QResNet: a variational entanglement skipping algorithm

Giulio Crognaletti [1 2], Michele Grossi [3], Sofia Vallecorsa[3], Angelo Bassi [1 2]

[1]Department of Physics, University of Trieste, Strada Costiera 11, 34151 Trieste, Italy

[2]INFN, Istituto Nazionale di Fisica Nucleare, Trieste Section, Via Valerio 2, 34127 Trieste, Italy

[3]CERN, Conseil Européen pour la Recherche Nucléaire, Espl. des Particules 1, 1211 Meyrin, Switzerland

QTML Conference, November 21st, 2023

Trainability

Variational Quantum Algorithms (VQA) are among the most promising NISQ algorithms.

- **Ansatz**: $U(\boldsymbol{\theta}) = \prod_{l=1}^{L} W_l S_l(\boldsymbol{\theta}_l)$ where $S(\boldsymbol{\theta}_l) = \bigotimes_k \exp(i\theta_{kl} P_{kl})$
- **Initial state**: ρ_{in}
- **Cost function**: $C(\boldsymbol{\theta}) = \mathrm{Tr}\left[\rho_{in} U^\dagger(\boldsymbol{\theta}) O U(\boldsymbol{\theta})\right]$
- **Observable**: $O = \sum_j c_j o_j$

Barren Plateau

A VQA is said to suffer from *barren plateau* if the scaling of $\mathrm{Var}_{\boldsymbol{\theta}}\left[\partial_k C\right]$ with respect to number of qubits q is exponentially decreasing, i.e.

$$\mathrm{Var}_{\boldsymbol{\theta}}\left[\partial_k C\right] \in O\left(\frac{1}{b^q}\right) \quad b > 1, \ \forall k$$

Most common causes include **locality of O** (Cerezo 21), **high expressivity** (Hölmes 22) and **noise** (Wang 21)

Why Residual Networks?

Classical *ResNets* have many desirable properties

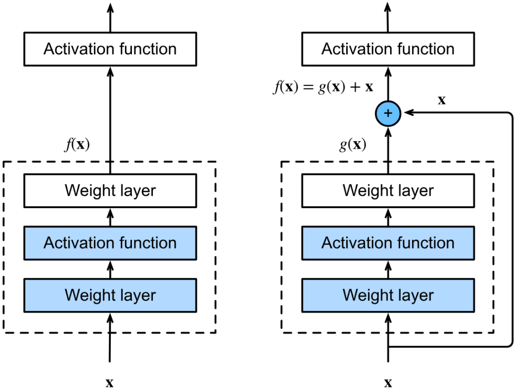


Figure 1: Regular and Residual blocks

A shift in the bias:

The use of *skip connections* in classical Neural Networks (NN) allow to shift the **bias** of the model:

- Improve trainability (better "gradient flow")
- Ease *transferability* between shallow and deep models

QResNet

Critical points:

- **No cloning** of quantum states
- **Unitarity** of the transformation

Quantum residual block

Classical skip-connection $\Rightarrow$ Quantum skip-connection

$$f(x) \rightarrow g(x) = x + f(x) \quad \Rightarrow \quad W \rightarrow \tilde{W} = \alpha \mathbb{1} + \beta W$$

Example of a paramterization of the *quantum skip connection* for each layer l:

$$\tilde{W}_l = \alpha \mathbb{1} + \beta W_l \rightarrow \tilde{W}_l(\phi) \equiv W_l^\dagger \exp(i\phi H_l) W_l = \cos(\phi)\mathbb{1} + i \sin(\phi) W_l^\dagger H_l W_l \quad (1)$$

This leads to the ansatz

$$G(\boldsymbol{\theta}, \phi) = \prod_{l=1}^{L} G_l(\boldsymbol{\theta}_l, \phi_l) \quad , \quad G_l(\boldsymbol{\theta}_l, \phi_l) = \tilde{W}_l(\phi_l) S(\boldsymbol{\theta}_l) \quad (2)$$

Initialization strategy

Furthermore, Eq. 1 can be readily generalized to $\tilde{W}_l(\phi_l) \equiv W_l^\dagger \left[\bigotimes_{j=1}^q \exp(i\phi_{jl}P_{jl})\right] W_l$

In general, the model $G(\boldsymbol{\theta}, \phi)$ can be very expressive, especially for *large L*

How does this help with vanishing gradients?

