Parameterized Quantum Circuits and Bayesian Optimization for violating the Bell Inequality

Giovanni Bindi

Università degli Studi di Firenze

19 January 2021

Introduction

Aim

Win a formulation of the CHSH game

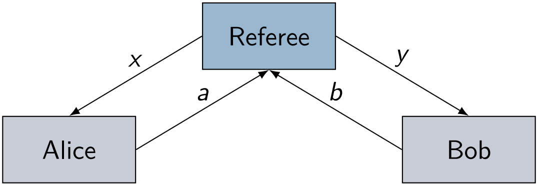
Methodology

Training a simple parameterized quantum circuit

Data

Outputs of quantum circuits simulations

Problem Formulation: CHSH Game [1]



1. The referee sends a bit x to Alice and a bit y to Bob. These bits are drawn from two Bernoulli-distributed random variables:

$$X \sim \mathcal{B}(0.5), \quad Y \sim \mathcal{B}(0.5)$$

2. Alice and Bob send back, respectively, bits a and b.

Alice and Bob **win the game** if[1]:

$$\boxed{x \wedge y = a \oplus b} \tag{1}$$

[1]Where $\wedge$ denotes the logical *AND* and $\oplus$ the logical *XOR* operations.

Strategy [1]

Naming $p_{AB|XY}(a, b|x, y)$ the conditional distribution which relates to the **strategy** operated by Alice and Bob, we define the winning function $V : \{0,1\}^4 \to \{0,1\}$ as:

$$V(x, y, a, b) := \begin{cases} 1 & \text{if } x \wedge y = a \oplus b, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

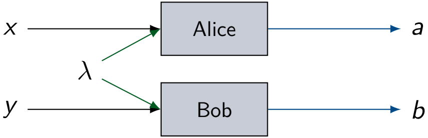
and, hence, express the probability of winning the CHSH game as

$$\begin{aligned} \mathbf{Pr}\{\text{win}\} &= \sum_{a,b,x,y} V(x, y, a, b) p_{AB|XY}(a, b|x, y) p_{XY}(x, y) = \\ &= \frac{1}{4} \sum_{a,b,x,y} V(x, y, a, b) p_{AB|XY}(a, b|x, y) \end{aligned} \tag{3}$$

Assumption: there is a random variable Λ that describes the strategy:

$$p_{AB|XY}(a, b|x, y) = \int \mathrm{d}\lambda\, p_{AB|\Lambda XY}(a, b|\lambda, x, y) p_{\Lambda|XY}(\lambda|x, y) \tag{4}$$

Classical Strategy [1]



We have conditional independence:

$$\begin{cases} p_{\Lambda|XY} = p_{\Lambda}(\lambda) \\ p_{AB|\Lambda XY} = p_{A|\Lambda X}(a|\lambda, x)p_{B|\Lambda Y}(b|\lambda, y) \end{cases}$$

Also it is always possibile to write:

$$\begin{cases} p_{A|\Lambda X}(a|\lambda, x) = \int \mathrm{d}n\, f(a|\lambda, x, n)p_N(n) \\ p_{B|\Lambda Y}(b|\lambda, y) = \int \mathrm{d}m\, g(b|\lambda, y, m)p_M(m) \end{cases}$$

$$\begin{aligned} p_{AB|XY} &= \int\int\int \mathrm{d}\lambda\, \mathrm{d}n\, \mathrm{d}m\, f(a|\lambda, x, n)g(b|\lambda, y, m)p_{\Lambda}(\lambda)p_N(n)p_M(m) \\ &= \int \mathrm{d}\lambda\, f'(a|\lambda, x)g'(b|\lambda, y)p_{\Lambda}(\lambda) \end{aligned} \tag{5}$$

where f, g, f' and g' are deterministic (binary-valued) functions.

