

Learning quantum states efficiently with the shadow overlap

Douaa Salah^{1,2} Sacha Cerf¹ Jonas Landman¹

¹DIENS, École normale supérieure, PSL University, CNRS, INRIA, 45 rue d’Ulm, 75005 Paris, France

²École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

1 Motivation and contribution

Training a quantum circuit to reproduce an n -qubit target state $|\psi\rangle$ can be cast as optimizing a parametrized circuit $|\psi_\theta\rangle = U(\theta)|0\rangle^{\otimes n}$ to match the target. A natural objective to maximize would be the fidelity $F(\theta) = |\langle\psi|\psi_\theta\rangle|^2$, but two obstacles stand in the way. (i) *No quantum access*. Fidelity cannot even be evaluated without coherent access to copies of $|\psi\rangle$; in some settings one has only measurement data or a classical representation of $|\psi\rangle$. (ii) *Barren plateaus*. Even with quantum access, optimizing the fidelity is not efficient, as *global* observables have exponentially vanishing gradients as n grows, $\text{Var}_\theta[F(\theta)] = 2^{-\Omega(n)}$, so gradient training stalls at scale [1, 2].

Certification offers a way past the first obstacle: Huang, Preskill and Soleimanifar (HPS) [3] introduced the *shadow overlap*, a measurable surrogate for the fidelity between a lab state and $|\psi\rangle$, estimated from single-qubit measurements of the lab state and classical query access to the amplitudes of $|\psi\rangle$. HPS moreover observed that for a specific $|\psi\rangle$ the underlying observable L is *local* and suggested it could therefore be an efficient loss function, without barren plateaus. Yet this claim was left without a general proof, and supported only by numerics in a basic setting (a gate-by-gate circuit construction), and not with a parametrized quantum circuit, leaving open whether barren plateaus arise. In this work we generalize, prove and simulate this conjecture.

Main results. We use the shadow overlap as the cost function of parametrized circuits to learn a target state $|\psi\rangle$, and give an exact criterion for when this is free of barren plateaus (whereas the fidelity is not), for shallow parametrized circuits. This criterion is based on a single property of the target state: we introduce in Section 3 the *Walsh–Hadamard* (WHT) degree $d_\star$ of $|\psi\rangle$: **the loss is free of barren plateaus if the WHT degree grows at most logarithmically, $d_\star = \mathcal{O}(\log n)$** . Informally, it quantifies the locality of the shadow overlap observable, namely the maximal Pauli weight in its decomposition. This quantitative and analytic method allows one to predict in advance for which states our method is free of barren plateaus, confirmed by numerical simulations.

Concretely, our contributions are: i) an *exact* Pauli decomposition of L , tying its locality to the WHT degree of the target, linked to the *conditional correlation*. ii) a rigorous gradient-variance formula and the resulting exact criterion above. iii) a taxonomy of which families of quantum states can be learned efficiently or not. iv) trainability is provably *distinct* from certifiability.

2 The shadow overlap loss

We train the circuit to produce $|\psi_\theta\rangle = U(\theta)|0\rangle^{\otimes n}$ close to $|\psi\rangle$, by maximizing their *shadow overlap*. To evaluate it, given $|\psi_\theta\rangle$, we follow [3] by repeatedly picking a random qubit k , then 1) we read out the other qubits in the Z basis with outcome $z \in \{0, 1\}^{n-1}$, and 2) we measure qubit k with a random Pauli σ_k^a , $a \in \{x, y, z\}$. A value ω can be obtained from an overlap between the post-measurement state of qubit k and another single-qubit state $|\Psi_{k,z}\rangle$ made from the amplitudes of $|\psi\rangle$ and conditioned on z . $\mathbb{E}[\omega]$ is the *shadow overlap*, and this is the loss we optimize. We can rewrite $\mathbb{E}[\omega] = \mathcal{L}(\theta) = \langle\psi_\theta|L|\psi_\theta\rangle$, where the associated observable is $L = \frac{1}{n} \sum_{k=1}^n h_k$ where $h_k = \sum_{z \in \{0,1\}^{n-1}} |z\rangle\langle z| \otimes |\Psi_{k,z}\rangle\langle\Psi_{k,z}|$.

