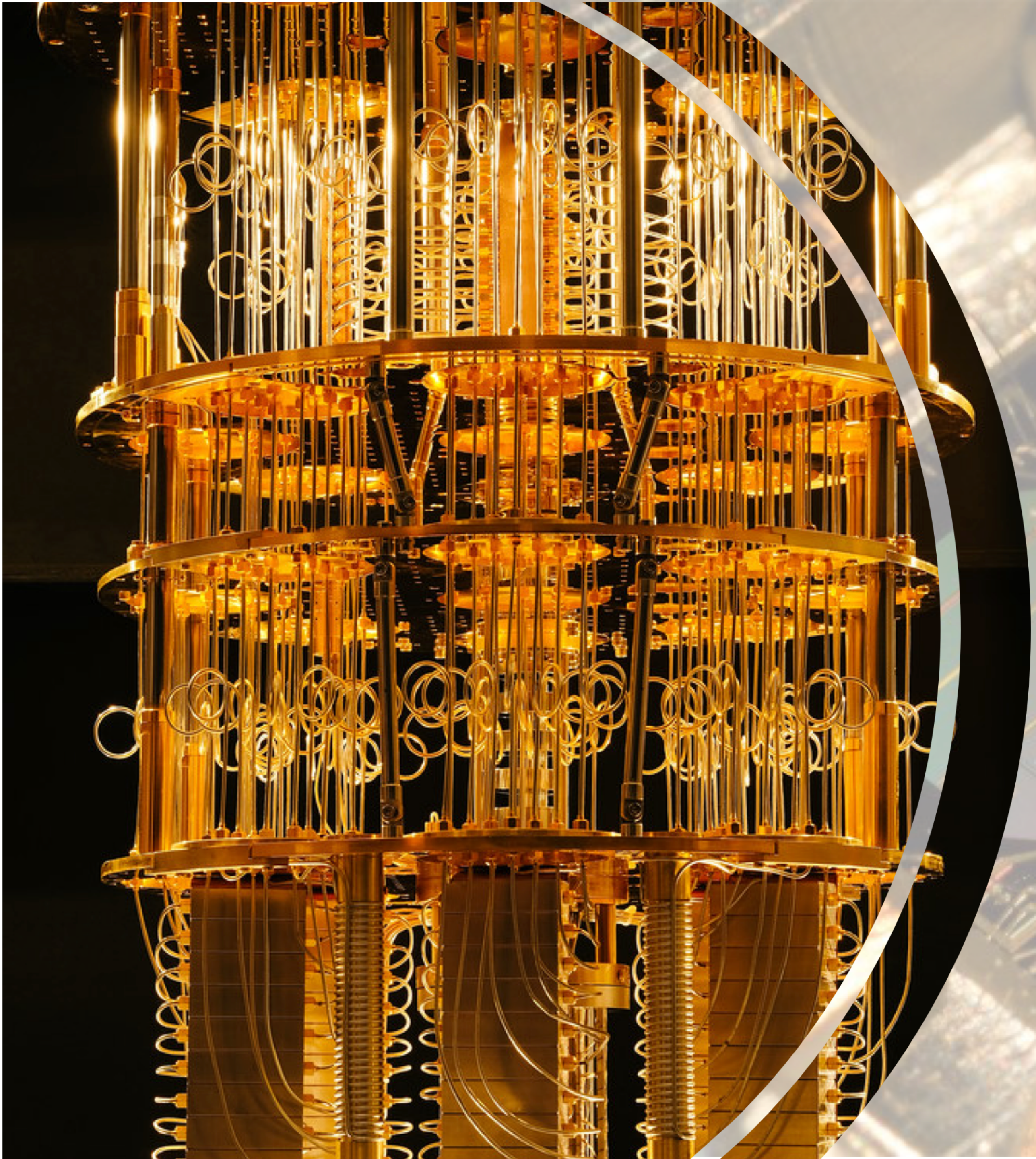


EPFL

Variational quantum simulation: a case study for understanding warm starts

R. Puig[*], M. Drudis[*], S. Thanaslip, Z. Holmes

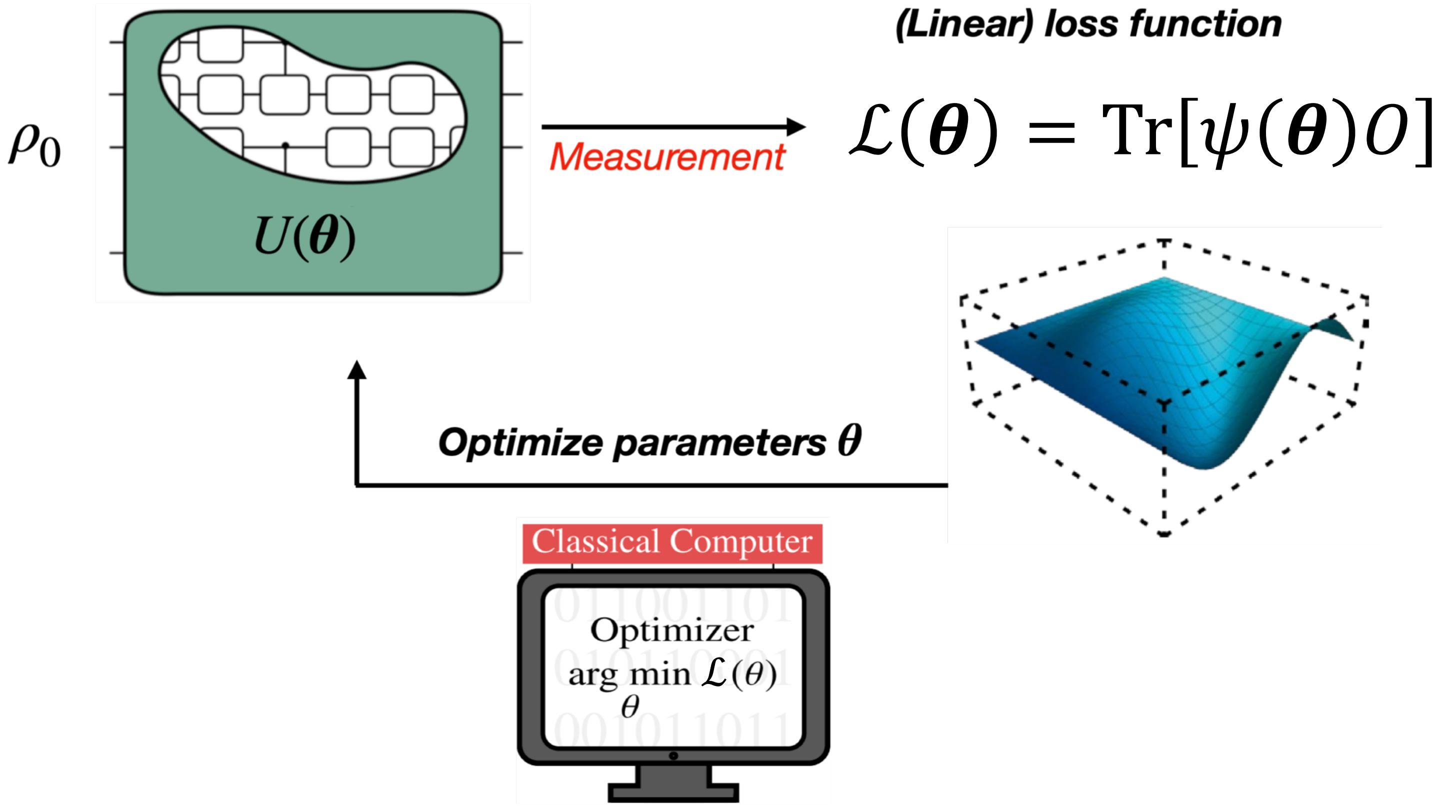
arXiv:2404.10044

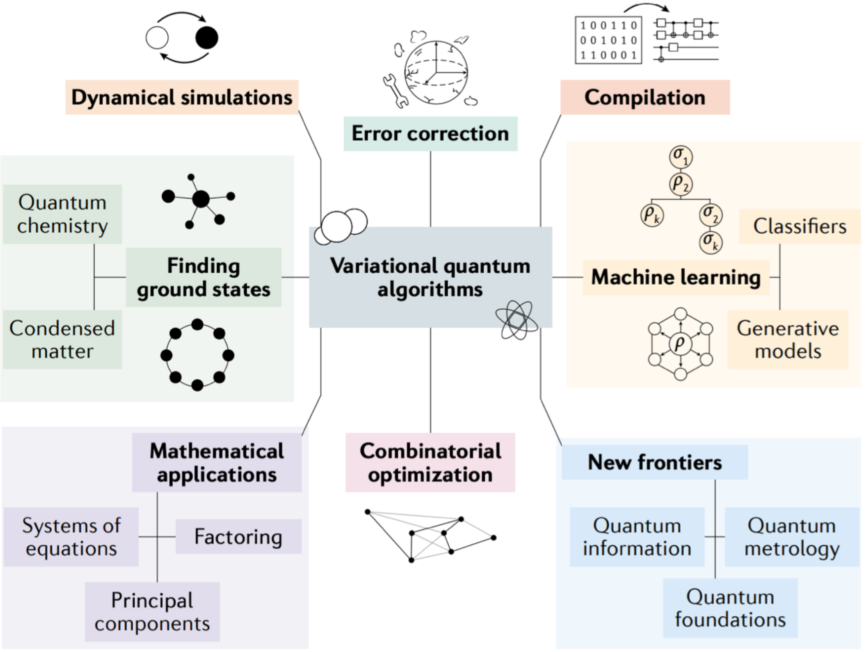
Marc Drudis

Supanut Thanasilp

Zoë Holmes

Variational Quantum Algorithms





M. Cerezo, et al. "Variational quantum algorithms." Nature Reviews Physics 3.9 (2021)

Barren plateau phenomena

$$\mathrm{Var}_{\mathrm{Unif}}(\mathcal{L}) \sim \frac{1}{2^n}$$

$$+$$

$$P(|\mathcal{L}| \geq \delta) \leq \frac{\mathrm{Var}(\mathcal{L})}{\delta^2}$$

↓

Probability of non-zero gradients vanishes exponentially with problem size.

↓

Shot required for training grows exponentially with problem size.

