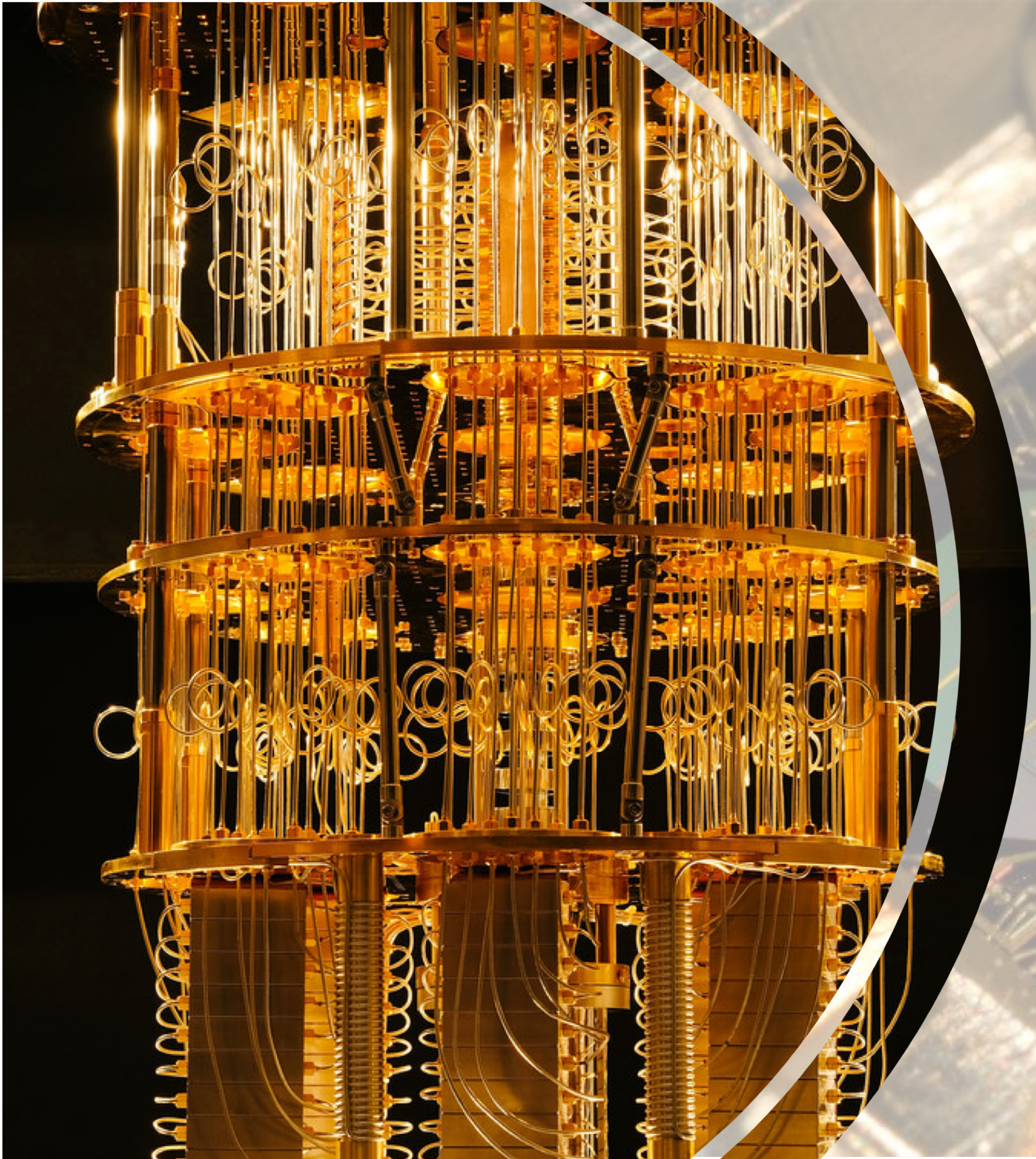


Guarantees and limitations for warm starts and iterative methods in variational quantum computing

Ricard Puig

Variational quantum simulation: a case study for understanding warm starts. **R Puig***, M Drudis*, et. al. arXiv:2404.10044

A unifying account of warm start guarantees for patches of quantum landscapes. H. Mhiri*, **R. Puig***, et. al. arXiv: 2501.xxxxx (in prep)

EPFL - Sorbonne

EPFL - IBM

EPFL

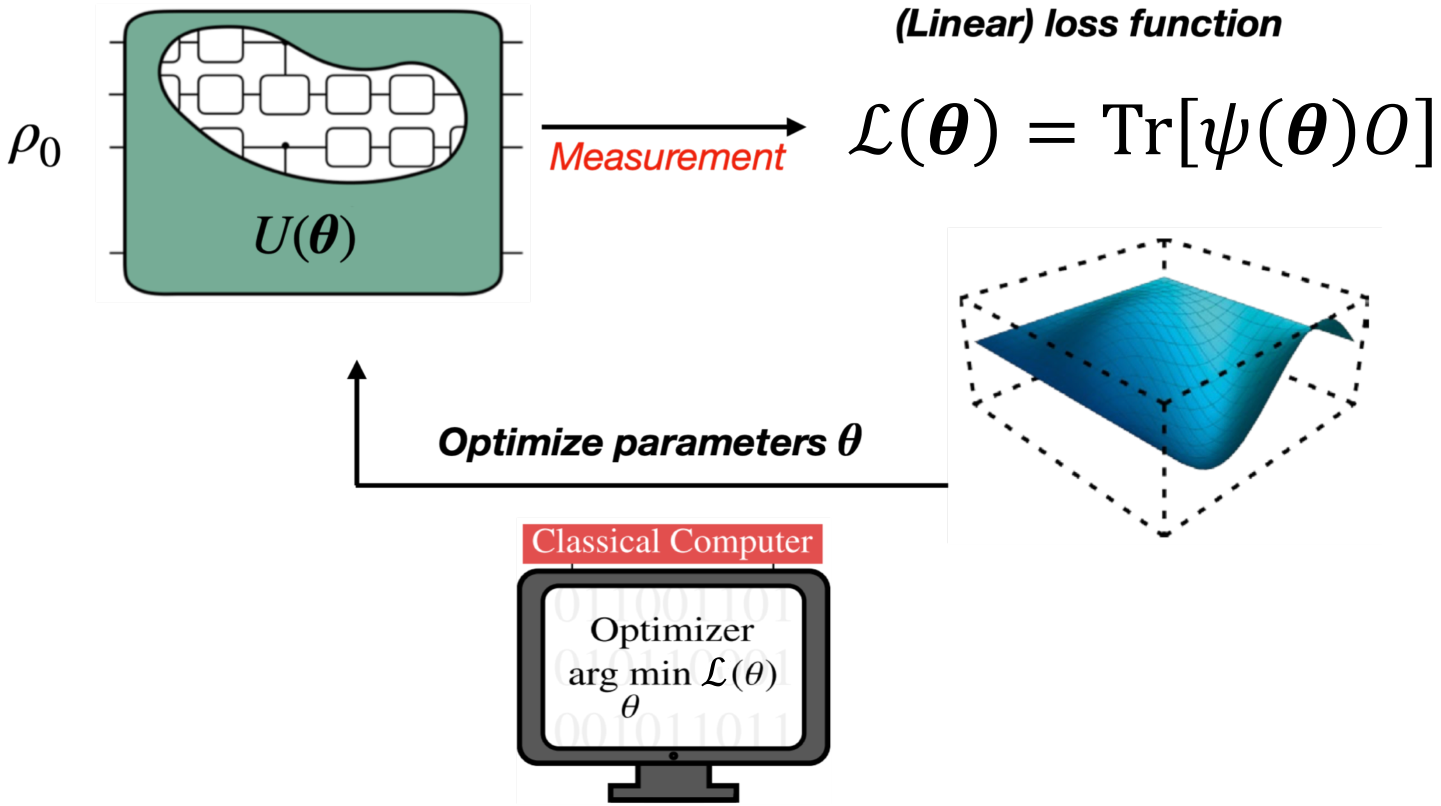
EPFL

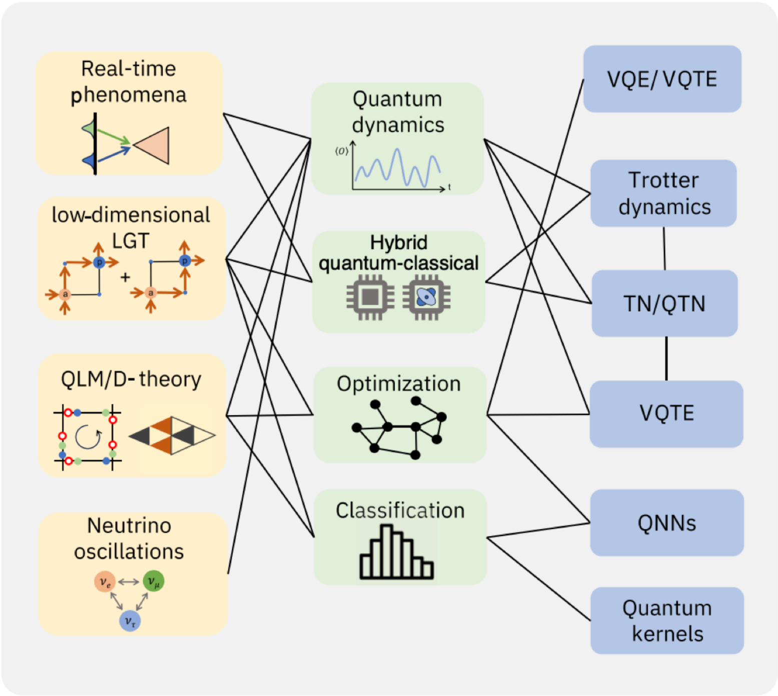
Chulalongkorn

EPFL - Chulalongkorn

EPFL

Variational Quantum Algorithms





Alberto Di Meglio, et al. "Quantum Computing for High-Energy Physics: State of the Art and Challenges." PRX Quantum **5**, 037001

Barren plateau phenomena

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

$+$

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

↓

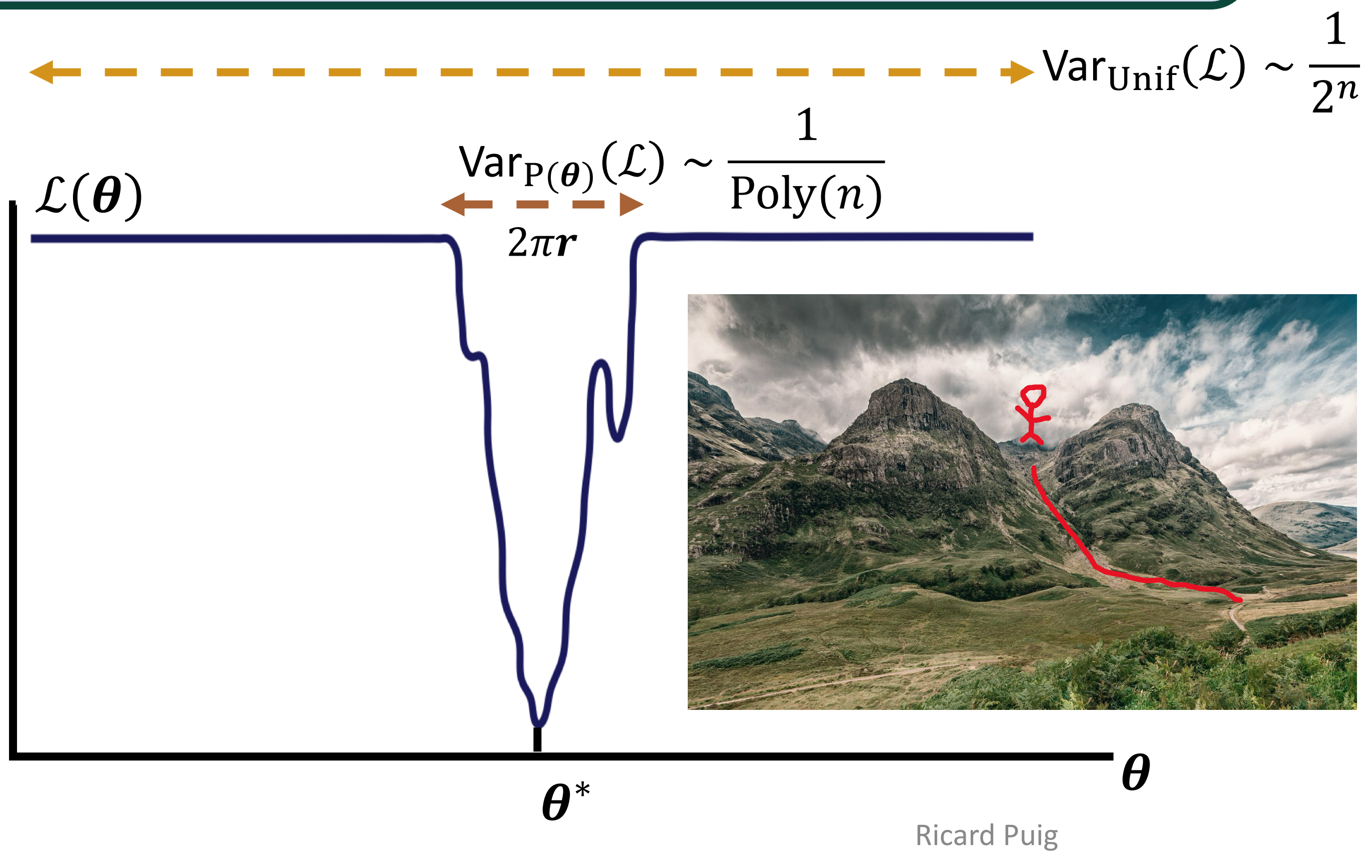
Probability of non-zero gradients vanishes exponentially with problem size.

