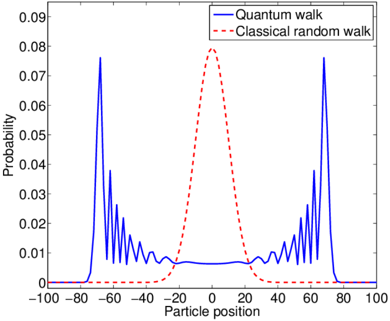
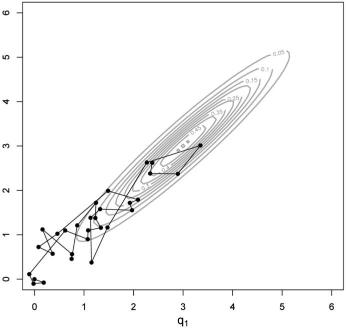
CLASSICAL & QUANTUM ALGORITHMS FOR LOGCONCAVE SAMPLING





Simon Apers
CNRS & IRIF, Paris
simonapers.github.io

EUROPT, Budapest, August ’23

LOGCONCAVE SAMPLING

QUANTUM AND CLASSICAL SPEEDUP

FUTURE DIRECTIONS

LOGCONCAVE SAMPLING

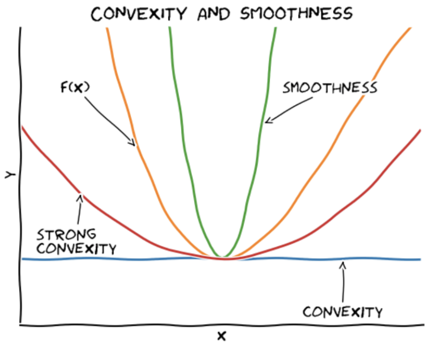
QUANTUM AND CLASSICAL SPEEDUP

FUTURE DIRECTIONS

CONVEX FUNCTIONS AND CONDITION NUMBER

Convex functions and condition number

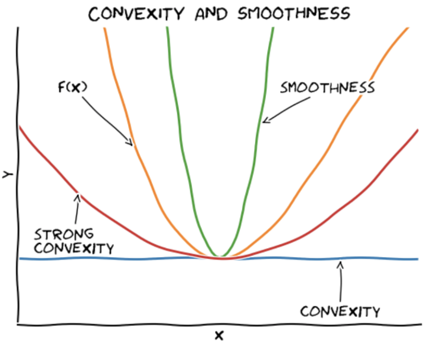
function $f : \mathbb{R}^d \to \mathbb{R}$



CONVEX FUNCTIONS AND CONDITION NUMBER

function $f : \mathbb{R}^d \to \mathbb{R}$

α-strongly convex, β-smooth:
$$\alpha I \preceq \nabla^2 f \preceq \beta I$$

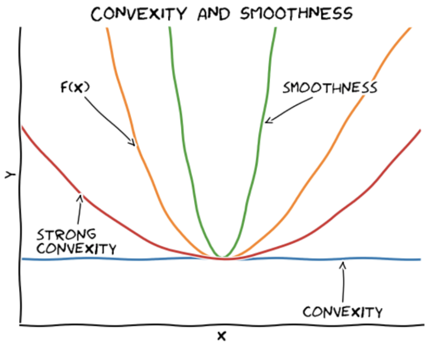


CONVEX FUNCTIONS AND CONDITION NUMBER

function $f : \mathbb{R}^d \to \mathbb{R}$

α-strongly convex, β-smooth:
$$\alpha I \preceq \nabla^2 f \preceq \beta I$$

condition number $\kappa = \frac{\beta}{\alpha}$



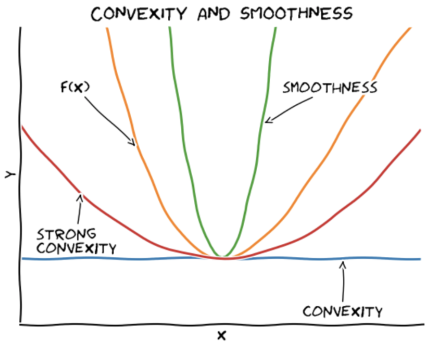
CONVEX FUNCTIONS AND CONDITION NUMBER

function $f : \mathbb{R}^d \to \mathbb{R}$

α-strongly convex, β-smooth:
$\alpha \boldsymbol{I} \preceq \nabla^2 f \preceq \beta \boldsymbol{I}$

