Main Question:** *Can we avoid size blowups by allowing some small approximation error?*

Allowing Approximation Error

- Approximate modeling: $D_f(P\|Q) < k\epsilon^2$ or $D_{\text{TV}}(P\|Q) < \epsilon$.

Allowing Approximation Error

- Approximate modeling: $D_f(P\|Q) < k\epsilon^2$ or $D_{TV}(P\|Q) < \epsilon$.
 - KL Divergence, Hellinger distance, Pearson divergence

Hardness of D_f -Approximation

Theorem: It is **NP-hard** to approximately model using an f -divergence for any model that can tractably marginalize.

Hardness of D_f -Approximation

Theorem: It is **NP-hard** to approximately model using an f -divergence for any model that can tractably marginalize.

- This is proved by a reduction from $\text{SAT}(f)$
- This result applies to **all TPMs!**

Large Dec. & Det. PCs for Approximate Modeling

Approximate modeling is NP-hard, but can we overcome the exponential gap caused by enforcing determinism?

- Enforcing determinism causes exponential blowups in the exact setting [2, 4].
- We show this extends to the approximate setting in both Boolean and Probabilistic Circuits.
- Connecting from approximating Boolean circuits to Probabilistic Circuits is non-trivial:
 - Approximation for PCs measured in terms of divergences.

Large Dec. & Det. PCs for Approximate Modeling

Approximate modeling is NP-hard, but can we overcome the exponential gap caused by enforcing determinism?

- Enforcing determinism causes exponential blowups in the exact setting [2, 4].
- We show this extends to the approximate setting in both Boolean and Probabilistic Circuits.
- Connecting from approximating Boolean circuits to Probabilistic Circuits is non-trivial:
 - Approximation for PCs measured in terms of divergences.
 - PCs can represent arbitrary distributions.

Large Dec. & Det. PCs for Approximate Modeling

- Showing there exists an exponential blowup in the approximate case:

Large Dec. & Det. PCs for Approximate Modeling

- Showing there exists an exponential blowup in the approximate case:
 - Take a Decomposable PC, P_n .

Large Dec. & Det. PCs for Approximate Modeling

- Showing there exists an exponential blowup in the approximate case:
 - Take a Decomposable PC, P_n .
 - Deterministic, decomposable $\mathcal{D}_{TV}$ -approximator $\implies$ weak approximator of Boolean function.

Large Dec. & Det. PCs for Approximate Modeling

- Showing there exists an exponential blowup in the approximate case:
 - Take a Decomposable PC, P_n .
 - Deterministic, decomposable $\mathcal{D}_{TV}$ -approximator $\implies$ weak approximator of Boolean function.
 - Exists a function such that P_n has size $O(n^2)$, but any deterministic, decomposable PC approximating P_n must be exponentially sized!

Cannot avoid the exponential gap in expressive efficiency by allowing for approximation!

Inferences Guarantees in Approximate Modeling

- Approximate Modeling is hard!

Inferences Guarantees in Approximate Modeling

- Approximate Modeling is hard!
- Can we at least say that inference using an approximate model is close to the true query answers?

Inferences Guarantees in Approximate Modeling

- Approximate Modeling is hard!
- Can we at least say that inference using an approximate model is close to the true query answers?

	$D_f(P\ Q)$	$D_{\text{TV}}(P\ Q)$
aMAR	✓	✓
aMAP	✓	✓
rMAR	○	○
rMAP	○	○
aCON	○	○
rCON	○	○

Thank you!

References

- [1] Borodin. Determinantal point processes, 2009.
- [2] Bova et al. Knowledge compilation meets communication complexity. In *IJCAI*, 2016.
- [3] Choi et al. Probabilistic circuits: A unifying framework for tractable probabilistic models, 2020.
- [4] Arthur Choi and Adnan Darwiche. On relaxing determinism in arithmetic circuits. In *ICML*, 2017.
- [5] Gal Elidan and Stephen Gould. Learning bounded treewidth bayesian networks. *Advances in neural information processing systems*, 21, 2008.