Computational Explorations of Total Variation Distance

Arnab Bhattacharyya[1], Sutanu Gayen[2], Kuldeep S. Meel[3], **Dimitrios Myrisiotis**[4], A. Pavan[5], and N. V. Vinodchandran[6]

[1]University of Warwick • [2]Indian Institute of Technology Kanpur • [3]University of Toronto; Georgia Institute of Technology • [4]CNRS@CREATE LTD. [5]Iowa State University • [6]University of Nebraska-Lincoln

March 22, 2025

What Is All About

What Is All About

Definition (Total variation distance)

For probability distributions P, Q over a common finite domain D, the *total variation (TV) distance between P and Q* is

$$d_{\mathrm{TV}}(P, Q) := \frac{1}{2} \sum_{x \in D} |P(x) - Q(x)| \, .$$

What Is All About

TV distance is important, because it has many desirable properties: It is a metric, it is bounded in $[0, 1]$, and is invariant with respect to bijections.

Related Work

Related Work

- Bhattacharyya, Gayen, Meel, Myrisiotis, Pavan, and Vinodchandran [BGM$^+$23] proved that exact computation of TV distance between products of Bernoulli distributions is #P-hard.

Related Work

- Bhattacharyya, Gayen, Meel, Myrisiotis, Pavan, and Vinodchandran [BGM+23] proved that exact computation of TV distance between products of Bernoulli distributions is #P-hard.
- Sahai and Vadhan [SV03] showed that TV distance is SZK-hard to additively estimate for distributions samplable by Boolean circuits.

Related Work

- On an algorithmic note, Bhattacharyya, Gayen, Meel, and Vinodchandran [BGMV20] designed efficient algorithms to additively estimate the TV distance between distributions that are efficiently samplable and efficiently computable.

Related Work

- Feng, Guo, Jerrum, and Wang [FGJW23] designed an FPRAS for estimating TV distance between products of Bernoulli distributions.

Related Work

- Feng, Guo, Jerrum, and Wang [FGJW23] designed an FPRAS for estimating TV distance between products of Bernoulli distributions.
- Later, Feng, Liu, and Liu [FLL24] gave an FPTAS for the same task.

Related Work

- Bhattacharyya, Gayen, Meel, Myrisiotis, Pavan, and Vinodchandran [BGM+24] gave a novel reduction from TV distance estimation to probabilistic inference.

  This connection yields an FPRAS for estimating the TV distance between Bayes nets of small treewidth.

Related Work

- Feng, Liu, and Yang [FLY25] studied TV distance between spin systems.

Related Work

- Feng, Liu, and Yang [FLY25] studied TV distance between spin systems.
- Bhattacharyya, Feng, Srivastava [BFS25] studied TV distance between Gaussians.

Our Results

Our Results: Mixtures of Product Distributions

Our Results

A *mixture of product distributions* is a probability distribution over products of Bernoulli distributions.

Our Results

Theorem

There is an efficient deterministic algorithm such that, given two mixtures of product distributions P and Q, decides whether $P = Q$ or not.

Our Results

This result is proved using elementary linear algebra techniques.

Our Results

The equality $P = Q$ can be seen as a system of equations, that may be captured by a small basis of coefficient vectors.

To check these equations, one may only check it for the basis vectors!

Our Results

This yields an efficient inductive algorithm to check equivalence.

Our Results:
Ising Models

Our Results

Ising models are probabilistic graphical models which are very popular in statistical physics.

Our Results

Theorem

If NP $\not\subseteq$ RP, *then there is no FPRAS that estimates the TV distance between Ising models.*

Our Results

This proof is based on the observation that

estimating the partition function of Ising models can be reduced to estimating TV distance between Ising models.

Our Results

The result then follows from a result by Jerrum and Sinclair [JS93], where they show that

if NP $\not\subseteq$ RP, then there is no FPRAS that estimates the partition function of Ising models.

Open Problems

Open Problems

- Can our ideas be extended to checking equivalence between mixtures of some subclass of Bayes net distributions, such as Bayes net distributions whereby their underlying graph is a tree or has small treewidth?

Open Problems

- Can our ideas be extended to checking equivalence between mixtures of some subclass of Bayes net distributions, such as Bayes net distributions whereby their underlying graph is a tree or has small treewidth?
- Can we extend this complexity-theoretic hardness of approximation result to other classes of probability distributions, such as factor graphs or general undirected probabilistic graphical models?

Our Work on arXiv

Our Work on arXiv

Thank You!

References

[BFS25] Arnab Bhattacharyya, Weiming Feng, and Piyush Srivastava.
Approximating the total variation distance between gaussians, 2025.

[BGM+23] Arnab Bhattacharyya, Sutanu Gayen, Kuldeep S. Meel, Dimitrios Myrisiotis, A. Pavan, and N. V. Vinodchandran.
On approximating total variation distance.
In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI 2023, 19th-25th August 2023, Macao, SAR, China*, pages 3479–3487. ijcai.org, 2023.

References

[BGM$^+$24] Arnab Bhattacharyya, Sutanu Gayen, Kuldeep S. Meel, Dimitrios Myrisiotis, A. Pavan, and N. V. Vinodchandran.
Total variation distance meets probabilistic inference.
In *Forty-first International Conference on Machine Learning, ICML 2024, Vienna, Austria, July 21-27, 2024*. OpenReview.net, 2024.

[BGMV20] Arnab Bhattacharyya, Sutanu Gayen, Kuldeep S. Meel, and N. V. Vinodchandran.
Efficient distance approximation for structured high-dimensional distributions via learning.
In *Proc. of NeurIPS*, 2020.

References

[FGJW23] Weiming Feng, Heng Guo, Mark Jerrum, and Jiaheng Wang.
A simple polynomial-time approximation algorithm for the total variation distance between two product distributions.
TheoretiCS, 2, 2023.

[FLL24] Weiming Feng, Liqiang Liu, and Tianren Liu.
On deterministically approximating total variation distance.
In *Proceedings of the 2024 ACM-SIAM Symposium on Discrete Algorithms, SODA 2024, Alexandria, VA, USA, January 7-10, 2024*, pages 1766–1791. SIAM, 2024.

References

[FLY25] Weiming Feng, Hongyang Liu, and Minji Yang.
Approximating the total variation distance between spin systems, 2025.

[JS93] Mark Jerrum and Alistair Sinclair.
Polynomial-time approximation algorithms for the ising model.
SIAM J. Comput., 22(5):1087–1116, 1993.

[SV03] Amit Sahai and Salil P. Vadhan.
A complete problem for Statistical Zero Knowledge.
J. ACM, 50(2):196–249, 2003.