Curse of Dimensionality

Choice in circuit

Barren plateaus in quantum neural network training landscapes

Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush & Hartmut Neven

Nature Communications **9**, Article number: 4812 (2018) | Cite this article

Curse of Dimensionality

Choice in circuit

- Too expressive

Barren plateaus in quantum neural network training landscapes

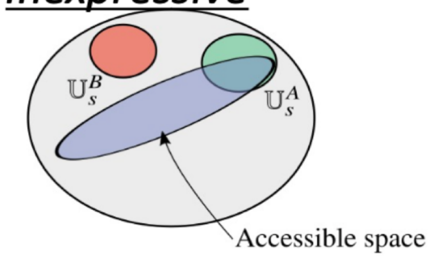
Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush & Hartmut Neven

Nature Communications **9**, Article number: 4812 (2018) | Cite this article

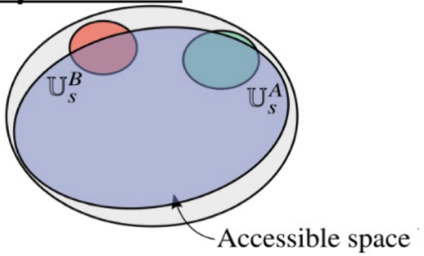
Connecting Ansatz Expressibility to Gradient Magnitudes and Barren Plateaus

Zoë Holmes, Kunal Sharma, M. Cerezo, and Patrick J. Coles
PRX Quantum **3**, 010313 – Published 24 January 2022

Inexpressive



Expressive



Curse of Dimensionality

Choice in circuit

- Too expressive
- Too entangling

Barren plateaus in quantum neural network training landscapes

Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush & Hartmut Neven

Nature Communications **9**, Article number: 4812 (2018) | Cite this article

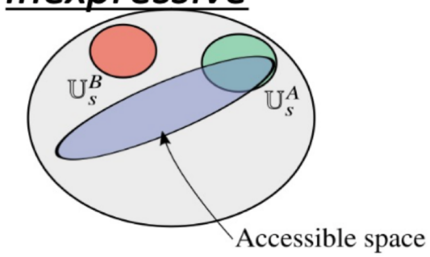
Connecting Ansatz Expressibility to Gradient Magnitudes and Barren Plateaus

Zoë Holmes, Kunal Sharma, M. Cerezo, and Patrick J. Coles
PRX Quantum **3**, 010313 – Published 24 January 2022

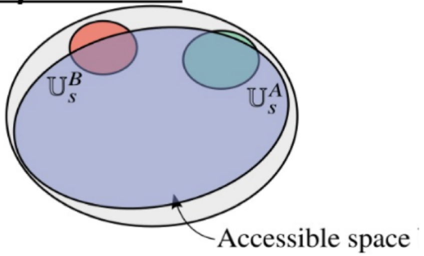
Entanglement-Induced Barren Plateaus

Carlos Ortiz Marrero, Mária Kieferová, and Nathan Wiebe
PRX Quantum **2**, 040316 – Published 25 October 2021

Inexpressive



Expressive



Curse of Dimensionality

Choice in circuit

- Too expressive
- Too entangling

Choice in target learning problem

Barren plateaus in quantum neural network training landscapes

Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush & Hartmut Neven

Nature Communications **9**, Article number: 4812 (2018) | Cite this article

Connecting Ansatz Expressibility to Gradient Magnitudes and Barren Plateaus

Zoë Holmes, Kunal Sharma, M. Cerezo, and Patrick J. Coles
PRX Quantum **3**, 010313 – Published 24 January 2022

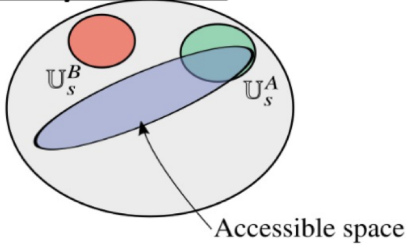
Entanglement-Induced Barren Plateaus

Carlos Ortiz Marrero, Mária Kieferová, and Nathan Wiebe
PRX Quantum **2**, 040316 – Published 25 October 2021

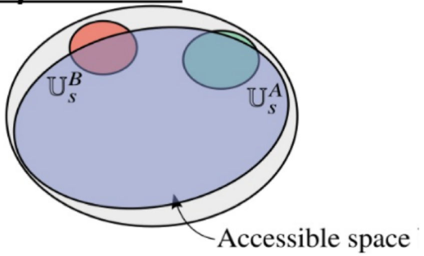
Barren Plateaus Preclude Learning Scramblers

Zoë Holmes, Andrew Arrasmith, Bin Yan, Patrick J. Coles, Andreas Albrecht, and Andr
Phys. Rev. Lett. **126**, 190501 – Published 12 May 2021

Inexpressive



Expressive



Curse of Dimensionality

Choice in circuit

Barren plateaus in quantum neural network training landscapes

Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush & Hartmut Neven

Nature Communications 9, Article number: 4812 (2018) | Cite this article

- Too expressive

Connecting Ansatz Expressibility to Gradient Magnitudes and Barren Plateaus

Zoë Holmes, Kunal Sharma, M. Cerezo, and Patrick J. Coles
PRX Quantum **3**, 010313 – Published 24 January 2022

- Too entangling

Entanglement-Induced Barren Plateaus

Carlos Ortiz Marrero, Mária Kieferová, and Nathan Wiebe
PRX Quantum **2**, 040316 – Published 25 October 2021

Choice in target learning problem

Barren Plateaus Preclude Learning Scramblers

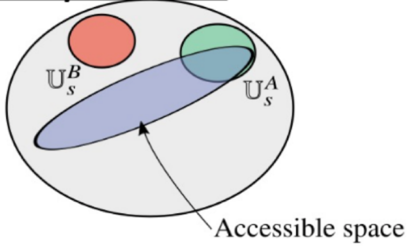
Zoë Holmes, Andrew Arrasmith, Bin Yan, Patrick J. Coles, Andreas Albrecht, and Andrew T. Sornbor
Phys. Rev. Lett. **126**, 190501 – Published 12 May 2021

Choice in cost function

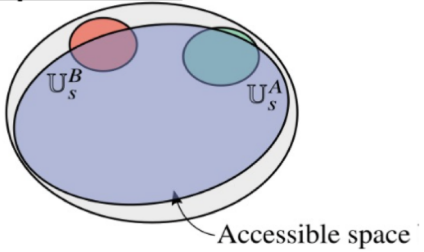
Cost function dependent barren plateaus in shallow parametrized quantum circuits

M. Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio & Patrick J. Coles

Inexpressive



Expressive



Global

$$H = \sigma_1^z \otimes \sigma_2^z \otimes \cdots \otimes \sigma_n^z$$

Local

$$H = \sigma_1^z \otimes \mathbb{I}_2 \otimes \cdots \otimes \mathbb{I}_n$$

Curse of Dimensionality

Choice in circuit

- Too expressive
- Too entangling

Choice in target learning problem

Choice in cost function

Noise

Barren plateaus in quantum neural network training landscapes

Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush & Hartmut Neven

Nature Communications **9**, Article number: 4812 (2018) | Cite this article

Connecting Ansatz Expressibility to Gradient Magnitudes and Barren Plateaus

Zoë Holmes, Kunal Sharma, M. Cerezo, and Patrick J. Coles
PRX Quantum **3**, 010313 – Published 24 January 2022

Entanglement-Induced Barren Plateaus

Carlos Ortiz Marrero, Mária Kieferová, and Nathan Wiebe
PRX Quantum **2**, 040316 – Published 25 October 2021

Barren Plateaus Preclude Learning Scramblers

Zoë Holmes, Andrew Arrasmith, Bin Yan, Patrick J. Coles, Andreas Albrecht, and Andrew T. Sornbor
Phys. Rev. Lett. **126**, 190501 – Published 12 May 2021

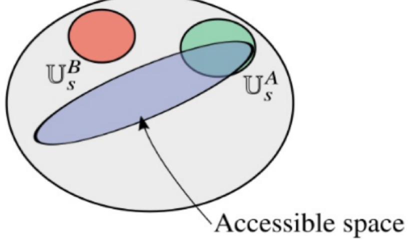
Cost function dependent barren plateaus in shallow parametrized quantum circuits

M. Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio & Patrick J. Coles

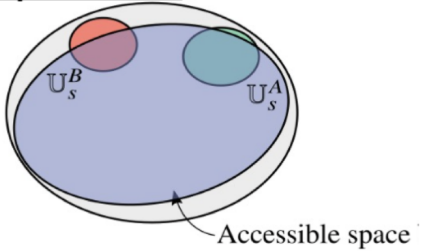
Noise-induced barren plateaus in variational quantum algorithms

Samson Wang, Enrico Fontana, M. Cerezo, Kunal Sharma, Akira Sone, Lukasz Cincio & Patrick J. Coles

