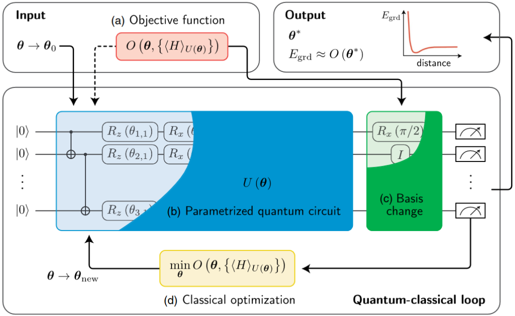
Analyzing variational quantum landscapes with information content

Adrián Pérez-Salinas
In collaboration with X. Bonet-Monroig, H. Wang

QTML 2023
November 24th 2023

Variational Quantum Algorithms

- Variational Quantum Algorithms are proposed to fit the current limitations of quantum hardware
- Quantum hardware is used to prepare a parameterized quantum circuit to construct a quantum state
- The quantum state is optimized with respect to certain metric (e.g. energy)
- Quantum advantage is an open question
- Several optimization issues
  - Circuit and sampling noise
  - Gradient not always available
  - Vanishing gradients / Barren Plateaus[1]
  - Plethora of local minima[2]



[3]

An open question

How can we train VQAs better?

[1] Jarrod R. McClean et al. Nature Communications 9.1 (Nov. 2018),
[2] Eric R. Anschuetz and Bobak T. Kiani. (Sept. 2022). eprint: arXiv:2205.05786.
[3] Kishor Bharti et al. Reviews of Modern Physics 94.1 (Feb. 2022),

Exploratory Landscape Analysis and Information Content

- Exploratory Landscape Analysis (ELA) is a numerical technique from classical optimization[4]
- ELA aims to efficiently extract properties of the landscape to be optimized by sampling
  - Efficient $=$ we need $\mathcal{O}(m)$ samples from the loss function, where m is the number of parameters
- ELA $\rightarrow$ ensure trainability, improve initialization, find suitable optimizer

Our approach

- We use Information Content (IC), a proxy for *variability* of the landscape
- We connect the expected norm of the gradient with features of the landscape
- Robust theoretical bounds and numerical checks

[4] Mario A. Muñoz, Michael Kirley, and Saman K. Halgamuge. IEEE Transactions on Evolutionary Computation 19.1 (2015).

Information Content

Computation

- Sample $\mathcal{O}(m)$ points and connect them randomly through random walks
- Compute in the random walk $\Delta C_i = \frac{C(\vec{\theta}_{i+1}) - C(\vec{\theta}_i)}{\left\| \vec{\theta}_{i+1} - \vec{\theta}_i \right\|}$.
- Discretize the walk to a sequence $\phi(\epsilon) = \begin{cases} \mathrm{sgn}(\Delta C_i) & \text{if } |\Delta C_i| > \epsilon \\ \odot & \text{if } |\Delta C_i| \leq \epsilon \end{cases}$
- Compute the IC as

$$H(\epsilon) = \sum_{a \neq b} -p_{ab} \log_6 (p_{ab}),$$

for $ab = \{-, \odot, +\}^2$, with p_{ab} the extimated probability of the sequence (ab) in $\phi(\epsilon)$

Interpretation

- Value of IC
  - Large IC means high variability in the landscape
  - Low IC means no change
  - IC gives insights in the probability of change
- Value of ϵ
  - If ϵ is large, the landscape is flat (to this scale)
  - If ϵ is smal, variability is enforced
  - ϵ provides insights in the value at which the landscape is variable
- Combining the values of IC and ϵ we can estimate the gradient norm, in average in all directions

Theoretical Results

Gradient norms are related to ϵ

$$\mathrm{Prob}\left(\mathbb{E}_{\vec{\theta}}\left(\nabla C(\vec{\theta})\cdot\vec{\delta}\right)\leq\epsilon\right)=\Phi_G\left(\frac{\epsilon\sqrt{m}}{\|\nabla C\|}\right), \tag{1}$$

for Φ_G the CDF of a normal distribution, and

$$\|\nabla C\|^2=\mathbb{E}\left(\left\|\nabla C(\vec{\theta})\right\|^2\right)$$