Classical Strategy - II [1]

We can then write

$$
\begin{aligned}
\mathbf{Pr}\{\text{win}\} &= \frac{1}{4}\sum_{a,b,x,y} V(x,y,a,b)\,p_{AB|XY}(a,b|x,y) \\
&= \frac{1}{4}\sum_{a,b,x,y} V(x,y,a,b)\int \mathrm{d}\lambda\, f'(a|\lambda,x)g'(b|\lambda,y)p_{\Lambda}(\lambda) \\
&= \int \mathrm{d}\lambda\, p_{\Lambda}(\lambda)\left[\frac{1}{4}\sum_{a,b,x,y} V(x,y,a,b)f'(a|\lambda,x)g'(b|\lambda,y)\right] \\
&\leq \frac{1}{4}\sum_{a,b,x,y} V(x,y,a,b)f'(a|\lambda^*,x)g'(b|\lambda^*,y) \qquad (6)
\end{aligned}
$$

and conclude that deterministic strategies are optimal among all classical strategies.

Deterministic Classical Strategy [1]

Naming a_x (b_y) the bit that Alice (Bob) sends once received the bit x (y), we can list all possible outcome of the CHSH game:

x	y	$x \wedge y$	$= a_x \oplus b_y$
0	0	0	$= a_0 \oplus b_0$
0	1	0	$= a_0 \oplus b_1$
1	0	0	$= a_1 \oplus b_1$
1	1	1	$= a_1 \oplus b_1$

We can see that, for a *deterministic classical strategy*, we have

$$\mathbf{Pr}\{\text{win}\} = \frac{1}{4} \sum_{a,b,x,y} V(x,y,a,b) p_{AB|XY}(a,b|x,y) \leq \frac{3}{4} \tag{7}$$

And that

$$\alpha := \mathbf{Pr}\{\text{win}\} - \mathbf{Pr}\{\text{lose}\} \implies \|4\alpha\| \leq 2 \tag{8}$$

Quantum Strategy [1]

Parameter λ can correspond to a shared quantum state $|\phi\rangle_{AB}$.
Alice's x-dependent local projective measurements: $\{\Pi_a^{(x)}\} : \sum_a \Pi_a^{(x)} = I$
Bob's y-dependent local projective measurements: $\{\Pi_b^{(y)}\} : \sum_b \Pi_b^{(y)} = I$
We then have

$$p_{AB|XY}(a, b|x, y) = \langle\phi|_{AB}\Pi_a^{(x)} \otimes \Pi_b^{(y)}|\phi\rangle_{AB} \tag{9}$$

and we can re-write Eq. 4 as

$$\mathbf{Pr}\{\text{win}\} = \frac{1}{4}\sum_{a,b,x,y} V(x, y, a, b)\langle\phi|_{AB}\Pi_a^{(x)} \otimes \Pi_b^{(y)}|\phi\rangle_{AB} \tag{10}$$

We define $A^{(x)} := \Pi_0^{(x)} - \Pi_1^{(x)}$ and $B^{(y)} := \Pi_0^{(y)} - \Pi_1^{(y)}$ and:

$$C_{AB} := A^{(0)} \otimes B^{(0)} + A^{(0)} \otimes B^{(1)} + A^{(1)} \otimes B^{(0)} - A^{(1)} \otimes B^{(1)} \tag{11}$$

Quantum Strategy - II [1], [2]

It is possible to show that

$$\alpha = \mathbf{Pr}\{\text{win}\} - \mathbf{Pr}\{\text{lose}\} = \frac{1}{4} \langle\phi|_{AB}\, C_{AB}\, |\phi\rangle_{AB} \tag{12}$$

and also that

$$\|C_{AB}\|_\infty \leq 2\sqrt{2} \tag{13}$$

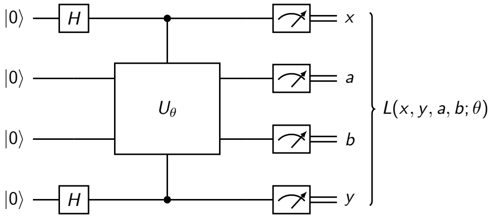
This implies that

$$\|4\alpha\|_\infty \leq 2\sqrt{2} \tag{14}$$

which is known as the *Tsirelson's bound*. This result translates to a larger maximum winning probability:

$$\mathbf{Pr}\{\text{win}\} \leq \frac{1}{2} + \frac{\sqrt{2}}{4} \approx 0.835 \tag{15}$$

General Idea



- The idea is to learn an optimal unitary such that some loss function is maximized.
- Relate the loss function to the possibility of winning the CHSH game.