Inexpressive



Expressive



Global

$$H = \sigma_1^z \otimes \sigma_2^z \otimes \cdots \otimes \sigma_n^z$$

Local

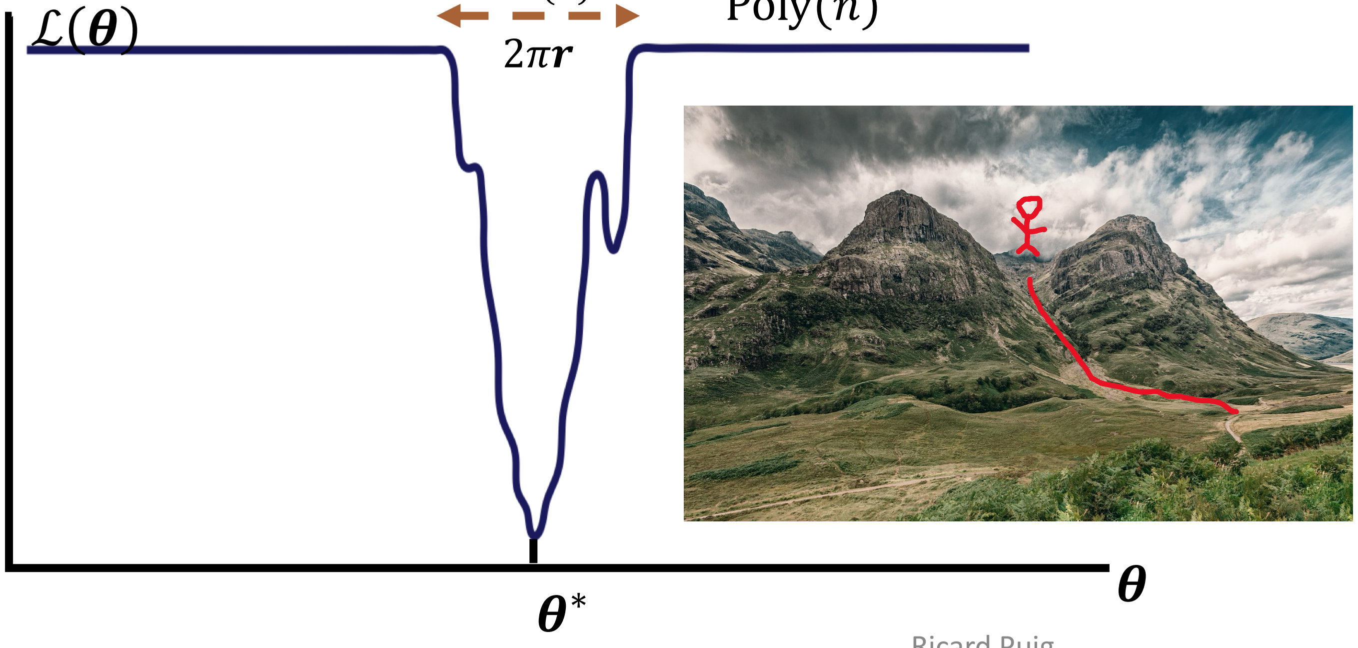
$$H = \sigma_1^z \otimes \mathbb{I}_2 \otimes \cdots \otimes \mathbb{I}_n$$

Average statement!

What happens if we know the solution to be close to a point??

$\mathrm{Var}_{\mathrm{Unif}}(\mathcal{L}) \sim \frac{1}{2\^{}n}$

$\mathrm{Var}_{\mathrm{P}(\boldsymbol{\theta})}(\mathcal{L}) \sim \frac{1}{\mathrm{Poly}(n)}$



Identity initialization and similar

$$\mathrm{Var}_{\mathrm{Unif}}(\mathcal{L}) \sim \frac{1}{2^n} \xrightarrow[\mathrm{Unif} \to \boldsymbol{\theta} \in [-\pi \boldsymbol{r}, +\pi \boldsymbol{r}]]{\text{They get}} \mathrm{Var}_{\mathrm{P}(\boldsymbol{\theta})}(\mathcal{L}) \sim \frac{1}{\mathrm{Poly}(n)}$$

Avoiding barren plateaus via Gaussian Mixture Model

Xiao Shi[1,2] and Yun Shang[1,3,*]

[1]Institute of Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
[2]School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
[3]NCMIS, MDIS, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190,China

(Dated: February 22, 2024)

Hardware-efficient ansatz without barren plateaus in any depth

Chae-Yeun Park,[1] Minhyeok Kang,[2,3,4] and Joonsuk Huh[1,2,3,4]

[1]Xanadu, Toronto, ON, M5G 2C8, Canada
[2]Department of Chemistry, Sungkyunkwan University, Suwon 16419, Korea
[3]SKKU Advanced Institute of Nanotechnology (SAINT), Sungkyunkwan University, Suwon 16419, Korea
[4]Institute of Quantum Biophysics, Sungkyunkwan University, Suwon 16419, Korea

(Dated: March 11, 2024)

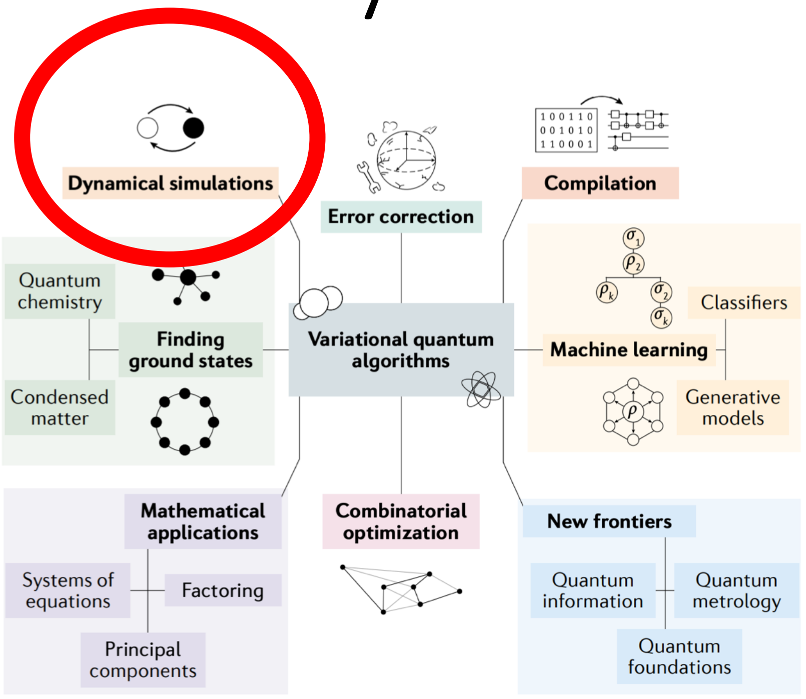
Trainability Enhancement of Parameterized Quantum Circuits via Reduced-Domain Parameter Initialization

Yabo Wang[1,2], Bo Qi[1,2], Chris Ferrie[3], and Daoyi Dong[4]

[1]Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, P. R. China
[2]University of Chinese Academy of Sciences, Beijing 100049, P. R. China
[3]Center for Quantum Software and Information, University of Technology Sydney, Ultimo NSW 2007, Australia
[4]School of Engineering and Information Technology, University of New South Wales, Canberra ACT 2600, Australia

Only look around zero.
The strategy does not take into account the problem structure

Let's try!



Noise-Resilient Quantum Dynamics Using Symmetry-Preserving Ansatzes

Matthew Otten,* Cristian L. Cortes, and Stephen K. Gray
Center for Nanoscale Materials, Argonne National Laboratory, Lemont, Illinois, 60439
(Dated: October 15, 2019)

An efficient quantum algorithm for the time evolution of parameterized circuits

Stefano Barison, Filippo Vicentini, and Giuseppe Carleo

Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

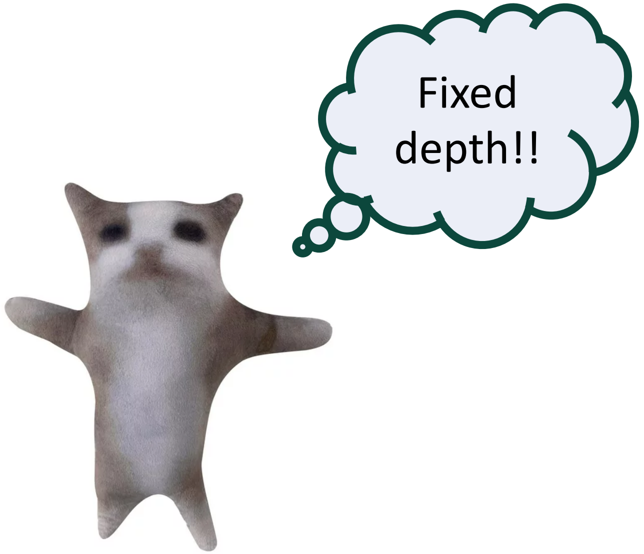
Quantum dynamics simulations beyond the coherence time on noisy intermediate-scale quantum hardware by variational Trotter compression

Noah F. Berthusen,[1,2,*] Thaís V. Trevisan,[1,3] Thomas Iadecola,[1,3,†] and Peter P. Orth[1,3,‡]
[1]*Ames Laboratory, Ames, Iowa 50011, USA*
[2]*Department of Electrical and Computer Engineering, Iowa State University, Ames, Iowa 50011, USA*
[3]*Department of Physics and Astronomy, Iowa State University, Ames, Iowa 50011, USA*

Variational Quantum Simulation

$$\mathcal{L}(\boldsymbol{\theta}) = 1 - \mathrm{Tr}\big[U(\boldsymbol{\theta})\psi_0 U^{\dagger}(\boldsymbol{\theta})\underbrace{\mathrm{e}^{-it\,H}\psi_0\mathrm{e}^{it\,H}}_{\text{Target state}}\big]$$

$$U(\boldsymbol{\theta}) = \prod_{j=1}^{M} e^{-i\theta_j \sigma_j} V_j$$

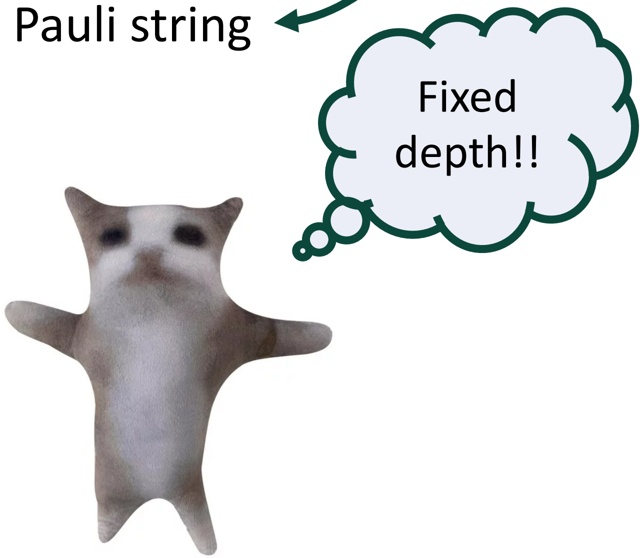


Variational Quantum Simulation

$$\mathcal{L}(\boldsymbol{\theta}) = 1 - \mathrm{Tr}\left[U(\boldsymbol{\theta})\psi_0 U^{\dagger}(\boldsymbol{\theta})\underbrace{\mathrm{e}^{-it\,H}\psi_0\mathrm{e}^{it\,H}}_{\text{Target state}}\right]$$