Maximal IC

Let $\epsilon_M = \mathrm{argmax}_\epsilon(H(\epsilon))$, then

$$\|\nabla C\|\in\Omega\left(\epsilon_{\mathrm{M}}\sqrt{m}\right) \tag{2}$$

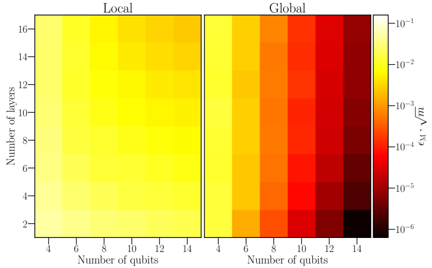
Sensitive IC

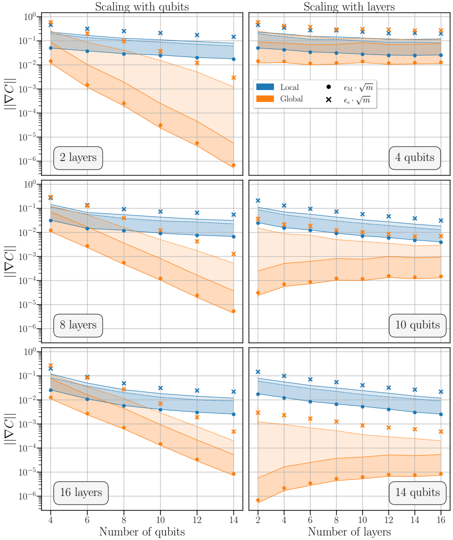
Let $\epsilon_S = \min\{\epsilon > 0 | H(\epsilon)\leq\eta\}$, then

$$\|\nabla C\|\leq\frac{\epsilon_{\mathrm{S}}\sqrt{m}}{\Phi_G^{-1}\left(1-3\eta/2\right)}. \tag{3}$$

Numerical experiments

- Experiments in known problems for barren plateaus[a] (up to shot noise).
- Same circuit, different (global / local) observables show different regimes
- We match the theoretical results, while numerical scaling factors are also available

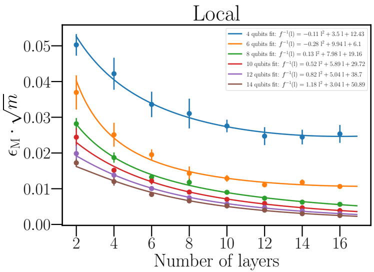


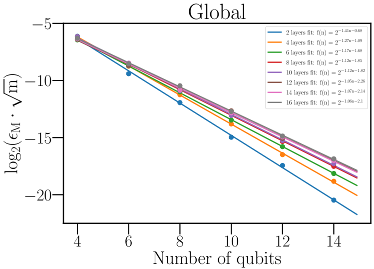


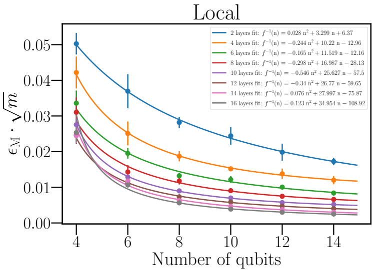
[a] M. Cerezo et al. Nature Communications 12.1 (Mar. 2021),

Numerical experiments

- Experiments in known problems for barren plateaus.
- Same circuit, different (global / local) observables show different regimes
- We match the theoretical results, while numerical scaling factors are also available







Conclusions

- We propose a data-driven method to study variational quantum algorithms
- Data-driven methods have a broader range of applicability than analytical methods
- We connect information content to the average norm of the gradient with analytical bounds
- We can characterize barren plateaus easily with remarkable accuracy
- Scaling prefactors are accessible for the first time
- Hopes for VQAs
  - Learn landscapes before optimization
  - Use suitable optimizers
  - Estimate resources for successful optimization

Paper accessible here

The end

Thank you for your attention

Analyzing variational quantum landscapes with information content

Adrián Pérez-Salinas
In collaboration with X. Bonet-Monroig, H. Wang

QTML 2023
November 24th 2023

Adrián Pérez-Salinas, Hao Wang, and Xavier Bonet-Monroig. (2023). arXiv: 2303.16893 [quant-ph]