condition number $\kappa = \frac{\beta}{\alpha}$

first order query access:
$f(x)$ and $\nabla f(x)$



CONVEX OPTIMIZATION

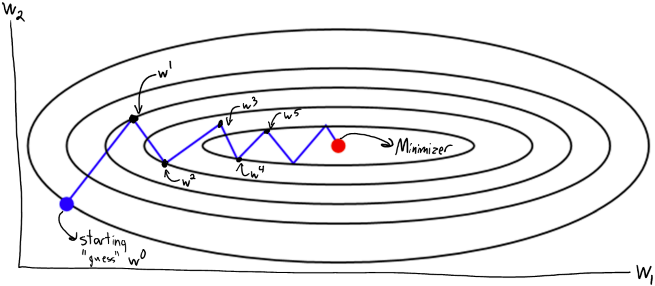
CONVEX OPTIMIZATION

find ε-approximation to optimum $x^* = \min_x f(x)$

CONVEX OPTIMIZATION

find ε-approximation to optimum $x^* = \min_x f(x)$

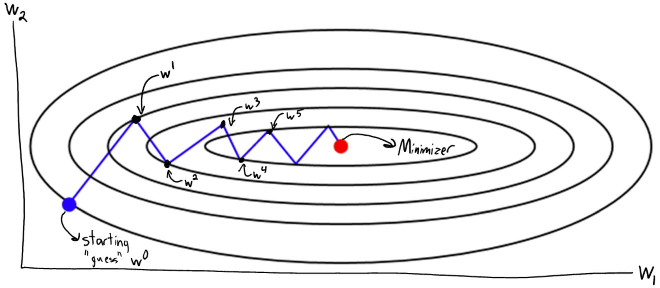
gradient descent: $x' = x - \eta \nabla f(x)$



CONVEX OPTIMIZATION

find ε-approximation to optimum $x^* = \min_x f(x)$

gradient descent: $x' = x - \eta \nabla f(x)$

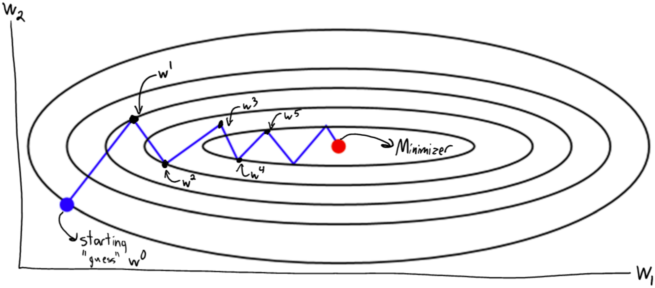


requires $O(\kappa \log(1/\varepsilon))$ queries

CONVEX OPTIMIZATION

find ε-approximation to optimum $x^* = \min_x f(x)$

gradient descent: $x' = x - \eta \nabla f(x)$



requires $O(\kappa \log(1/\varepsilon))$ queries

! Nesterov acceleration: $O(\sqrt{\kappa} \log(1/\varepsilon))$ queries

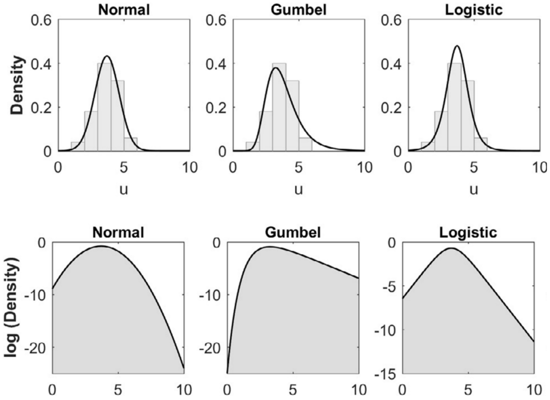
LOGCONCAVE SAMPLING

LOGCONCAVE SAMPLING

sample ε-close to $\pi(x) \propto e^{-f(x)}$ (in TV distance)

