On the query complexity of sampling from non-log-concave distributions

Yuchen He, Chihao Zhang
Shanghai Jiao Tong University

2025.7

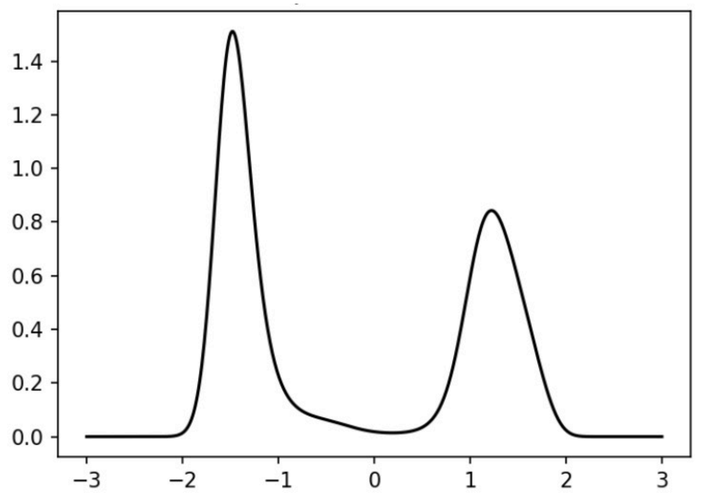
Introduction to sampling

Introduction to sampling

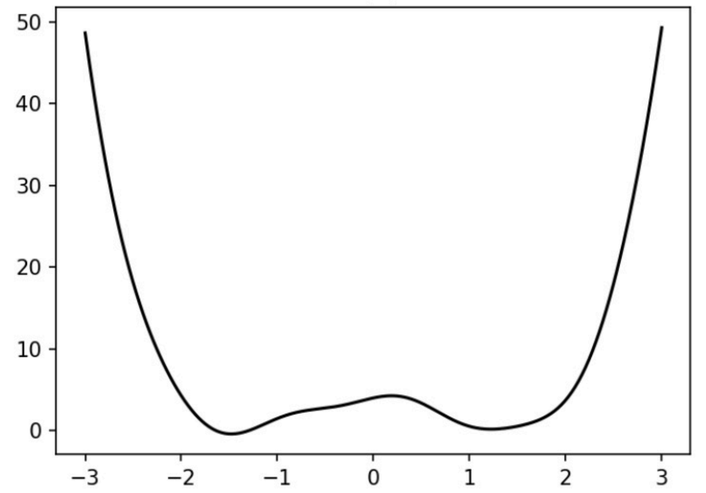
μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Introduction to sampling

μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$



$\mu(x) \propto e^{-f(x)}$

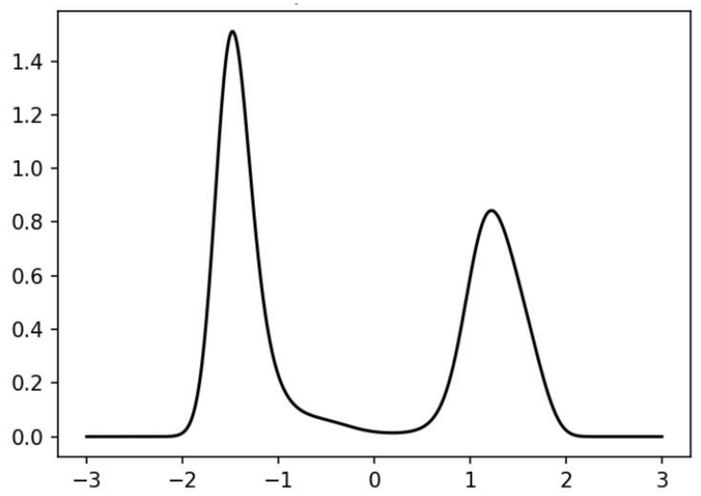


$f(x)$

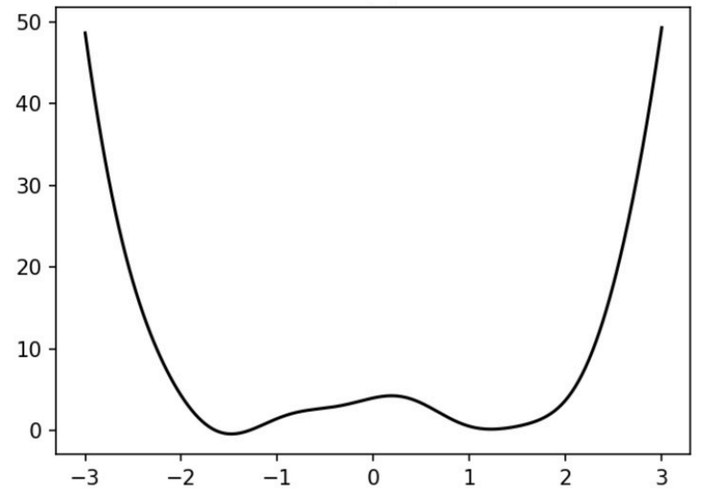
Introduction to sampling

μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Target : given query access to f and ∇f, generate a sample from μ with ϵ error in total variation distance within as few queries as possible.



$\mu(x) \propto e^{-f(x)}$



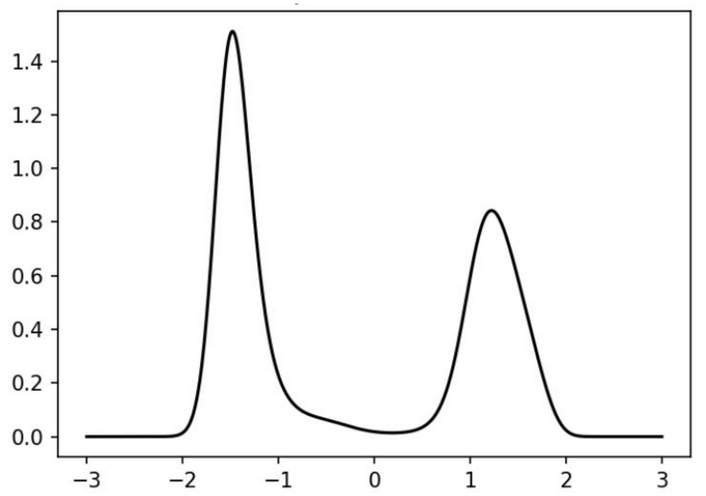
$f(x)$

Introduction to sampling

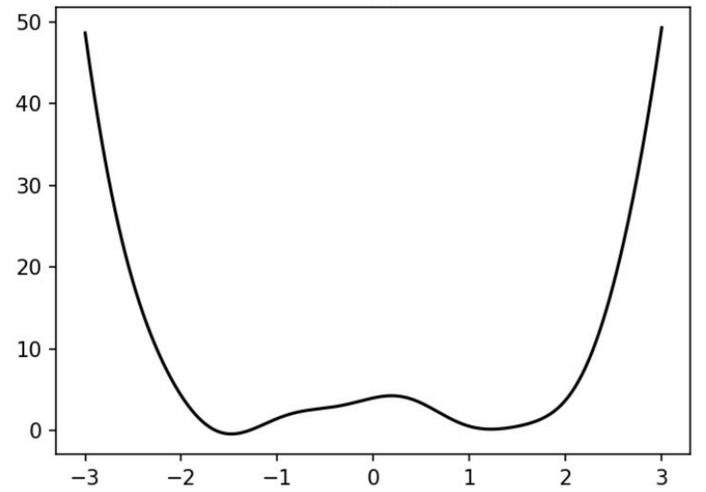
μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Target : given query access to f and ∇f, generate a sample from μ with ϵ error in total variation distance within as few queries as possible.

Assumptions :



$\mu(x) \propto e^{-f(x)}$



$f(x)$

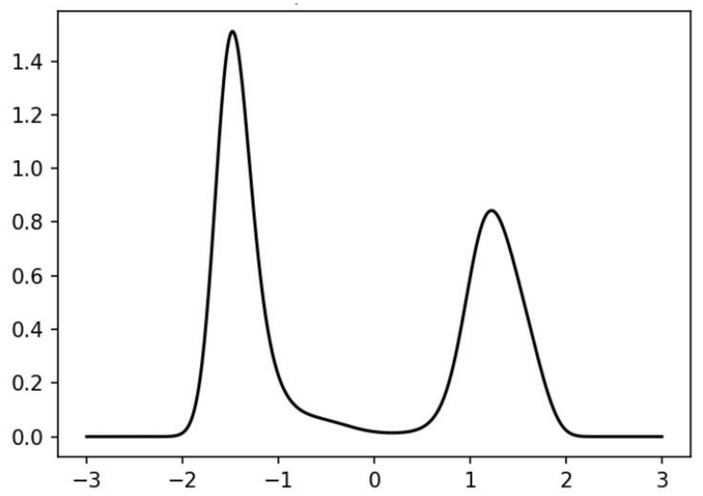
Introduction to sampling

μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

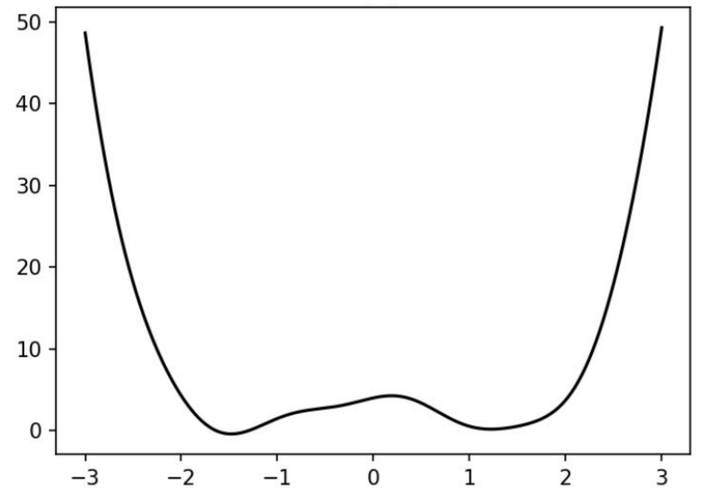
Target : given query access to f and ∇f, generate a sample from μ with ϵ error in total variation distance within as few queries as possible.

Assumptions :

(1) L-log-smooth: $\forall x, y \in \mathbb{R}^d, \|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$



$\mu(x) \propto e^{-f(x)}$



$f(x)$

Introduction to sampling

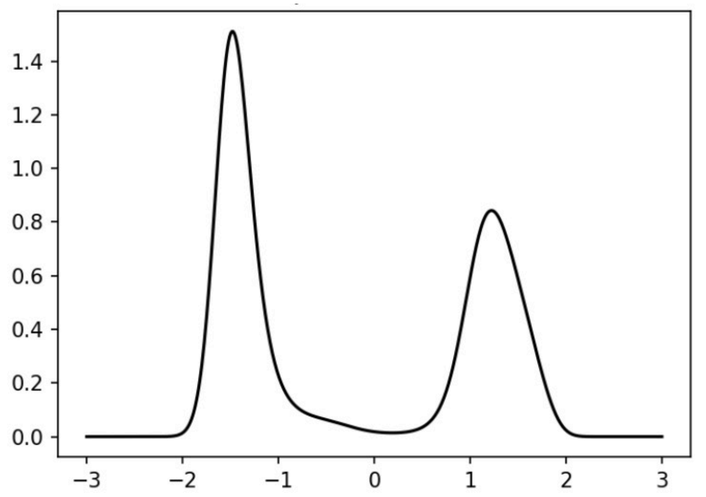
μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Target : given query access to f and ∇f, generate a sample from μ with ϵ error in total variation distance within as few queries as possible.

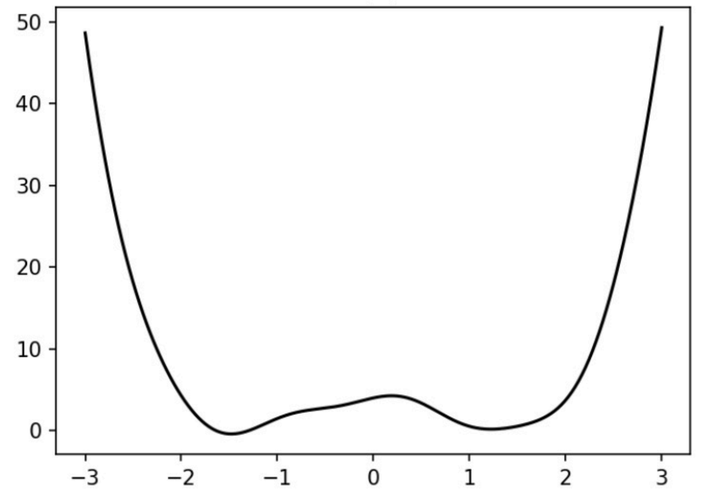
Assumptions :

(1) L-log-smooth: $\forall x, y \in \mathbb{R}^d, \|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$

(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \leq M$



$\mu(x) \propto e^{-f(x)}$



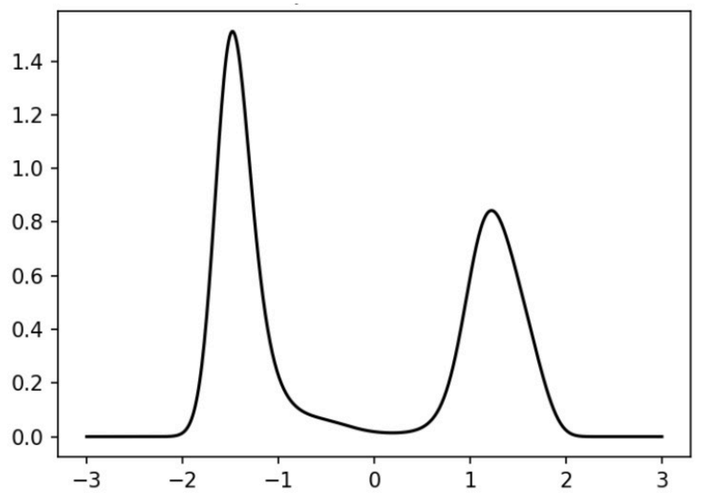
$f(x)$

Introduce to non-log-concave sampling

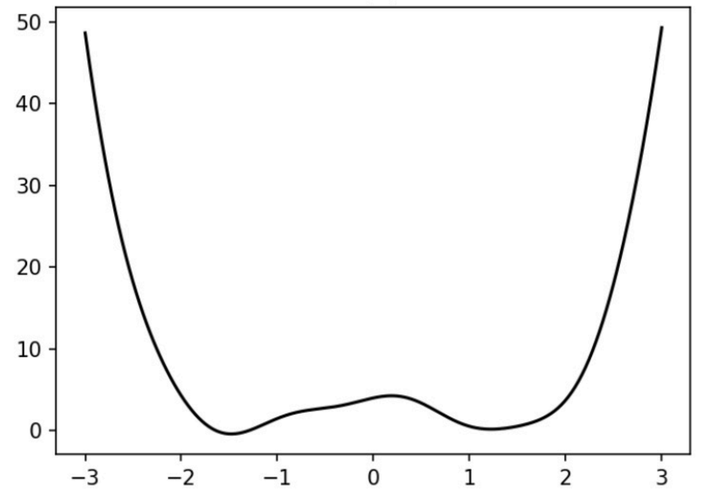
μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Assumptions: (1) L-log-smooth: ∇f is L-Lipschitz

