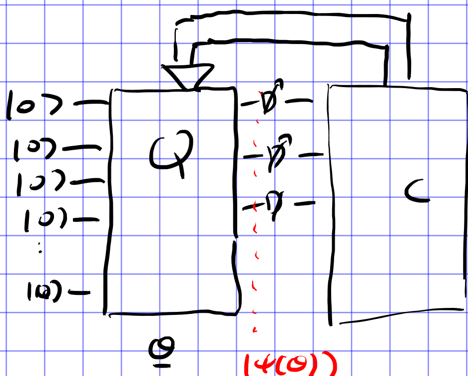
VARIATIONAL QUANTUM ALGORITHMS II

leonardo.banchi@unifi.it

- Gradients
- Barren Plateau
- Applications



$$\hat{U} = T \exp\left(-i \int_0^t \hat{H}(t)\, dt\right)$$

$$\hat{U} = e^{i\theta \hat{\sigma}_\alpha} \qquad \hat{\sigma}_\alpha \text{ Pauli operator}$$

$$|\psi(\theta)\rangle = \underbrace{U_L(\theta)\, W_L \cdots U_1(\theta)\, W_1}\, |0 \cdots 0\rangle$$

Parametric Quantum Circuit
"
Quantum N.N.

L = # Layers

$\theta = (\theta_1, \ldots, \theta_m)$

<u>Cost Function</u>

$$C(\theta) = \langle \psi(\theta) | \hat{A} | \psi(\theta) \rangle$$

E.g. $\hat{A} = \hat{H} \Rightarrow C(\theta) = E(\theta)$

$$\frac{\partial C(\theta)}{\partial \theta_i} \simeq \frac{C(\theta + \varepsilon e_i) - C(\theta)}{\varepsilon}$$

$e_i = (0\ 0\ \ldots\ 1\ 0)$

$\uparrow$ iTh position

$$\hat{A} = \sum_{\ell} a_\ell\, |a_\ell\rangle\langle a_\ell|$$

results a_ℓ prob

$$P(a_\ell|\theta) = |\langle \psi(\theta) | a_\ell \rangle|^2$$

$$C(\theta) \simeq \frac{1}{M} \sum_{l=1}^{M} a_{\theta_l} = \langle \hat{A} \rangle_\theta$$

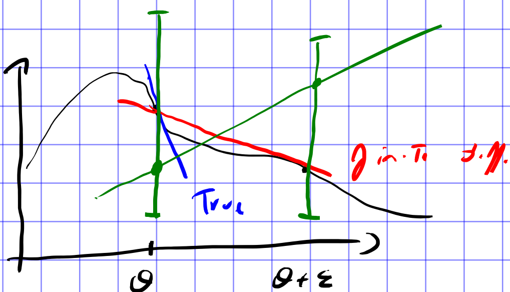
outcome l

\# shots

Variance

$$\frac{\langle \hat{A}^2 \rangle_\theta - \langle \hat{A} \rangle_\theta^2}{M} = \delta^2$$

small



$\delta \ll \varepsilon$

$$\hat{A} = \sum \lambda_l \hat{P}_l$$

Eg. $\hat{H} = \sum J_{ij} \hat{\sigma}_i^x \hat{\sigma}_j^x + h_i^\alpha \hat{\sigma}_i^\alpha$

$$\langle \hat{A} \rangle = \sum_l \lambda_l \langle \hat{P}_l \rangle$$

$$|\psi(\theta)\rangle = \hat{W} e^{i\theta_j \hat{X}_j} \hat{V} |0 \cdots 0\rangle$$

independent on θ_j / possibly dependent on $\theta_i \rightarrow i \neq j$

$\hat{x}_j^2 = \pi$ $\quad ! \quad \frac{\pi}{4}$

$$\frac{\partial C}{\partial \theta_j} = \frac{C(\theta + \frac{u}{2} e_j) - C(\theta - \frac{u}{2} e_j)}{\sin(u)}$$

2 different circuits

Let's call $a^{\pm}_{\theta_{s,j}}$ outcome of shot s parameter j $\pm$ $(\theta \pm \frac{u}{2} e_j)$

$$\frac{\partial C}{\partial \theta_j} = \frac{1}{n} \sum_{s=1}^{n} \frac{a^{+}_{\theta_s j} - a^{-}_{\theta_s j}}{\sin u}$$

$$g_j = \frac{a^{+}_{\theta j} - a^{-}_{\theta j}}{\sin u}$$

Unbiased estimator of the gradient

$$\frac{\partial C}{\partial \theta_j} = \mathbb{E}(g_j)$$

Expectation value

Variance $\simeq (\sin u)^{-2}$

opt. $u = \frac{\pi}{2}$

Adaptive strategies — Hübner et al. arxiv 1909.09083

$$e^{it(\hat{H} + \theta \hat{V})}$$

$$\exp\left(i \sum J_{ij} \sigma_j^x \sigma_i^x + \underline{\theta_j} \sigma_j^z \right)$$

arxiv 2005.10299

$$U_j = \exp\left(i \left(H_j + \theta_j \hat{V}_j \right)\right)$$

$$\frac{\partial e^{\hat{z}}}{\partial \theta_j} = \int_0^1 ds \; e^{s\hat{z}} \frac{\partial \hat{z}}{\partial \theta_j} e^{(1-s)\hat{z}}$$

$$C(\theta) = \langle \psi(\theta) | \hat{A} | \psi(\theta) \rangle$$

$$\frac{\partial C}{\partial \theta} = \int_0^1 ds \left[C_+(\theta, s) - C_-(\theta, s) \right]$$

$$C_\pm(\theta, s) = \langle \psi(\theta, s) | \hat{A} | \psi(\theta, s) \rangle$$

$$|\psi(\theta, s)\rangle = W \exp\left(is \left(\hat{H} + \theta_j \hat{V} \right)\right) e^{\pm i \frac{\pi}{4} \hat{V}} \exp\left(i(1-s) \left(\hat{H} + \theta_j \hat{V} \right)\right) V |0 \ldots 0\rangle$$