Logconcave sampling

sample ε-close to $\pi(x) \propto e^{-f(x)}$ (in TV distance)



Logconcave sampling

Langevin algorithm:

$$x' = x + \eta \nabla f(x) + \zeta, \quad \zeta \sim \mathcal{N}(0, \eta I)$$

Logconcave sampling

Langevin algorithm:

$$x' = x + \eta \nabla f(x) + \zeta, \quad \zeta \sim \mathcal{N}(0, \eta I)$$

+ Metropolis-Hastings correction

LOGCONCAVE SAMPLING

Langevin algorithm:

$$x' = x + \eta \nabla f(x) + \zeta, \quad \zeta \sim \mathcal{N}(0, \eta I)$$

+ Metropolis-Hastings correction

Faster high-accuracy log-concave sampling via algorithmic warm starts

Jason M. Altschuler
NYU
ja4775@nyu.edu

Sinho Chewi
MIT
schewi@mit.edu

February 22, 2023

Langevin requires $O(\kappa \log(1/\varepsilon))$ steps

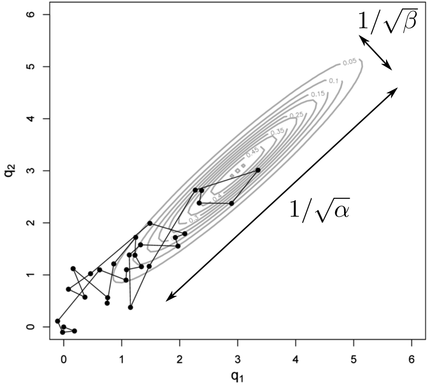
Logconcave sampling

κ-scaling even from "warm start" (e.g., $x_0 = x^*$)

LOGCONCAVE SAMPLING

κ-scaling even from "warm start" (e.g., $x_0 = x^*$)

slowdown caused by *diffusion*

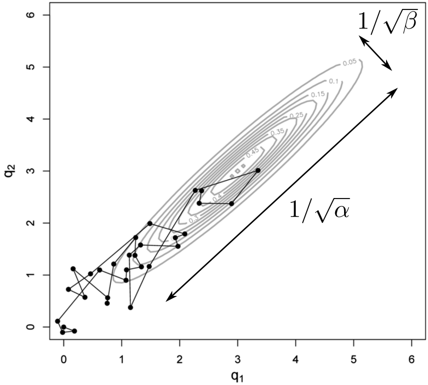


LOGCONCAVE SAMPLING

κ-scaling even from "warm start" (e.g., $x_0 = x^*$)

slowdown caused by *diffusion*

step length $\sim 1/\sqrt{\beta}$



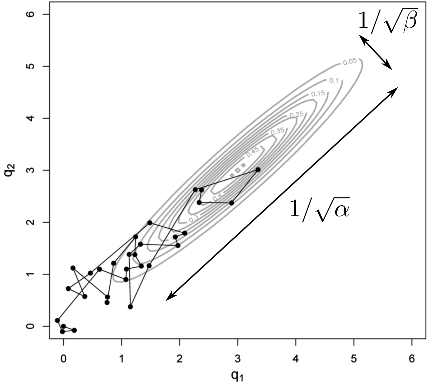
Logconcave sampling

κ-scaling even from "warm start" (e.g., $x_0 = x^*$)

slowdown caused by *diffusion*

step length $\sim 1/\sqrt{\beta}$

t steps distance $\sim \sqrt{t}/\sqrt{\beta}$
$\sim 1/\sqrt{\alpha}$ requires $t \sim \beta/\alpha = \kappa$



Logconcave sampling

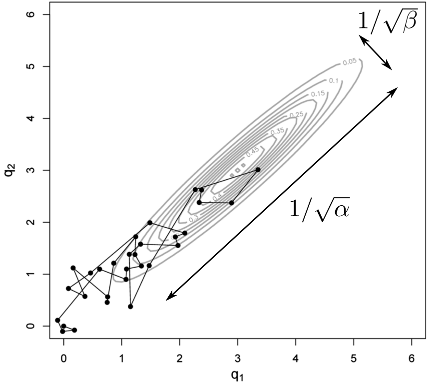
κ-scaling even from "warm start" (e.g., $x_0 = x^*$)

slowdown caused by *diffusion*

step length $\sim 1/\sqrt{\beta}$

t steps distance $\sim \sqrt{t}/\sqrt{\beta}$
$\sim 1/\sqrt{\alpha}$ requires $t \sim \beta/\alpha = \kappa$

improve to $\sqrt{\kappa}$?