$$U(\boldsymbol{\theta}) = \prod_{j=1}^{M} e^{-i\theta_j \sigma_j} V_j$$

Pauli string



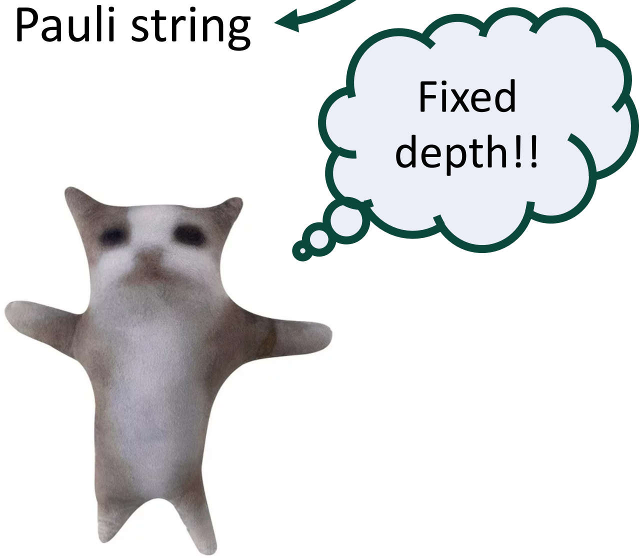
Variational Quantum Simulation

$$\mathcal{L}(\boldsymbol{\theta}) = 1 - \mathrm{Tr}\left[U(\boldsymbol{\theta})\psi_0 U^{\dagger}(\boldsymbol{\theta}) \underbrace{\mathrm{e}^{-it\,H}\psi_0 \mathrm{e}^{it\,H}}_{\text{Target state}}\right]$$

Non-parametrized

$$U(\boldsymbol{\theta}) = \prod_{j=1}^{M} e^{-i\theta_j \sigma_j} V_j$$

Pauli string



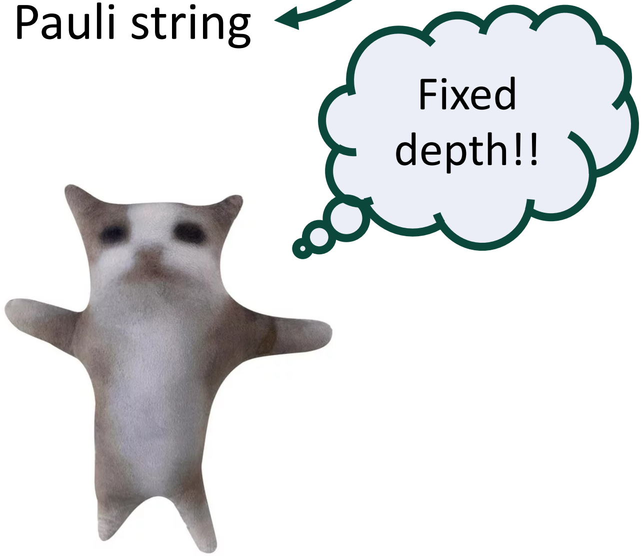
Variational Quantum Simulation

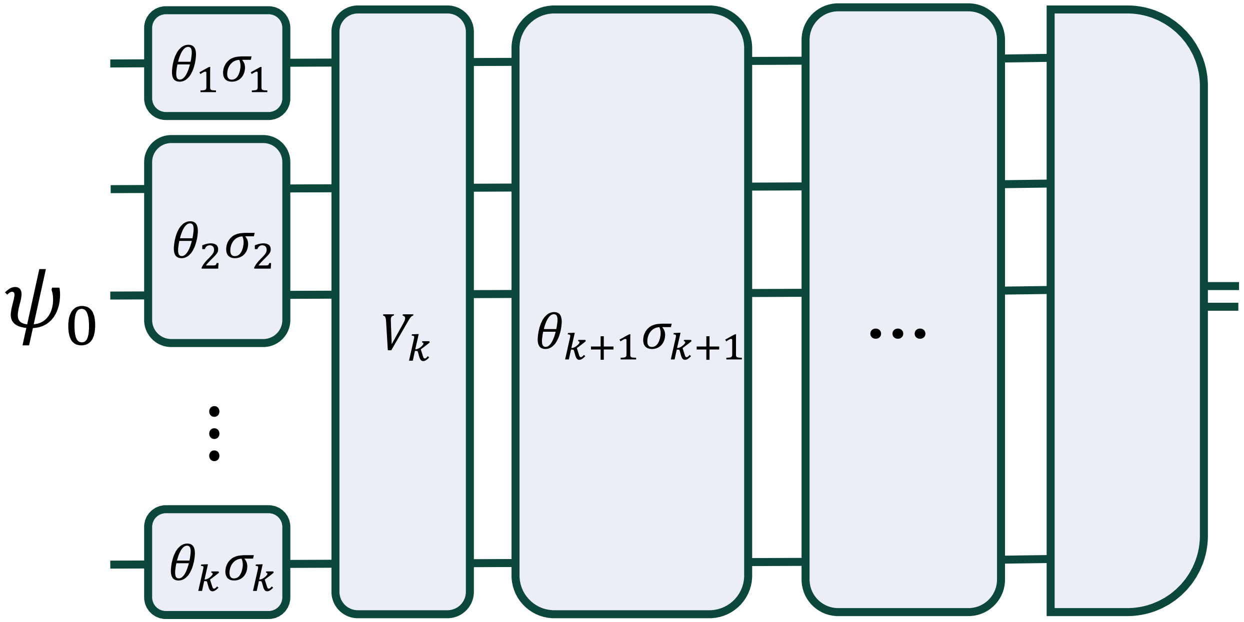
$$\mathcal{L}(\boldsymbol{\theta}) = 1 - \mathrm{Tr}\left[U(\boldsymbol{\theta})\psi_0 U^\dagger(\boldsymbol{\theta})\underbrace{\mathrm{e}^{-it\,H}\psi_0\mathrm{e}^{it\,H}}_{\text{Target state}}\right]$$

