

Task 2

February 22, 2020

1 The Problem

Implement a circuit which returns $|00\rangle$ and $|11\rangle$ with equal probability.

Requirements:

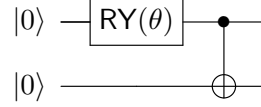
- Circuit should consist only of CNOTs, RXs and RYs. Start from all parameters in parametric gates being equal to 0. You should find the right set of parameters using gradient descent (you might use more advanced optimization methods if you like).
- Simulations must be done with sampling - i.e. limited number of measurements per iteration and noise.
- Compare the results for different numbers of measurements: 1, 10, 100, 1000.

Bonus question:

How to make sure you produce state $|00\rangle + |11\rangle$ and not $|00\rangle - |11\rangle$?

2 Solution

Here I present a solution using two qubits. The basic idea is to use a parameterized RY(θ) gate and use a controlled-not gate to copy the phase-shifted state.



The parameterized gate is defined as:

$$\text{RY}(\theta) = \exp\left(-\frac{i\theta}{2}\text{Y}\right) = \cos(\theta/2) \mathbb{I} - i\sin(\theta/2)\text{Y} \quad (1)$$

so that when we apply it to one of the initial qubits, we get:

$$\text{RY}(\theta) |0\rangle = \cos(\theta/2) |0\rangle + \sin(\theta/2) |1\rangle \quad (2)$$

and thus the total state so far is:

$$|\psi(\theta)\rangle = \cos(\theta/2) |00\rangle + \sin(\theta/2) |10\rangle \quad (3)$$

Next, the CNOT gate will flip the second qubit only in the second term (because of linearity). Our output state will be:

$$|\psi_{\text{out}}(\theta)\rangle = \text{CNOT} |\psi(\theta)\rangle = \cos(\theta/2) |00\rangle + \sin(\theta/2) |11\rangle \quad (4)$$

We have now control over the output state by varying the parameter θ : when $\theta = 0$ we can produce the $|00\rangle$ state only and if $\theta = \pi/2$, we get:

$$|\psi_{\text{out}}(\pi/2)\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \quad (5)$$

which is a state with equal probabilities for $|00\rangle$ and $|11\rangle$:

$$|\langle 00|\psi_{\text{out}}\rangle|^2 = |\langle 11|\psi_{\text{out}}\rangle|^2 = 0.5 \quad (6)$$

Notice that if we set $\theta = -\pi/2$, we will obtain the antisymmetric state $|\psi_{\text{out}}(-\pi/2)\rangle = \frac{1}{\sqrt{2}} (|00\rangle - |11\rangle)$, which still gives equal probabilities.

3 Implementation

When implementing this solution on software, let's forget for a moment we know the desired values of the parameter and let's look for it using gradient descent. I have used the open source library *Cirq*, and its basic *Controlled-not* and *Rotate-Y* gates. Also, I have introduced noise channels that apply some Pauli gate with a probability p and the identity with probability $1 - p$. The code is in a [Jupyter Notebook](#) done with coLab, because I think it's easier to read it in blocks and it should be quick to access online.

The gradient descent method allows us to vary the parameter in such a way that a desired function has zero derivative. Here I used a function that adds the different amplitudes as measured by the sampling:

$$\epsilon(\text{counts}) = \text{counts}(|01\rangle) + \text{counts}(|10\rangle) + (\text{counts}(|00\rangle) - \text{counts}(|11\rangle))^2 \quad (7)$$

where I divided $\text{counts}(|\cdot\rangle) = \text{frequencies}(|\cdot\rangle)/N_{\text{samples}}$ to make the errors commensurate with the probabilities, and where there's an additional term to try to reduce the difference between the desired states.

The strategy now is to update the parameter θ (which we can initialize randomly between $(-\pi, \pi)$) with the variation in the function ϵ :

$$\Delta\epsilon = \epsilon(\text{new counts}) - \epsilon(\text{prev. counts}) \quad (8)$$

$$\theta_{\text{new}} = \theta_{\text{old}} - \gamma\Delta\epsilon - \text{sign}(\Delta\epsilon)\epsilon \quad (9)$$

where the last term allows us to avoid a local maximum by penalizing (according to the sign of the variation) a high error. We choose a small step γ which here we set by hand (e.g. 0.01), and iterate the procedure until $\Delta\epsilon$ is below a desired threshold.