↓

Shot required for training grows exponentially with problem size.

Average statement!

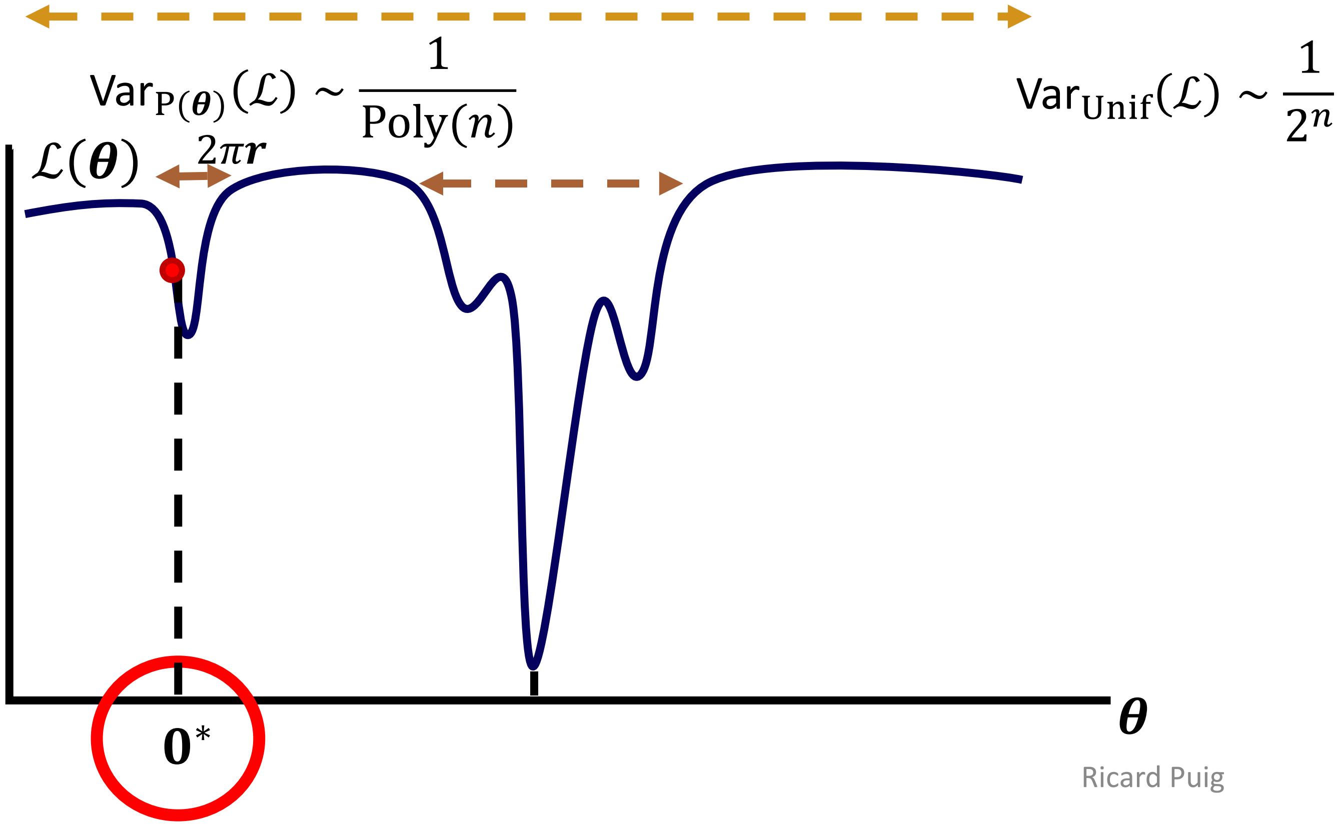
But what if we look around a region with curvature?



Identity initialization

$\text{Var}_{\text{Unif}}(\mathcal{L}) \sim \frac{1}{2^n}$ —— They get ——→ $\text{Var}_{\text{P}(\boldsymbol{\theta})}(\mathcal{L}) \sim \frac{1}{\text{Poly}(n)}$

$\text{Unif} \rightarrow \boldsymbol{\theta} \in [-\pi \boldsymbol{r}, +\pi \boldsymbol{r}]$



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

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

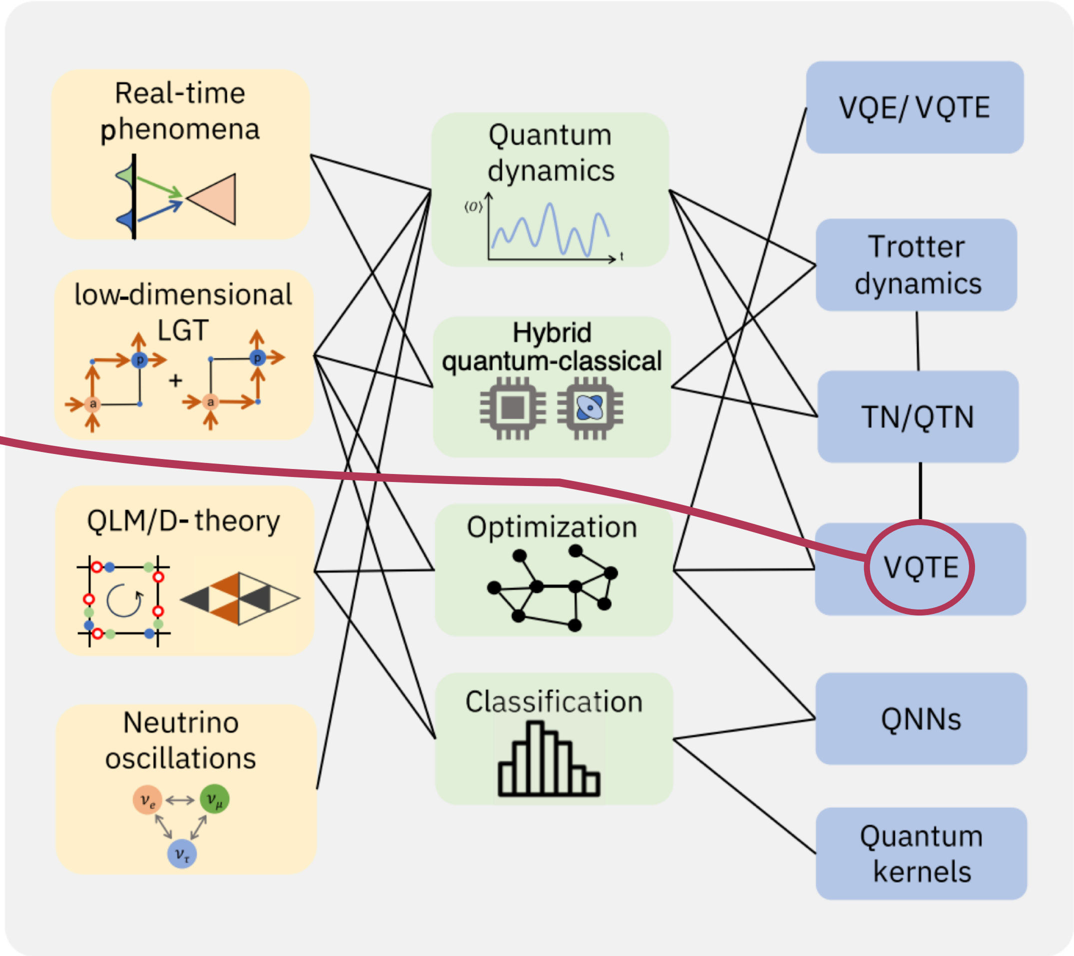
Chae-Yeun Park,[1] Minhyeok Kang,[2,3,4] and Joonsuk Huh[1,2,3,4]

[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)

Understanding warm starts and regions of attraction

Variational quantum simulation: a case study for understanding warm starts. **R Puig***, M Drudis*, et. al. arXiv:2404.10044

Understand and study warm-starts via a case example.
Iterative simulation of time dynamics



Alberto Di Meglio, et al. "Quantum Computing for High-Energy Physics: State of the Art and Challenges." PRX Quantum **5**, 037001

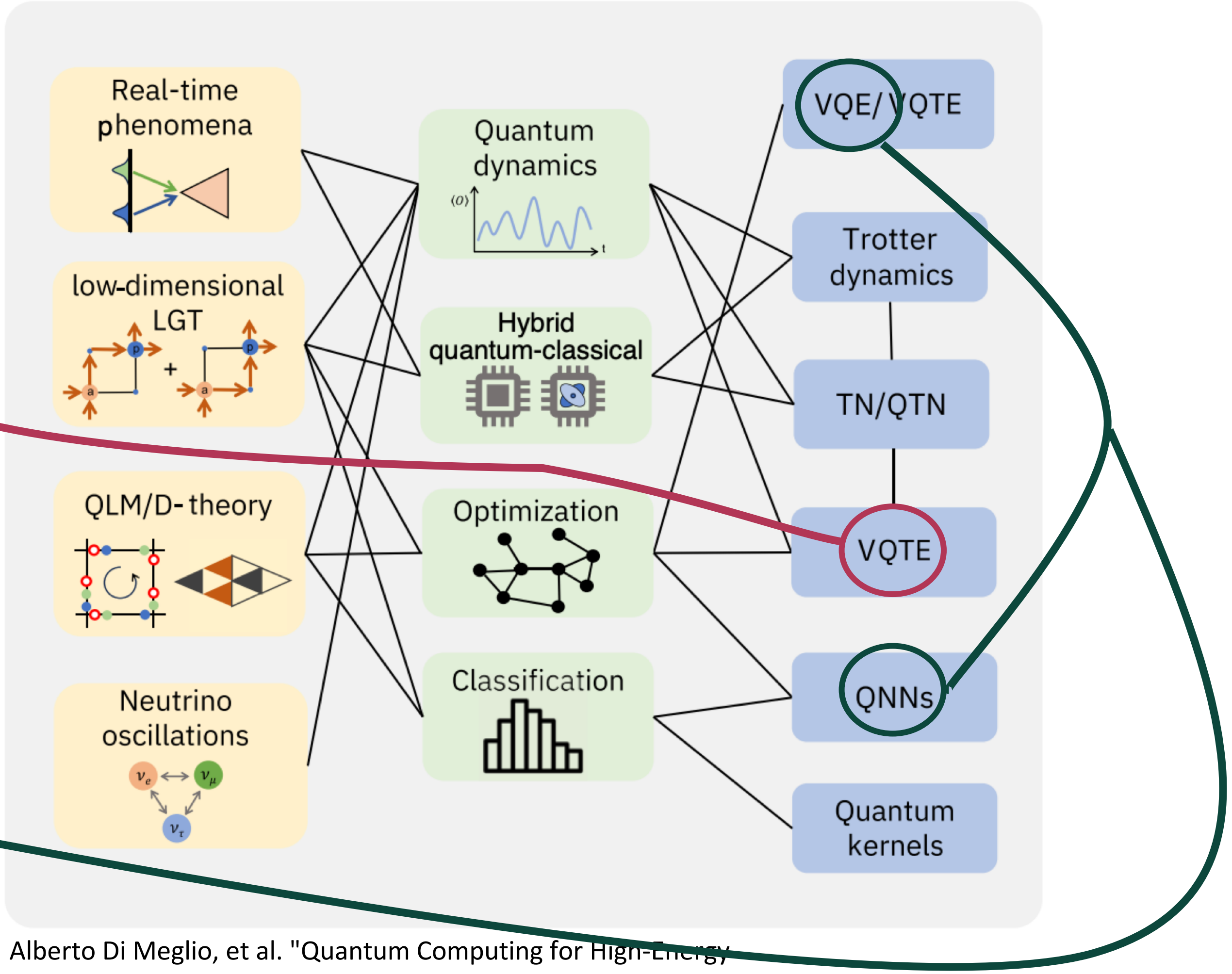
Understand warm starts and regions of attraction

Variational quantum simulation: a case study for understanding warm starts. **R Puig***, M Drudis*, et. al. arXiv:2404.10044

Understand and study warm-starts via a case example.
Iterative simulation of time dynamics

A unifying account of warm start guarantees for patches of quantum landscapes. H. Mhiri*, **R. Puig***, et. al. arXiv: 2501.xxxxx (in prep)

General bound to study patches of gradients.

