Towards a Bayesian Variational Quantum Eigensolver

Arthur G. Rattew
Roman Garnett
Washington University in St. Louis

Introduction

- Quantum hardware is currently in the noisy intermediate-scale quantum (NISQ) era.
- NISQ hardware has several limitations which require algorithms to be specifically designed for the paradigm (so no breaking SHA256 yet).
- However, we are just entering a period where these devices may have practical applications.
- Google's (justifiably disputed) "quantum supremacy" demonstration around a month ago underscores how these devices can outperform the best supercomputers.
- We will discuss a sub-problem of a specific paradigm of NISQ-algorithms, called variational quantum eigensolvers.
- Specifically, we will concern ourselves with the optimization of the parameters in a quantum circuit to optimize a more expensive objective function.

Background

You are given a cost function of the form:

$$f : \mathbb{C}^N \mapsto \mathbb{R}.$$

You are also given a parameterized quantum circuit with parameters given by $\boldsymbol{\theta} \in \mathbb{R}^d$, with $d << N$. The circuit outputs a normalized vector, $\boldsymbol{x_\theta} \in \mathbb{C}^N$ which only depends on the input parameters. Then,

the cost function is defined as:

$$f(\boldsymbol{\theta}) \equiv f(\boldsymbol{x_\theta}) = \boldsymbol{x_\theta}^\dagger H \boldsymbol{x_\theta},$$

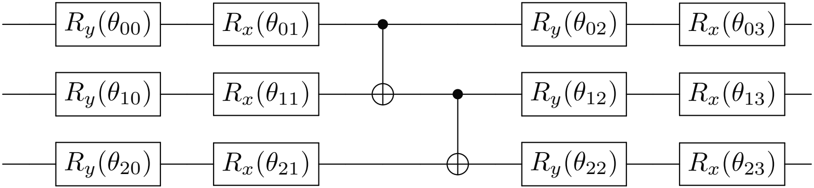
where H encodes a desired optimization problem and is Hermitian ($H = H^\dagger$). Here we are simply making explicit that the cost function only depends on the parameter vector. The problem may then be

formally stated as identifying $\boldsymbol{\theta}^*$ such that:

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \left(f(\boldsymbol{x_\theta}) \right).$$

Background – Example

For instance, a simple quantum parameterized circuit acting on a 2^3-dimensional complex vector may look like:



For the sake of this talk, you may pretend that the above is simply a neural-network.

Background

- Essentially, this problem reduces to the optimization of a parameter vector.
- This is a non-convex, noisy optimization problem. Fortunately, a quantum property, entanglement, can help reduce the difficulty of this task.
- Our preliminary approach extends the algorithm RotoSolve, as presented by Ostaszewski *et al.* in 2019 [2].

RotoSolve

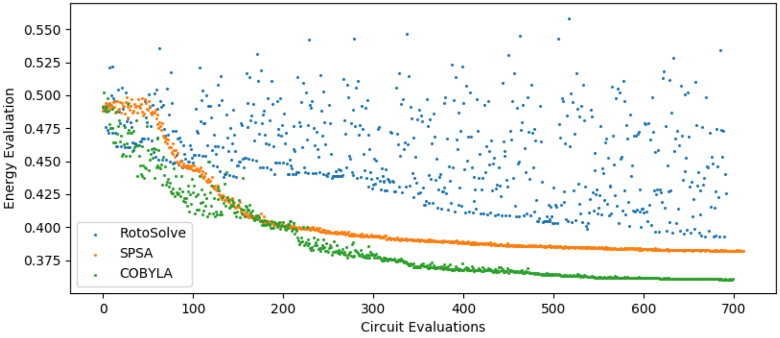
- RotoSolve relies on the optimal configuration of a single parameter in the circuit at a time.
- The algorithm then sequentially cycles through all the parameters in a given circuit, optimally setting each one (while holding all other parameters fixed) until convergence.
- For reasons related to the definitions of quantum gates, the cost function of a circuit with all parameters apart from θ_i held constant may be written in the form:

  $$f(\theta_i) = a \sin(\theta_i + b) + c,$$

  where a, b and c are all real.
- This equation assumes all gates in the circuit are from the set $\{R_x, R_y, R_z\}$.

RotoSolve – In Defense of Greedy Optimization

- Although RotoSolve is greedy, it does not perform substantially worse than other frequently used optimizers.



RotoSolve – Single Parameter Optimization

$$f(\theta_i) = a\sin(\theta_i + b) + c \tag{1}$$

- As shown by Ostaszewski *et al.*, the optimal value for θ_i in Eq. 1 may be calculated analytically with three calls to the cost function (assuming no noise is present):

$$\theta_i^* = -\frac{\pi}{2} - \operatorname{arctan2}\left(2f(0) - f(\frac{\pi}{2}) - f(-\frac{\pi}{2}), f(\frac{\pi}{2}) - f(-\frac{\pi}{2})\right). \tag{2}$$

Enter Bayesian Optimization

- Actual quantum computers are noisy and so the analytical expression shown in Eq. 2 rapidly loses accuracy when the noise level is increased.
- However, we believe that the ability to optimally set a single parameter in a circuit at a time, although greedy, may be a useful subroutine in an algorithm optimizing all of the parameters in a quantum circuit.
- Consequently, we built a noise resistant (probably optimal?) single-parameter Bayesian optimizer (NRSPBO) to find θ_i^* when the evaluation of $f(\theta_i)$ is noisy.
- Assumed additive Gaussian noise.

