Divergence Frontiers for Generative Models: Sample Complexity, Quantization Level, and Frontier Integral

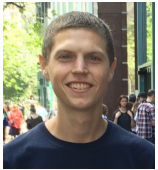
Lang Liu, Krishna Pillutla, Sean Welleck, Sewoong Oh, Yejin Choi, Zaid Harchaoui

University of Washington

Team

Lang Liu

Krishna Pillutla



Sean Welleck

Sewoong Oh

Yejin Choi

Zaid Harchaoui

Image and Text Generation

High quality but low variety

Kynkäänniemi et al. (2019)

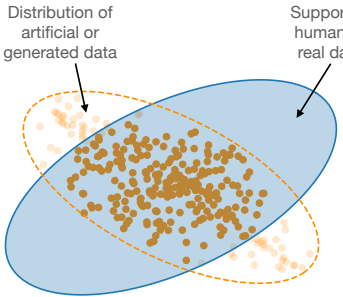
Several people have asked about the techniques we used when cleaning out my mom's fabric stash last week..... Next, you need to get **a small, sharp knife**. I like to use **a small, sharp knife**. I like to use **a small, sharp knife**.

Pillutla et al. (2021)

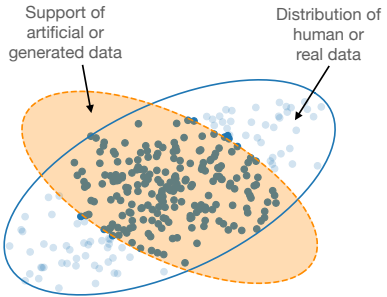
Low quality but high variety

Several people have asked about the techniques we used when cleaning out my mom's fabric stash last week..... I had a great deal of **décor management** and was able to **stash the excess items away for safekeeping**.

Type I and Type II Errors in Generative Modeling



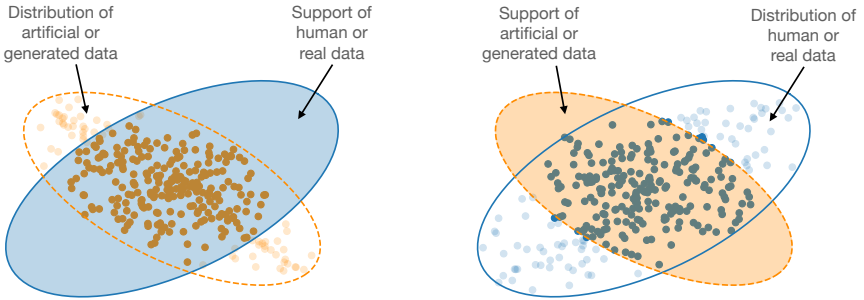
Type I error



Type II error

How to quantify them?

Type I and Type II Errors in Generative Modeling



Type I error

$\mathrm{KL}(Q\|P)$

P: real data distribution

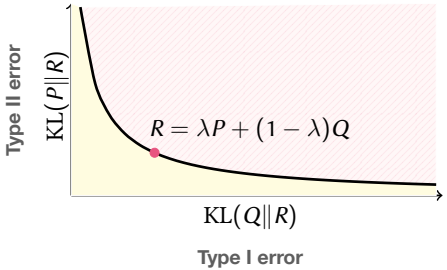
Q: generated data distribution

Type II error

$\mathrm{KL}(P\|Q)$

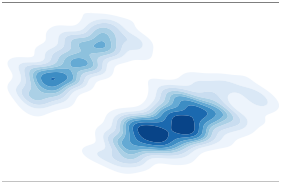
Divergence Frontiers for Generative Models

- Divergence frontiers for data distribution P and model distribution Q.
- Applications in vision (Sajjadi et al. '18, Kynkäänniemi et al. '19, Djolonga et al. '20).
- Applications in NLP (Pillutla et al. '21; this NeurIPS).