(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \leq M$



$\mu(x) \propto e^{-f(x)}$



$f(x)$

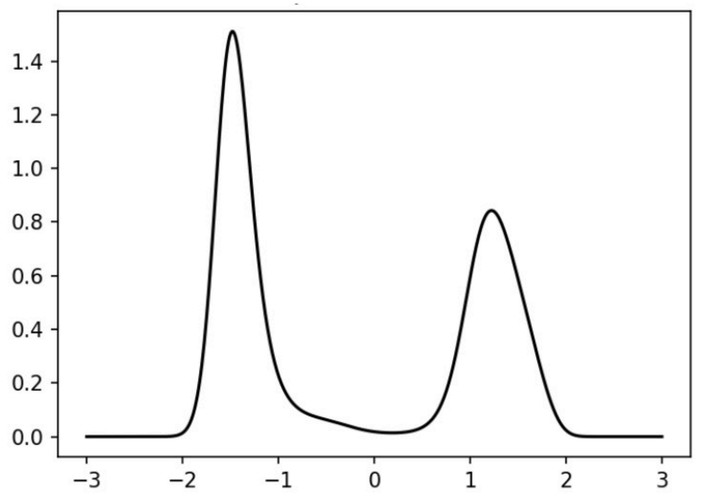
Introduce to non-log-concave sampling

μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

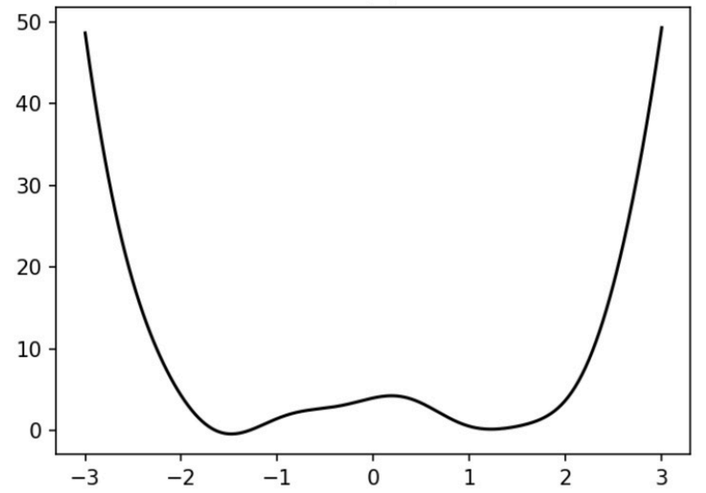
Assumptions: (1) L-log-smooth: ∇f is L-Lipschitz

(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \leq M$

Previous results:



$\mu(x) \propto e^{-f(x)}$



$f(x)$

Introduce to non-log-concave sampling

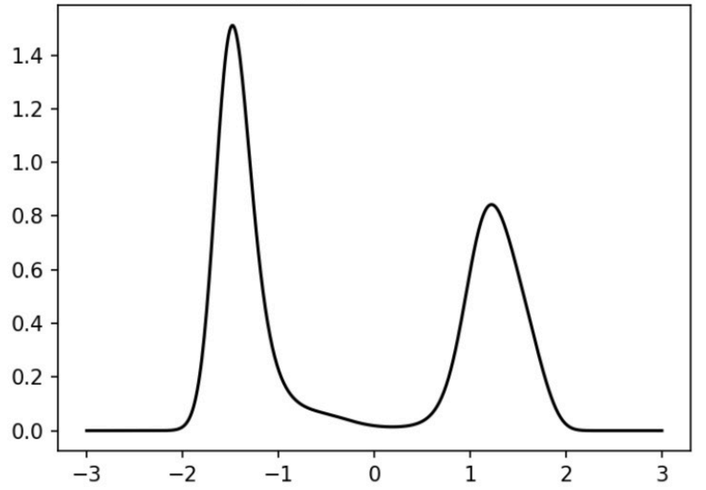
μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Assumptions: (1) L-log-smooth: ∇f is L-Lipschitz

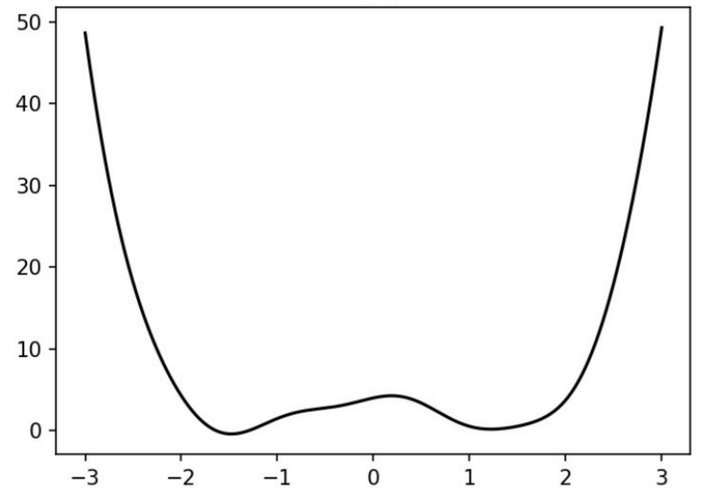
(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \leq M$

Previous results:

- $e^{\Omega(d)}$ queries is inevitable in general (Ge, Lee and Risteski (2018))



$\mu(x) \propto e^{-f(x)}$



$f(x)$

Introduce to non-log-concave sampling

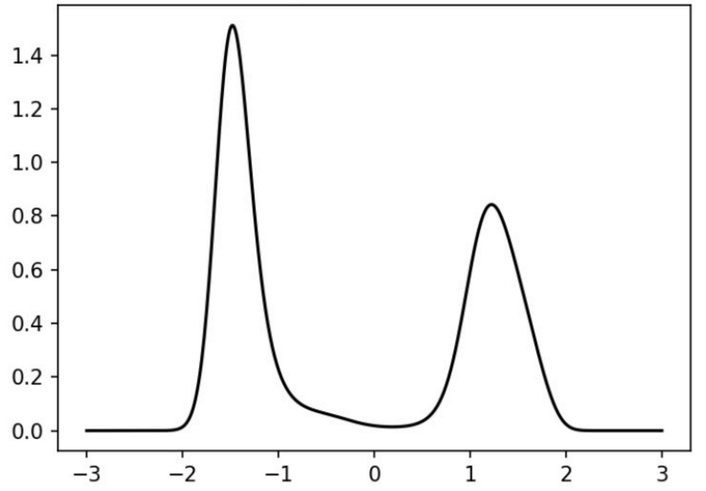
μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Assumptions: (1) L-log-smooth: ∇f is L-Lipschitz

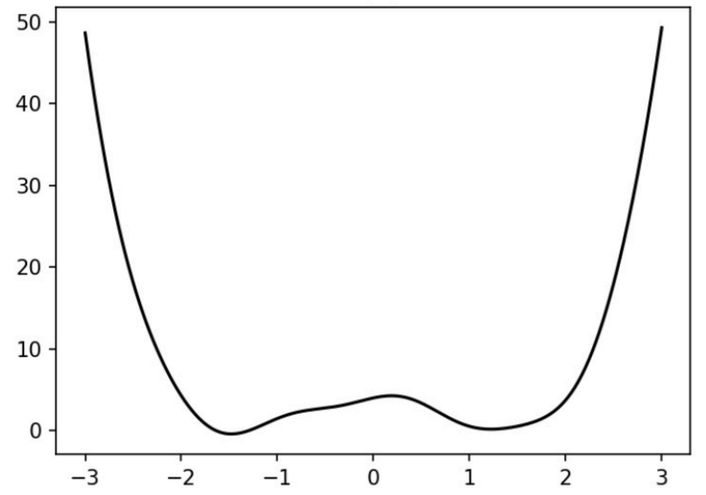
(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \le M$

Previous results:

- $e^{\Omega(d)}$ queries is inevitable in general (Ge, Lee and Risteski (2018))
- $\exp\left\{\mathcal{O}\left(d\log\frac{L(d+M)}{\epsilon}\right) + \max_k \|x_k\|^2\right\}$ (He, Rojas and Tao (2024))



$\mu(x) \propto e^{-f(x)}$



$f(x)$

Introduce to non-log-concave sampling

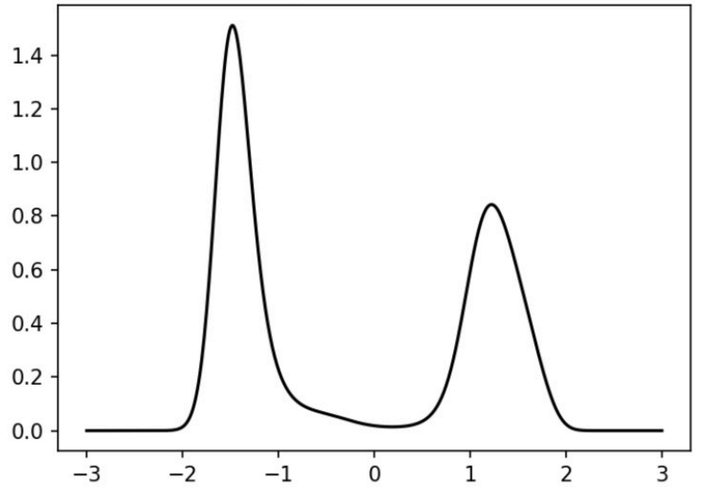
μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Assumptions: (1) L-log-smooth: ∇f is L-Lipschitz

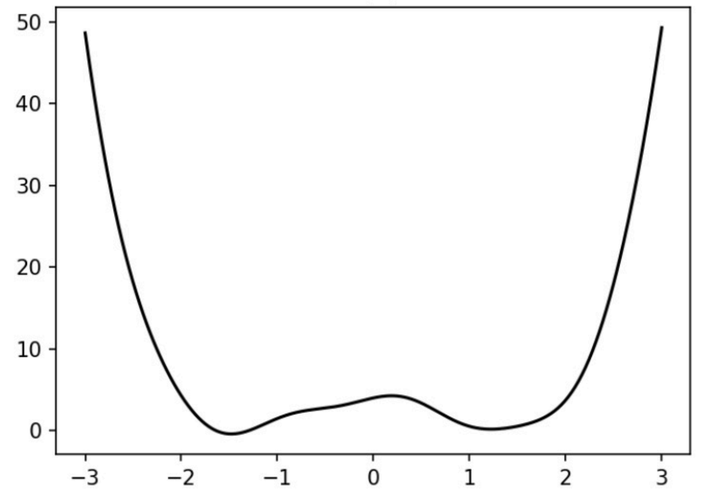
(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \leq M$

Previous results:

- $e^{\Omega(d)}$ queries is inevitable in general (Ge, Lee and Risteski (2018))
- $\exp\left\{\mathcal{O}\left(d\log\frac{L(d+M)}{\epsilon}\right)+\max_k\|x_k\|^2\right\}$ (He, Rojas and Tao (2024))
- $\exp\left\{\mathcal{O}\left(L^3\cdot\log^3\left(\frac{Ld+M}{\epsilon}\right)\right)\right\}$ (Huang, Zou, Dong, Ma and Zhang (2024))



$\mu(x) \propto e^{-f(x)}$



$f(x)$

Introduce to non-log-concave sampling

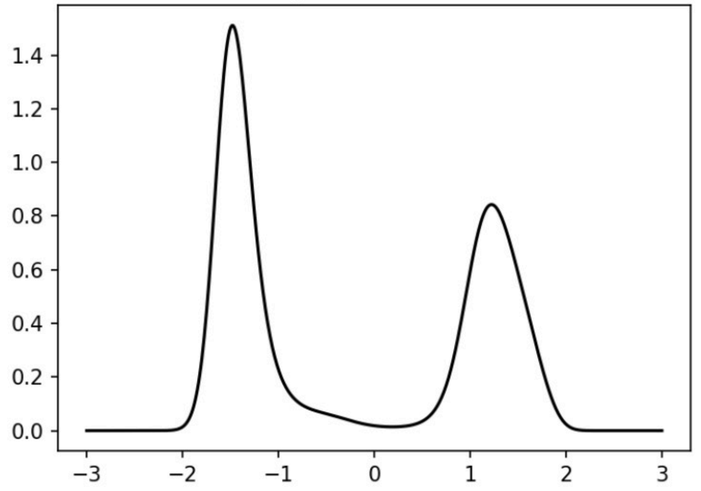
μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Assumptions: (1) L-log-smooth: ∇f is L-Lipschitz

(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \le M$

Previous results:

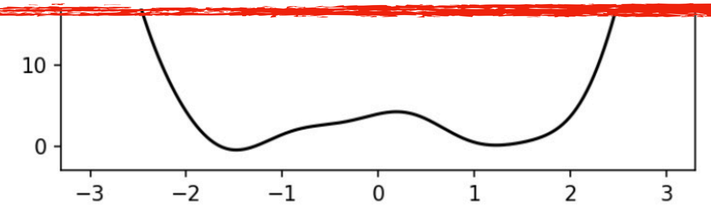
- $e^{\Omega(d)}$ queries is inevitable in general (Ge, Lee and Risteski (2018))
- $\exp\left\{\mathcal{O}\left(d\log\frac{L(d+M)}{\epsilon}\right)+\max_k\|x_k\|^2\right\}$ (He, Rojas and Tao (2024))
- $\exp\left\{\mathcal{O}\left(L^3\cdot\log^3\left(\frac{Ld+M}{\epsilon}\right)\right)\right\}$ (Huang, Zou, Dong, Ma and Zhang (2024))