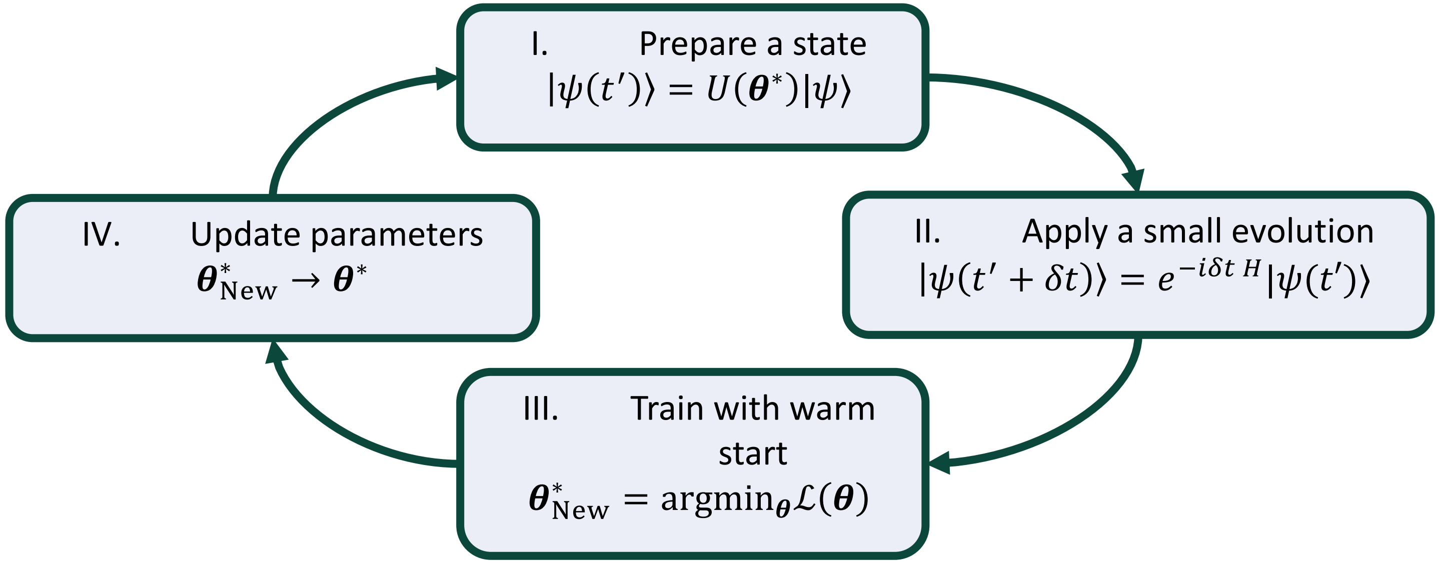


Alberto Di Meglio, et al. "Quantum Computing for High-Energy Physics: State of the Art and Challenges." PRX Quantum **5**, 037001

Variational quantum simulation

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

$U(\boldsymbol{\theta})$ →Pauli rotations and non parametrized gates



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

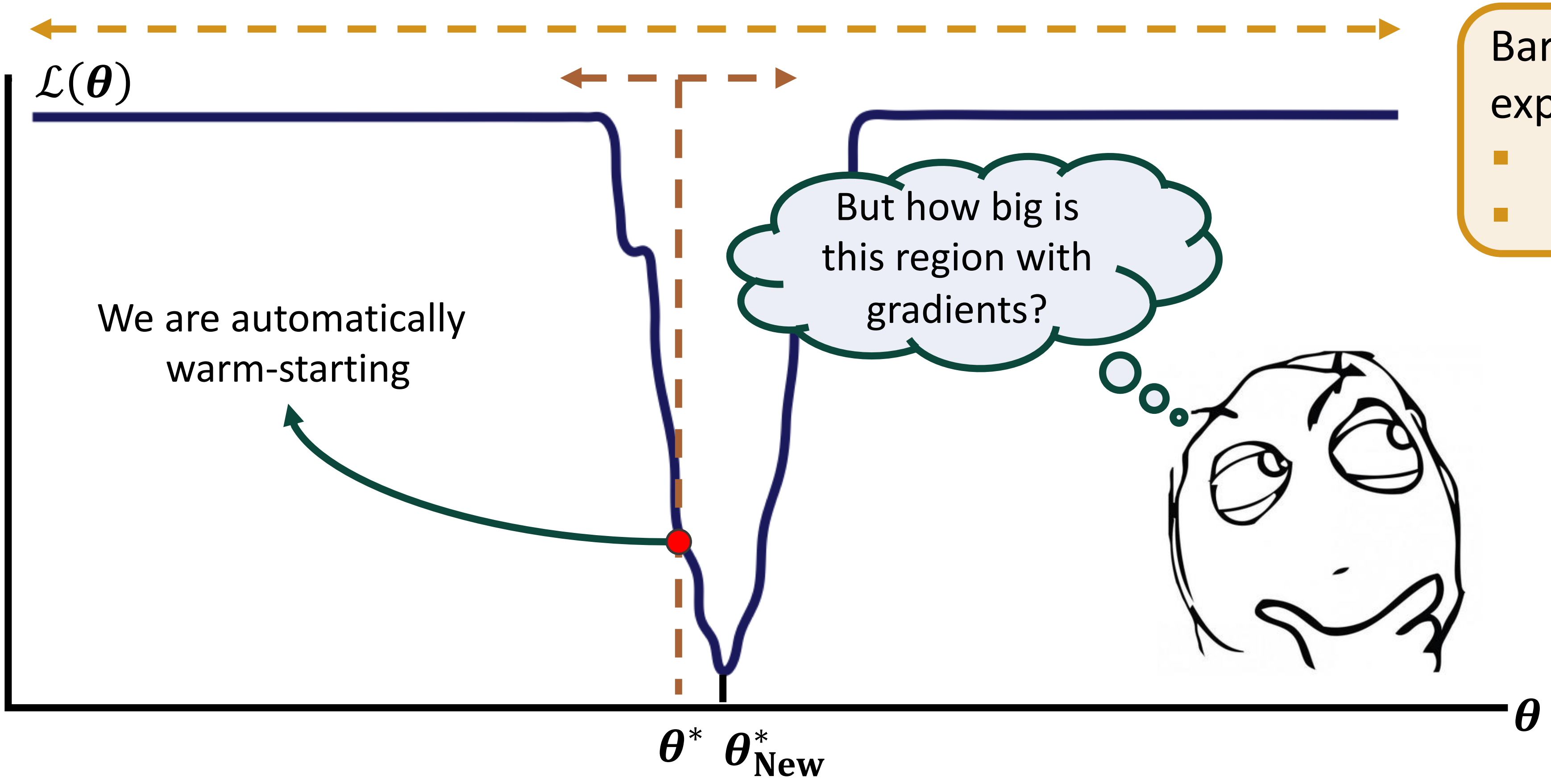
Quantum dynamics simulations beyond the coherence time on noisy intermediate-scale quantum hardware by variational Trotter compression

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: a case study for understanding warm starts. **R Puig***, M Drudis*, et. al. arXiv:2404.10044

So what are we trying to get to?

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



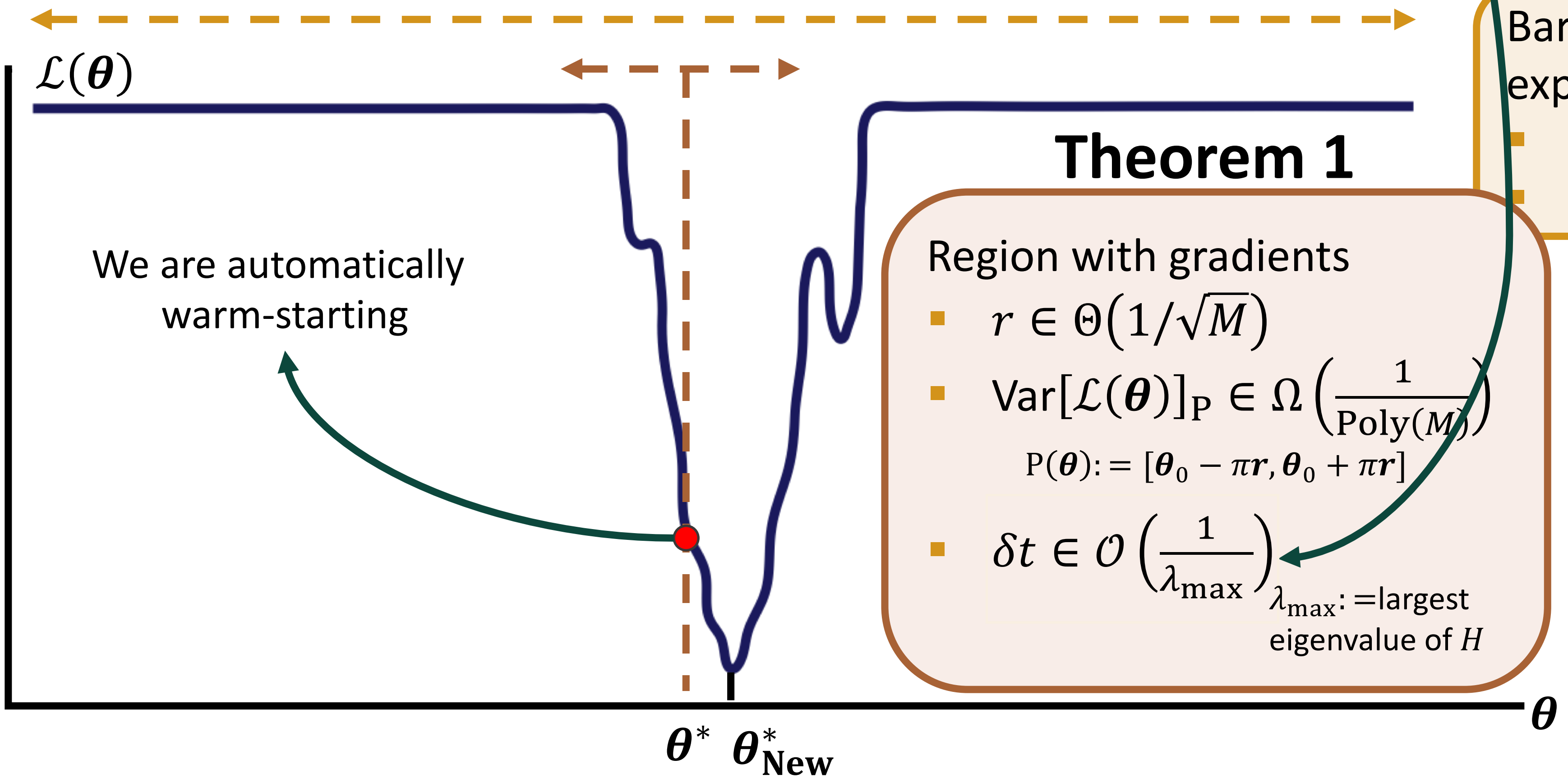
Barren Plateau: deep + expressive

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

Variational quantum simulation: a case study for understanding warm starts. **R Puig***, M Drudis*, et. al. arXiv:2404.10044