An obvious choice is just the winning frequency on N simulations:

$$L(x, y, a, b; \theta) = \frac{1}{N} \sum_{i=1}^{N} V_\theta(x, y, a, b) \tag{16}$$

Parameterized Quantum Circuits [3]

$$|\psi_0\rangle \equiv \boxed{U_\theta} \equiv |\psi_\theta\rangle$$

We define a parameterized quantum circuit (PQC) as a parameterized unitary operation on n qubits, U_θ:

$$|\psi_\theta\rangle = U_\theta\,|\psi_0\rangle \tag{17}$$

where, usually, $|\psi_0\rangle = |0\rangle^{\otimes n}$. Sometimes compared to neural networks:

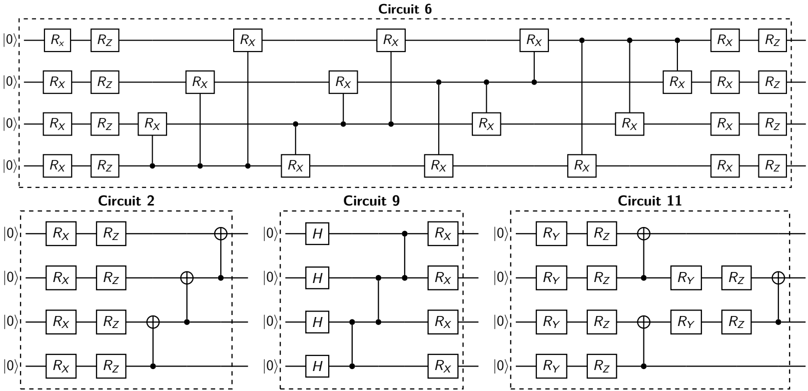
Similarities

- Universal approximator.
- Number of parameters scales polinomially in n.

Differences

- Unitary.
- Internal state not accessible.

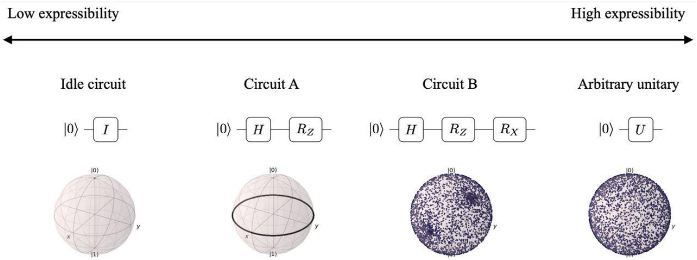
Parameterized Quantum Circuits: Some Examples [4]



Expressibility and **Entangling Capability** studied on 19 (4-qubits) circuits wrt the increasing number of consecutive layers (from 1 to 4).

Circuit Descriptors : Expressibility [4]

Idea: Quantify the ability of a PQC to generate states of an Hilbert space. The simplest case being $\mathcal{H} = \mathbb{C}^2$:



Estimated through the KL divergence between two distributions of fidelities: the first one coming from the PQC while the latter from the Haar ensemble:

$$\mathrm{Exp} := D_{\mathrm{KL}}(P_{\mathrm{PQC}}(F;\theta) \| P_{\mathrm{Haar}}(F)) \tag{18}$$

Circuit Descriptors: Entangling Capability [4], [5]

The entangling capability of a PQC is estimated through the Meyer-Wallach measure, here presented in the form given by Brennen. For a state $|\psi\rangle \in \mathbb{C}^{2n}$ we define