$\mu(x) \propto e^{-f(x)}$

With an extra requirement: log-smooth during the Ornstein-Uhlenbeck process (OU process)

$$\mathrm{d}X_t = -X_t\mathrm{d}t + \mathrm{d}B_t\ ,\ X_0 \sim \mu$$



$f(x)$

Introduce to non-log-concave sampling

μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

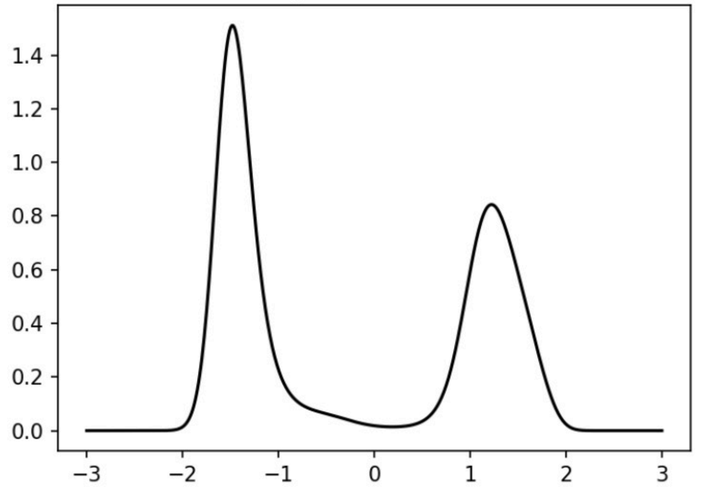
Assumptions: (1) L-log-smooth: ∇f is L-Lipschitz

(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \le M$

Previous results:

- $e^{\Omega(d)}$ queries is inevitable in general (Ge, Lee and Risteski (2018))
- $\exp\left\{\mathcal{O}\left(d\log\frac{L(d+M)}{\epsilon}\right)+\max_k\|x_k\|^2\right\}$ (He, Rojas and Tao (2024))
- $\exp\left\{\mathcal{O}\left(L^3\cdot\log^3\left(\frac{Ld+M}{\epsilon}\right)\right)\right\}$ (Huang, Zou, Dong, Ma and Zhang (2024))

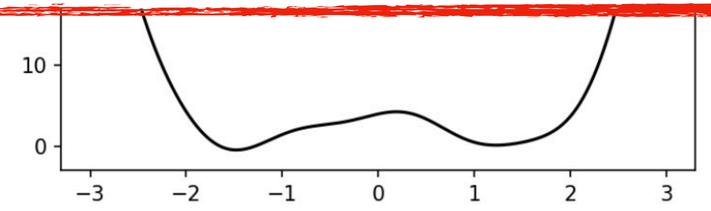
Motivation: understanding the query complexity under Assumptions (1) and (2).



$\mu(x) \propto e^{-f(x)}$

With an extra requirement: log-smooth during the Ornstein-Uhlenbeck process (OU process)

$$\mathrm{d}X_t = -X_t\mathrm{d}t + \mathrm{d}B_t\ ,\ X_0 \sim \mu$$



$f(x)$

Our results

μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Assumptions: (1) L-log-smooth: ∇f is L-Lipschitz

(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \leq M$

Our results

μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Assumptions: (1) L-log-smooth: ∇f is L-Lipschitz

(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \leq M$

- An almost tight query complexity bound $\left(\frac{LM}{d\epsilon}\right)^{\Theta(d)}$ to guarantee ϵ error of total variation distance

Our results

μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Assumptions: (1) L-log-smooth: ∇f is L-Lipschitz

(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \le M$

- An almost tight query complexity bound $\left(\frac{LM}{d\epsilon}\right)^{\Theta(d)}$ to guarantee ϵ error of total variation distance

  → This implies $\mathrm{poly}\left(\log 1/\epsilon\right)$ samplers do not exist.

Our results

μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Assumptions: (1) L-log-smooth: ∇f is L-Lipschitz

(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \le M$

- An almost tight query complexity bound $\left(\frac{LM}{d\epsilon}\right)^{\Theta(d)}$ to guarantee ϵ error of total variation distance
  - → This implies $\mathrm{poly}\left(\log 1/\epsilon\right)$ samplers do not exist.
  - → Functional inequalities like Poincaré inequality do not hold.

Our results

μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Assumptions: (1) L-log-smooth: ∇f is L-Lipschitz

(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \le M$

- An almost tight query complexity bound $\left(\frac{LM}{d\epsilon}\right)^{\Theta(d)}$ to guarantee ϵ error of total variation distance
  - This implies $\mathrm{poly}\left(\log 1/\epsilon\right)$ samplers do not exist.
  - Functional inequalities like Poincaré inequality do not hold.
- The smoothness during OU process is strictly stronger than the smoothness of μ

Our results

μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Assumptions: (1) L-log-smooth: ∇f is L-Lipschitz

(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \le M$

- An almost tight query complexity bound $\left(\frac{LM}{d\epsilon}\right)^{\Theta(d)}$ to guarantee ϵ error of total variation distance
  - This implies $\mathrm{poly}\left(\log 1/\epsilon\right)$ samplers do not exist.
  - Functional inequalities like Poincaré inequality do not hold.
- The smoothness during OU process is strictly stronger than the smoothness of μ

  We construct a group of distribution such that they are $O(1)$-log-smooth but $\Omega(d)$-log-smooth for $t = \Theta(1)$ during the OU process.

Our results

μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Assumptions: (1) L-log-smooth: ∇f is L-Lipschitz

(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \le M$

- An almost tight query complexity bound $\left(\frac{LM}{d\epsilon}\right)^{\Theta(d)}$ to guarantee ϵ error of total variation distance
  - This implies $\mathrm{poly}\left(\log 1/\epsilon\right)$ samplers do not exist.
  - Functional inequalities like Poincaré inequality do not hold.
- The smoothness during OU process is strictly stronger than the smoothness of μ

  We construct a group of distribution such that they are $O(1)$-log-smooth but $\Omega(d)$-log-smooth for $t = \Theta(1)$ during the OU process.
- The lower bound to approximate the global minimizer of f with constant error is $(c \cdot LM)^{\frac{d}{2}}$

Our results

μ —— a distribution over $\Omega = \mathbb{R}^d$, with density $\mu(x) \propto e^{-f(x)}$

Assumptions: (1) L-log-smooth: ∇f is L-Lipschitz

(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \le M$

- An almost tight query complexity bound $\left(\frac{LM}{d\epsilon}\right)^{\Theta(d)}$ to guarantee ϵ error of total variation distance
  - → This implies $\mathrm{poly}\left(\log 1/\epsilon\right)$ samplers do not exist.
  - → Functional inequalities like Poincaré inequality do not hold.
- The smoothness during OU process is strictly stronger than the smoothness of μ

  We construct a group of distribution such that they are $O(1)$-log-smooth but $\Omega(d)$-log-smooth for $t = \Theta(1)$ during the OU process.
- The lower bound to approximate the global minimizer of f with constant error is $(c \cdot LM)^{\frac{d}{2}}$
  - → Sampling is simpler than optimization.

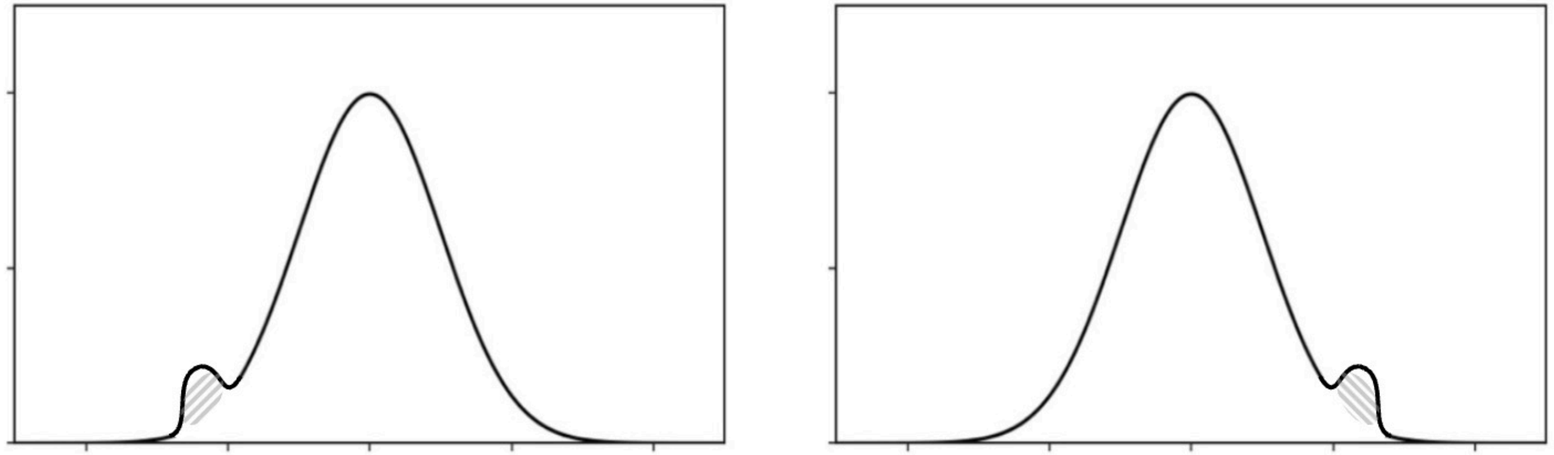
Lower bound: construct hard instances

Lower bound: construct hard instances

The hardness lies in distinguishing a group of similar distributions.

Lower bound: construct hard instances

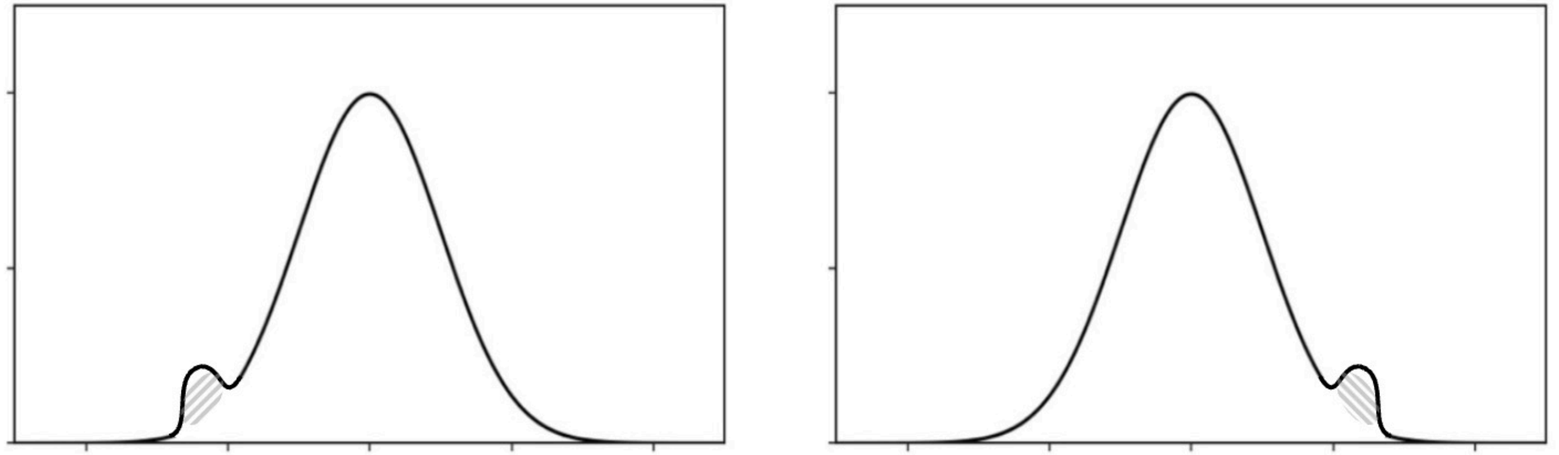
The hardness lies in distinguishing a group of similar distributions.



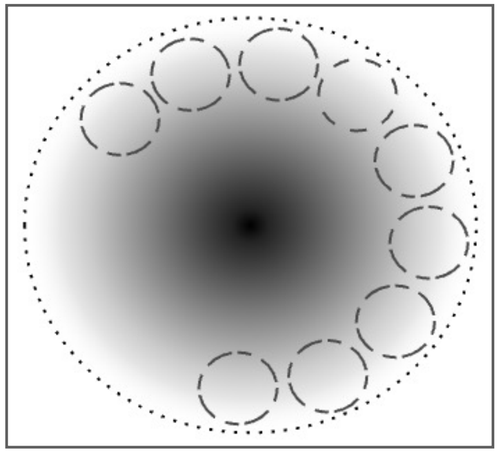
Disturb the Gaussian with a “needle” of ϵ mass

Lower bound: construct hard instances

The hardness lies in distinguishing a group of similar distributions.