3.1 Bonus

If we wish to ensure we obtain the symmetric state $|\psi_{\text{out}}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, we have to make sure that we don't fall into the negative minimum $\theta = -\pi/2$. We enforce this by adding another penalizing term a negative parameter (which we denote here with a Heaviside Θ function, or unit step):

$$\theta_{\text{new}} = \theta_{\text{old}} - \gamma\Delta\epsilon - \text{sign}(\Delta\epsilon)\epsilon + \Theta(-\theta) \quad (10)$$

4 Results

Here I have chosen a small noise $p = 0.05$ which also gives an order of magnitude for the error threshold: the iterations stop when $\epsilon < 0.05$, or when they are more than 100. The number of samples has to be high enough to ensure that we get a meaningful error function: If we just take 1 sample, the frequency count is just 1 so the error is always 1 ($\Delta\epsilon = 0$ and no gradient descent). We expect that the averages of these counted quantities converge towards the expected value and so we begin to see more coherent results as we increase the number of samples.

Below are some examples of the observed behavior, for $N_{\text{samples}} \in \{10, 100, 1000\}$:

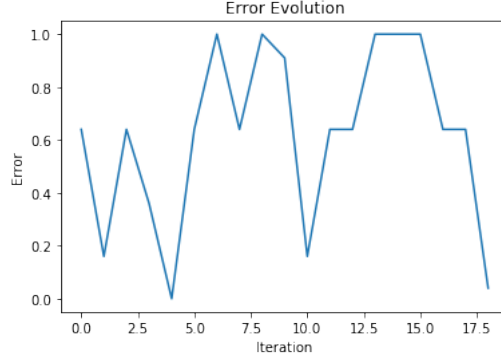


Figure 1: Evolution of the error function with the number of iterations. Here the threshold for success in the variation of the parameter is 0.05, but we also forced the code to run for at least 10 iterations before stopping. It eventually reached threshold again at 18 iterations.

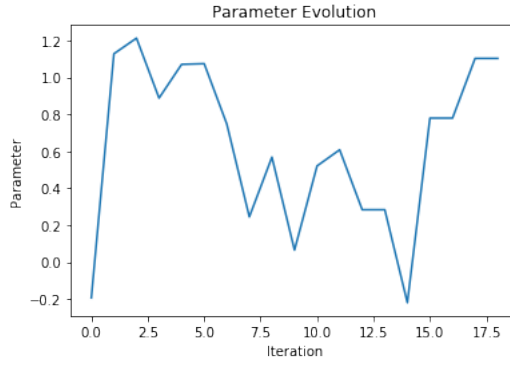


Figure 2: Evolution of the varying parameter θ . Since we are selecting the symmetric state, the evolution converges towards the expected value $\theta = \pi/2$, although the few number of samples can make the error very small without really reaching the best parameters.

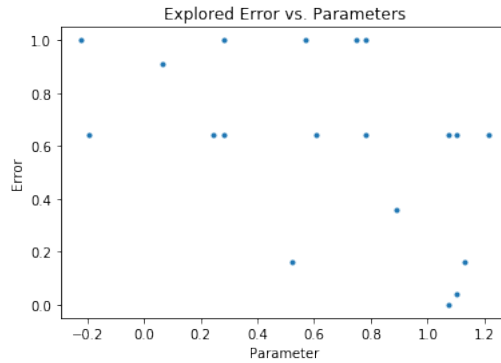


Figure 3: Relationship between the explored parameters and their calculated error, There is no clear relationship that can be inferred, again because of the small number of samples.

```
Counter({'[0 0]': 6, '[1 1]': 4})

Circuit:
(0, 0): —D(0.05)—Ry(0.351π)—@—M('result')—
                    |
(0, 1): —D(0.05)——————X—M—————
          Details:
```

Figure 4: Displayed frequency count and resulting circuit after the iterations. The parameter was close enough so that the counts are more or less similar. The final value for the RY gate was 0.351π .

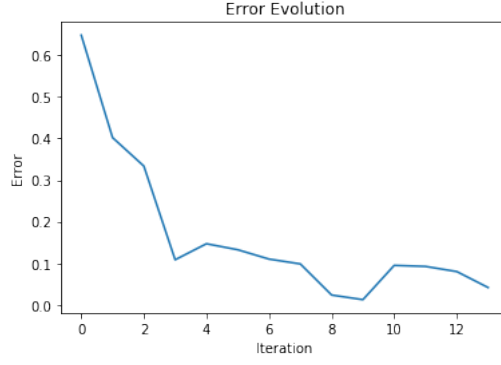


Figure 5: Evolution of the error function with the number of iterations. The threshold for success in the variation of the parameter is 0.05 and the higher number of samples (100) made this a smoother run towards some minimum.