Stochastic Parameter Shift Rule

For QuDits arxiv: 2107.12390

$$C' = \sum_{N=1}^{2R} C\left(\frac{2N-1}{2R}\pi\right) \frac{(-1)^{N-1}}{4R \sin^2\left(\frac{2N-1}{4R}\pi\right)}$$

R = # freq

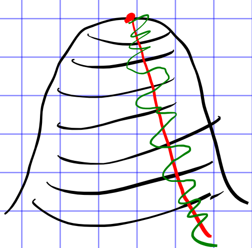
$e^{i\theta \hat{G}}$

$$\frac{\partial C}{\partial \theta_j} = \mathbb{E}(g_j)$$

Stochastic Gradient Descent

$$\vec{\theta}_{n+1} = \vec{\theta}_n - \gamma \vec{G}$$

gradient estimator



SGD converges to local optimum

$$C(\theta_m) - C(\theta_{opt}) \simeq \sqrt{\frac{\mathbb{E}(G^2)}{m}} \longrightarrow 0$$

local optimal

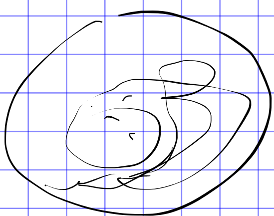
$\mathbb{E}(G^2)$ can be bounded by Quantum Fisher Information

$$F_{ij} = \langle \partial_i \psi | \partial_j \psi \rangle - \langle \partial_i \psi | \psi \rangle \langle \psi | \partial_j \psi \rangle$$

arxiv 1912.06744 Gentini

QFI measures the expressivity of $|\psi(\theta)\rangle$

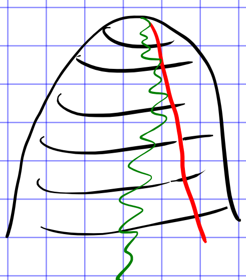
(see Haug et al 2102.01659)



Trade off between expressib. and "Trainability" of parametric quantum circuits

noise adds a bias

but reduces QFI



BARREN PLATEAU

How big is the gradient

$$|\psi(\theta)\rangle = \hat{W}\, e^{i\theta_j \hat{H}_j}\, \hat{V}\, |0 \dots 0\rangle$$

$$C(\theta) = \langle \hat{A} \rangle_\theta \qquad \hat{H}_j^2 = \mathbb{1}$$

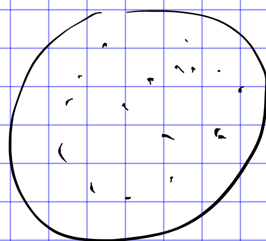
$$\frac{\partial C}{\partial \theta_j} = (\ ) \sin(2\theta_j) + (\ ) \cos(2\theta_j)$$

$$\mathbb{E}_{\theta_j} \frac{\partial C}{\partial \theta_j} = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\qquad) = 0$$

if circuit is highly expressive

=> changing θ_i

=> possible t-design



Hilbert

$$\mathbb{E}_\theta\, \hat{U}(\theta)^{\otimes t}\, \hat{O}\, \hat{U}(\theta)^{\otimes t\,\dagger}$$

$$\simeq \int \underbrace{dU}_{\text{Haar}}\; U^{\otimes t}\, \hat{O}\, U^{\otimes t\,\dagger}$$

$$C(\theta) = \langle 0 \cdots 0 | V^{\dagger} e^{-i\theta_j \hat{H}_j} W^{\dagger} \hat{A} W e^{+i\theta_j \hat{H}_j} V | 0 \cdots 0 \rangle$$

$$\frac{\partial C(\theta)}{\partial \theta_j} = \langle 0 \cdots 0 | V^{\dagger} \underbrace{\left(-i \left[H_1 , e^{-i\hat{H}\theta_j} W^{\dagger} A W \right] \right)}_{B} V | 0 \cdots 0 \rangle$$

suppose V is a 2-design

$$\mathrm{Var}\left[\frac{\partial C}{\partial \theta} \right] = \mathbb{E}_{\theta}\left[\left(\frac{\partial C}{\partial \theta} \right)^2 \right]$$

$$\simeq \int dV \left(\langle 0 \cdots 0 | V^{\dagger} B V | 0 \cdots 0 \rangle \right)^2$$

McClean et al. (Nat Comm 2018)

$$\mathrm{Var}\left[\frac{\partial C}{\partial \theta_j} \right] \simeq O\left(\mathrm{poly}\left(\frac{1}{2^n} \right) \right)$$

$n = \#$ qubits

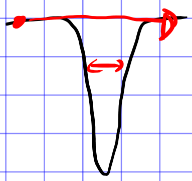
$\longrightarrow 0$ as $n \to \infty$

"self averaging" "concentration of measure"

single sample close to average

with high probability

$\frac{\partial C}{\partial \theta_j}$ is exponentially small



Cerezo "Cost Function dependent .. "

mitigation strategies