LOGCONCAVE SAMPLING

QUANTUM AND CLASSICAL SPEEDUP

FUTURE DIRECTIONS

QUANTUM SPEEDUP

Quantum Algorithms for Sampling Log-Concave Distributions and Estimating Normalizing Constants

Part of Advances in Neural Information Processing Systems 35 (NeurIPS 2022) Main Conference Track

Authors

Andrew M. Childs, Tongyang Li, Jin-Peng Liu, Chunhao Wang, Ruizhe Zhang

QUANTUM SPEEDUP

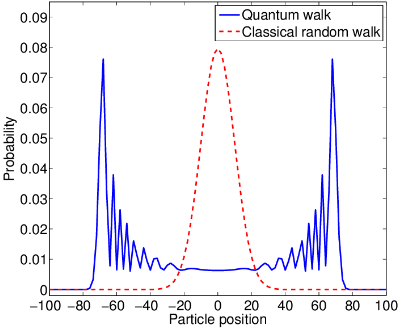
Quantum Algorithms for Sampling Log-Concave Distributions and Estimating Normalizing Constants

Part of Advances in Neural Information Processing Systems 35 (NeurIPS 2022) Main Conference Track

Authors

Andrew M. Childs, Tongyang Li, Jin-Peng Liu, Chunhao Wang, Ruizhe Zhang

quantum walks show “ballistic” wave-like behavior



Quantum speedup

Quantum Algorithms for Sampling Log-Concave Distributions and Estimating Normalizing Constants

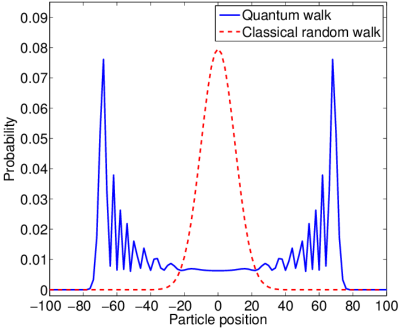
Part of Advances in Neural Information Processing Systems 35 (NeurIPS 2022) Main Conference Track

Authors

Andrew M. Childs, Tongyang Li, Jin-Peng Liu, Chunhao Wang, Ruizhe Zhang

quantum walks show "ballistic" wave-like behavior

quantum walk based on Langevin algorithm (+ simulated annealing) requires $O(\sqrt{\kappa}\log(1/\varepsilon))$ queries



Classical speedup

Hamiltonian Monte Carlo for efficient Gaussian sampling: long and random steps '22

Simon Apers* Sander Gribling* Dániel Szilágyi*

Classical speedup

Hamiltonian Monte Carlo for efficient Gaussian sampling: long and random steps '22

Simon Apers* Sander Gribling* Dániel Szilágyi*

for Gaussians:

"Hamiltonian Monte Carlo" algorithm
requires $O(\sqrt{\kappa}\log(1/\varepsilon))$ queries

CLASSICAL SPEEDUP

Hamiltonian Monte Carlo for efficient Gaussian sampling:
long and random steps '22

Simon Apers* Sander Gribling* Dániel Szilágyi*

for Gaussians:
"Hamiltonian Monte Carlo" algorithm
requires $O(\sqrt{\kappa}\log(1/\varepsilon))$ queries

conjecture:
$O(\sqrt{\kappa}\log(1/\varepsilon))$ queries for *all* logconcave distributions

HAMILTONIAN MONTE CARLO

Hamiltonian Monte Carlo

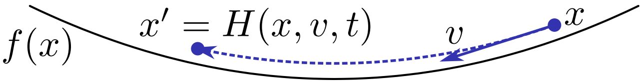
$H(x, v, t)$ for $x, v \in \mathbb{R}^d$, $t \in \mathbb{R}$