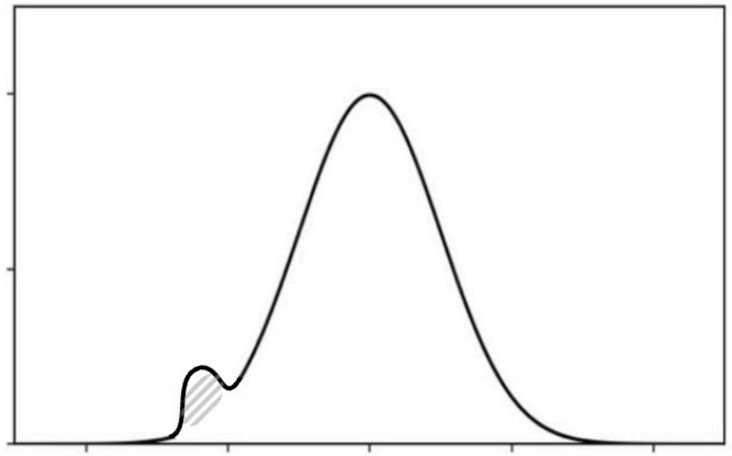
Disturb the Gaussian with a “needle” of ϵ mass

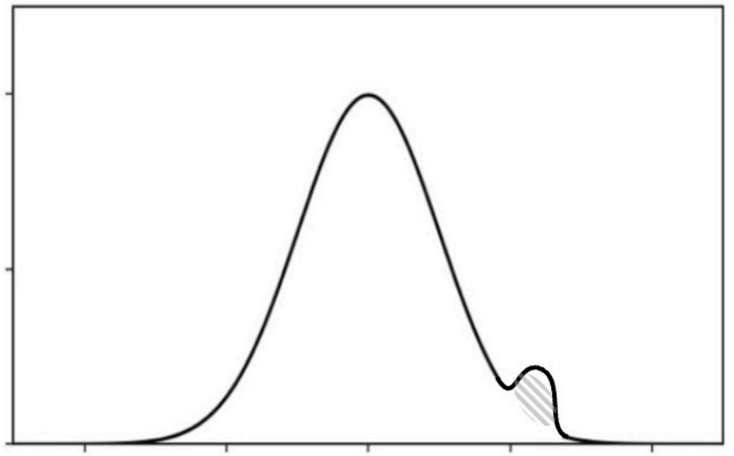


2-dim Gaussian

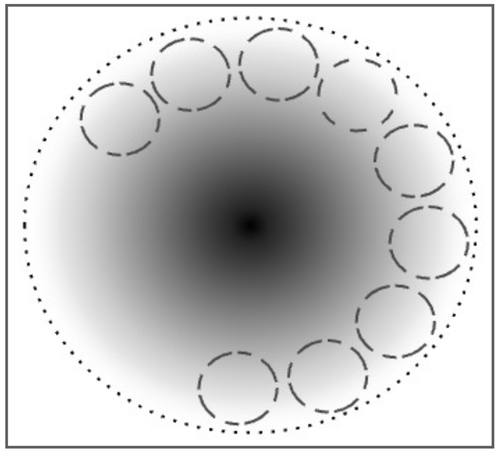
Lower bound: construct hard instances

The hardness lies in distinguishing a group of similar distributions.





Disturb the Gaussian with a “needle” of ϵ mass

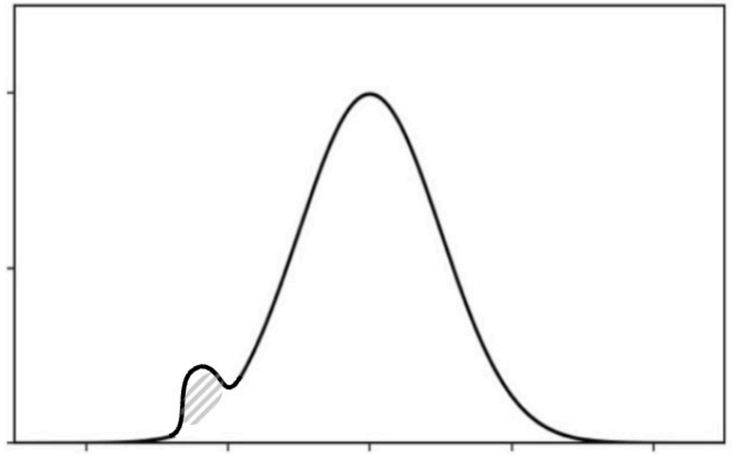


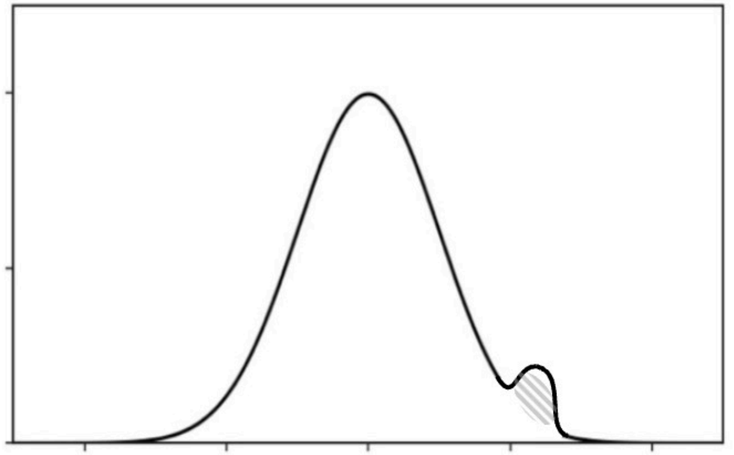
2-dim Gaussian

Reduce the lower bound of sampling to a packing problem under two constraints:

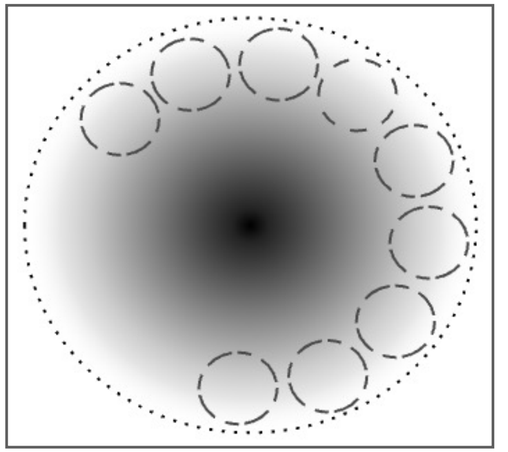
Lower bound: construct hard instances

The hardness lies in distinguishing a group of similar distributions.





Disturb the Gaussian with a "needle" of ϵ mass



2-dim Gaussian

Reduce the lower bound of sampling to a packing problem under two constraints:

(1) L-log-smooth: ∇f is L-Lipschitz

(2) Bounded second moment: $\mathbf{E}_{X\sim\mu}\left[\|X\|^2\right] \leq M$

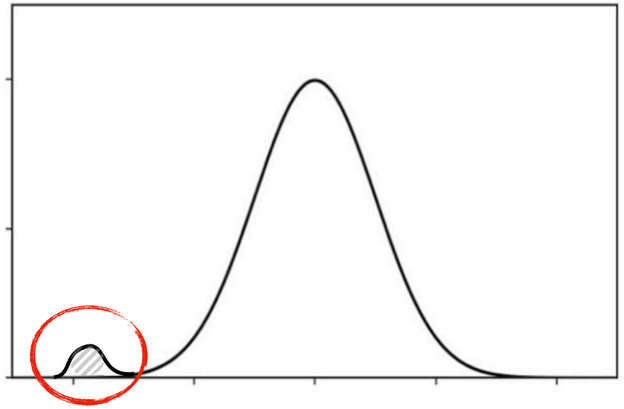
Lower bound: construct hard instances

Lower bound: construct hard instances

To pack more candidate circles to perturb:

Lower bound: construct hard instances

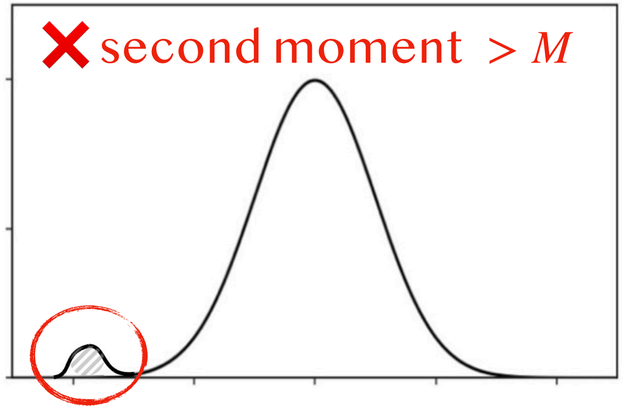
To pack more candidate circles to perturb:



Put the “needle” farther away

Lower bound: construct hard instances

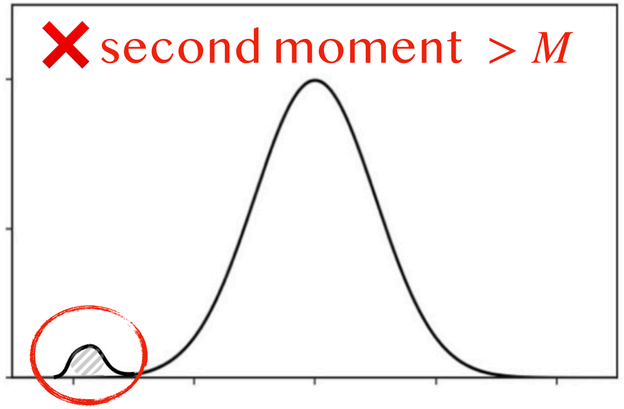
To pack more candidate circles to perturb:



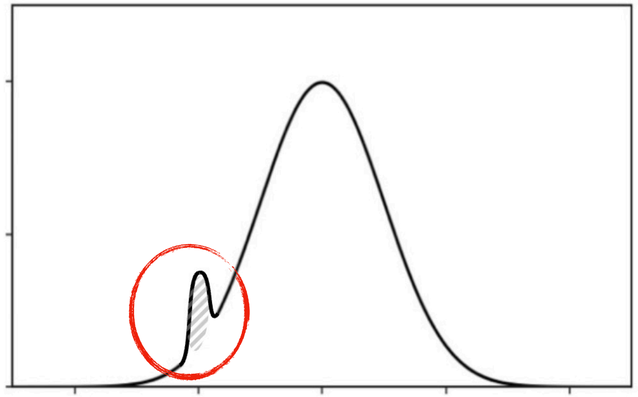
Put the “needle” farther away

Lower bound: construct hard instances

To pack more candidate circles to perturb:



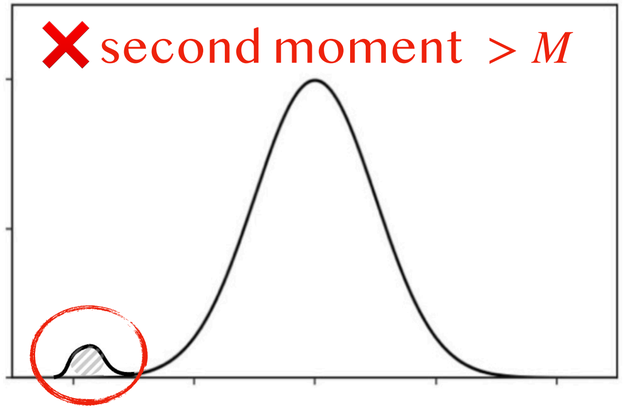
Put the “needle” farther away



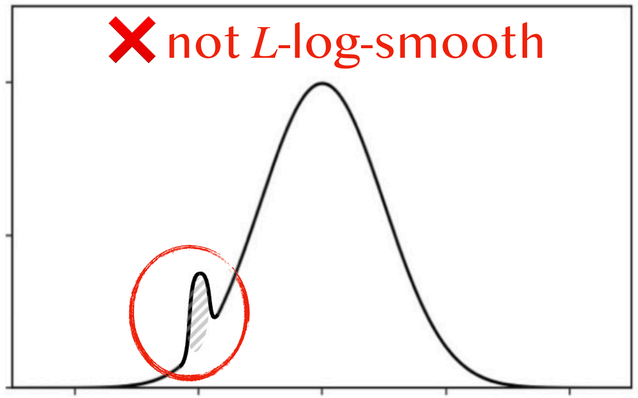
Use a taller, narrower “needle”

Lower bound: construct hard instances

To pack more candidate circles to perturb:



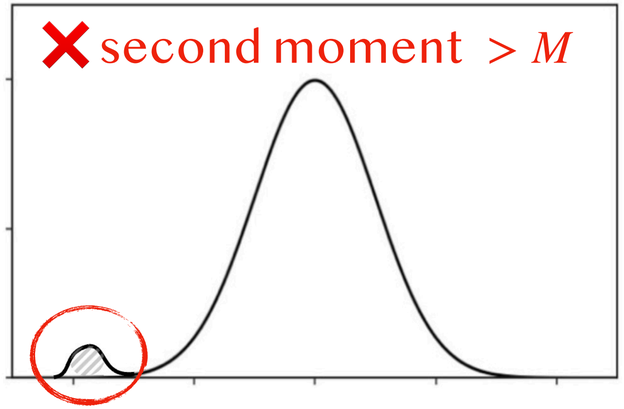
Put the “needle” farther away



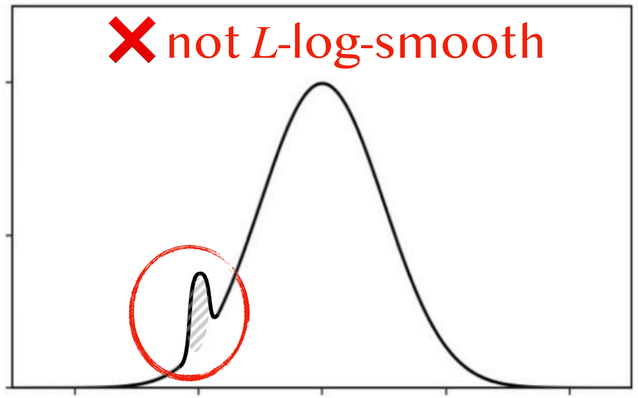
Use a taller, narrower “needle”

Lower bound: construct hard instances

To pack more candidate circles to perturb:



Put the “needle” farther away

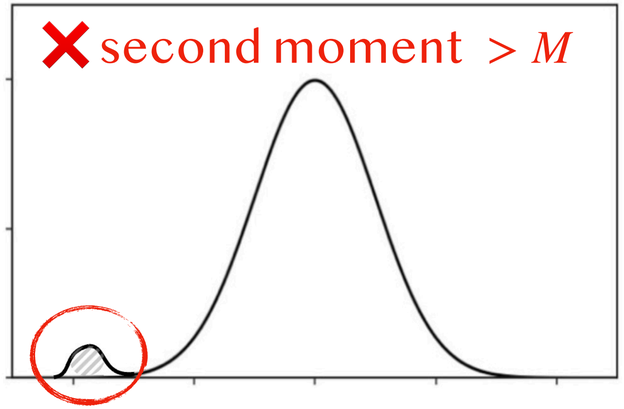


Use a taller, narrower “needle”