Region with gradients

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



Barren Plateau : deep + expressive

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

Theorem 1

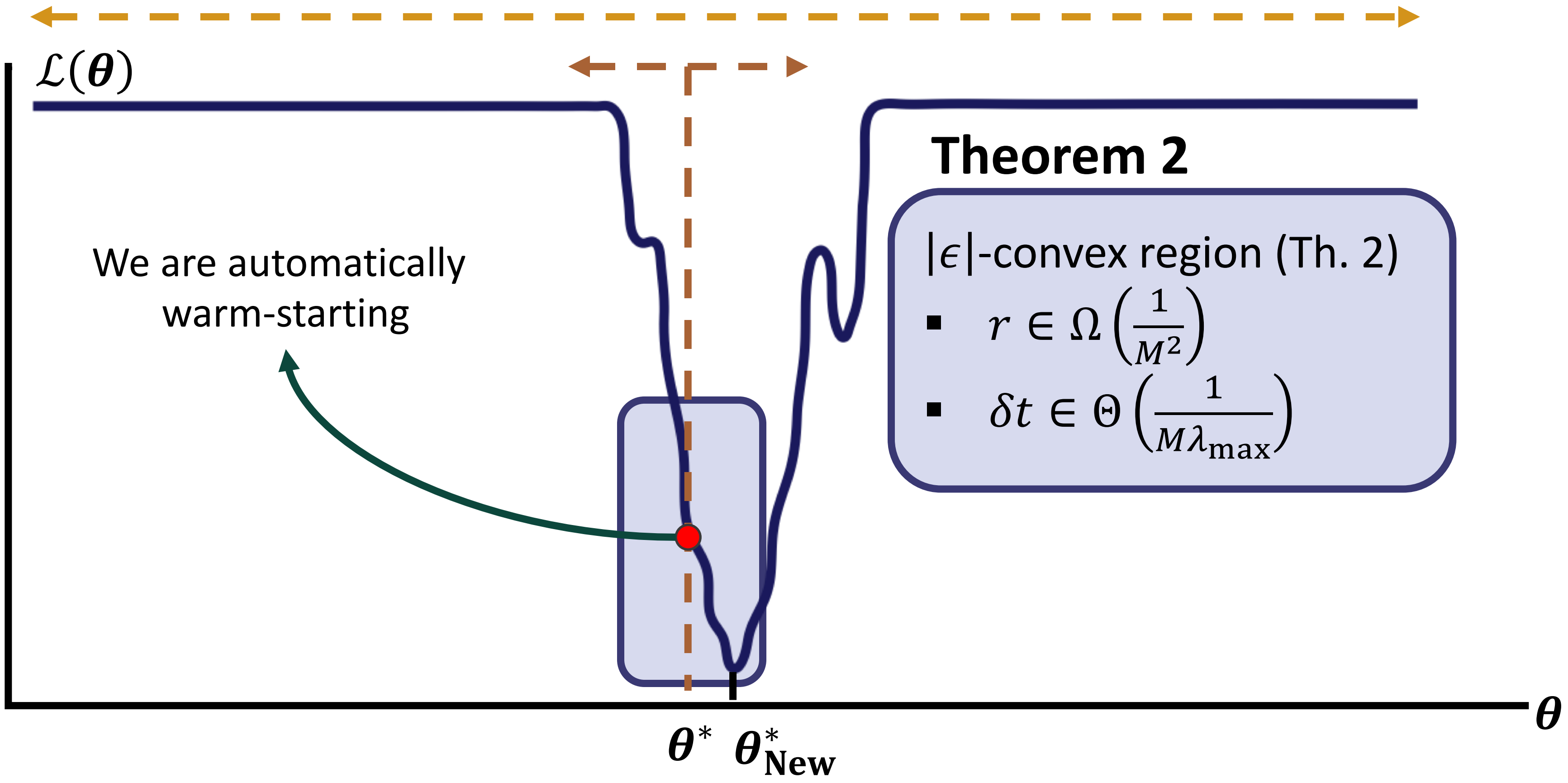
Region with gradients

- $r \in \Theta\left(1/\sqrt{M}\right)$
- 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

ϵ-convex region around the starting point

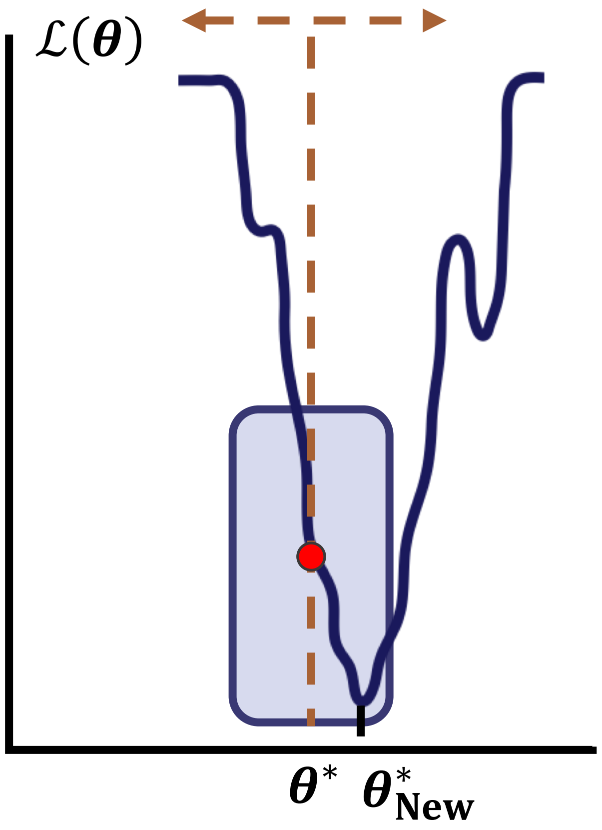


Barren Plateau
- $r \in \Theta(1)$

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

Variational quantum simulation: a case study for understanding warm starts. **R Puig***, M Drudis*, et. al. arXiv:2404.10044

What does this mean in practice?



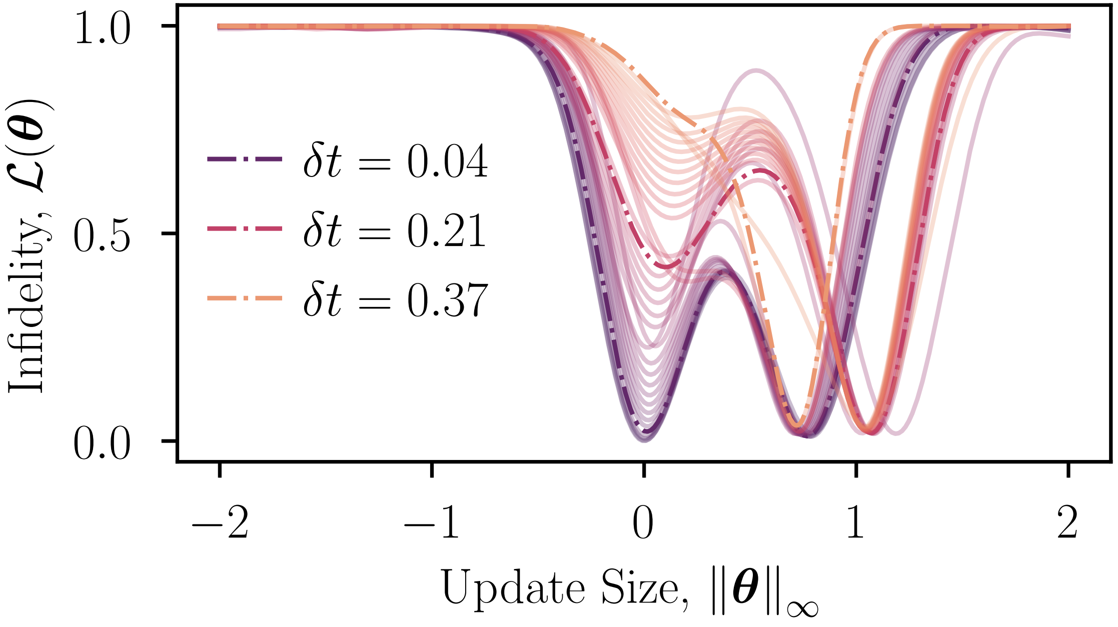
Gradients (Th. 1): poly scaling in r and δt to get poly variances

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

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

Variational quantum simulation: a case study for understanding warm starts. **R Puig***, M Drudis*, et. al. arXiv:2404.10044

Minima can jump

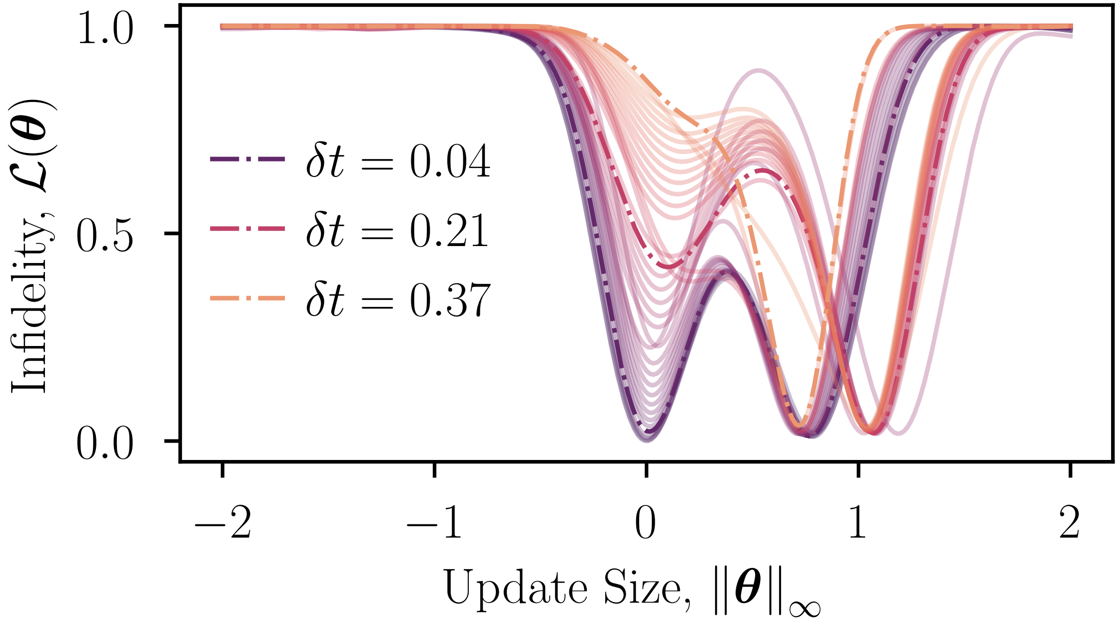


1-D cut. 10 qubit Ising Hamiltonian $H = \sum X_i X_{i+1} - 0.95 \sum Y_i$

We use a 2-layered Hamiltonian Variational Ansatz.

Variational quantum simulation: a case study for understanding warm starts. **R Puig***, M Drudis*, et. al. arXiv:2404.10044

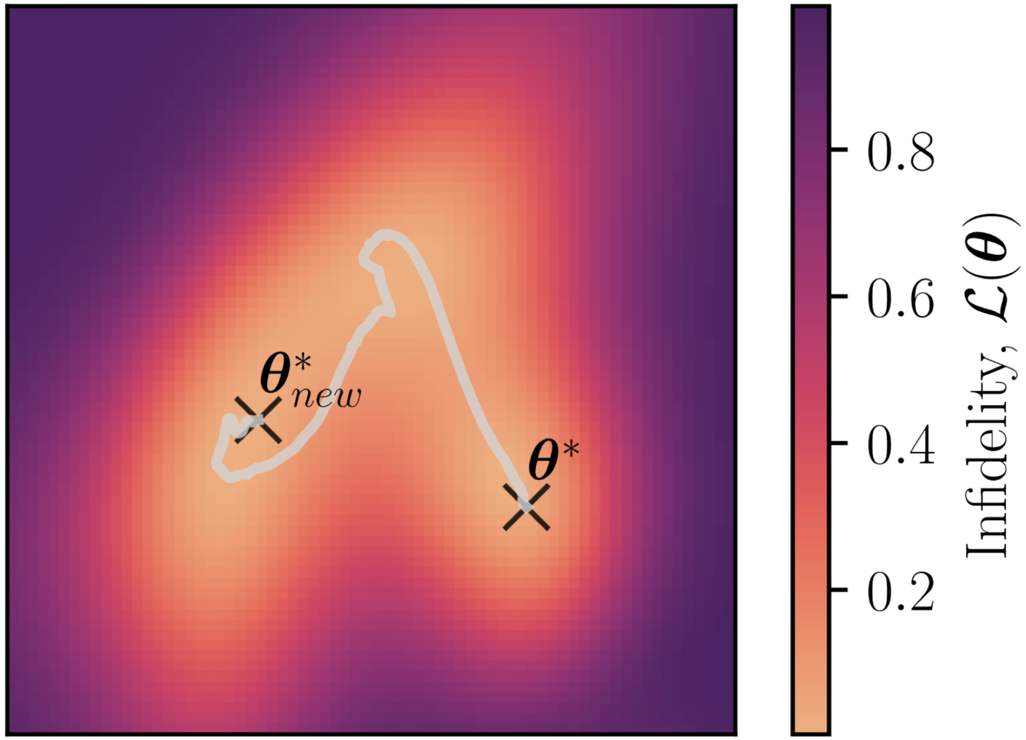
Minima can jump



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.

BUT?