Hamiltonian Monte Carlo

$H(x, v, t)$ for $x, v \in \mathbb{R}^d$, $t \in \mathbb{R}$
= position after Hamiltonian dynamics
for time t, from (x, v), with potential energy $f(x)$

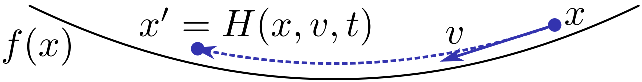
Hamiltonian Monte Carlo

$H(x, v, t)$ for $x, v \in \mathbb{R}^d$, $t \in \mathbb{R}$
= position after Hamiltonian dynamics
for time t, from (x, v), with potential energy $f(x)$



Hamiltonian Monte Carlo

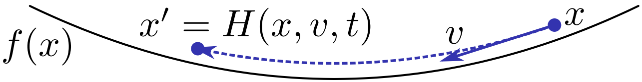
$H(x, v, t)$ for $x, v \in \mathbb{R}^d$, $t \in \mathbb{R}$
= position after Hamiltonian dynamics
for time t, from (x, v), with potential energy $f(x)$



HMC step: from $x \in \mathbb{R}^d$ do

Hamiltonian Monte Carlo

$H(x, v, t)$ for $x, v \in \mathbb{R}^d$, $t \in \mathbb{R}$
= position after Hamiltonian dynamics
for time t, from (x, v), with potential energy $f(x)$

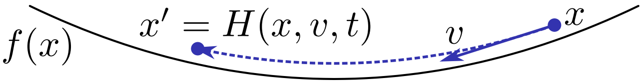


HMC step: from $x \in \mathbb{R}^d$ do

1. pick random time $t \sim 1/\sqrt{\alpha}$ and velocity $v \sim \mathcal{N}(0, I_d)$

Hamiltonian Monte Carlo

$H(x, v, t)$ for $x, v \in \mathbb{R}^d$, $t \in \mathbb{R}$
= position after Hamiltonian dynamics
for time t, from (x, v), with potential energy $f(x)$

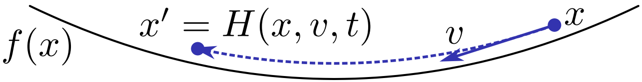


HMC step: from $x \in \mathbb{R}^d$ do

1. pick random time $t \sim 1/\sqrt{\alpha}$ and velocity $v \sim \mathcal{N}(0, I_d)$
2. set $x' = H(x, v, t)$

HAMILTONIAN MONTE CARLO

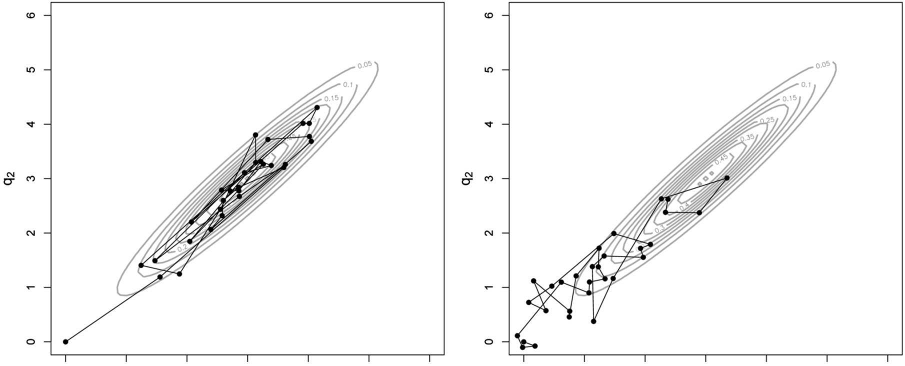
$H(x, v, t)$ for $x, v \in \mathbb{R}^d$, $t \in \mathbb{R}$
= position after Hamiltonian dynamics
for time t, from (x, v), with potential energy $f(x)$



HMC step: from $x \in \mathbb{R}^d$ do

1. pick random time $t \sim 1/\sqrt{\alpha}$ and velocity $v \sim \mathcal{N}(0, I_d)$
2. set $x' = H(x, v, t)$
3. Metropolis-Hastings correction