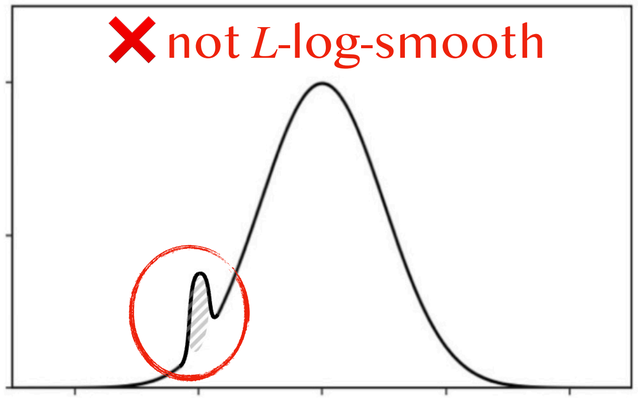
Main contribution: use better base distribution

Lower bound: construct hard instances

To pack more candidate circles to perturb:

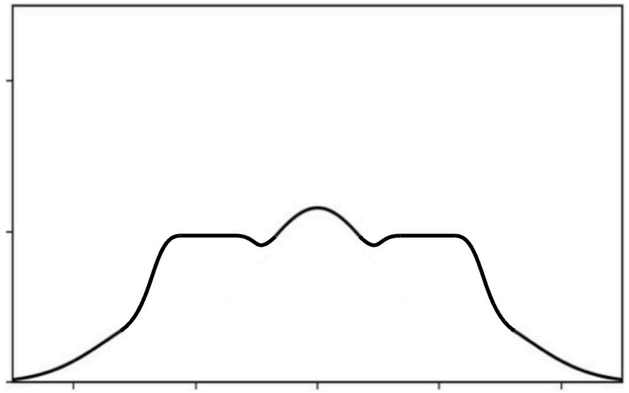


Put the “needle” farther away



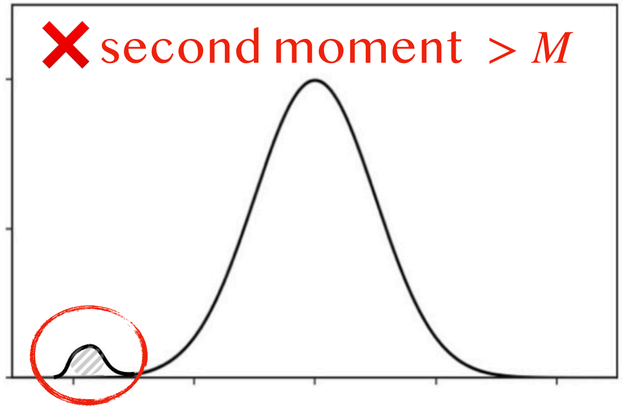
Use a taller, narrower “needle”

Main contribution: use better base distribution

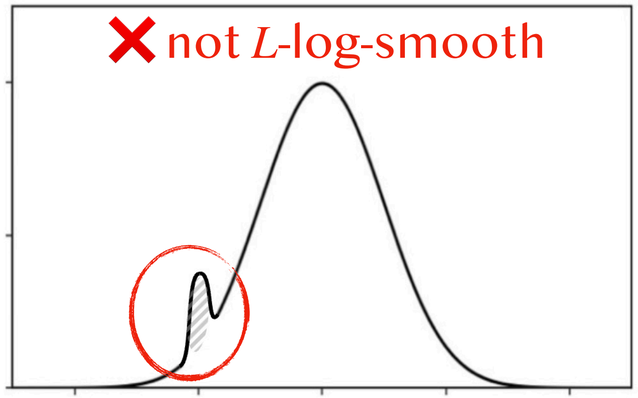


Lower bound: construct hard instances

To pack more candidate circles to perturb:

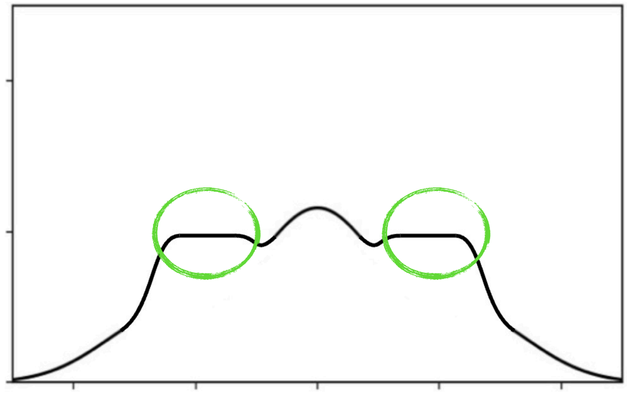


Put the “needle” farther away



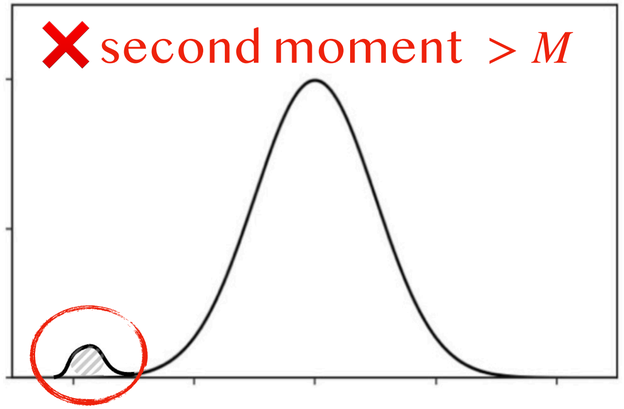
Use a taller, narrower “needle”

Main contribution: use better base distribution

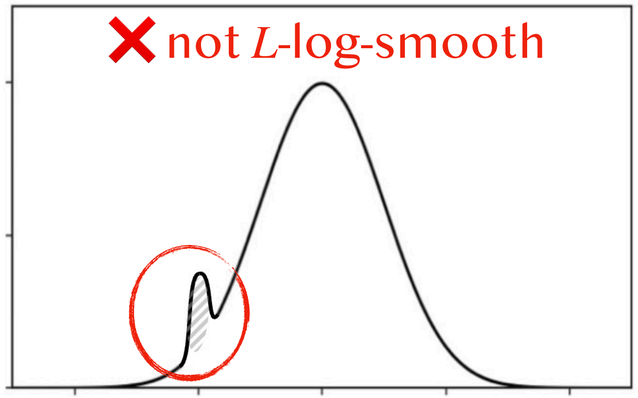


Lower bound: construct hard instances

To pack more candidate circles to perturb:

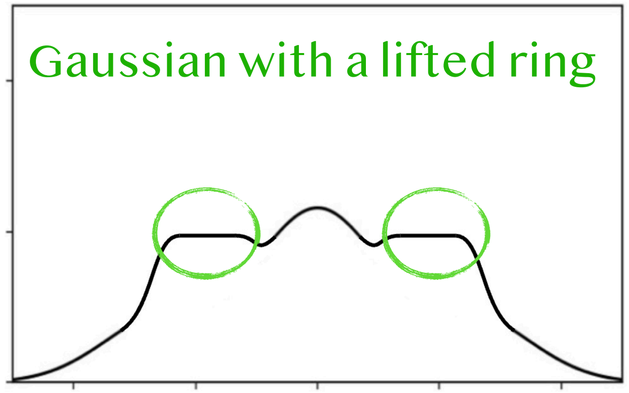


Put the “needle” farther away



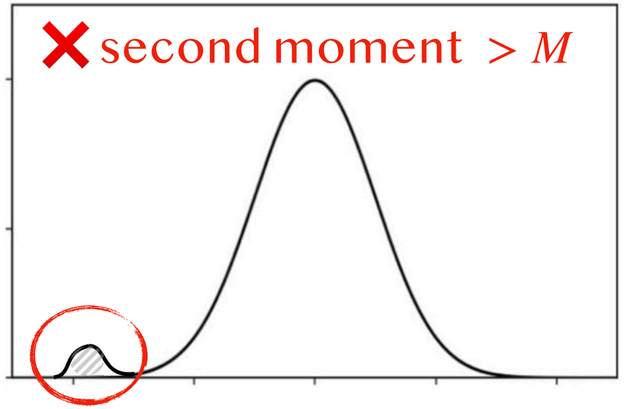
Use a taller, narrower “needle”

Main contribution: use better base distribution

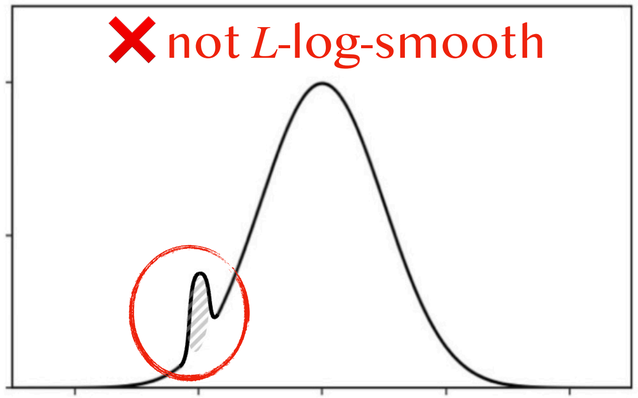


Lower bound: construct hard instances

To pack more candidate circles to perturb:

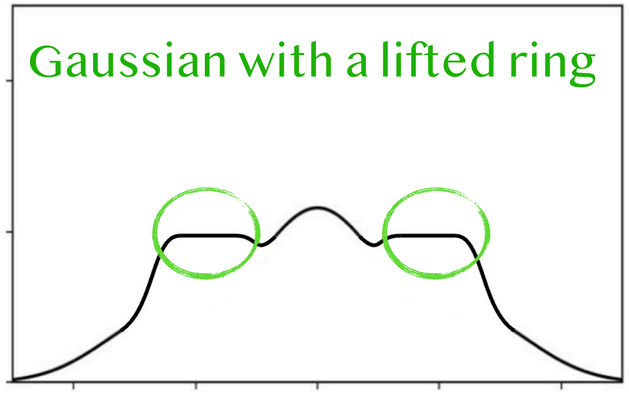


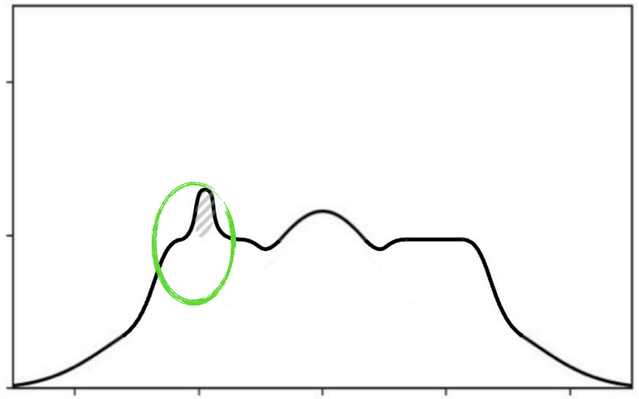
Put the “needle” farther away



Use a taller, narrower “needle”

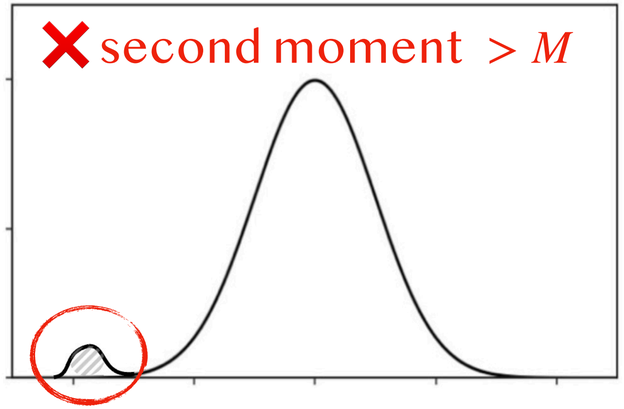
Main contribution: use better base distribution



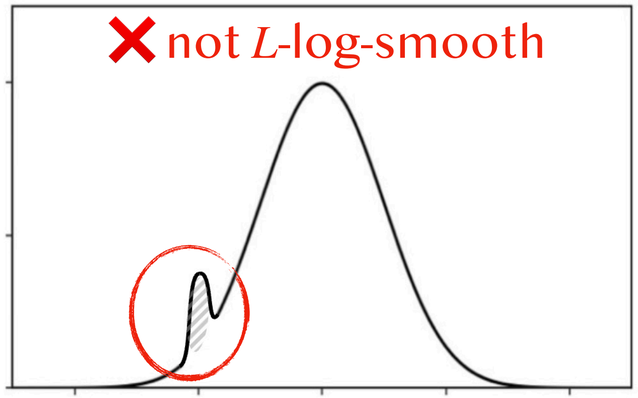


Lower bound: construct hard instances

To pack more candidate circles to perturb:

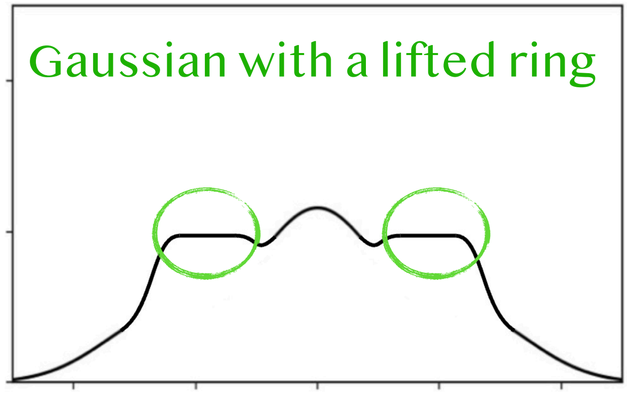


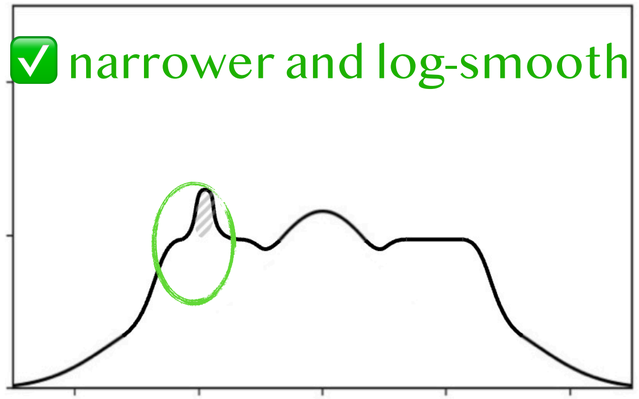
Put the “needle” farther away



Use a taller, narrower “needle”