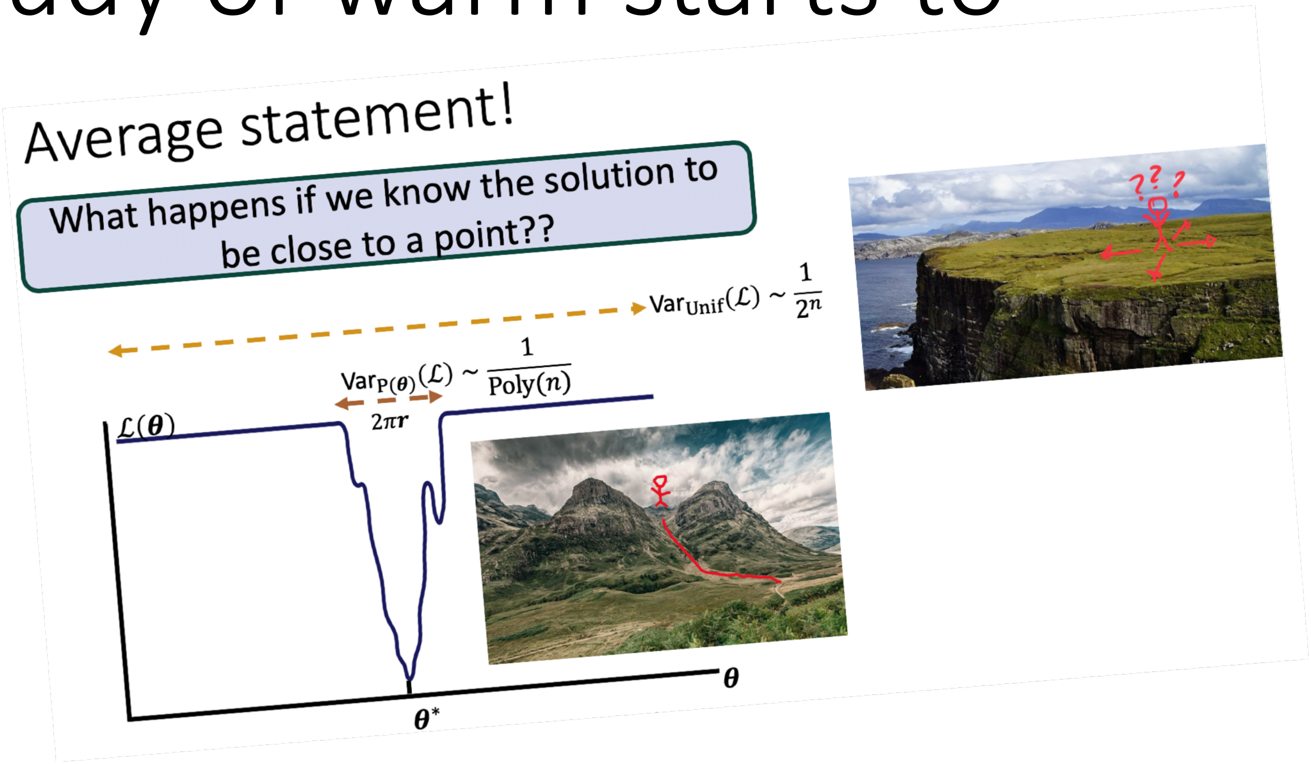


Plotted with ORQVIZ
M. S. Rudolph et al, arXiv:2111.04695 (2021).

Variational quantum simulation: a case study for understanding warm starts. **R Puig***, M Drudis*, et. al. arXiv:2404.10044

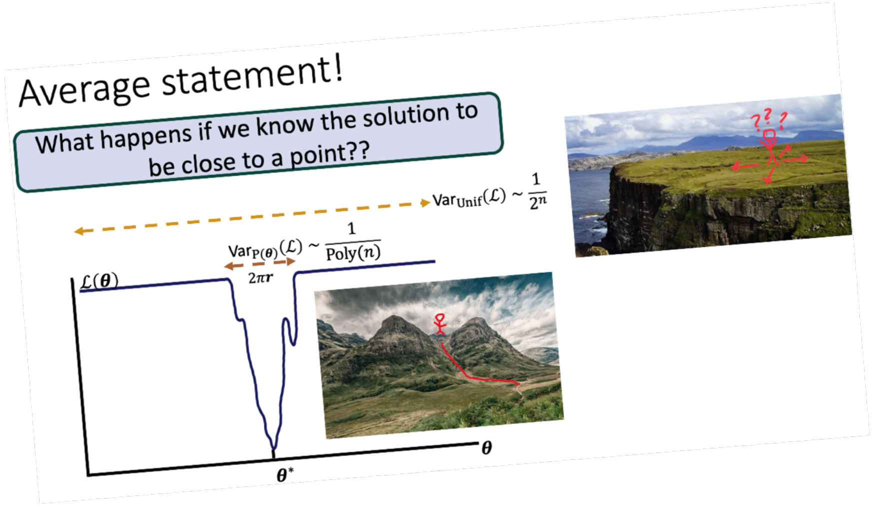
Can we extend the study of warm starts to more general VQA?

A unifying account of warm start guarantees for patches of quantum landscapes. H. Mhiri*, **R. Puig***, et. al. arXiv: 2501.xxxxx (in prep)



Can we extend the study of warm starts to more general VQA?

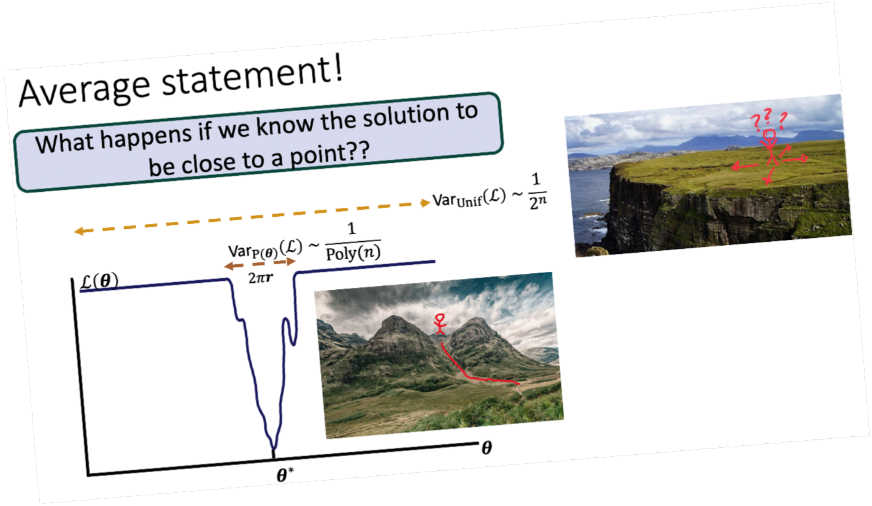
$$\mathcal{L}(\boldsymbol{\theta}) = 1 - \mathrm{Tr}\Big[U(\boldsymbol{\theta})\psi_0 U^{\dagger}(\boldsymbol{\theta})\underbrace{\mathrm{e}^{-it\,H}\psi_0\mathrm{e}^{it\,H}}_{\text{Target state}}\Big]$$



A unifying account of warm start guarantees for patches of quantum landscapes. H. Mhiri*, **R. Puig***, et. al. arXiv: 2501.xxxxx (in prep)

Can we extend the study of warm starts to more general VQA?

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

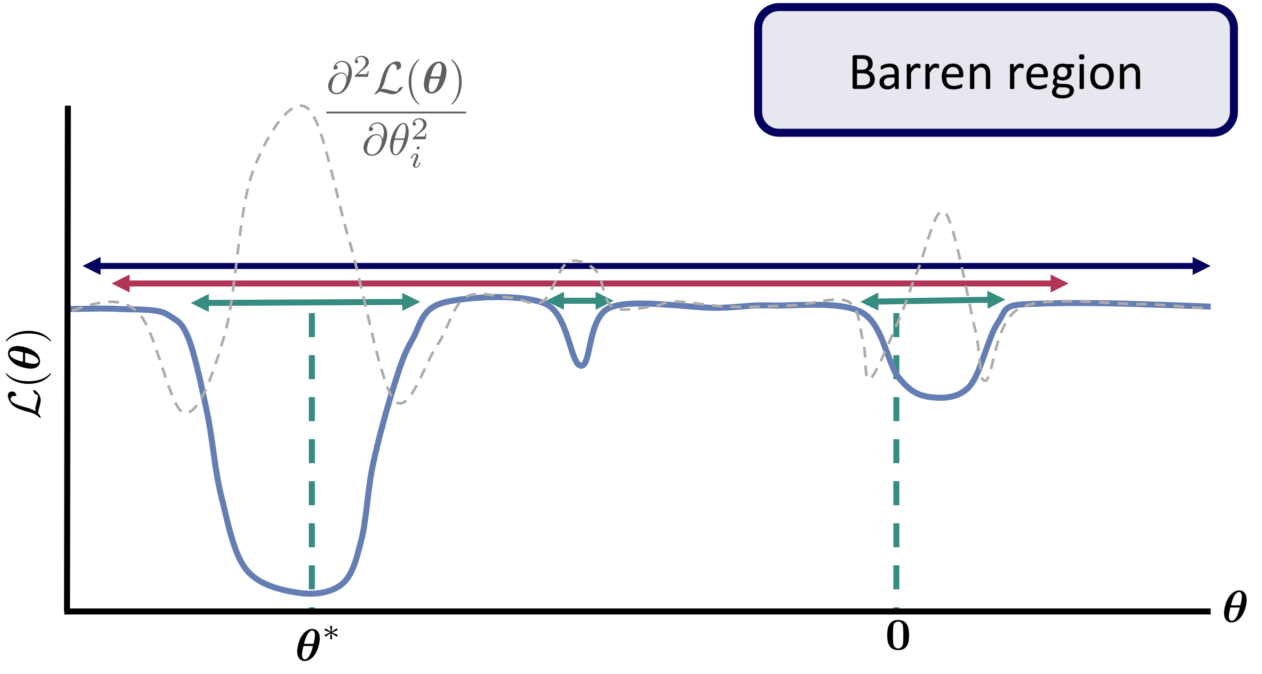


$$\mathcal{L}(\boldsymbol{\theta}) = \mathrm{Tr}\left[U(\boldsymbol{\theta})\psi_0 U^{\dagger}(\boldsymbol{\theta}) O\right]$$

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

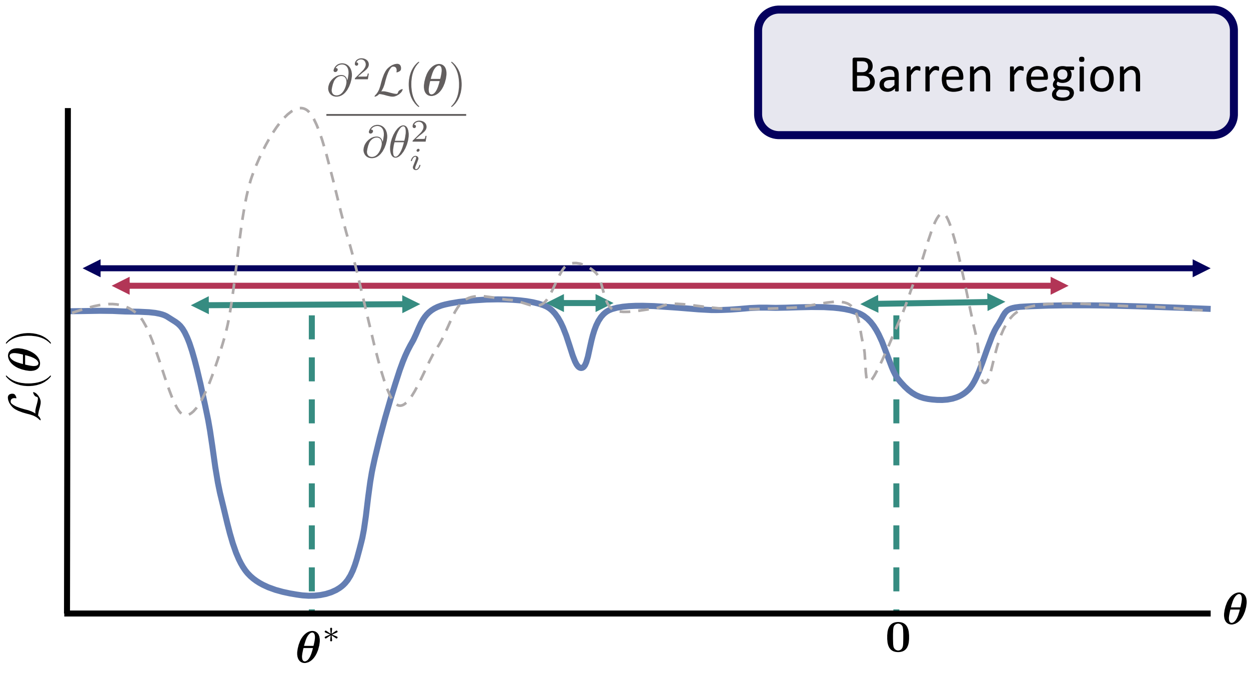
A unifying account of warm start guarantees for patches of quantum landscapes. H. Mhiri*, **R. Puig***, et. al. arXiv: 2501.xxxxx (in prep)

YES!
Capture previous bounds and generalize them.