Non-parametrized

$$U(\boldsymbol{\theta}) = \prod_{j=1}^{M} e^{-i\theta_j \sigma_j} V_j$$

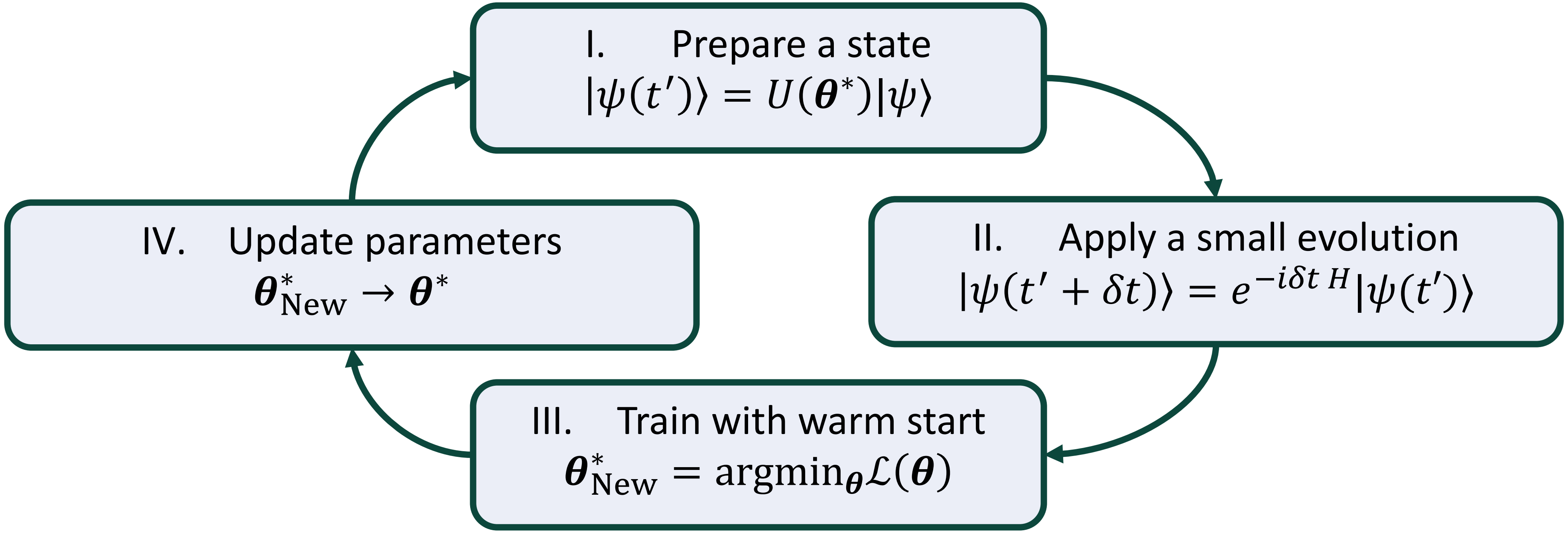
Pauli string





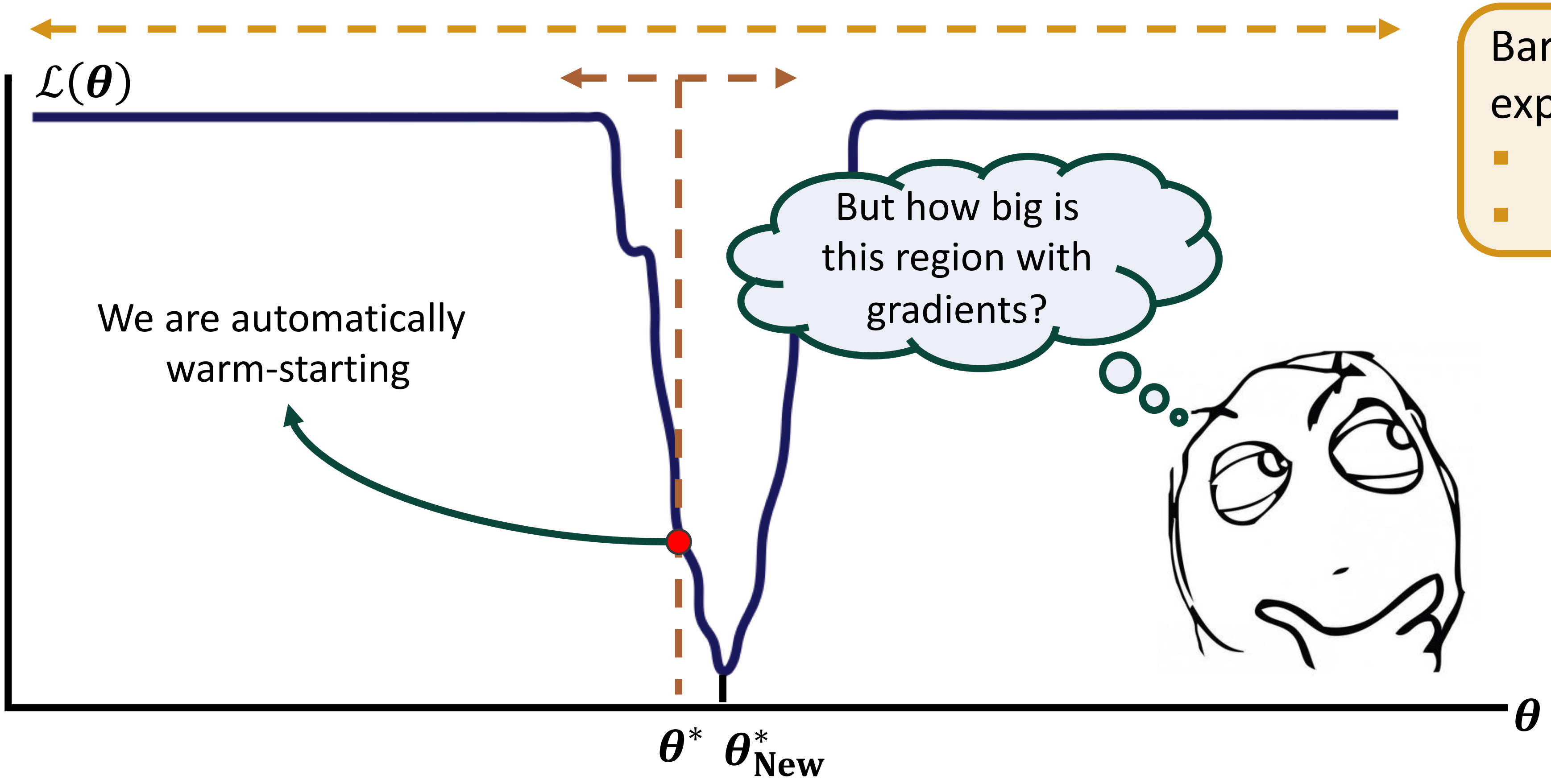
Iterative variational trotter compression

$$\mathcal{L}(\delta\boldsymbol{\theta}) = \mathrm{Tr}\left[U(\boldsymbol{\theta}^* + \delta\boldsymbol{\theta})\psi(t')U^\dagger(\boldsymbol{\theta}^* + \delta\boldsymbol{\theta})\mathrm{e}^{-i\delta t\, H}\psi(t')\mathrm{e}^{i\delta t\, H}\right]$$



So what are we trying to get to?

$$\mathcal{L}(\delta\boldsymbol{\theta}) = 1 - \mathrm{Tr}\left[U(\boldsymbol{\theta}^* + \delta\boldsymbol{\theta})\psi(t')U^\dagger(\boldsymbol{\theta}^* + \delta\boldsymbol{\theta})\mathrm{e}^{-i\delta t\, H}\psi(t')\mathrm{e}^{i\delta t\, H}\right]$$

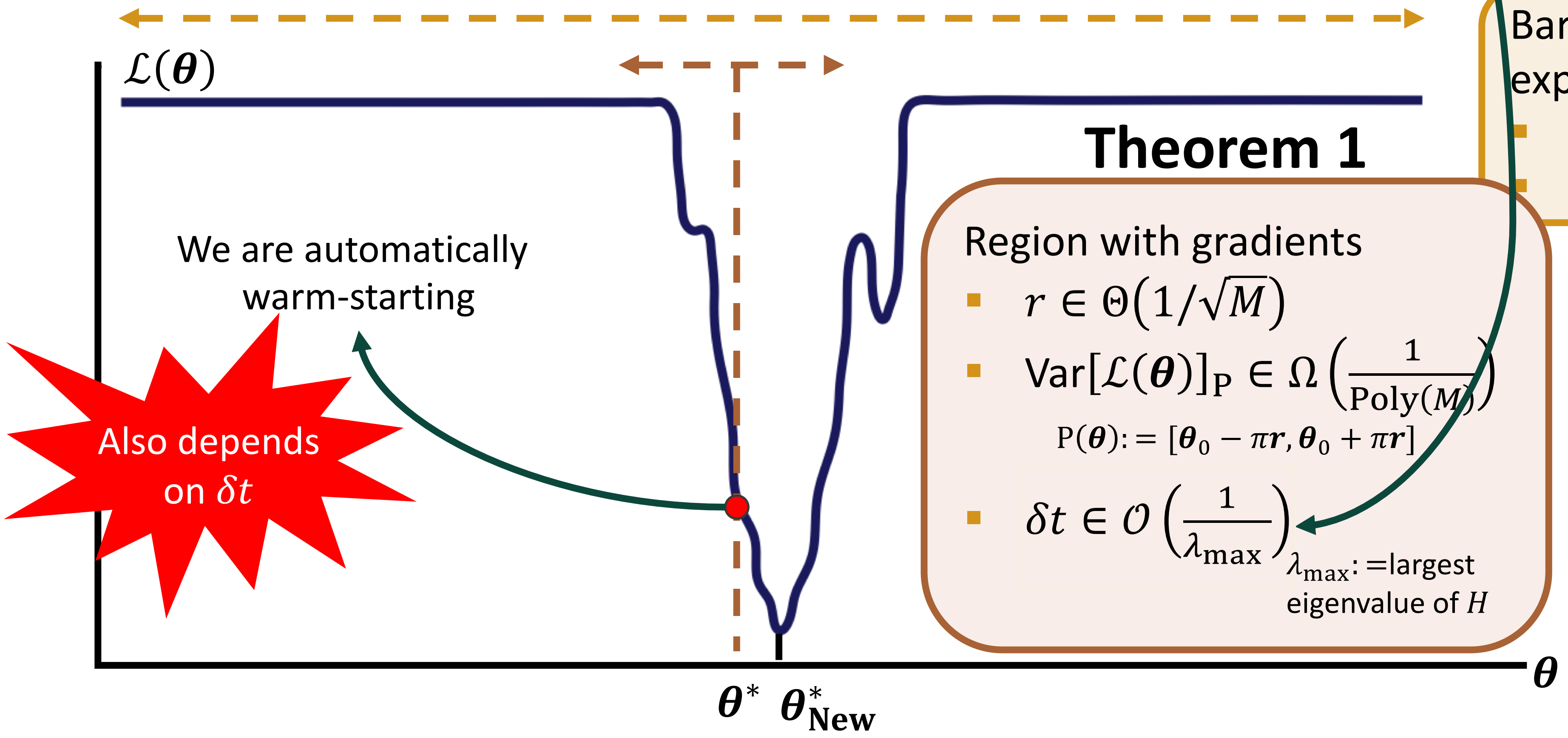


Barren Plateau: deep + expressive

- $r \in \Theta(1)$
- Var$[\mathcal{L}(\boldsymbol{\theta})]_{\mathrm{Unif}} \in \mathcal{O}(c^{-n})$

Region with gradients

$$\mathcal{L}(\delta\boldsymbol{\theta}) = 1 - \mathrm{Tr}\left[U(\boldsymbol{\theta}^* + \delta\boldsymbol{\theta})\psi(t')U^\dagger(\boldsymbol{\theta}^* + \delta\boldsymbol{\theta})\mathrm{e}^{-i\delta t\, H}\psi(t')\mathrm{e}^{i\delta t\, H}\right]$$



Theorem 1

Region with gradients

- $r \in \Theta(1/\sqrt{M})$
- $\mathrm{Var}[\mathcal{L}(\boldsymbol{\theta})]_{\mathrm{P}} \in \Omega\left(\frac{1}{\mathrm{Poly}(M)}\right)$

  $\mathrm{P}(\boldsymbol{\theta}) := [\boldsymbol{\theta}_0 - \pi \boldsymbol{r}, \boldsymbol{\theta}_0 + \pi \boldsymbol{r}]$
- $\delta t \in \mathcal{O}\left(\frac{1}{\lambda_{\max}}\right)$

  $\lambda_{\max}$: =largest eigenvalue of H

Barren Plateau : deep + expressive

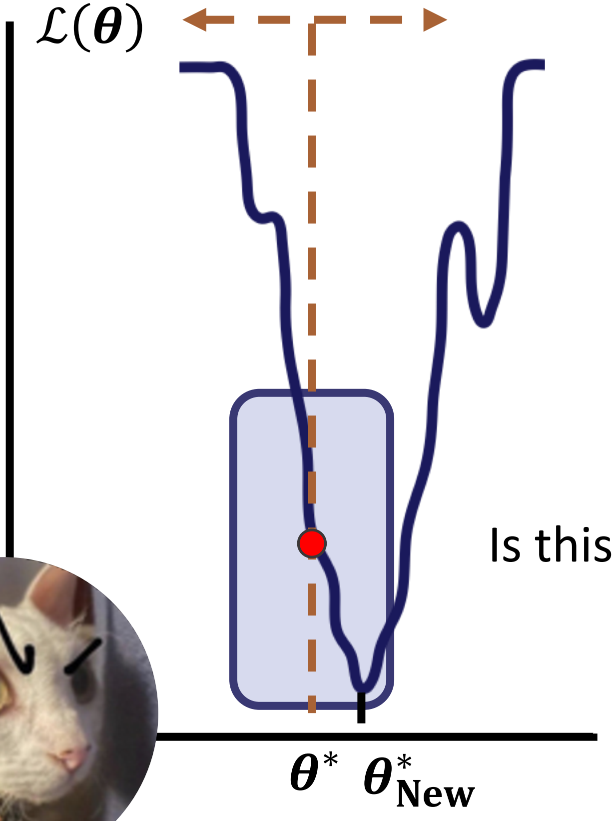
- $r \in \Theta(1)$
- $\mathrm{Var}[\mathcal{L}(\boldsymbol{\theta})]_{\mathrm{Unif}} \in \mathcal{O}(c^{-n})$