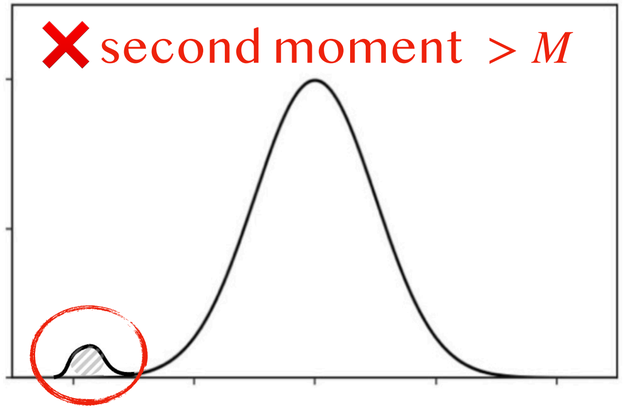
Main contribution: use better base distribution



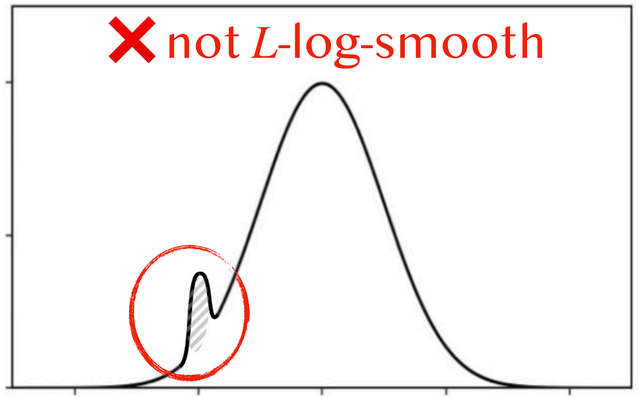


Lower bound: construct hard instances

To pack more candidate circles to perturb:

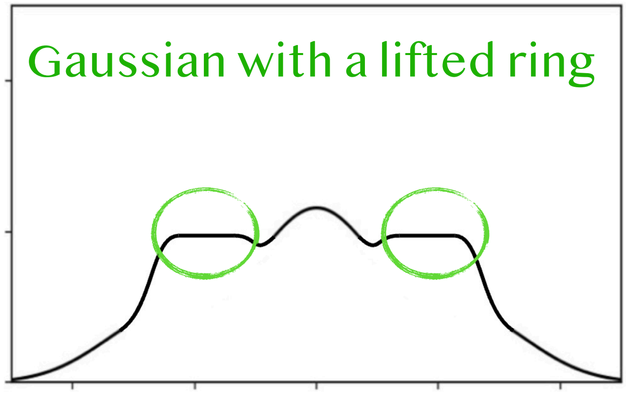


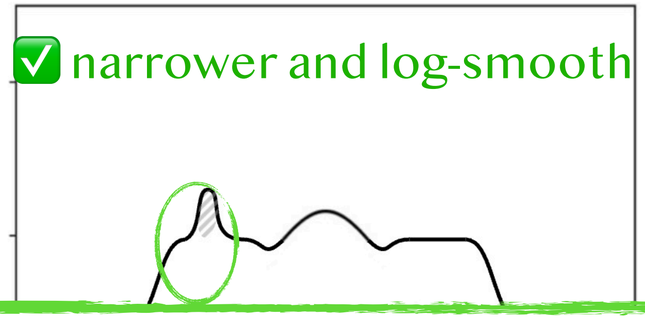
Put the “needle” farther away



Use a taller, narrower “needle”

Main contribution: use better base distribution

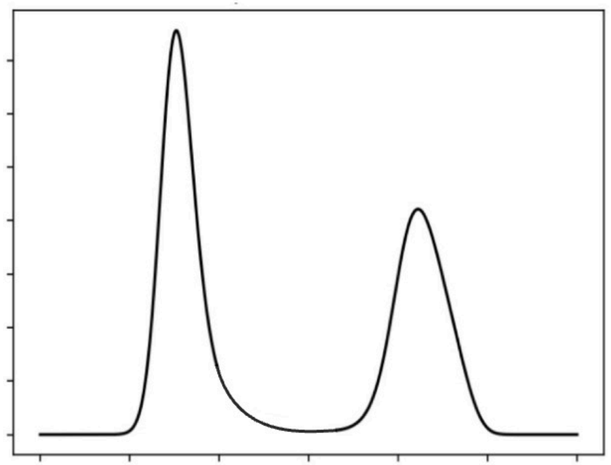




As a result, we can construct $\left(\frac{LM}{d\epsilon}\right)^{\Omega(d)}$ such distributions.

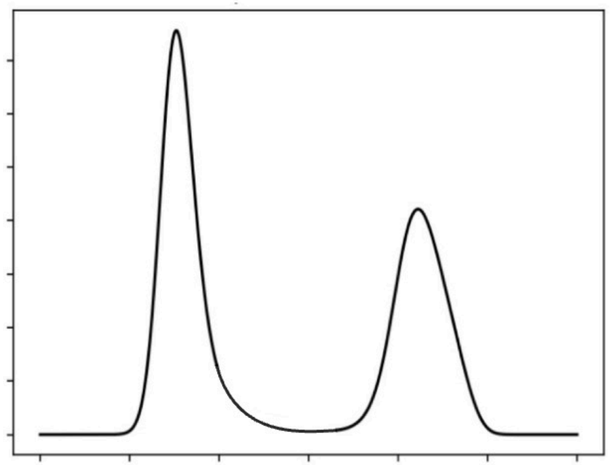
Upper bound: truncation scheme

Upper bound: truncation scheme

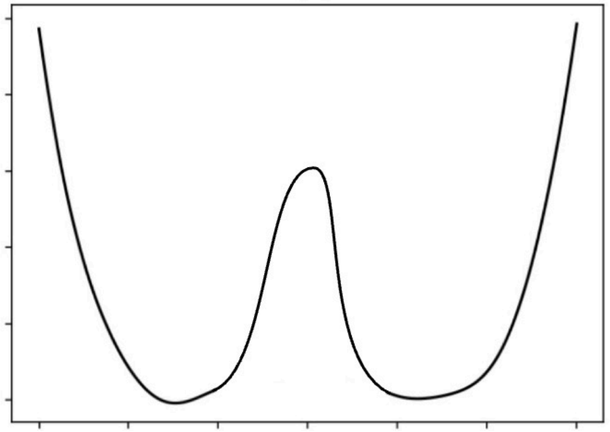


$$\mu(x) \propto e^{-f(x)}$$

Upper bound: truncation scheme



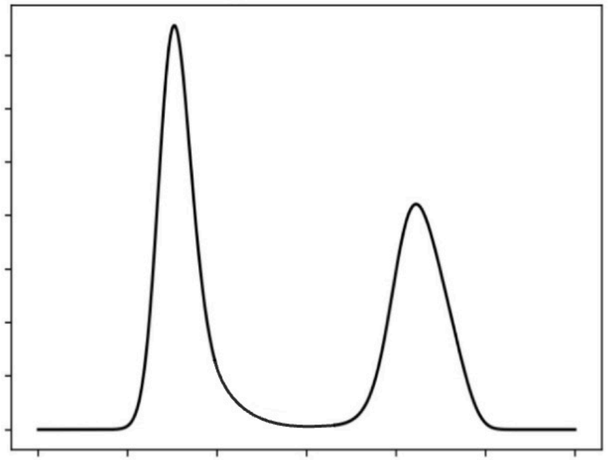
$\mu(x) \propto e^{-f(x)}$



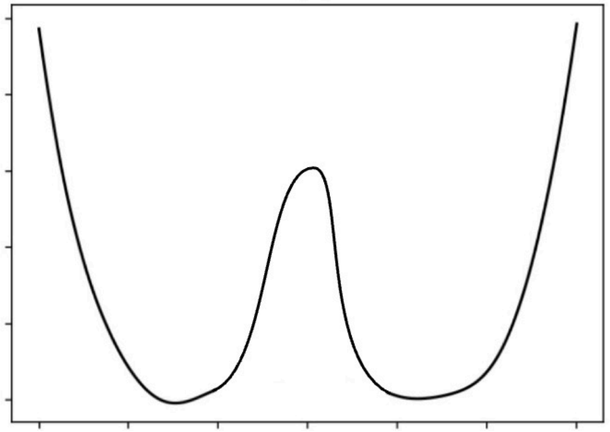
$f(x)$

Upper bound: truncation scheme

Difficulty: The low valley that separate the modes



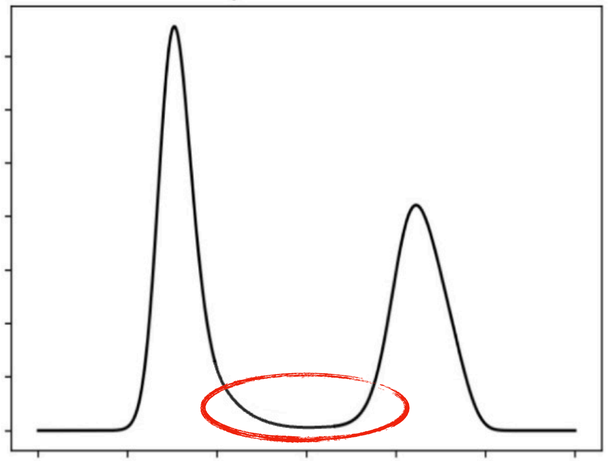
$\mu(x) \propto e^{-f(x)}$



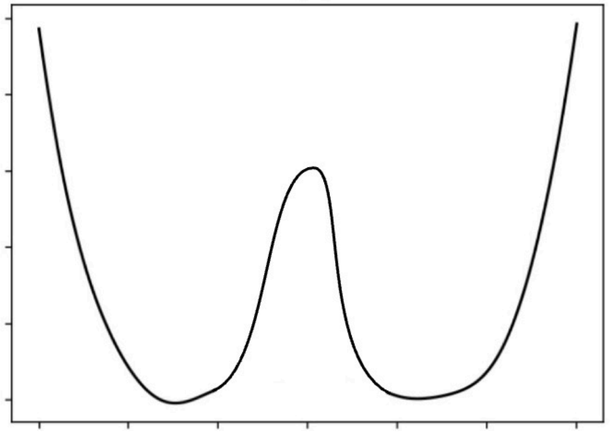
$f(x)$

Upper bound: truncation scheme

Difficulty: The low valley that separate the modes



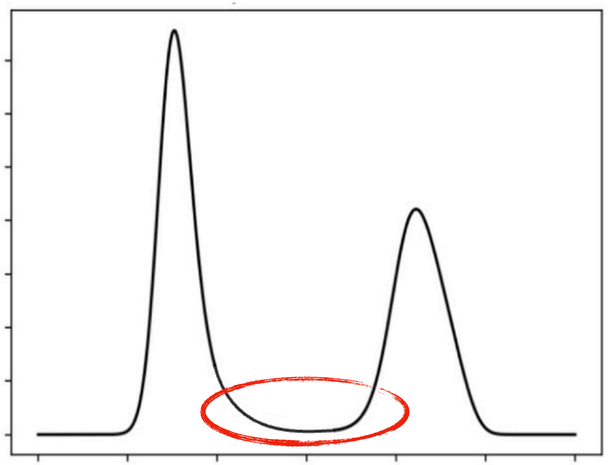
$\mu(x) \propto e^{-f(x)}$



$f(x)$

Upper bound: truncation scheme

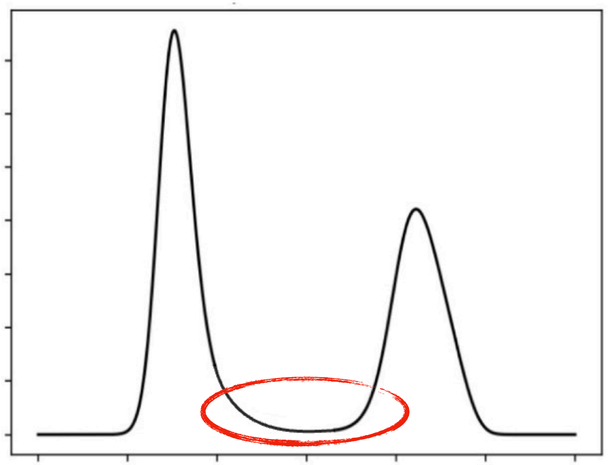
Difficulty: The low valley that separate the modes