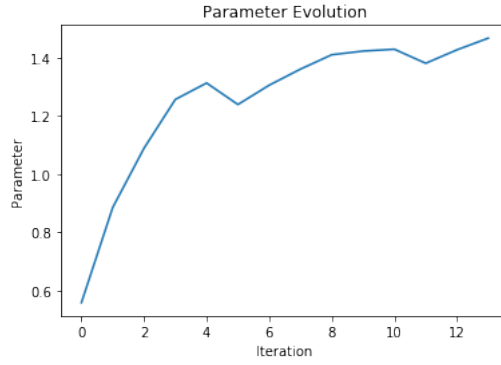


Figure 6: Evolution of the varying parameter θ . Since we are selecting the symmetric state, the evolution converges towards the expected value $\theta = \pi/2$

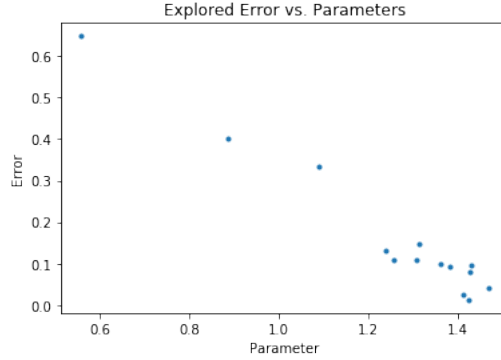


Figure 7: Relationship between the explored parameters and their calculated error, showing both the minimum of convergence and also the maximum error at $\theta = \pi$ that we want to avoid.

```
Counter({'[1 1]': 50, '[0 0]': 46, '[0 1]': 2, '[1 0]': 2})

Circuit:
(0, 0): —D(0.05)—Ry(0.467π)—@—M('result')—
                |
(0, 1): —D(0.05)——————X—M—————
```

Figure 8: Final frequency count and resulting circuit. Notice the counter of states $|01\rangle$ and $|10\rangle$ is not zero because this time the noise in the model (shown by the “D” gates at 0.05 percent) added some extra hits in some samples. The final value for the RY gate was 0.467π , which is what we expect for the symmetric state.

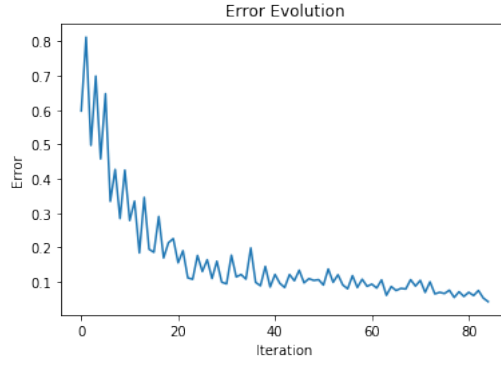


Figure 9: Evolution of the error function with the number of iterations, now for 1000 samples on each run. The threshold for success in the variation of the parameter is 0.05 and the convergence is quite clear, although the noise can cause some wiggling in the measured error.

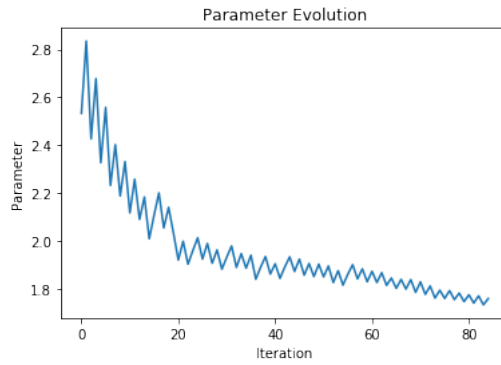


Figure 10: Evolution of the varying parameter θ . Since we are selecting the symmetric state, the evolution converges towards the expected value $\theta = \pi/2$

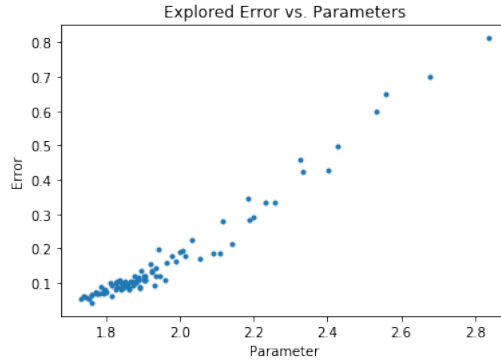


Figure 11: Relationship between the explored parameters and their calculated error, showing both the minimum of convergence and also the maximum error at $\theta = \pi$ that should be avoided, in a different region, since we started at $\theta \sim 2.6$

```
Counter({'[1 1]': 542, '[0 0]': 429, '[1 0]': 15, '[0 1]': 14})

Circuit:
(0, 0): —D(0.05)—Ry(0.56π)—@—M('result')—
(0, 1): —D(0.05)—————X—M————
```

Figure 12: Displayed frequency count and resulting circuit after the iterations. Notice the counter of states $|01\rangle$ and $|10\rangle$ is not zero because we have added noise to the model, shown by the “D” gates at 0.05 percent. The final value for the RY gate was 0.56π , which is what we expect for the symmetric state, and the noise we put into the model.