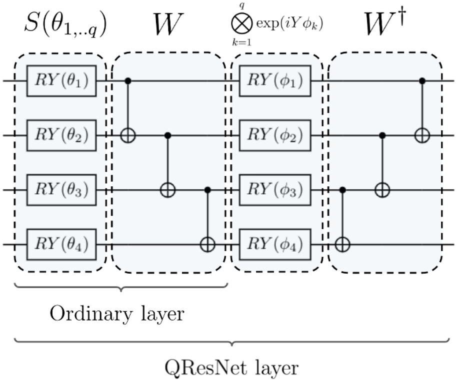
$\Rightarrow$ The expressive power can be easily adjusted by controlling the parameter

$$\alpha_l = \frac{1}{2^q}\mathrm{Tr}\left[\tilde{W}_l(\phi_l)\right] = \prod_{j=1}^q cos(\phi_{jl}) \approx 1 - \frac{1}{2}\sum_{j=1}^q \phi_{jl}^2 + \dots$$

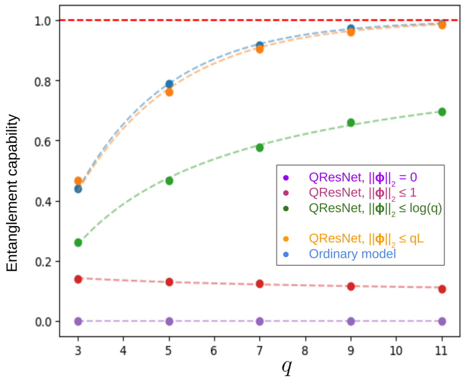
It is easy to identify an initialization strategy that ensures trainability:

- If $||\phi_l||_2^2 = 0$, then $G(\boldsymbol{\theta}, \mathbf{0}) = \prod_{l=1}^L S(\boldsymbol{\theta}_l)$ which is *separable at initialization* regardless of the scaling $L(q)$
- If $||\phi_l||_2^2$ but *small enough*, we may have a *low enough entanglement initialization* to avoid concentration phenomena. Intuitively (and empirically), one expects $||\phi_l||_2^2 \in O(\log(q)/L)$ to suffice for this.

A practical example



(a) Single layer of the considered models.

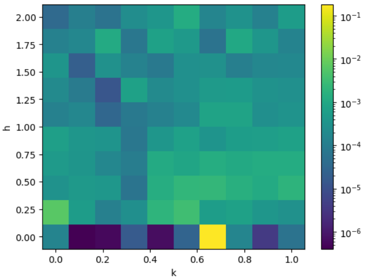


(b) Entanglement capability measure.

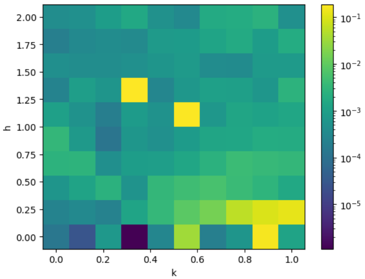
Figure 2: Setting of the numerical tests.

ANNNI model ground state energy estimation

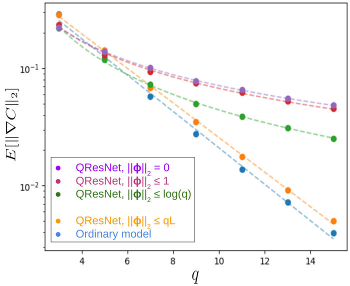
$$H_{\kappa,h} = \sum_{i=1}^{N} \sigma_x^i \sigma_x^{i+1} - \kappa \sigma_x^i \sigma_x^{i+2} + h\sigma_z^i \tag{3}$$



(a) $q = 6$



(b) $q = 12$



(c) Gradient norm scaling

Figure 3: ANNNI VQE, numerical results.

Thank you for your attention!

Any questions?

Contact: giulio.crognaletti@phd.units.it