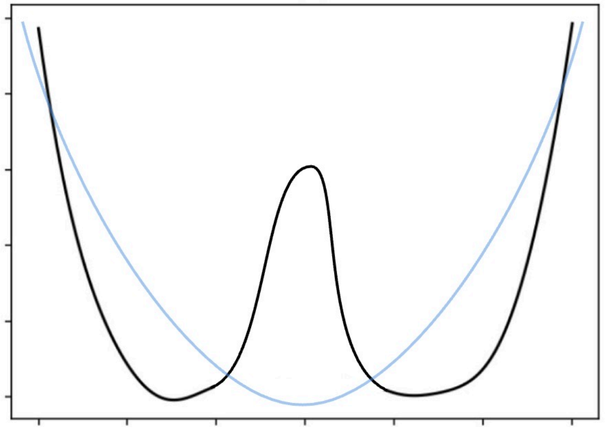
$$\mu(x) \propto e^{-f(x)}$$

Upper bound: truncation scheme

Difficulty: The low valley that separate the modes



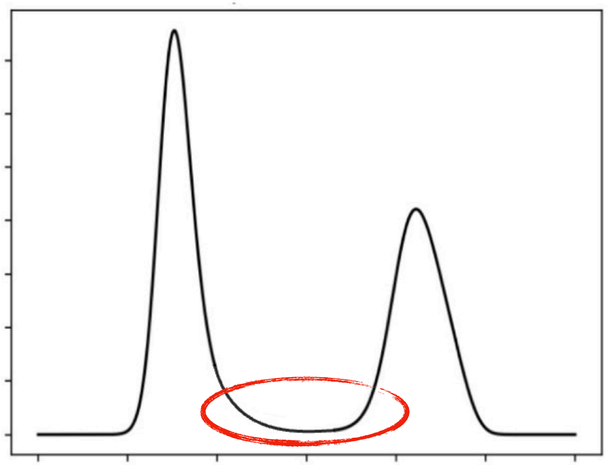
$\mu(x) \propto e^{-f(x)}$



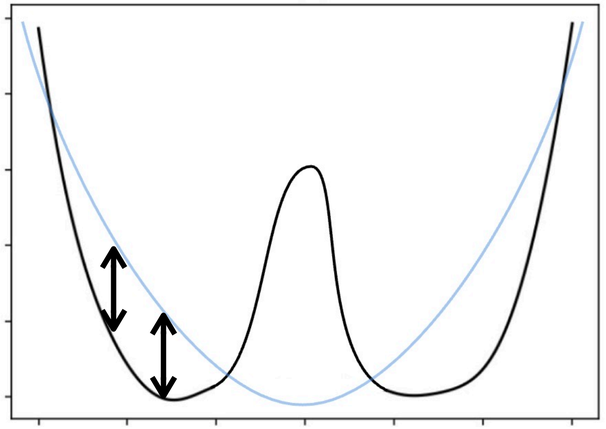
$f(x)$ and $c\|x\|^2$

Upper bound: truncation scheme

Difficulty: The low valley that separate the modes



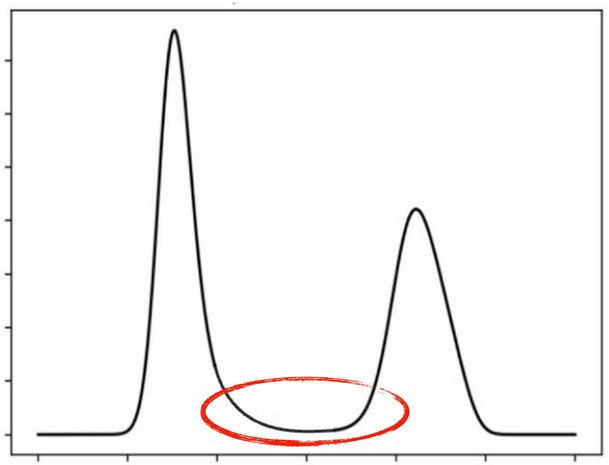
$\mu(x) \propto e^{-f(x)}$



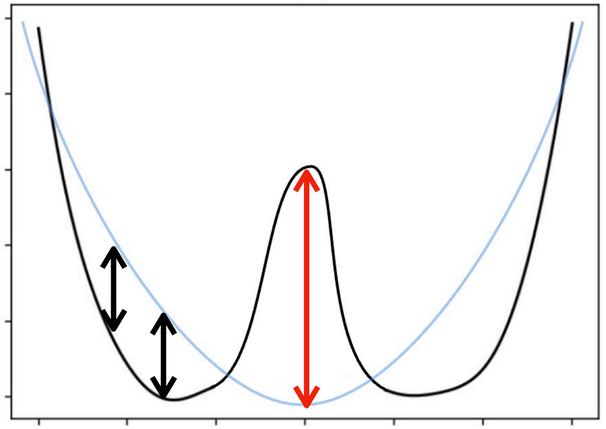
$f(x)$ and $c\|x\|^2$

Upper bound: truncation scheme

Difficulty: The low valley that separate the modes



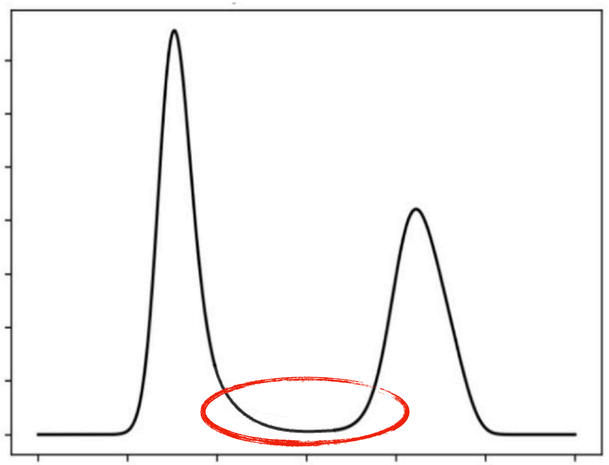
$\mu(x) \propto e^{-f(x)}$



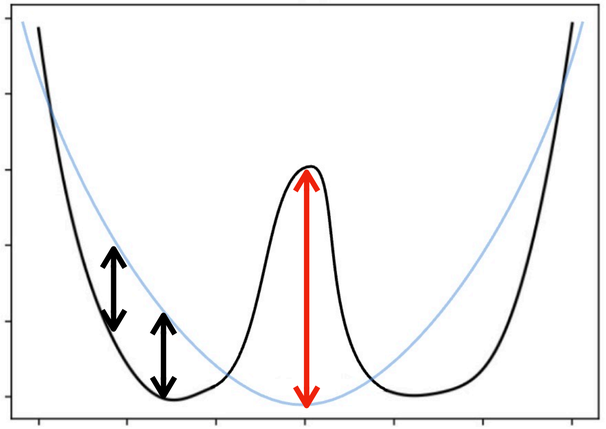
$f(x)$ and $c\|x\|^2$

Upper bound: truncation scheme

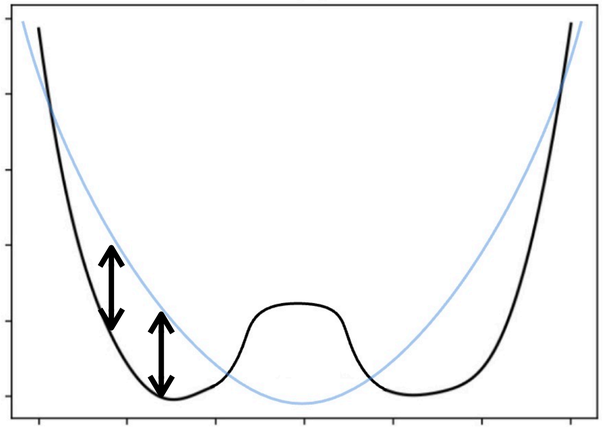
Difficulty: The low valley that separate the modes



$\mu(x) \propto e^{-f(x)}$



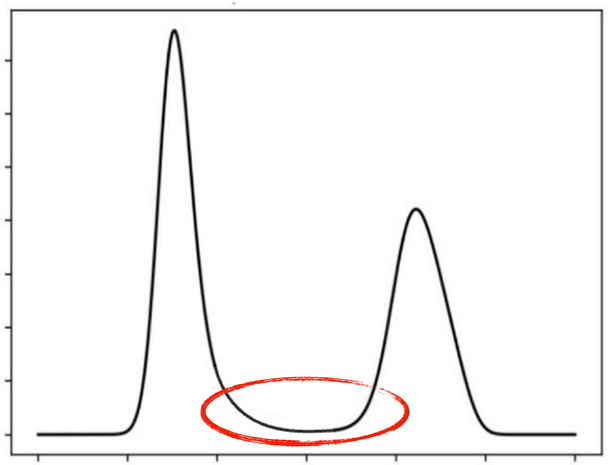
$f(x)$ and $c\|x\|^2$



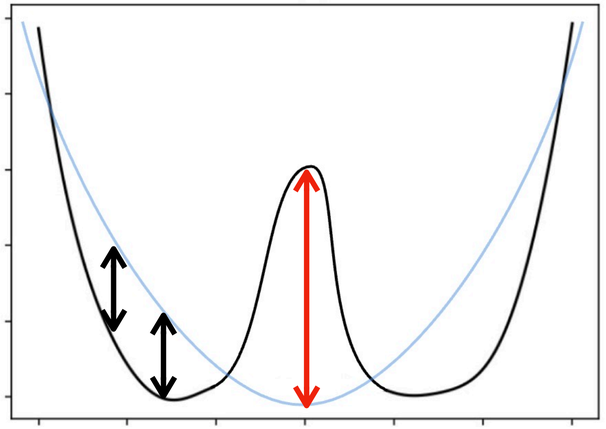
$\hat{f}(x)$ and $c\|x\|^2$

Upper bound: truncation scheme

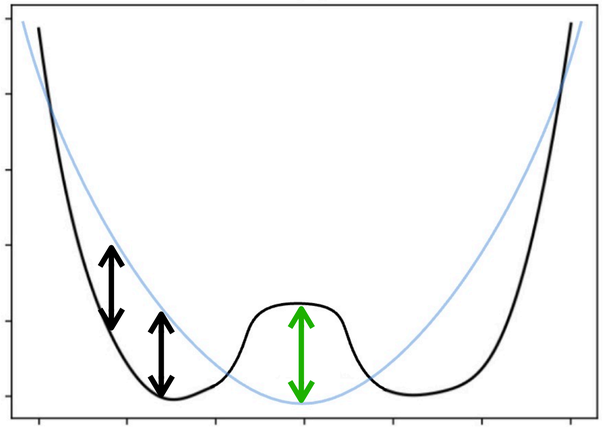
Difficulty: The low valley that separate the modes



$\mu(x) \propto e^{-f(x)}$



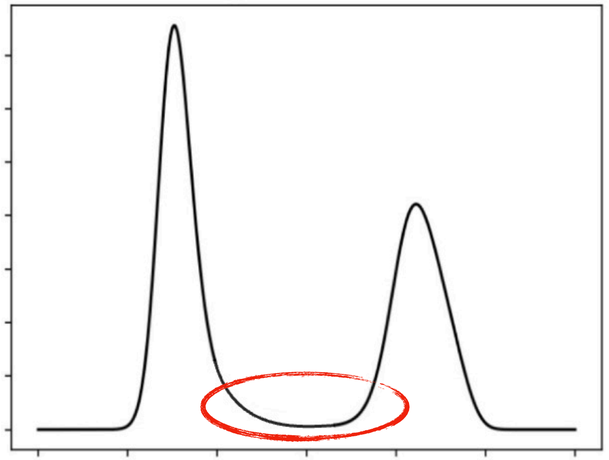
$f(x)$ and $c\|x\|^2$



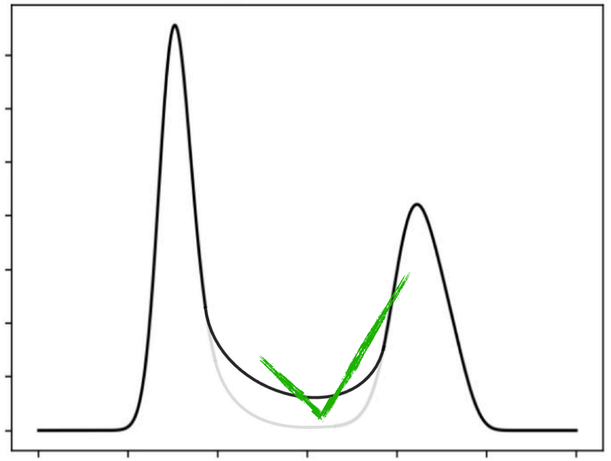
$\hat{f}(x)$ and $c\|x\|^2$

Upper bound: truncation scheme

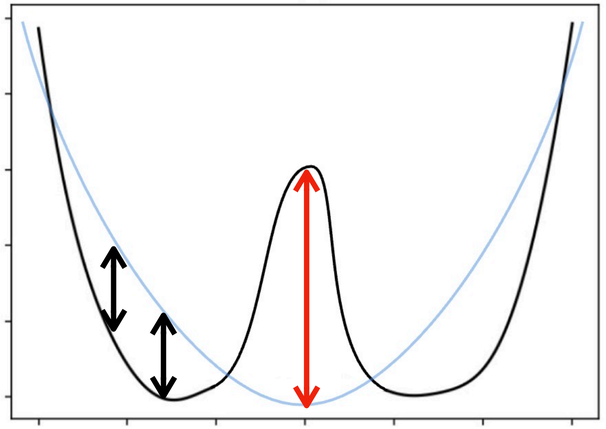
Difficulty: The low valley that separate the modes