NRSPBO

- Without loss of generality, we will assume the parameter vector only contains the element θ_i, and so we will simply write it as θ.
- We use a GP to model the underlying sinusoidal function.
- Following the work presented by Adler in 1981 we define a kernel of the form:

$$K(\theta, \theta'; \alpha^2, s^2) = \alpha^2 \cos(\theta - \theta') + s^2,$$

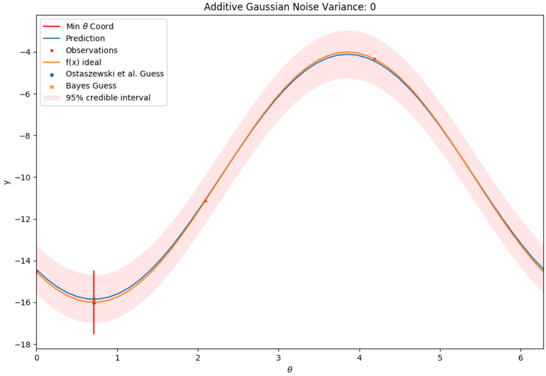
  noting that the s^2 term accounts for the mean of the function, and so we simply use a zero mean [1].
- Thus we obtain the GP:

$$p(f(\theta) \mid \alpha^2, s^2) = \mathcal{GP}(f(\theta), \mathbf{0}, \alpha^2 \cos(\theta - \theta') + s^2),$$

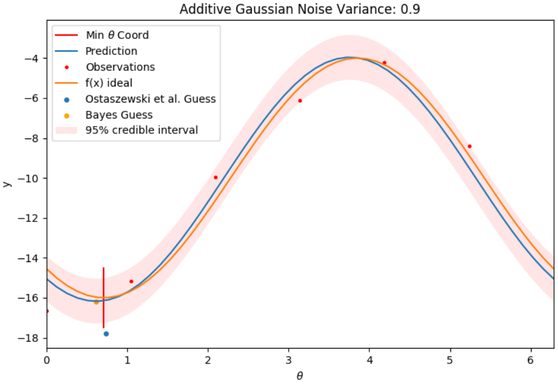
NRSPBO – Note

- We have yet to implement an acquisition function to select which points we should test next, so all of the following results are obtained by using the predictive distribution generated by fitting the model to a pre-selected training dataset.

NRSPBO – Noiseless

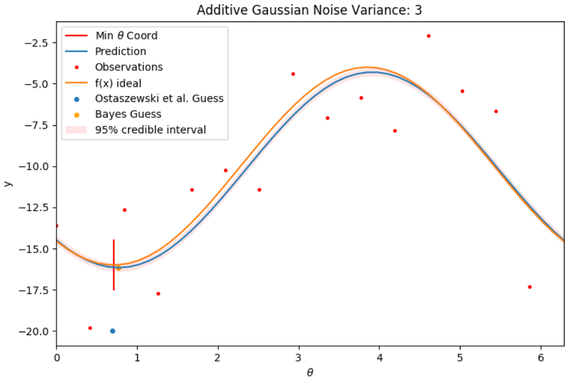


NRSPBO – Noise



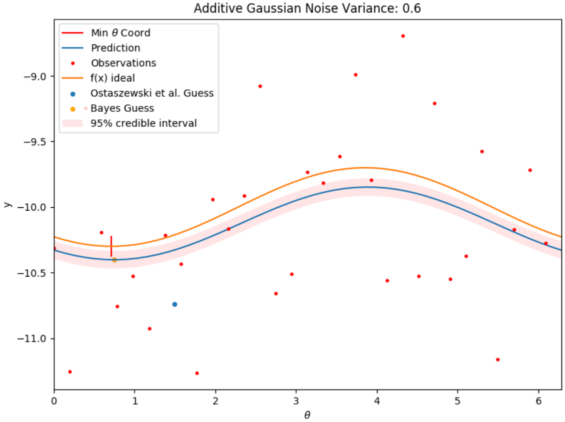
Fairly high signal-to-noise ratio, 6 samples.

NRSPBO – Noise



Moderate signal-to-noise ratio, 15 samples.

NRSPBO – Noise



Very low signal-to-noise ratio, 32 samples.

A Bayesian Variational Quantum Eigensolver

- The ultimate goal is to adaptively construct and optimize a variational quantum circuit.
- The component just discussed will be used as a subroutine.
- Will adaptively grow quantum circuits, to do so, we need to define a kernel giving the “similarity” of any two circuits.

Thanks

R.J. Adler. *The Geometry of Random Fields*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 1981. ISBN: 9780898718980. URL: https://books.google.com/books?id=8a_-VEL7GUYC.

Mateusz Ostaszewski, Edward Grant, and Marcello Benedetti. "Quantum circuit structure learning". In: *arXiv preprint arXiv:1905.09692* (2019).