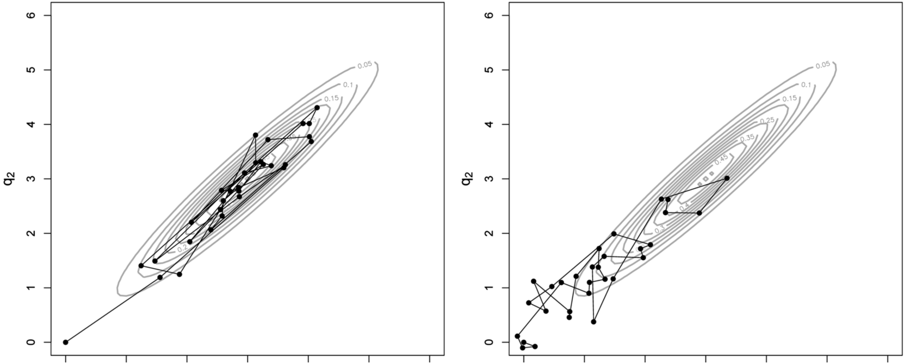
Hamiltonian dynamics yield “ballistic” motion



(left: HMC, right: Langevin)

HAMILTONIAN MONTE CARLO

Hamiltonian dynamics yield "ballistic" motion



(left: HMC, right: Langevin)

= source of $\kappa \to \sqrt{\kappa}$ speedup

HAMILTONIAN INTEGRATION

? how to compute $H(x, v, t)$?

Hamiltonian integration

? how to compute $H(x, v, t)$?

$\rightarrow$ leapfrog integrator:

$$v_{t+\delta/2} = v_t - \frac{\delta}{2}\nabla f(x_t)$$
$$x_{t+\delta} = x_t + \delta v_{t+\delta/2}$$
$$v_{t+\delta} = v_{t+\delta/2} - \frac{\delta}{2}\nabla f(x_{t+\delta})$$

Hamiltonian integration

? how to compute $H(x, v, t)$?

$\rightarrow$ leapfrog integrator:

$$
\begin{aligned}
v_{t+\delta/2} &= v_t - \frac{\delta}{2}\nabla f(x_t) \\
x_{t+\delta} &= x_t + \delta v_{t+\delta/2} \\
v_{t+\delta} &= v_{t+\delta/2} - \frac{\delta}{2}\nabla f(x_{t+\delta})
\end{aligned}
$$

= symplectic, second-order integrator

Hamiltonian integration

? how to compute $H(x, v, t)$?

$\rightarrow$ leapfrog integrator:

$$\begin{aligned} v_{t+\delta/2} &= v_t - \frac{\delta}{2}\nabla f(x_t) \\ x_{t+\delta} &= x_t + \delta v_{t+\delta/2} \\ v_{t+\delta} &= v_{t+\delta/2} - \frac{\delta}{2}\nabla f(x_{t+\delta}) \end{aligned}$$

= symplectic, second-order integrator

bounded error for step size $\delta \in O\left(\frac{1}{\sqrt{\beta} d^{1/4}}\right)$

Hamiltonian integration

? how to compute $H(x, v, t)$?

$\rightarrow$ leapfrog integrator:

$$\begin{aligned} v_{t+\delta/2} &= v_t - \frac{\delta}{2}\nabla f(x_t) \\ x_{t+\delta} &= x_t + \delta v_{t+\delta/2} \\ v_{t+\delta} &= v_{t+\delta/2} - \frac{\delta}{2}\nabla f(x_{t+\delta}) \end{aligned}$$

= symplectic, second-order integrator

bounded error for step size $\delta \in O\left(\frac{1}{\sqrt{\beta} d^{1/4}}\right)$

integration time $t \in O(1/\sqrt{\alpha})$
requires $t/\delta \in O(\sqrt{\kappa} d^{1/4})$ gradient queries

HMC comparison

HMC comparison

Langevin:

$O(\kappa\sqrt{d}\log(1/\varepsilon))$

HMC comparison

Langevin:

$$O(\kappa\sqrt{d}\log(1/\varepsilon))$$

quantum walk:

$$O(\sqrt{\kappa}d\log(1/\varepsilon))$$

$O(\sqrt{\kappa}d^{1/4}\log(1/\varepsilon))$ (with warm start)