Comes from a loose bound (and does not go ''too far away'')

$$\delta t \in \mathcal{O}\left(\frac{1}{\lambda_{\max}}\right)$$

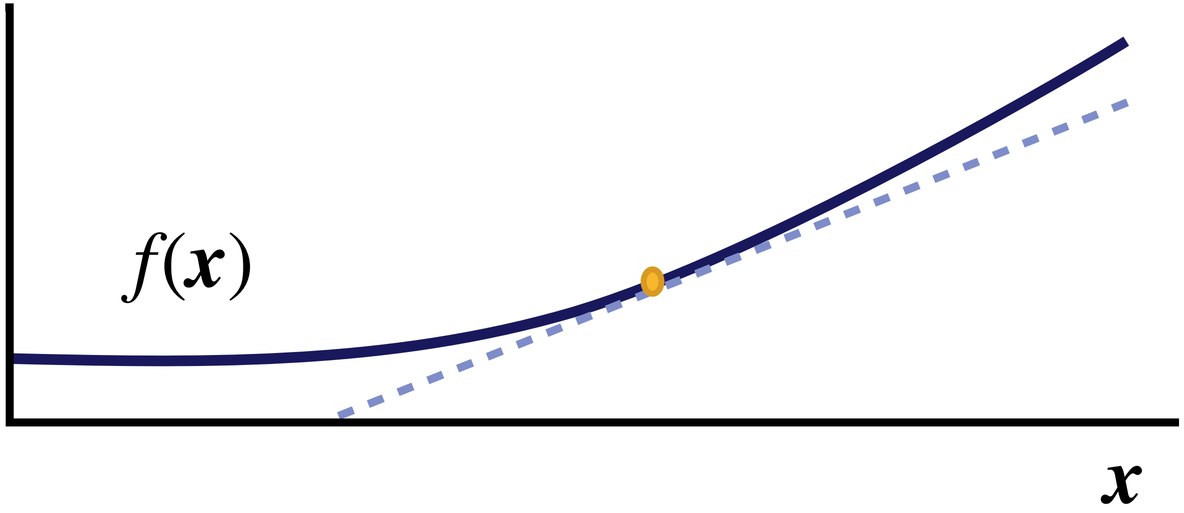
ϵ-convex region around the starting point

$$\mathcal{L}(\delta\boldsymbol{\theta}) = 1 - \mathrm{Tr}\left[U(\boldsymbol{\theta}^* + \delta\boldsymbol{\theta})\psi(t')U^\dagger(\boldsymbol{\theta}^* + \delta\boldsymbol{\theta})\mathrm{e}^{-i\delta t\,H}\psi(t')\mathrm{e}^{i\delta t\,H}\right]$$



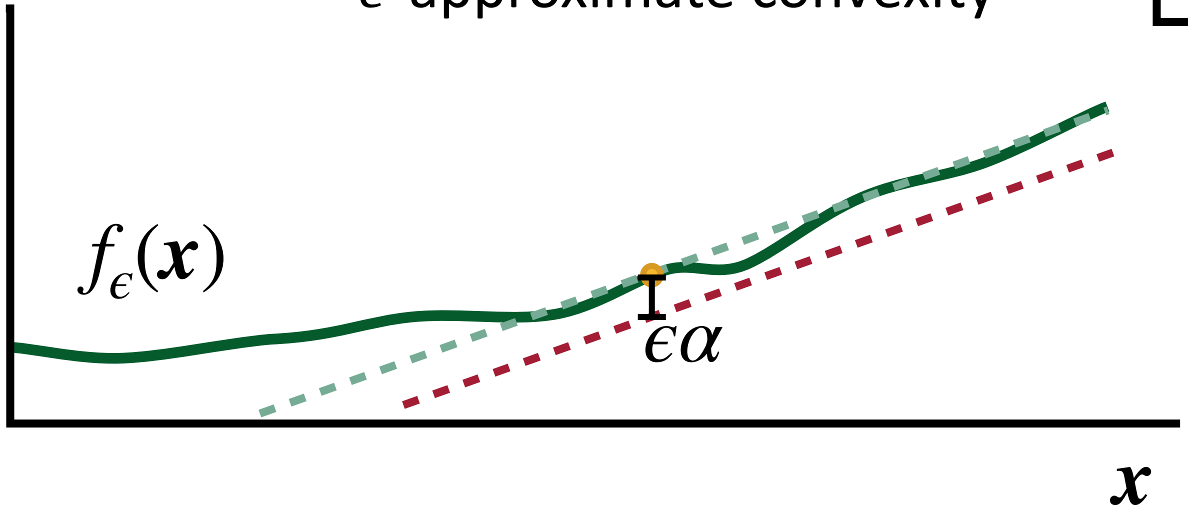
Is this convex?

ϵ-approximate convexity

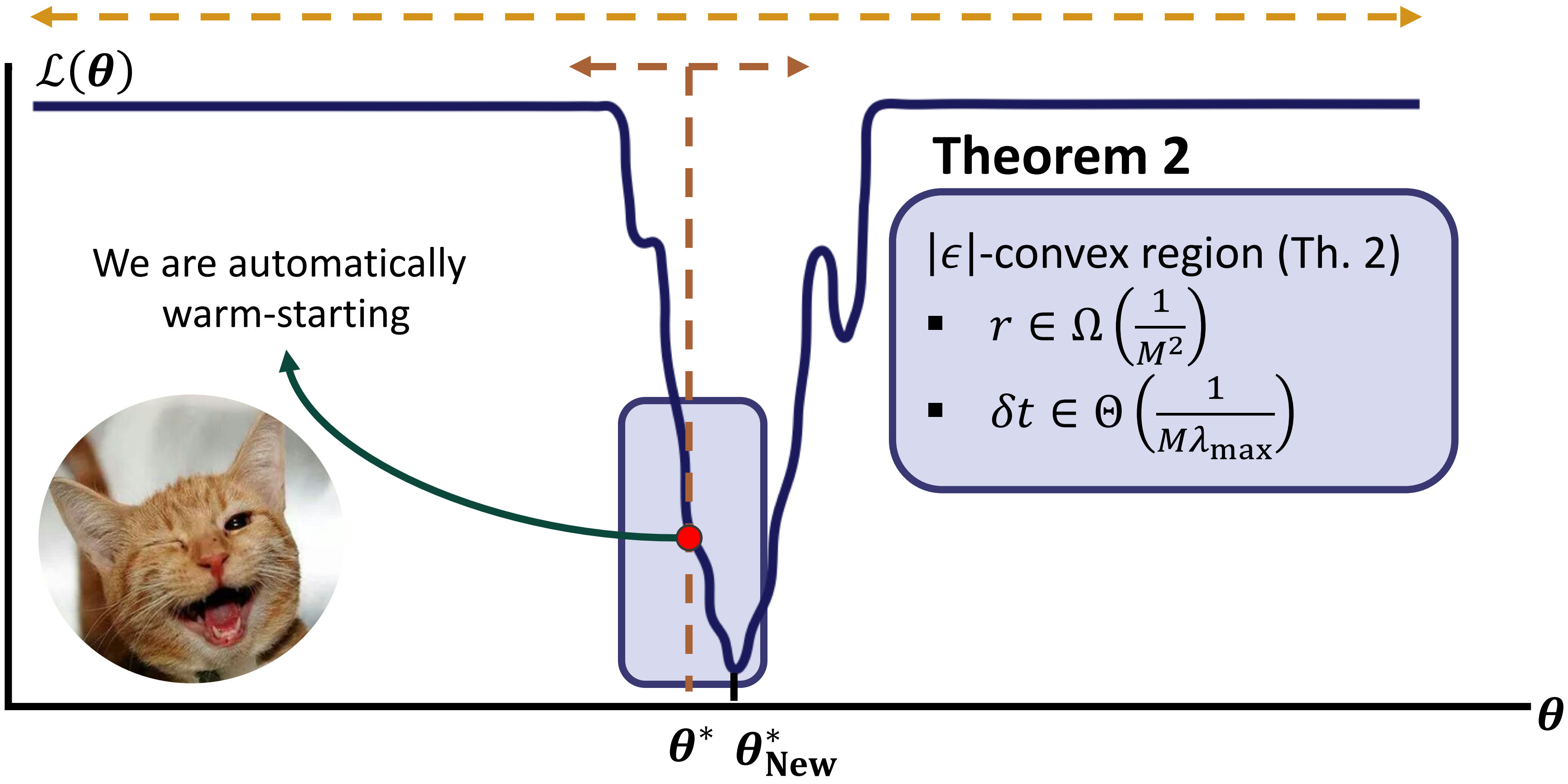


convexity

ϵ-approximate convexity



ϵ-convex region around the starting point



Theorem 2

$|\epsilon|$-convex region (Th. 2)

- $r \in \Omega\left(\frac{1}{M^2}\right)$
- $\delta t \in \Theta\left(\frac{1}{M\lambda_{\max}}\right)$