3 Exact Pauli decomposition & locality

For each qubit $k \in [n]$ and outcome $z \in \{0, 1\}^{n-1}$, we can decompose the 1-qubit density matrix $|\Psi_{k,z}\rangle\langle\Psi_{k,z}| = c_0 \mathbb{I}_k + \sum_a r_k^a(z) \sigma_k^a$, where $r_k^a(z)$, $a \in \{x, y, z\}$, are its *Bloch components*. Each $z \mapsto r_k^a(z)$ is a real-valued function on $\{0, 1\}^{n-1}$, on which one can perform a Walsh–Hadamard transform (Boolean Fourier decomposition) which gives corresponding WHT coefficients $\hat{r}_k^a(S)$, for each subset $S \subseteq \bar{k} := [n] \setminus \{k\}$. For each subset we also define $Z_S = \prod_{i \in S} \sigma_i^z$, acting on $|S|$ qubits. We prove the following lemma:

Lemma 1 (Exact decomposition). *For every target $|\psi\rangle$, its shadow overlap’s observable can be written*

$$L = \frac{1}{2} \mathbb{I} + \frac{1}{2n} \sum_{k=1}^n \sum_{S \subseteq \bar{k}} \sum_a \hat{r}_k^a(S) Z_S \otimes \sigma_k^a \quad (1)$$

L is decomposed as a sum of Pauli strings $Z_S \otimes \sigma_k^a$ of *weight* $1 + |S|$, *i.e.*, acting on $1 + |S|$ qubits. If some Pauli strings act on a large ($\Theta(n)$) number of qubits, L is considered *global*. For each target state $|\psi\rangle$ we define its *max WHT degree* as $d_\star = \max_{k,a} \max\{|S| : \hat{r}_k^a(S) \neq 0\}$. Thus the *locality* of L is controlled entirely by the WHT coefficients of the Bloch components. We now have a way to link the target state and the locality of the shadow overlap’s observable.

We also introduce the *WHT mass* for a given Pauli weight d as $w_k(d) := \sum_a \sum_{|S|=d} \hat{r}_k^a(S)^2$.

4 The trainability criterion

So far this decomposition required no assumption on the circuit $U(\theta)$. We now consider the general, block-wise, parametrized circuit from [2], with D layers. It was shown that a Pauli string P with weight w has gradient variance $\text{Var}_\theta[\langle P \rangle] := v(w) \sim c^{-w \mathcal{O}(D)}$. We prove in our case:

Proposition 1 (Variance formula). $\text{Var}_\theta[\mathcal{L}(\theta)] \approx \frac{1}{4n^2} \sum_{k=1}^n \sum_{d \geq 0} w_k(d) v(1+d)$.

Since $v(1+d) = \Omega(1/\text{poly } n)$ up to $d = \mathcal{O}(\log n/D)$ and reaches $2^{-\Omega(n)}$ only at $d = \Theta(n)$, the WHT mass $w_k(d)$ decides the variance. We prove the following, noting that $\sum_d w_k(d) = 1$:

Theorem 1 (Barren Plateau criterion). *Fix the circuit depth. Every Pauli string in L has weight at most $1 + d_\star$, so*

$$\text{Var}_\theta[\mathcal{L}(\theta)] \geq \frac{v(1+d_\star)}{4n}, \quad v(1+d_\star) \sim c^{-(1+d_\star)\mathcal{O}(D)}.$$

Hence the loss is **trainable**, $\text{Var}_\theta[\mathcal{L}(\theta)] = \Omega(1/\text{poly}(n))$, when the max WHT degree stays low, $d_\star = \mathcal{O}(\log n)$ (*local*); and it suffers a **barren plateau**, $\text{Var}_\theta[\mathcal{L}(\theta)] = 2^{-\Omega(n)}$, when $d_\star$ grows as $\Theta(n)$ (*global*).

This proves and sharpens the Huang, Preskill and Soleimanifar [3] observation: **the shadow overlap is not universally plateau-free**, and trainability is exactly the low-degree condition. We show that the target states $|\psi\rangle$ for which the shadow overlap is a trainable loss are exactly those with **short-range correlations**. We also prove that **Trainability** $\neq$ **Certifiability** (in the sense of the mixing time τ in [3]).

$$|\psi\rangle \xrightarrow{\text{random meas.}} |\Psi_{k,z}\rangle \xrightarrow{\text{Bloch}} r_k^a(z) \xrightarrow{\text{WHT}} \hat{r}_k^a(S) \xrightarrow{\max |S|} d_\star \xrightarrow{\leq \mathcal{O}(\log n)?} \text{trainable}$$

5 Types of quantum states efficiently learnable

Given a target state $|\psi\rangle$, the Bloch component functions and their WHT can be obtained from ratios of the amplitudes of $|\psi\rangle$. Using our method, one can predict if the shadow overlap will be a good loss to learn a circuit for producing $|\psi\rangle$, while for most $|\psi\rangle$ using fidelity as a loss would lead to a barren plateau. We can therefore predict the trainability of this method for several families of target states.