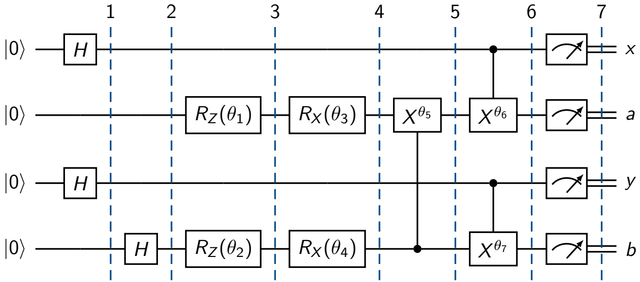
$$Q(|\psi\rangle) := \frac{1}{n} \sum_{k=1}^{n} 2(1 - \mathrm{Tr}[\rho_k^2]) \tag{19}$$

where ρ_k is the one-qubit reduced density operator of the k-th qubit after tracing out the rest. Being $|\psi_\theta\rangle = U_\theta |\psi_0\rangle$ we define the entangling capability as

$$\mathrm{Ent} := \frac{1}{|S|} \sum_{\theta \in S} Q(|\psi_\theta\rangle) \tag{20}$$

where $S = \{\theta_i\}_{i=0}^{N}$ is the set of sampled circuit parameters.

The PQC in detail



Parameters $\theta_1, \theta_2, \theta_3$ and θ_4 have been constrained inside $[0, 2\pi]$ while θ_5, θ_6 and θ_7 inside $[-2, 2]$.

Cost: There are 10 gates: 7 single-qubit gates and 3 two-qubit gates. The parametric gates are 4 rotations and 3 CNOTs.

Expressibility: Exp ≈ 0.03.

Entangling capability: Ent ≈ 0.35.

Bayesian Optimization [6]

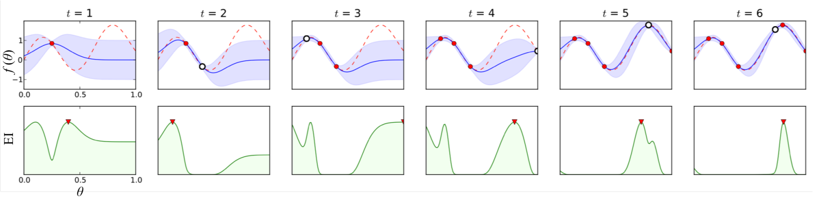
Aim: Find

$$\theta^* = \arg\max_{\theta \in \Theta} L(x, y, a, b; \theta) \tag{21}$$

Idea: Attack the problem with Bayesian Optimization:

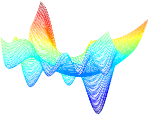
- Objective function is costly to evaluate.
- We don't have access to derivatives.
- The parameter space is "small" (usually $dim(\Theta) < 20$).

Actors involved: a prior distribution, a posterior distribution and an *acquisition function*.



Simulation Setup

colab Cirq

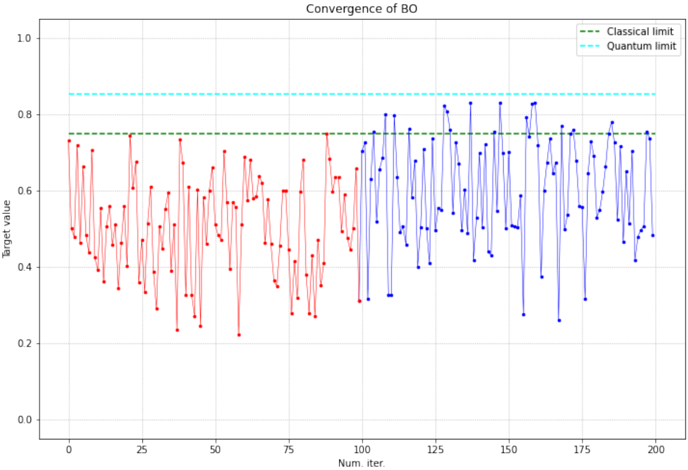


Pseudo-code

```
Init PQC
Init Simulator
for i = 1, ..., N do
    Sample θ_i ∈ Θ according to the acquisition function.
    for j = 1, ..., M do
        Run the simulation and collect data.
    end for
    Update prior and posterior based on data.
end for
return Best set of parameters θ*.
```