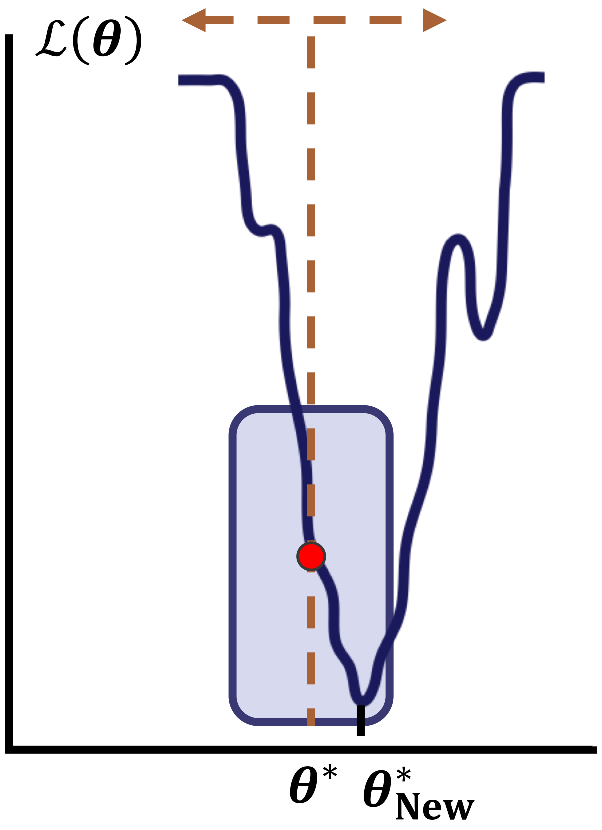
Barren Plateau

- $r \in \Theta(1)$

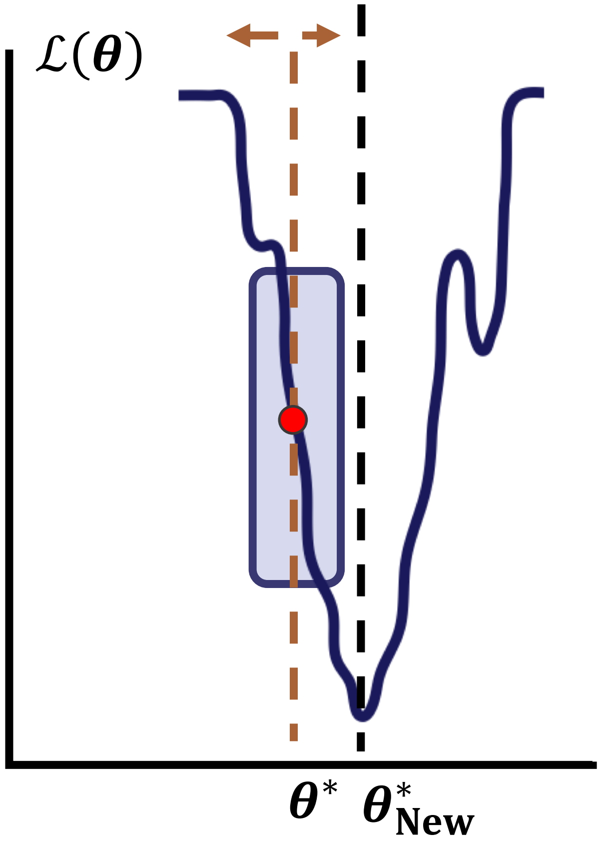
Gradient region (Th. 1)

- $r \in \Theta\left(\frac{1}{\sqrt{M}}\right)$
- $\delta t \in \mathcal{O}\left(\frac{1}{\lambda_{\max}}\right)$

Adiabatic minima



or

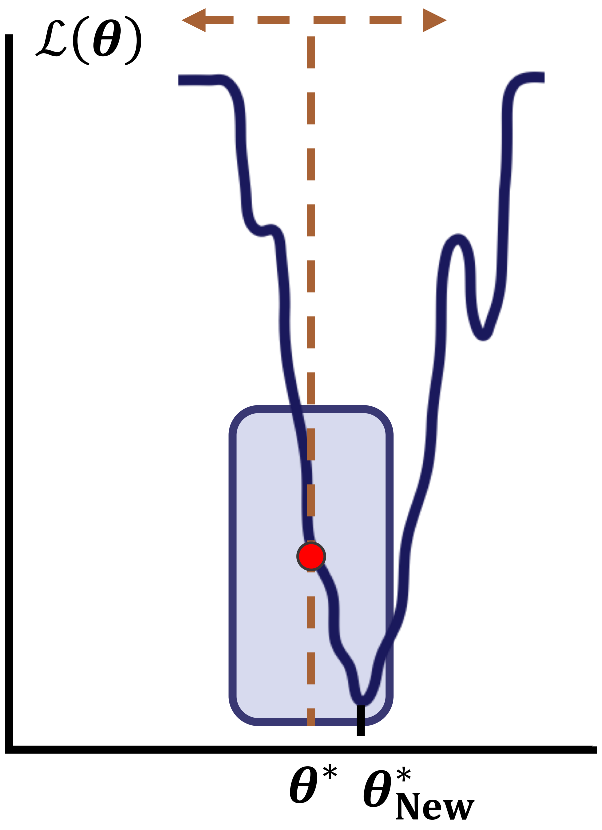


Barren Plateau
- $r \in \Theta(1)$

Gradient region (Th. 1)
- $r \in \Theta\left(\frac{1}{\sqrt{M}}\right)$
- $\delta t \in \mathcal{O}\left(\frac{1}{\lambda_{\max}}\right)$

$|\epsilon|$-convex region (Th. 2)
- $\delta t \in \Theta\left(\frac{1}{M\lambda_{\max}}\right)$
- $r \in \Omega\left(\frac{1}{M^2}\right)$

Adiabatic minima



Theorem 3: to ensure the **adiabatic minima** is in the region we need

- Theorem 1: $\delta t \in \mathcal{O}\left(\frac{1}{M\lambda_{\max}}\right)$
- Theorem2: $\delta t \in \mathcal{O}\left(\frac{1}{M^{5/2}\lambda_{\max}}\right)$

Barren Plateau

- $r \in \Theta(1)$

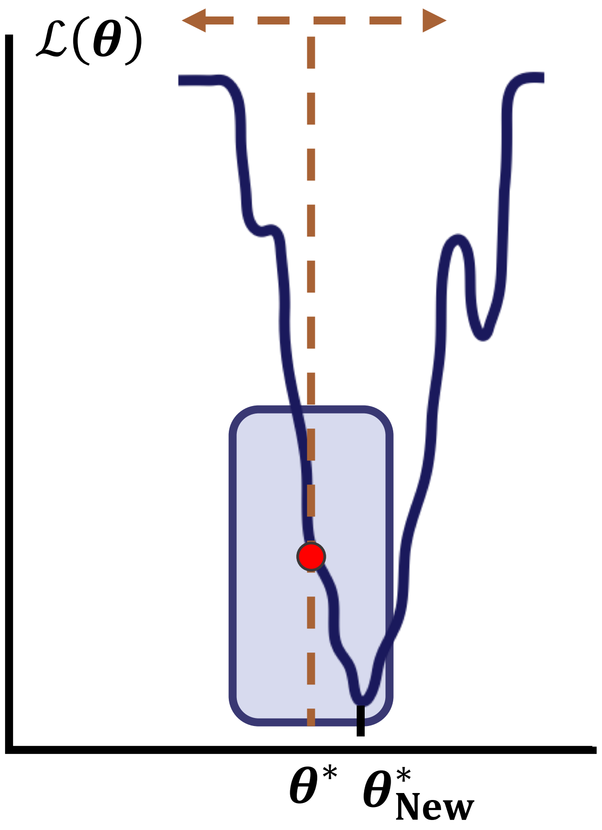
Gradient region (Th. 1)

- $r \in \Theta\left(\frac{1}{\sqrt{M}}\right)$
- $\delta t \in \mathcal{O}\left(\frac{1}{\lambda_{\max}}\right)$

$|\epsilon|$-convex region (Th. 2)

- $\delta t \in \Theta\left(\frac{1}{M\lambda_{\max}}\right)$
- $r \in \Omega\left(\frac{1}{M^2}\right)$

What does this mean in practice?



Gradients (Th. 1): poly scaling in r and δt to get poly variances

ϵ-convexity (Th. 2): poly scaling in r and δt to get poly variances

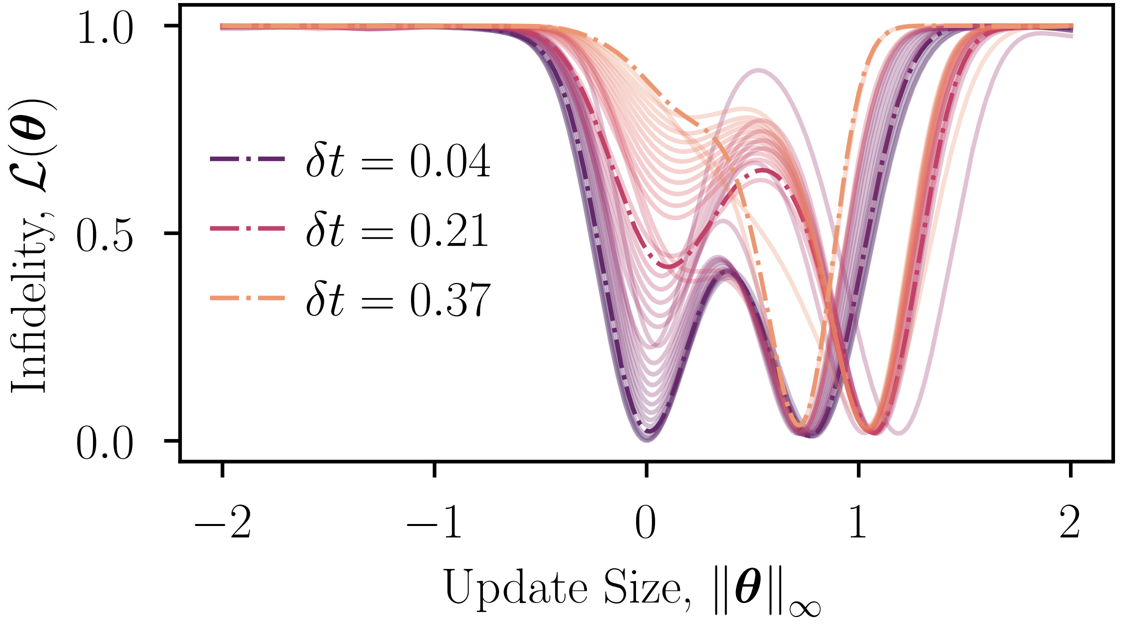
To ensure we are inside the minima (Th. 3)

- Theorem 1: poly scaling in δt
- Theorem2: poly scaling in δt

However, might still be too small in practice as the variance is roughly $\frac{1}{M^2}$

Minima can jump

The good (or just better) minimum, can have a finite jump for an infinitely small δt