A unifying account of warm start guarantees for patches of quantum landscapes. H. Mhiri*, **R. Puig***, et. al. arXiv: 2501.xxxxx (in prep)

YES!
Capture previous bounds and generalize them.



Theorem 1

Region with gradients

- $r \in \Theta\left(\frac{1}{\sqrt{M}\mathrm{Poly}(n)}\right)$
- $\mathrm{Var}[\mathcal{L}(\boldsymbol{\theta})]_{\mathrm{P}} \in \Omega(r^4)$

  $\mathrm{P}(\boldsymbol{\theta}) := [\boldsymbol{\theta}_0 - \pi \boldsymbol{r}, \boldsymbol{\theta}_0 + \pi \boldsymbol{r}]$

ARROUND a point with substantial curvature

Large second derivative

A unifying account of warm start guarantees for patches of quantum landscapes. H. Mhiri*, **R. Puig***, et. al. arXiv: 2501.xxxxx (in prep)

When does this bound apply?

Theorem 1

Region with gradients

- $r \in \Theta\left(\frac{1}{\sqrt{M}\mathrm{Poly}(n)}\right)$
- $\mathrm{Var}[\mathcal{L}(\boldsymbol{\theta})]_{\mathrm{P}} \in \Omega(r^4)$

$\mathrm{P}(\boldsymbol{\theta}) := [\boldsymbol{\theta}_0 - \pi \boldsymbol{r}, \boldsymbol{\theta}_0 + \pi \boldsymbol{r}]$

ARROUND a point with substantial curvature
Large second derivative

1. Englobe previous results
2. Treat correlated parameters, something that could not be done before.
3. Region around the solution (Corollary 2)

A unifying account of warm start guarantees for patches of quantum landscapes. H. Mhiri*, **R. Puig***, et. al. arXiv: 2501.xxxxx (in prep)

Ansätze of interest

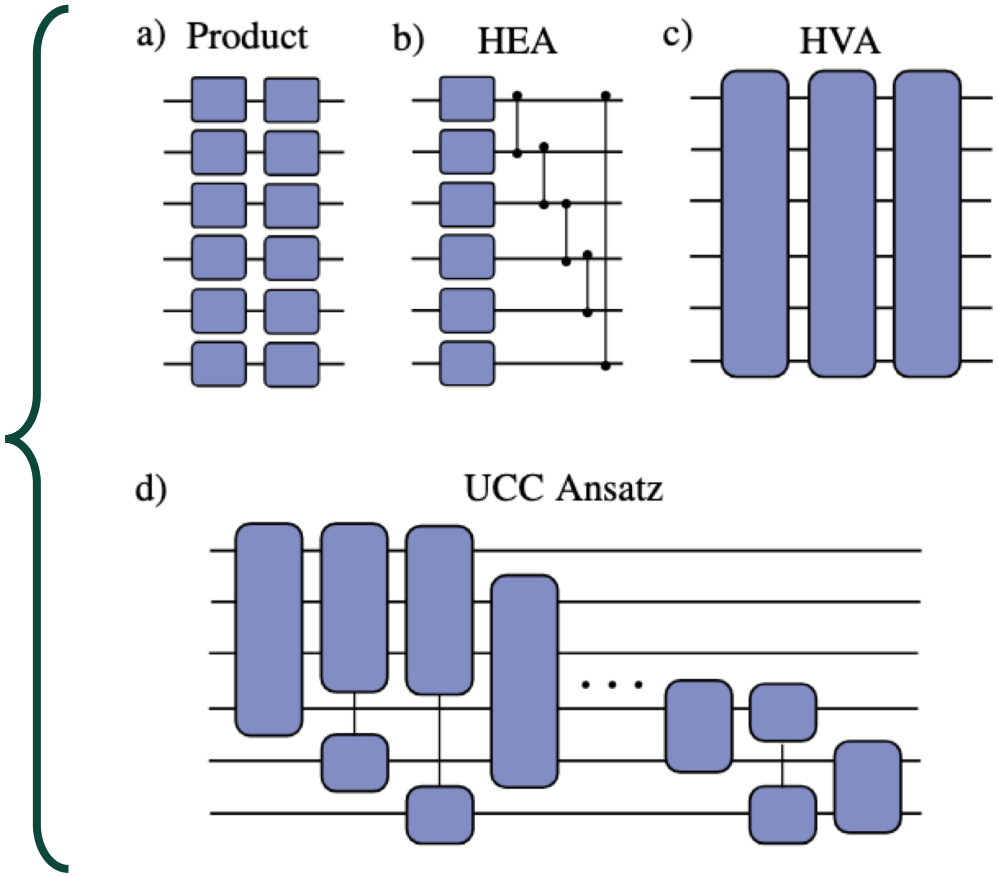
Theorem 1

Region with gradients

- $r \in \Theta\left(\frac{1}{\sqrt{M}\mathrm{Poly}(n)}\right)$
- $\mathrm{Var}[\mathcal{L}(\boldsymbol{\theta})]_{\mathrm{P}} \in \Omega(r^4)$

$\mathrm{P}(\boldsymbol{\theta}) := [\boldsymbol{\theta}_0 - \pi \boldsymbol{r}, \boldsymbol{\theta}_0 + \pi \boldsymbol{r}]$

ARROUND a point with substantial curvature
Large second derivative



A unifying account of warm start guarantees for patches of quantum landscapes. H. Mhiri*, **R. Puig***, et. al. arXiv: 2501.xxxxx (in prep)

Ansätze of interest

Theorem 1

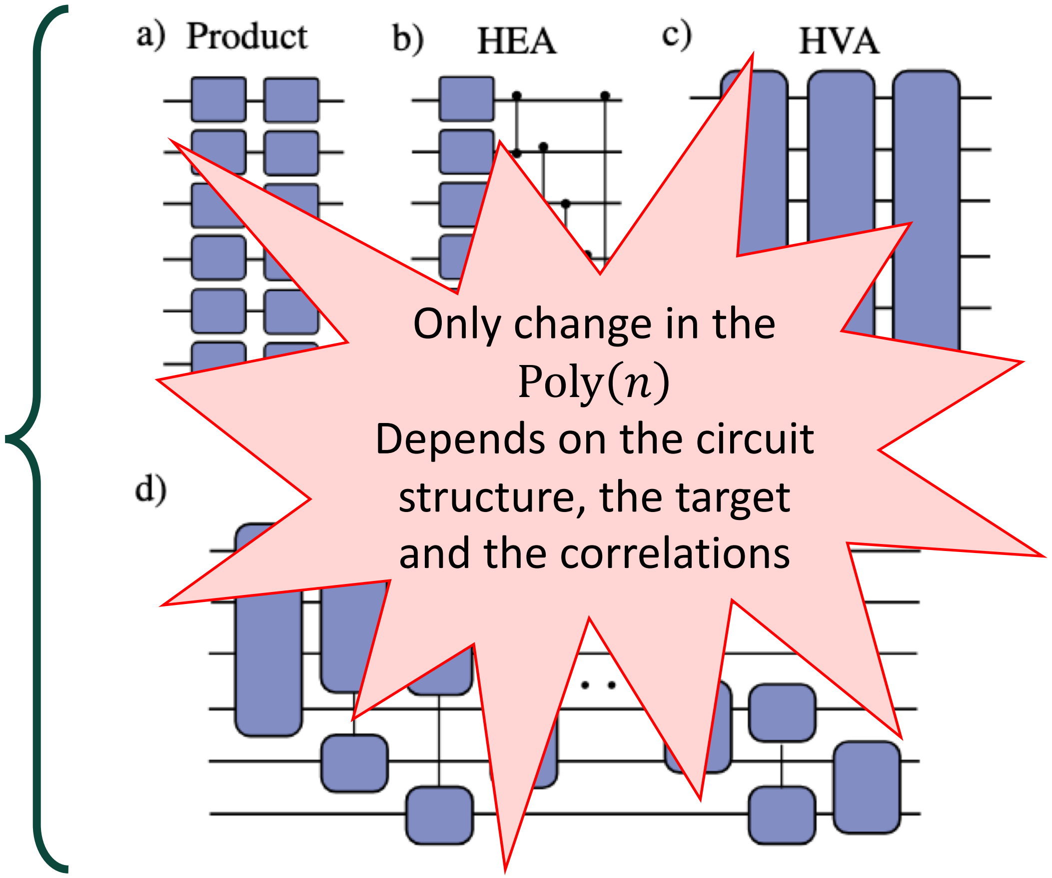
Region with gradients

- $r \in \Theta\left(\frac{1}{\sqrt{M}\mathrm{Poly}(n)}\right)$
- $\mathrm{Var}[\mathcal{L}(\boldsymbol{\theta})]_{\mathrm{P}} \in \Omega(r^4)$

$\mathrm{P}(\boldsymbol{\theta}) := [\boldsymbol{\theta}_0 - \pi \boldsymbol{r}, \boldsymbol{\theta}_0 + \pi \boldsymbol{r}]$