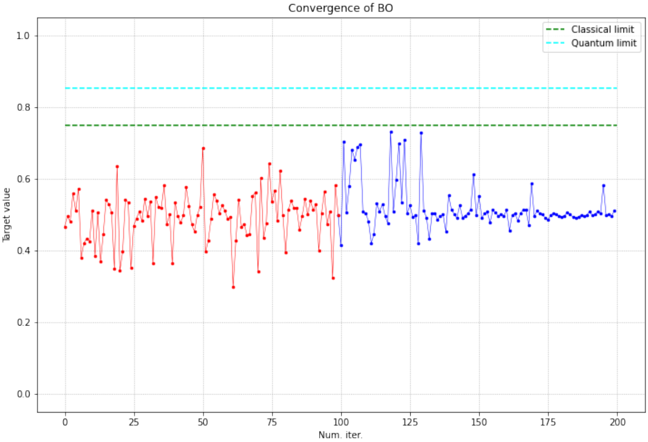
Results

Using $N = 200$ optimization steps and $M = 5000$ circuit simulations per step:



Results - II

Run the previous algorithm with a circuit with no entangling capability (Ent $= 0.$) and comparable expressibility (Exp ≈ 0.04). This PCQ has 22 gates: 20 one-qubit gates and 2 two-qubit gates.

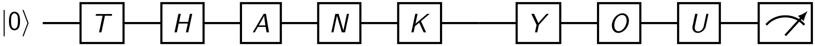


Conclusions

- Joint work of PQC and classical optimization $\rightarrow$ win the CHSH game
- Goal: show that entangle can emerge naturally from the setup
- More of a PoC than a real experiment
- Provide a simple setup for future and more complex ideas

Look at the code

https://github.com/w00zie/pqc_chsh


|0⟩
T
H
A
N
K
Y
O
U

References I

[1] M. M. Wilde, "Preface to the second edition," *Quantum Information Theory*, pp. xi–xii, DOI: 10.1017/9781316809976.001. [Online]. Available: http://dx.doi.org/10.1017/9781316809976.001.

[2] B. S. Cirel'Son, "Quantum generalizations of Bell's inequality," *Letters in Mathematical Physics*, vol. 4, no. 2, pp. 93–100, Mar. 1980. DOI: 10.1007/BF00417500.

[3] M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, "Parameterized quantum circuits as machine learning models," *Quantum Science and Technology*, vol. 4, no. 4, p. 043 001, Nov. 2019, ISSN: 2058-9565. DOI: 10.1088/2058-9565/ab4eb5. [Online]. Available: http://dx.doi.org/10.1088/2058-9565/ab4eb5.

References II

[4] S. Sim, P. D. Johnson, and A. Aspuru-Guzik, "Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms," *Advanced Quantum Technologies*, vol. 2, no. 12, p. 1 900 070, Oct. 2019, ISSN: 2511-9044. DOI: 10.1002/qute.201900070. [Online]. Available: http://dx.doi.org/10.1002/qute.201900070.

[5] G. K. Brennen, "An observable measure of entanglement for pure states of multi-qubit systems," *Quantum Info. Comput.*, vol. 3, no. 6, pp. 619–626, Nov. 2003, ISSN: 1533-7146.

[6] P. I. Frazier, *A tutorial on bayesian optimization*, 2018. arXiv: 1807.02811 [stat.ML].