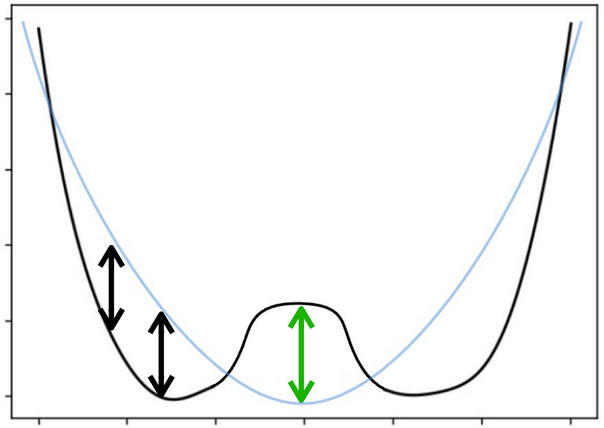
$\mu(x) \propto e^{-f(x)}$



$\hat{\mu}(x) \propto e^{-\hat{f}(x)}$



$f(x)$ and $c\|x\|^2$

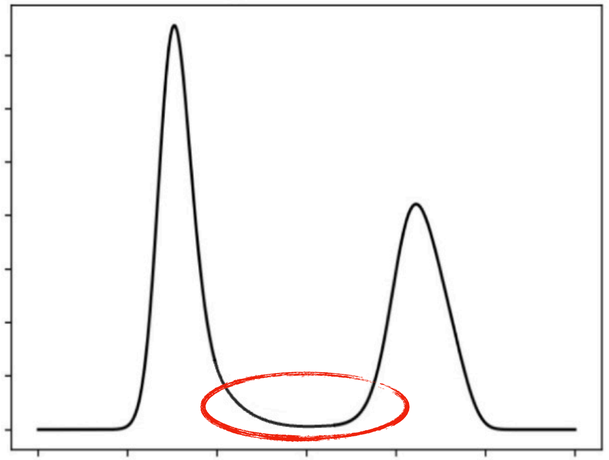


$\hat{f}(x)$ and $c\|x\|^2$

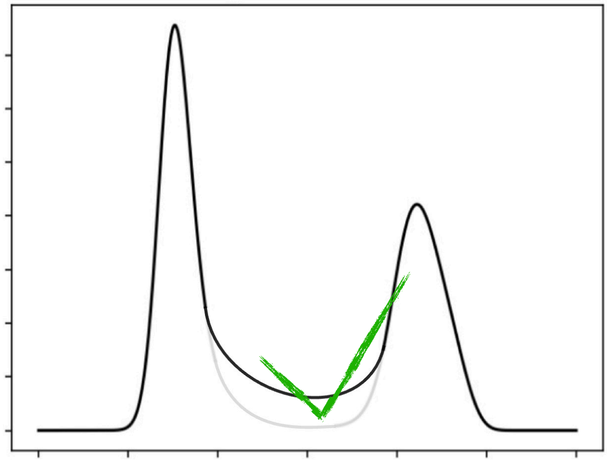
Upper bound: truncation scheme

Difficulty: The low valley that separate the modes

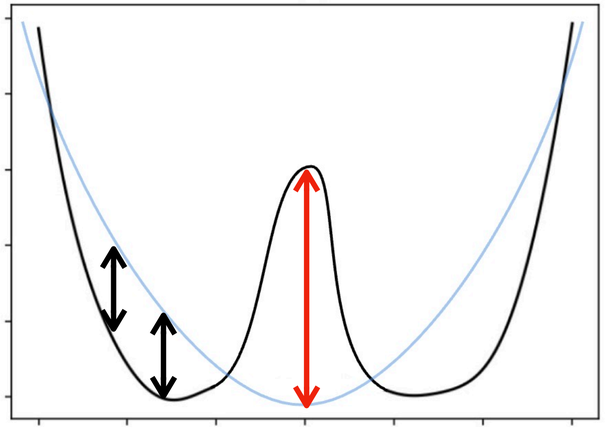
Solution: Truncation



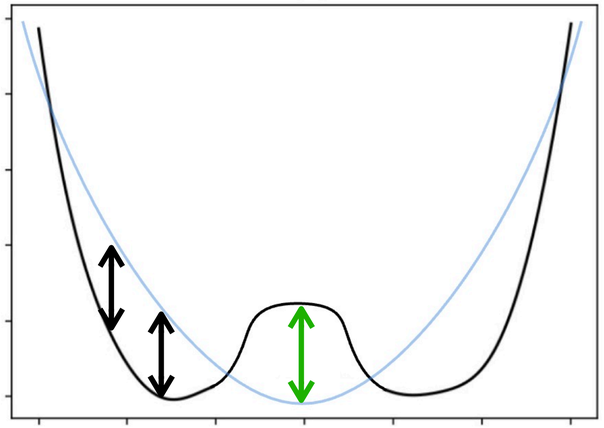
$\mu(x) \propto e^{-f(x)}$



$\hat{\mu}(x) \propto e^{-\hat{f}(x)}$



$f(x)$ and $c\|x\|^2$



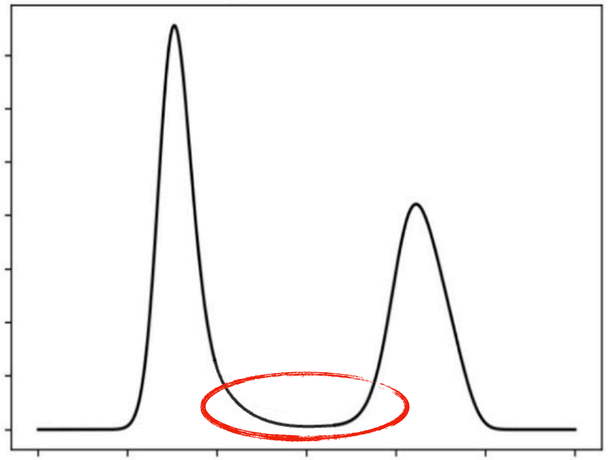
$\hat{f}(x)$ and $c\|x\|^2$

Upper bound: truncation scheme

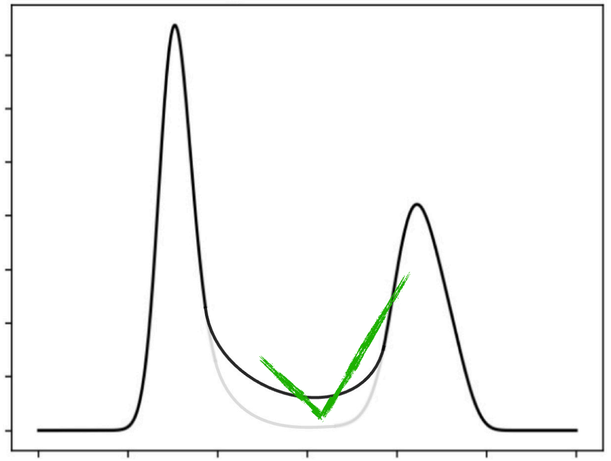
Difficulty: The low valley that separate the modes

Solution: Truncation

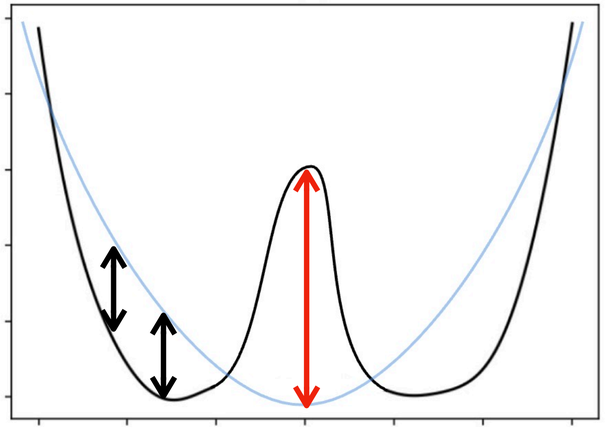
The distribution $\hat{\mu}$ satisfies a Poincaré inequality with $\gamma = \left(\frac{LM}{d\epsilon}\right)^{-O(d)}$.



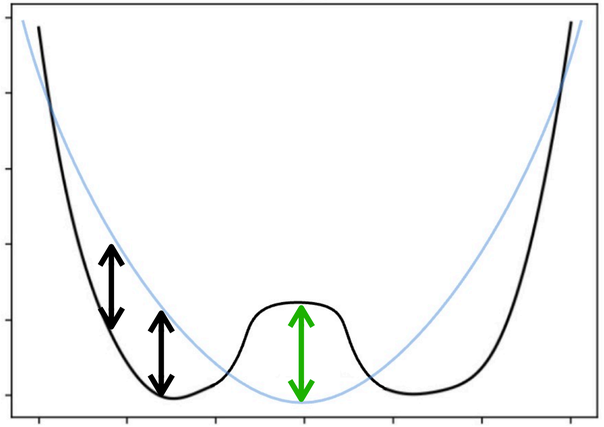
$\mu(x) \propto e^{-f(x)}$



$\hat{\mu}(x) \propto e^{-\hat{f}(x)}$



$f(x)$ and $c\|x\|^2$



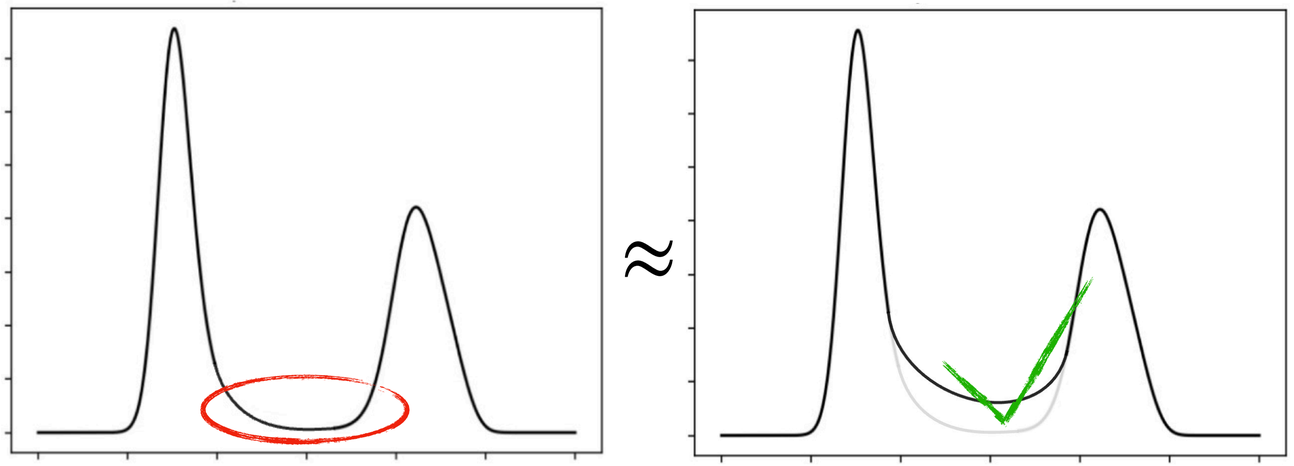
$\hat{f}(x)$ and $c\|x\|^2$

Upper bound: truncation scheme

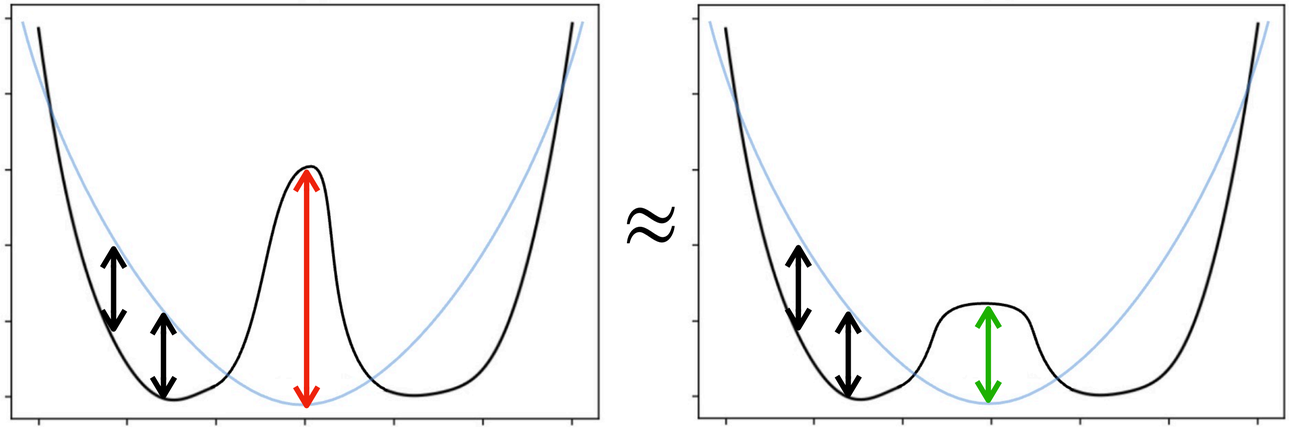
Difficulty: The low valley that separate the modes

Solution: Truncation

The distribution $\hat{\mu}$ satisfies a Poincaré inequality with $\gamma = \left(\frac{LM}{d\epsilon}\right)^{-O(d)}$.



$\mu(x) \propto e^{-f(x)}$ $\quad \approx \quad \hat{\mu}(x) \propto e^{-\hat{f}(x)}$



$f(x)$ and $c\|x\|^2$ $\quad \approx \quad \hat{f}(x)$ and $c\|x\|^2$

Future directions

Future directions

What is the tight query complexity bound?

Future directions

What is the tight query complexity bound?

- We know it is $\left(\frac{LM}{d\epsilon}\right)^{c \cdot d}$, we guess $c = \frac{1}{2}$.

Future directions

What is the tight query complexity bound?

- We know it is $\left(\frac{LM}{d\epsilon}\right)^{c\cdot d}$, we guess $c = \frac{1}{2}$.

Understand to what extent a diffusion process can maintain the smoothness. Can this condition result in efficient algorithm as well?

Future directions

What is the tight query complexity bound?

- We know it is $\left(\frac{LM}{d\epsilon}\right)^{c\cdot d}$, we guess $c = \frac{1}{2}$.

Understand to what extent a diffusion process can maintain the smoothness. Can this condition result in efficient algorithm as well?

- $\exp\left\{\mathcal{O}\left(L^3 \cdot \log^3\left(\frac{Ld+M}{\epsilon}\right)\right)\right\} = \left(\frac{Ld+M}{\epsilon}\right)^{\mathcal{O}\left(L^3 \cdot \log^2\left(\frac{Ld+M}{\epsilon}\right)\right)}$ (Huang, Zou, Dong, Ma and Zhang (2024))

Future directions

What is the tight query complexity bound?

- We know it is $\left(\frac{LM}{d\epsilon}\right)^{c\cdot d}$, we guess $c = \frac{1}{2}$.

Understand to what extent a diffusion process can maintain the smoothness. Can this condition result in efficient algorithm as well?

- $\exp\left\{\mathcal{O}\left(L^3 \cdot \log^3\left(\frac{Ld+M}{\epsilon}\right)\right)\right\} = \left(\frac{Ld+M}{\epsilon}\right)^{\mathcal{O}\left(L^3 \cdot \log^2\left(\frac{Ld+M}{\epsilon}\right)\right)}$ (Huang, Zou, Dong, Ma and Zhang (2024))
- $Ld^3 \cdot \left(\frac{Ld+M}{\epsilon^2}\right)^{\mathcal{O}(L)}$ (coming soon)

Future directions

What is the tight query complexity bound?

- We know it is $\left(\frac{LM}{d\epsilon}\right)^{c\cdot d}$, we guess $c = \frac{1}{2}$.

Understand to what extent a diffusion process can maintain the smoothness. Can this condition result in efficient algorithm as well?

- $\exp\left\{\mathcal{O}\left(L^3 \cdot \log^3\left(\frac{Ld+M}{\epsilon}\right)\right)\right\} = \left(\frac{Ld+M}{\epsilon}\right)^{\mathcal{O}\left(L^3 \cdot \log^2\left(\frac{Ld+M}{\epsilon}\right)\right)}$ (Huang, Zou, Dong, Ma and Zhang (2024))

- $Ld^3 \cdot \left(\frac{Ld+M}{\epsilon^2}\right)^{\mathcal{O}(L)}$ (coming soon)

Thank you!