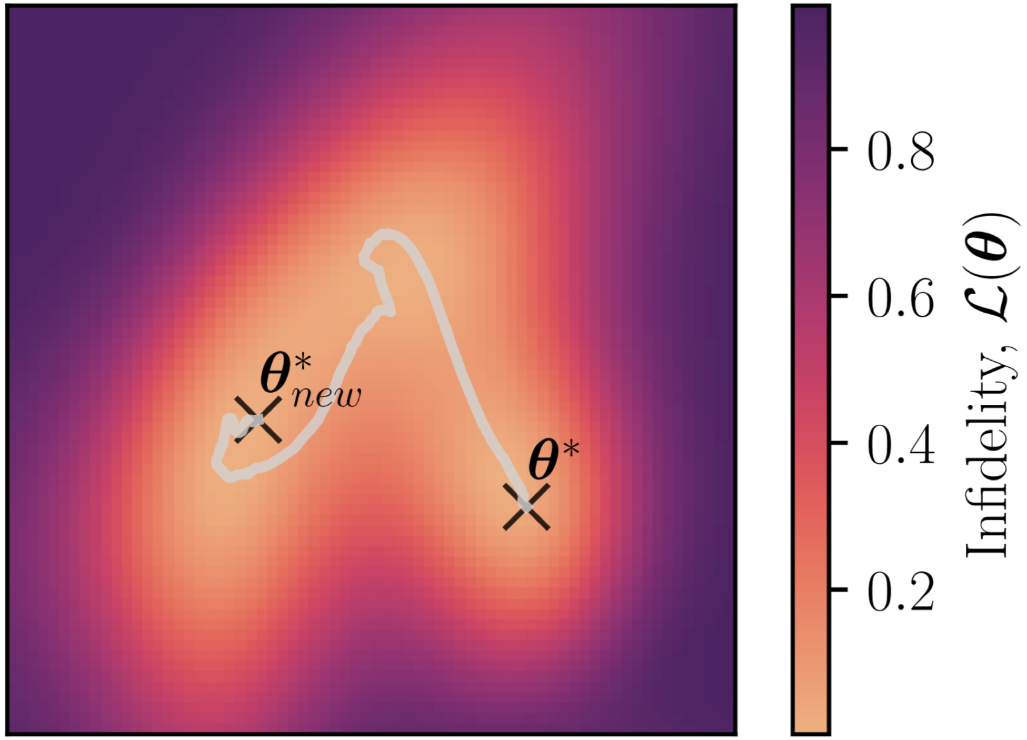
1-D cut. 10 qubit Ising Hamiltonian $H = \sum X_i X_{i+1} - 0.95 \sum Y_i$

We use a 2-layered Hamiltonian Variational Ansatz.

Fertile valleys?

10 qubit Ising Hamiltonian $H = \sum X_i X_{i+1} - 0.95 \sum Y_i$

We use a 2-layered Hamiltonian Variational Ansatz.



Plotted with ORQVIZ
M. S. Rudolph et al, arXiv:2111.04695 (2021).

Outlook on other iterative approaches

Reinterpret the results!

Imaginary time evolution $e^{i\delta t H} \rightarrow e^{\delta \tau H}$

Learning an unknown unitary (independent of the state) $\left|\mathrm{Tr}\left[U(\theta)V^{\dagger}\right]\right|^{2}$

C. Cirstoiu et. al., npj Quantum Information 6, 1 (2020)

Outlook on other iterative approaches

Reinterpret the results!

Imaginary time evolution $e^{i\delta t} \rightarrow e^{\delta\tau}$

Learning an unknown unitary (independent of the state) $\left|\mathbf{Tr}\left[U(\theta)V^{\dagger}\right]\right|$

C. Cirstoiu et. al., npj Quantum Information 6, 1 (2020)

FIDELITY TYPE LOSS

$$\langle\psi(\theta)|\psi_{\mathrm{Target}}\rangle \geq \frac{1}{2}$$

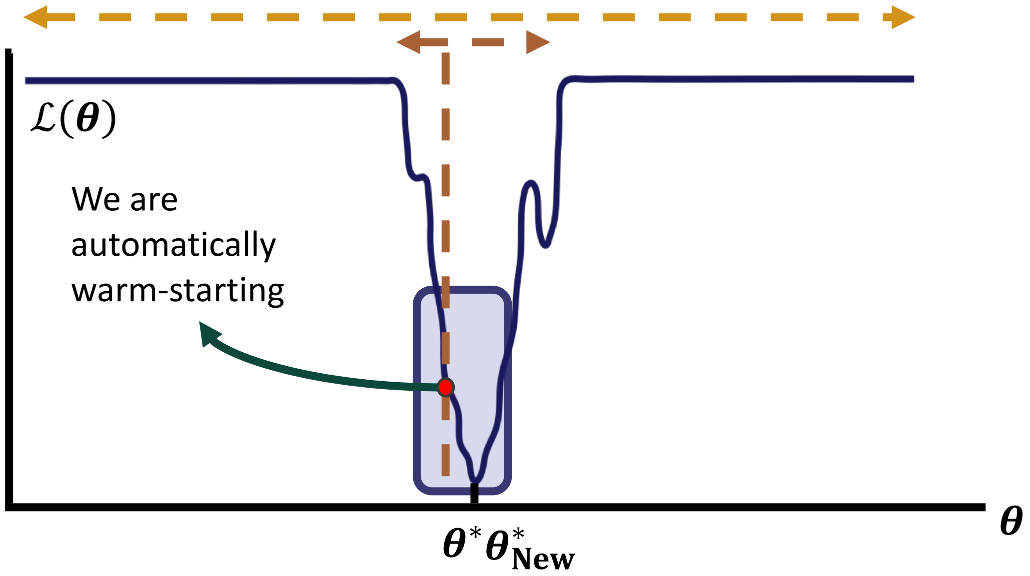
To sum up

Gradients: poly scaling in r and δt to get poly variances

ϵ-convexity: poly scaling in r and δt to get poly variances

To ensure we are inside the minima

- Theorem 1: poly scaling in δt
- Theorem2: poly scaling in δt



Still difficult in practice

Poly scaling: $=\sim \frac{1}{\mathrm{Poly}(M)}$

To sum up

Gradients: poly scaling in r and δt to get poly variances

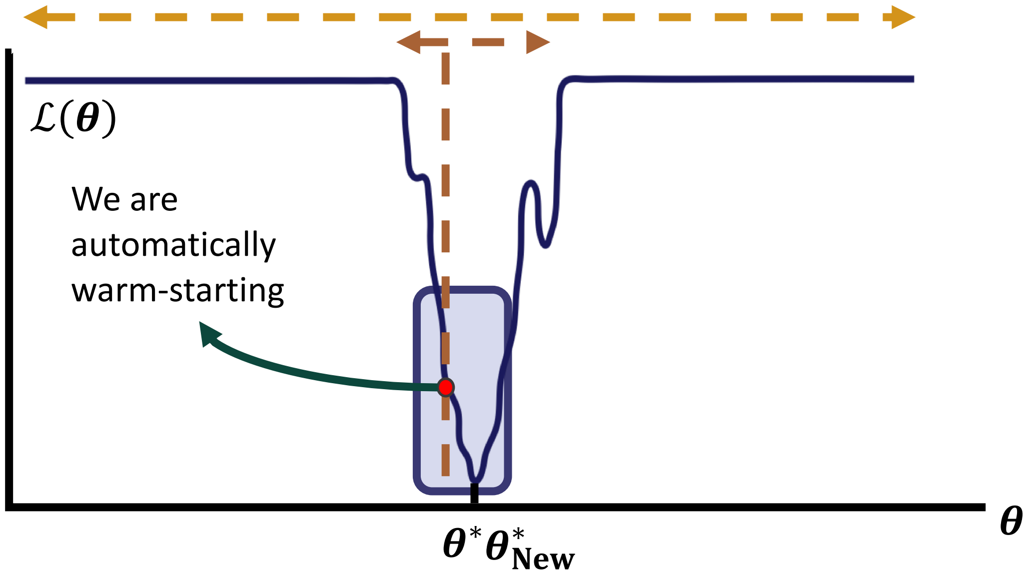
ϵ-convexity: poly scaling in r and δt to get poly variances

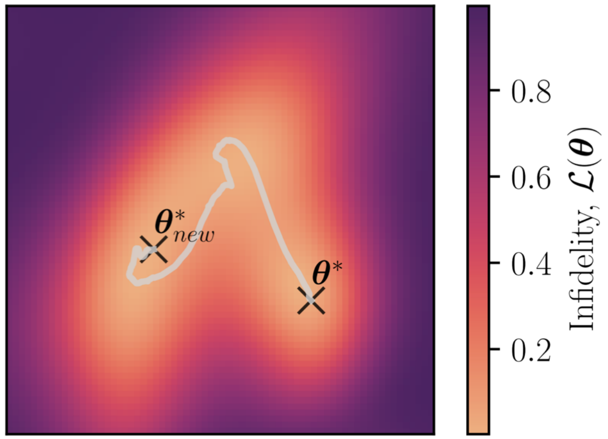
To ensure we are inside the minima

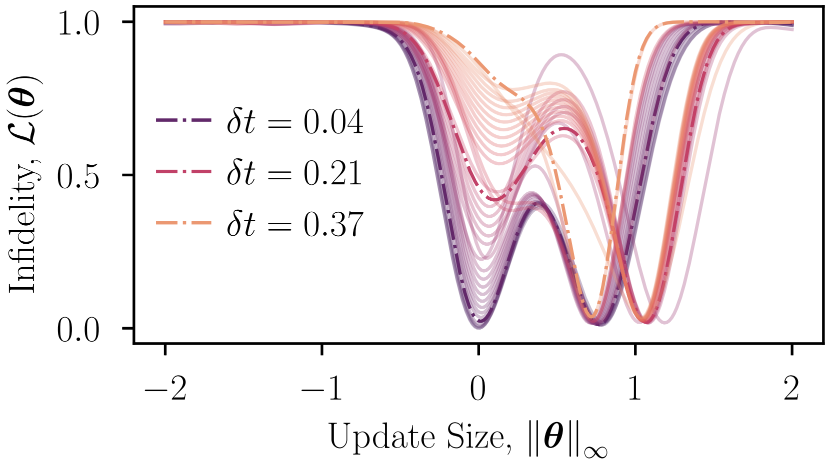
- Theorem 1: poly scaling in δt
- Theorem2: poly scaling in δt

Still difficult in practice

Poly scaling: $=\sim \frac{1}{\text{Poly}(M)}$







Further analysis

- Beyond conventional average.
- This is still an average case analysis (Jumps and valleys).
- Most results here carry over to (iterative) variational approaches.
- Generalize for arbitrary generators and correlated angles.
- Surrogability?