ARROUND a point with substantial curvature

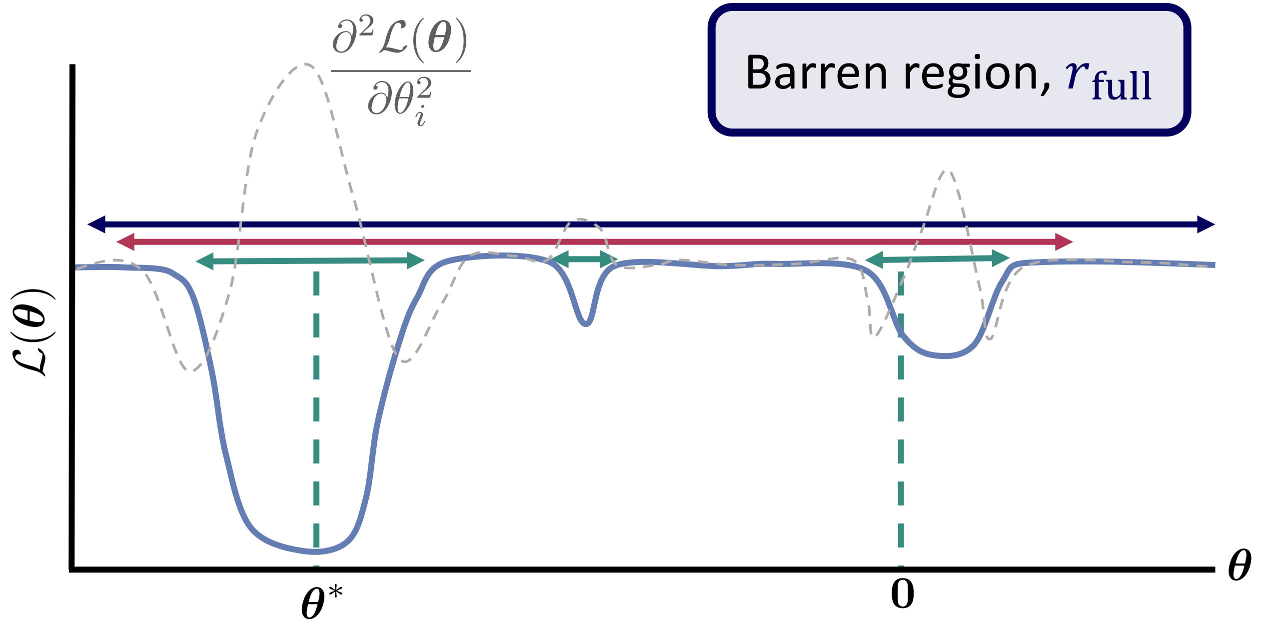
Large second derivative

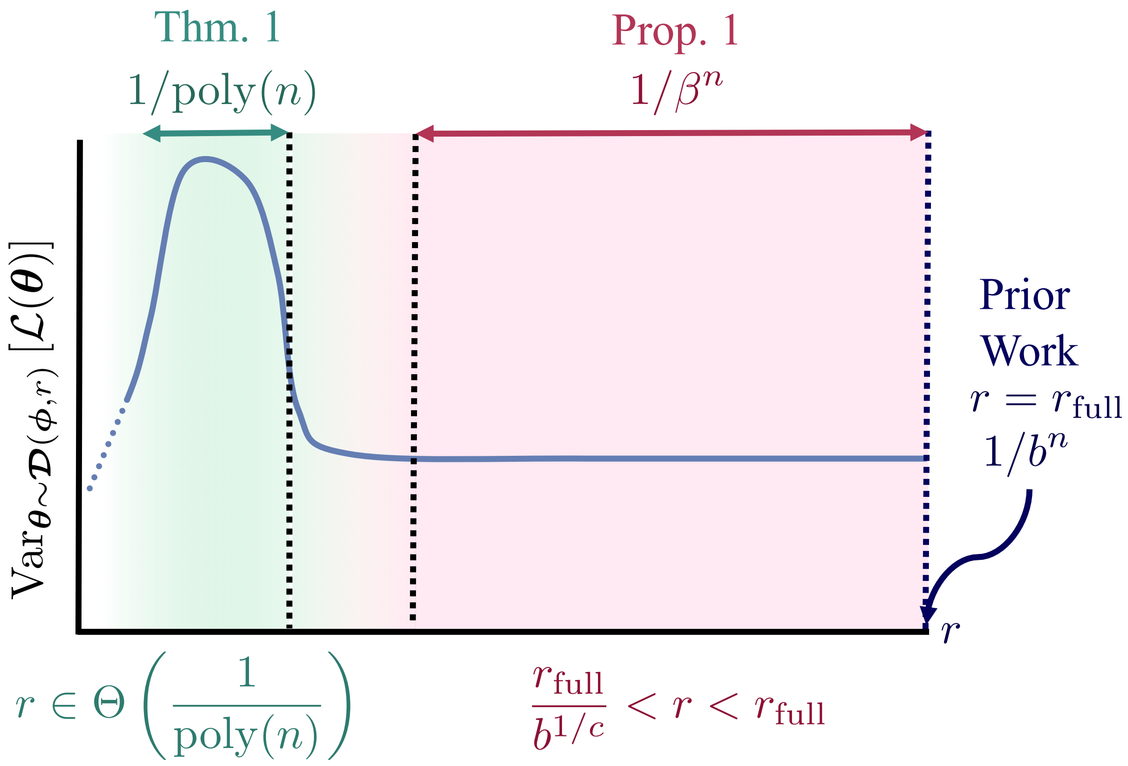


A unifying account of warm start guarantees for patches of quantum landscapes. H. Mhiri*, **R. Puig***, et. al. arXiv: 2501.xxxxx (in prep)

A unifying account of warm start guarantees for patches of quantum landscapes. H. Mhiri*, **R. Puig***, et. al. arXiv: 2501.xxxxx (in prep)

Regions with BP?





Theorem 1

Region with gradients

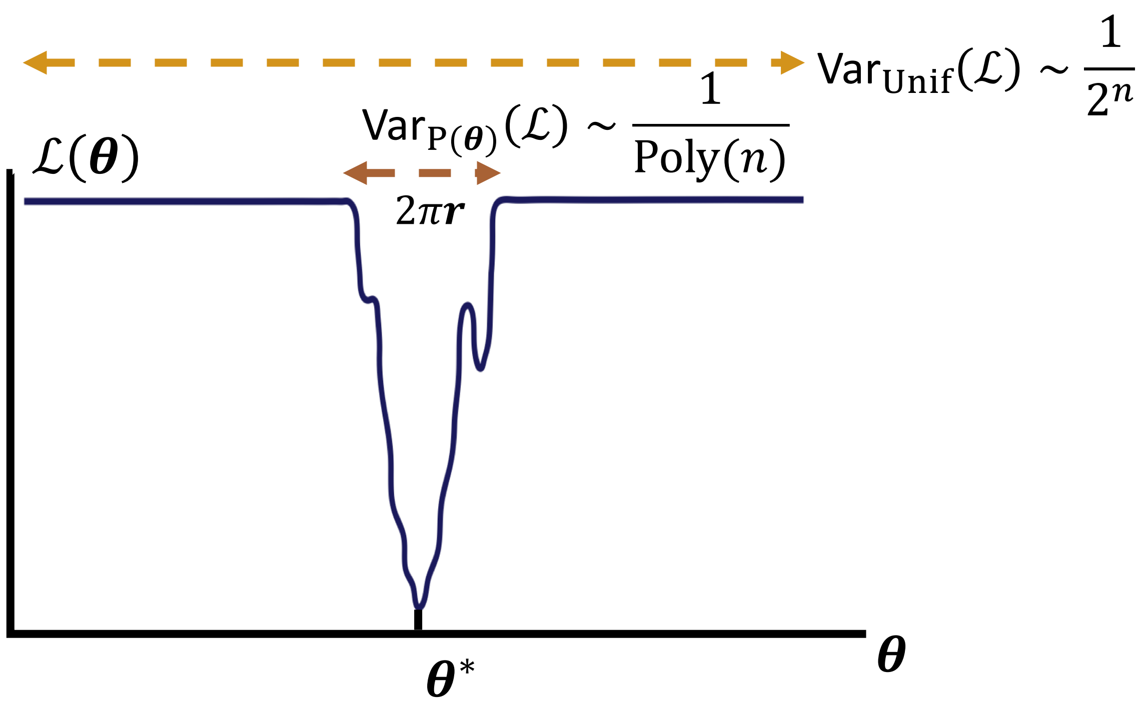
- $r \in \Theta\left(\frac{1}{\sqrt{M}\text{Poly}(n)}\right)$
- $\text{Var}[\mathcal{L}(\boldsymbol{\theta})]_{\text{P}} \in \Omega(r^4)$

$\text{P}(\boldsymbol{\theta}) := [\boldsymbol{\theta}_0 - \pi \boldsymbol{r}, \boldsymbol{\theta}_0 + \pi \boldsymbol{r}]$

Proposition 1

If $\mathcal{L}(\boldsymbol{\theta})$ has BP at r_{full}, it has BP at some $r \in \Theta(1)$

Conclusions

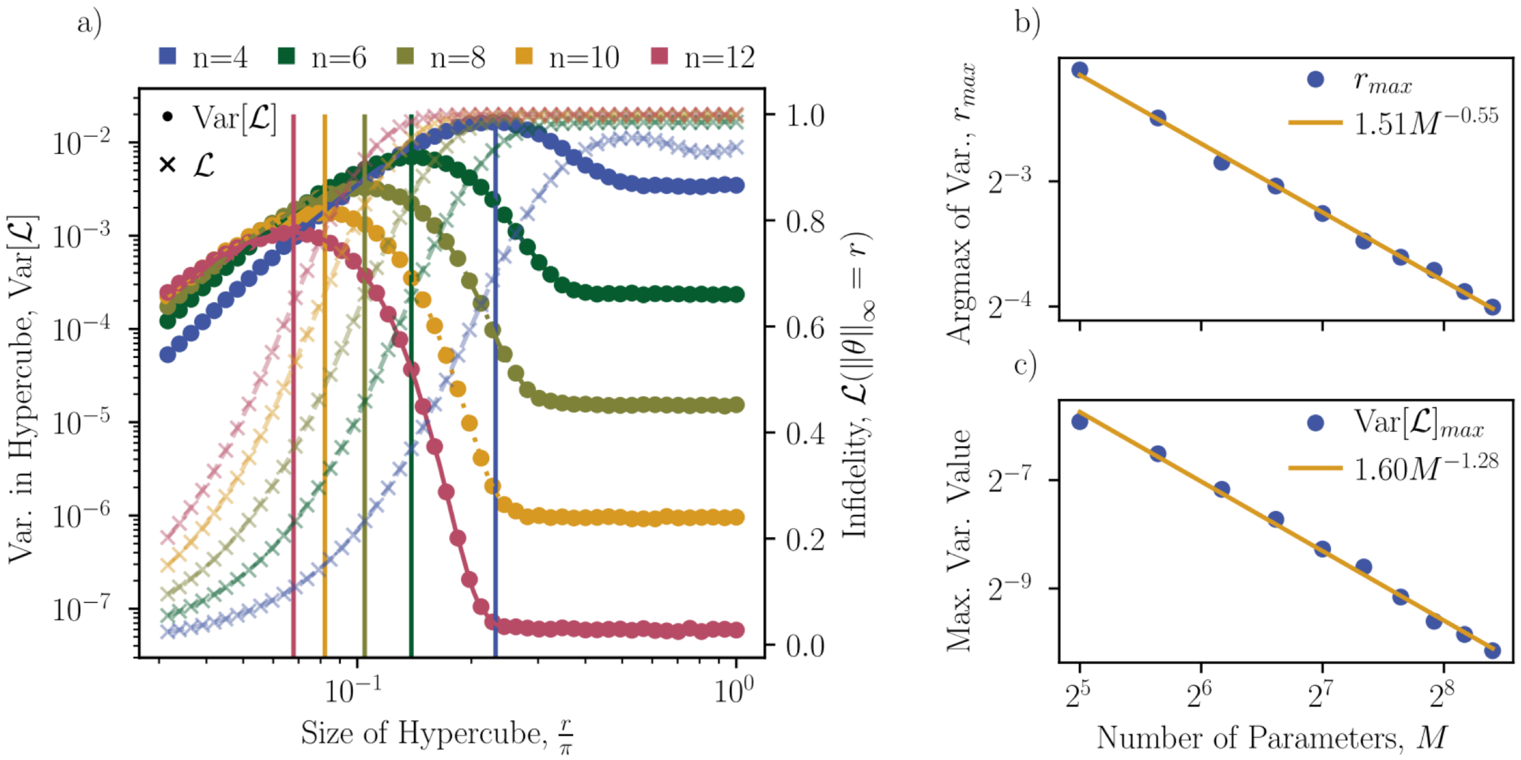


1. Guarantees for VQS
2. Possibly not sufficient in practice
3. Jumps and fertile valleys?
4. Study warm starts and regions of attraction in general
5. Connection with classical stimulability.

Efficient quantum-enhanced classical simulation for patches of quantum landscapes. S Lerch*, **R Puig***, MS Rudolph*, et. al. arXiv: 2411.19896

Extra-slides :)

Region with gradients



$|\psi_0\rangle = U(\boldsymbol{\theta}_0)|\psi\rangle$

$e^{-i\delta t H}|\psi_0\rangle = U(\boldsymbol{\theta}_0)|\psi\rangle$

Compute the experimental variance

U has linear depth

$$\mathcal{L}(\delta\boldsymbol{\theta}) = 1 - \mathrm{Tr}\left[U\left(\boldsymbol{\theta}_0 + \frac{r}{\pi}\right)\psi_0 U^\dagger\left(\boldsymbol{\theta}_0 + \frac{r}{\pi}\right) \mathrm{e}^{-i\delta t\, H} U(\boldsymbol{\theta}_0)\psi_0 U^\dagger(\boldsymbol{\theta}_0)\mathrm{e}^{i\delta t\, H}\right]$$