We have proven that for a wide variety of quantum states for which the fidelity loss is untrainable, our method is efficient and barren-plateau-free. To that end we prove that the WHT degree $d_\star$ is $\mathcal{O}(1)$ or $\mathcal{O}(\log n)$. This is the case for

- *product states*: $|\psi\rangle = \bigotimes_{i=1}^n |\phi_i\rangle$, with the canonical example of $|+\rangle$;
- *local phase states*: $|\psi\rangle = 2^{-n/2} \sum_{x \in \{0,1\}^n} e^{i\phi(x)} |x\rangle$ with $\phi(x) = \sum_{|i-j| \leq R} g(x_i, x_j)$ and bounded interaction range $R = \mathcal{O}(1)$. As R grows, more qubits contribute and the interactions become less local.
- *graph / cluster states*: $|\psi\rangle = \prod_{(i,j) \in E} CZ_{ij} |+\rangle^{\otimes n}$ for a graph $G = (V, E)$
- *gapped ground states*: $|\psi\rangle$ is the ground state of the transverse-field Ising Hamiltonian $H_{\text{TFIM}} = -J \sum_i \sigma_i^z \sigma_{i+1}^z - h \sum_i \sigma_i^x$ in the *disordered/gapped* phase ($h > J$).

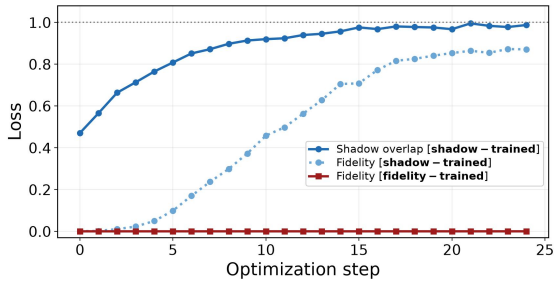
On the other hand, we can prove that *global phase states* such as the global parity state $|\psi\rangle = 2^{-n/2} \sum_{x \in \{0,1\}^n} (-1)^{\phi(x)} |x\rangle$ with $\phi(x) = \bigoplus_{i=1}^n x_i$ have a high WHT degree $d_\star = \Theta(n)$, which induces barren plateaus (even though a fast mixing time τ [3] keeps them certifiable).

6 Numerical simulations

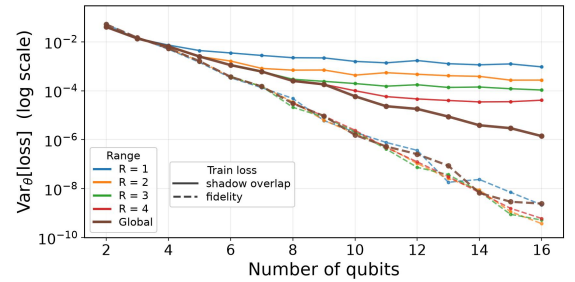
We have performed extensive simulations to verify our theory. We display an excerpt here, using PennyLane's *SimplifiedTwoDesign* ansatz with $D = 2$ layers [4].

a) *Learning Curves*: For the ground state of a TFIM Hamiltonian, we observe as predicted that training a PQC with the shadow overlap loss (blue) allows one to learn the ground state. Both fidelity (dotted) and shadow overlap (solid) increase during training. The same learning with a direct fidelity loss is not successful (red).

b) *Loss Variance Scaling*: For several local (real) phase states¹ with increasing range R , up to a maximally global version with $R = \lfloor n/2 \rfloor$, we observe as expected that the variance of the shadow overlap loss scales with n more favorably than the fidelity loss. For small R , no barren plateaus are observed. As predicted, as R increases, the scaling is less favorable, until it displays a barren plateau for the global case.



(a) Learning curves for a TFIM ground state



(b) Loss variance scaling for local/global phase states

¹ $|\psi\rangle = 2^{-n/2} \sum_{x \in \{0,1\}^n} (-1)^{\phi(x)} |x\rangle$, $\phi(x) = \sum_{|i-j| \leq R} x_i x_j$

References

- [1] Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 9(1):4812, 2018. doi: 10.1038/s41467-018-07090-4.
- [2] M. Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio, and Patrick J. Coles. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nature Communications*, 12(1):1791, 2021. doi: 10.1038/s41467-021-21728-w.
- [3] Hsin-Yuan Huang, John Preskill, and Mehdi Soleimanifar. Certifying almost all quantum states with few single-qubit measurements. *arXiv preprint arXiv:2404.07281*, 2024.
- [4] Ville Bergholm, Josh Izaac, Maria Schuld, Christian Gogolin, Shahnawaz Ahmed, Vishnu Ajith, M Sohaib Alam, Guillermo Alonso-Linaje, Bharath AkashNarayanan, Ali Asadi, et al. PennyLane: Automatic differentiation of hybrid quantum-classical computations. *arXiv preprint arXiv:1811.04968*, 2018.