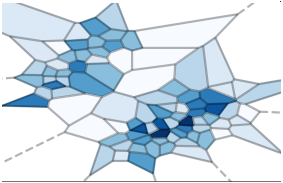


Estimation Procedure of Divergence Frontiers

Continuous Distribution



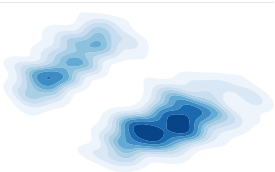
Quantized Distribution



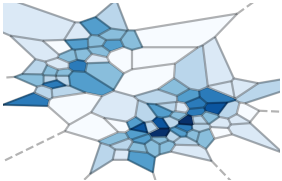
Empirical Estimator

Estimation Procedure of Divergence Frontiers

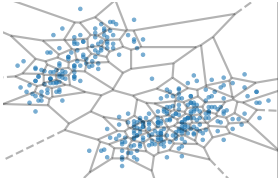
Continuous Distribution



Quantized Distribution



Empirical Estimator



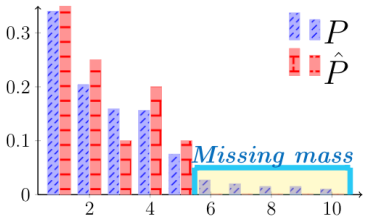
1. How to select the quantization level k?

2. Can we do better than the naïve empirical estimator?

3. How many data points are needed to achieve a good accuracy?

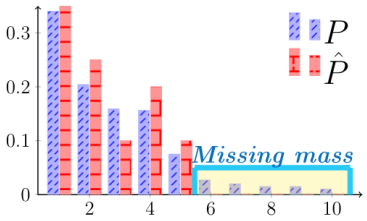
Main Results

- Finite-sample bounds.
  - ▷ **Quantization level** $k \propto O(n^{1/3})$.
  - ▷ **Missing-mass adaptive smoothing** improves the estimation accuracy (e.g., add-constant and Good-Turing).
  - ▷ **Sample complexity** $O(n^{-1/2} \log n)$.
-
-



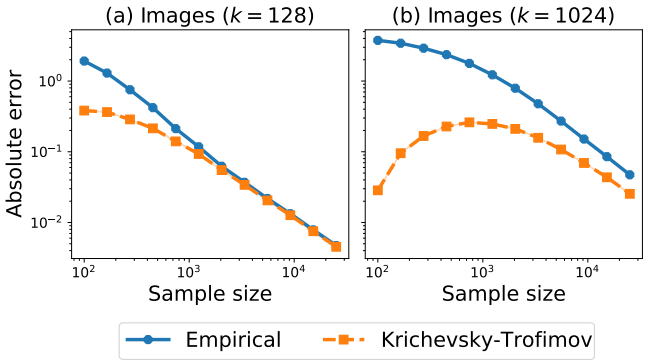
Main Results

- Finite-sample bounds.
  - ▷ **Quantization level** $k \propto O(n^{1/3})$.
  - ▷ **Missing-mass adaptive smoothing** improves the estimation accuracy (e.g., add-constant and Good-Turing).
  - ▷ **Sample complexity** $O(n^{-1/2} \log n)$.
- Statistical summary—**frontier integral**.
- Generalization to **f-divergences**.



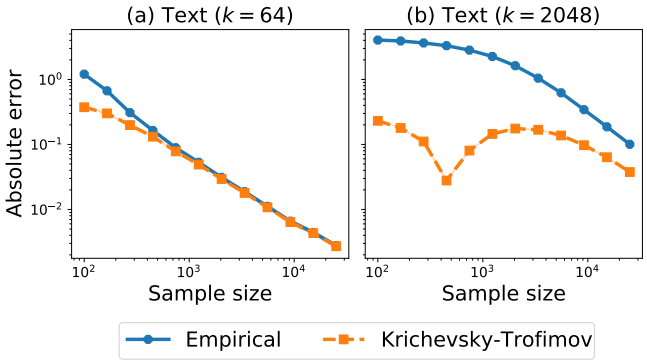
Experimental Results

Missing-mass adaptive smoothing improves the estimation accuracy.



Experimental Results

Missing-mass adaptive smoothing improves the estimation accuracy.



Thank You

Paper: arxiv.org/abs/2106.07898

Thank you!