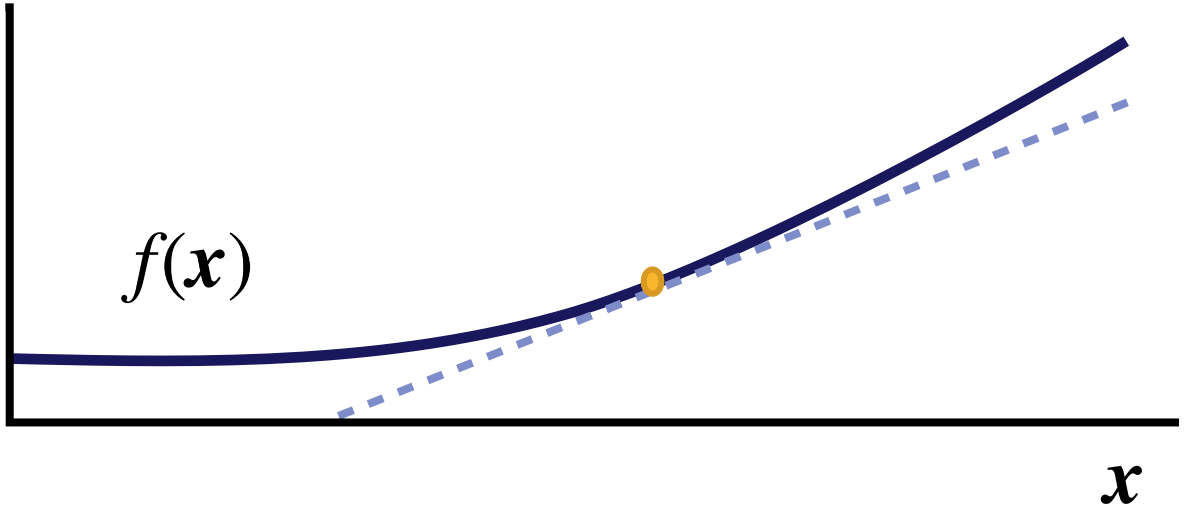
ϵ-approximate convexity

A differentiable function $f(\boldsymbol{x})$ of several variables $f : \mathbb{R}^N \to \mathbb{R}$ is ϵ-approximate convex in a region $\mathcal{R}$ if

$$[\nabla^2 f(\boldsymbol{x})]_{\min} \geq -|\epsilon|$$

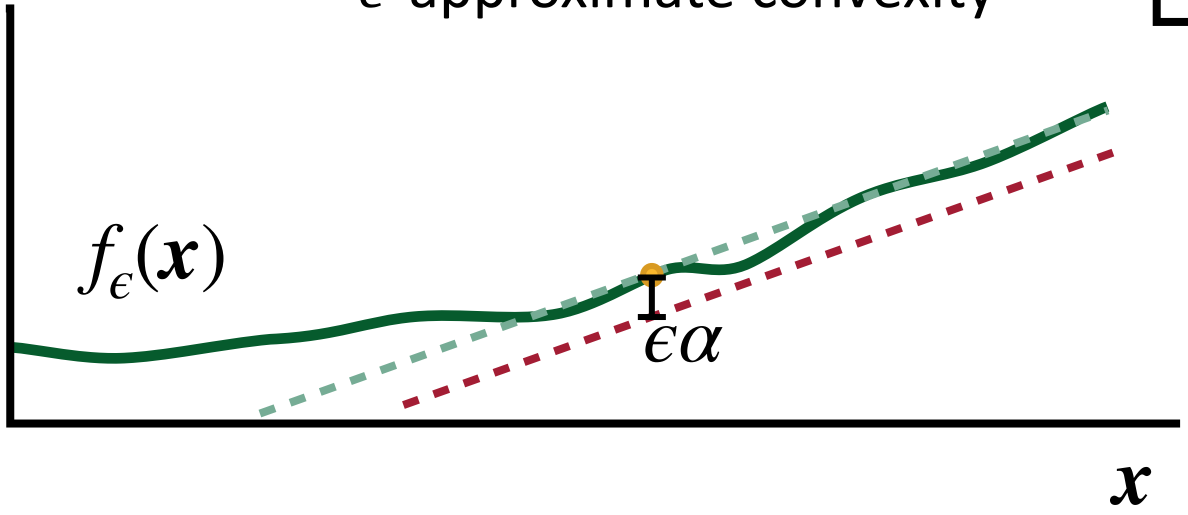
for all $\boldsymbol{x} \in \mathcal{R}$. Here $\nabla\hat{}2 f(\boldsymbol{x})$ denotes the Hessian of $f(\boldsymbol{x})$ and $[A]_{\min}$ is the smallest eigenvalue of A.

ϵ-approximate convexity

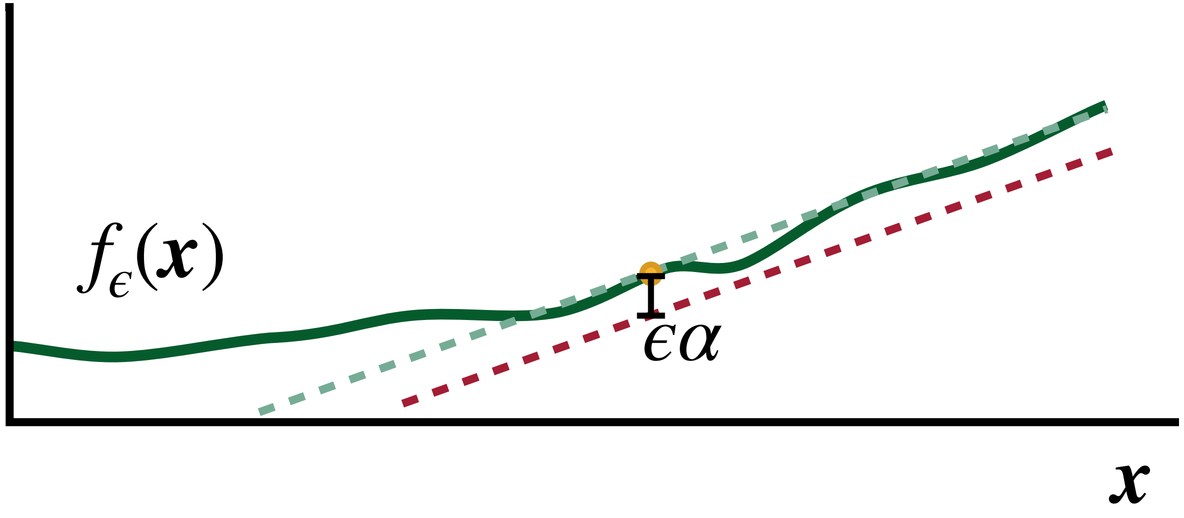


convexity

ϵ-approximate convexity



ϵ-approximate convexity

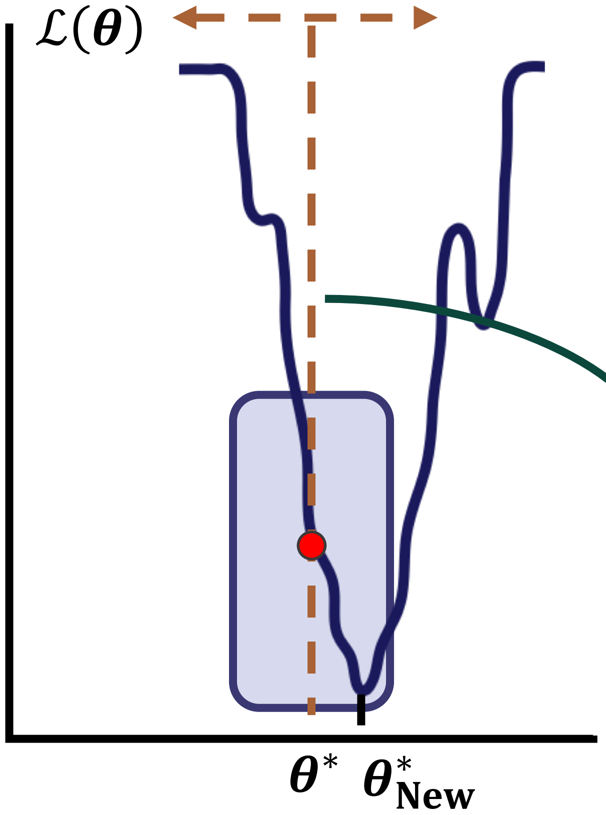


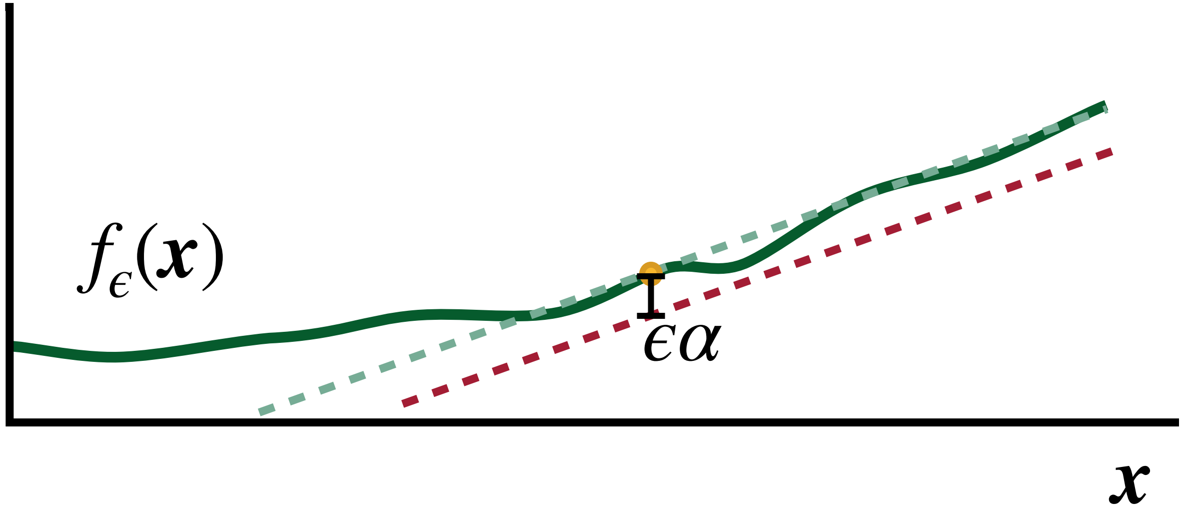
$$\alpha = \frac{1}{2} \max_{\boldsymbol{a},\boldsymbol{b} \in \mathcal{R}} |\boldsymbol{a} - \boldsymbol{b}|_2^2$$

But recall that in the gradient region

$$r \sim \frac{1}{\sqrt{M}} \Rightarrow \alpha \lesssim \frac{1}{\sqrt{M}}$$

ϵ-approximate convexity





But recall that in the gradient region

$$r \sim \frac{1}{\sqrt{M}} \Rightarrow \alpha \lesssim \frac{1}{\sqrt{M}}$$

$$\alpha = \frac{1}{2} \max_{\boldsymbol{a},\boldsymbol{b} \in \mathcal{R}} |\boldsymbol{a} - \boldsymbol{b}|_2^2$$

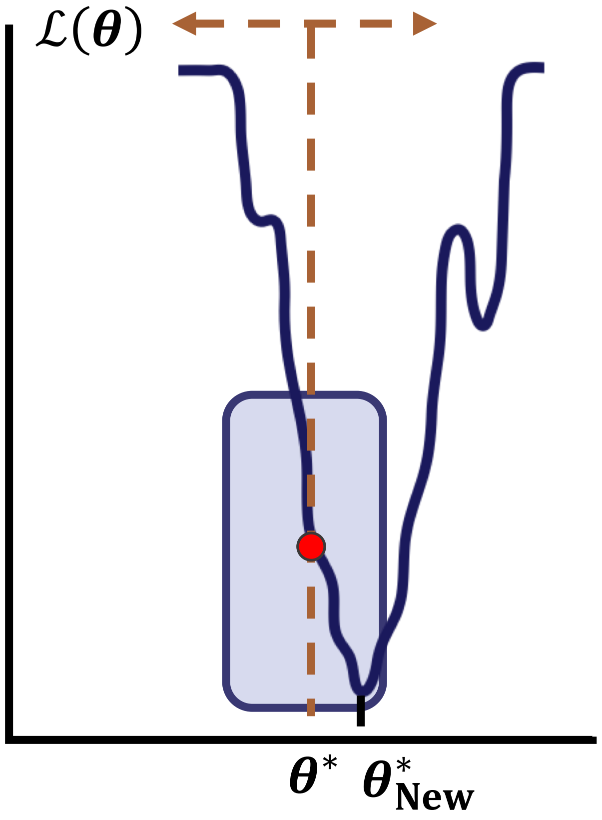
Adiabatic minima

For any time δt in the range $[0, T]$, a function corresponding to the evolution of the adiabatic minima for some initial minimum $\boldsymbol{\theta}^*$, is a continuous function $\boldsymbol{\theta}_{\boldsymbol{A}}(\delta t) \in C^\infty(\mathbb{R}, \mathbb{R}^M)$ such that $\boldsymbol{\theta}_{\boldsymbol{A}}(0) = \boldsymbol{\theta}^*$ and

$$\nabla \mathcal{L}(\boldsymbol{\theta}_{\boldsymbol{A}}(\delta t), \delta t) = 0$$

$\boldsymbol{\theta}_{\boldsymbol{A}}(\delta t)$ is the adiabatic minimum at δt.

Adiabatic minima



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

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

With

$$\beta_A = \frac{\dot{\boldsymbol{\theta}}_A^{\ T}(\delta t)\left(\nabla^2_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}_A)\right)\dot{\boldsymbol{\theta}}_A(\delta t)}{\left|\dot{\boldsymbol{\theta}}_A(\delta t)\right|_2^2}$$

$\beta_A \to 0$ corresponds to the to the curvature of the loss at the minimum being flat in the direction in which the adiabatic minimum move

Barren Plateau

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

Gradients (Th. 1)

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

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

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

ϵ-convexity (Th. 2)

- $\delta t \in \Theta\left(\frac{2|\epsilon|}{M\lambda_{\max}}\right)$
- $r \in \Omega\left(\frac{2|\epsilon|}{M^2} - \frac{\lambda_{\max}\delta t}{M}\right)$
- $[\nabla^2(\boldsymbol{\theta})]_{\min} \geq -|\epsilon|$

Correlations