HMC comparison

Langevin:

$O(\kappa\sqrt{d}\log(1/\varepsilon))$

quantum walk:

$O(\sqrt{\kappa}d\log(1/\varepsilon))$

$O(\sqrt{\kappa}d^{1/4}\log(1/\varepsilon))$ (with warm start)

HMC:

$O(\sqrt{\kappa}d^{1/4}\log(1/\varepsilon))$ (for Gaussians)

HMC comparison

Langevin:
$$O(\kappa\sqrt{d}\log(1/\varepsilon))$$

quantum walk:
$$O(\sqrt{\kappa}d\log(1/\varepsilon))$$
$O(\sqrt{\kappa}d^{1/4}\log(1/\varepsilon))$ (with warm start)

HMC:
$O(\sqrt{\kappa}d^{1/4}\log(1/\varepsilon))$ (for Gaussians)

! compare with "dimension-free" scaling of $O(\sqrt{\kappa})$ for optimization

LOGCONCAVE SAMPLING

QUANTUM AND CLASSICAL SPEEDUP

FUTURE DIRECTIONS

OPEN QUESTIONS

Open questions

$\sqrt{\kappa}$-scaling for general logconcave?

Open questions

$\sqrt{\kappa}$-scaling for general logconcave?

dimension-free scaling for Gaussian / general logconcave?

(e.g., use higher-order integrators)

Open questions

$\sqrt{\kappa}$-scaling for general logconcave?

dimension-free scaling for Gaussian / general logconcave?

(e.g., use higher-order integrators)

classical and quantum speedups in non-continuous setting?

! MCMC often in discrete graph setting

EXTRA: SCHRÖDINGER DYNAMICS

EXTRA: SCHRÖDINGER DYNAMICS

? replace Hamiltonian dynamics by Schrödinger dynamics ?

$$\boxed{\frac{\partial}{\partial t}\psi(t) = -i\hat{H}\psi(t)}$$

EXTRA: SCHRÖDINGER DYNAMICS

? replace Hamiltonian dynamics by Schrödinger dynamics ?

$$\frac{\partial}{\partial t}\psi(t) = -i\hat{H}\psi(t)$$

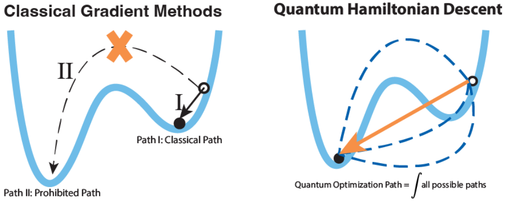
e.g., can use “quantum tunneling” for nonconvex optimization

EXTRA: SCHRÖDINGER DYNAMICS

? replace Hamiltonian dynamics by Schrödinger dynamics ?

$$\frac{\partial}{\partial t}\psi(t) = -i\hat{H}\psi(t)$$

e.g., can use "quantum tunneling" for nonconvex optimization

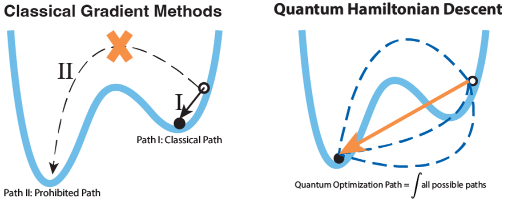


EXTRA: SCHRÖDINGER DYNAMICS

? replace Hamiltonian dynamics by Schrödinger dynamics ?

$$\frac{\partial}{\partial t}\psi(t) = -i\hat{H}\psi(t)$$

e.g., can use "quantum tunneling" for nonconvex optimization



On Quantum Speedups for Nonconvex Optimization via Quantum Tunneling Walks* '22

Yizhou Liu,† Weijie J. Su,‡ Tongyang Li§

THANK YOU!

figure references:

https://www.pokutta.com/blog/research/2018/12/06/cheatsheet-smooth-idealized.html
https://link.springer.com/article/10.1007/s11222-012-9373-1
https://www.cs.ubc.ca/ schmidtm/Courses/540-W19/L4.pdf
https://www.sciencedirect.com/science/article/abs/pii/S0167473017300693