Maximum Quantum Winning Probability - 1

When the inputs x and y are $00, 01, 10$ Alice and Bob win if they report back the same restult. The probability is

$$\langle\phi|_{AB}\,\Pi_0^{(x)} \otimes \Pi_0^{(y)}\,|\phi\rangle_{AB} + \langle\phi|_{AB}\,\Pi_1^{(x)} \otimes \Pi_1^{(y)}\,|\phi\rangle_{AB}$$

The probability for it not to happen is

$$\langle\phi|_{AB}\,\Pi_0^{(x)} \otimes \Pi_1^{(y)}\,|\phi\rangle_{AB} + \langle\phi|_{AB}\,\Pi_1^{(x)} \otimes \Pi_0^{(y)}\,|\phi\rangle_{AB}$$

So in this case $\mathbf{Pr}\{\text{win}\}$ - $\mathbf{Pr}\{\text{lose}\} = \langle\phi|_{AB}\,A^{(x)} \otimes B^{(y)}\,|\phi\rangle_{AB}$

When $x = 1$ and $y = 1$ we have that

$$\mathbf{Pr}\{\text{win}\} - \mathbf{Pr}\{\text{lose}\} = -\langle\phi|_{AB}\,A^{(1)} \otimes B^{(1)}\,|\phi\rangle_{AB}$$

So, averaging over all values of the input bits

$$\alpha := \mathbf{Pr}\{\text{win}\} - \mathbf{Pr}\{\text{lose}\} = \frac{1}{4}\,\langle\phi|_{AB}\,C_{AB}\,|\phi\rangle_{AB}$$

We have

$$C_{AB}^2 = 4I_{AB} - \left[A^{(0)}, A^{(1)}\right] \otimes \left[B^{(0)}, B^{(1)}\right]$$

and

$$\begin{aligned} \|C_{AB}^2\|_\infty &= \|4I_{AB} - \left[A^{(0)}, A^{(1)}\right] \otimes \left[B^{(0)}, B^{(1)}\right]\|_\infty \\ &\leq 4\|I_{AB}\|_\infty + \|\left[A^{(0)}, A^{(1)}\right] \otimes \left[B^{(0)}, B^{(1)}\right]\|_\infty \\ &= 4 + \|\left[A^{(0)}, A^{(1)}\right]\|_\infty \|\left[B^{(0)}, B^{(1)}\right]\|_\infty \\ &\leq 4 + 2 \cdot 2 = 8 \end{aligned}$$

So $\|C_{AB}\|_\infty \leq 2\sqrt{2}$ and

$$\|4\alpha\|_\infty = \| \langle\phi|_{AB}\, C_{AB}\, |\phi\rangle_{AB} \|_\infty \leq 2\sqrt{2}$$

Expressibility Calculation

Expressibility is

$$\text{Exp} := D_{\text{KL}}(P_{\text{PQC}}(F;\theta) \| P_{\text{Haar}}(F))$$

where, being $d = dim(\mathcal{H})$

$$P_{\text{Haar}}(F) = (d-1)(1-F)^{d-2} \tag{22}$$

Expressibility

1: Init PQC
2: Init Simulator
3: **for** $i = 1, \dots, M$ **do**
4: Sample two sets of parameters θ_i, θ_j uniformly in Θ.
5: Calculate fidelity between the two output states.
6: **end for**
7: Calculate KL between fidelities and (22)

Expressibility Calculation - 2

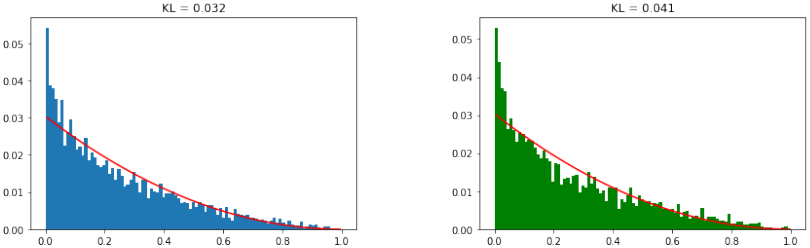


Figure 1: Entangling circuit (left) and non-entangling circuit (right)

Estimated with $M = 5000$ samples per circuit, compared to the theoretical density function (22) in red.