Quantum Amplitude Damping Channel Discrimination

Students: Albert Solana, Jesús Huerta Mentor: Matteo Rosati

We analyze and implement some strategies to discriminate between two different amplitude damping channels with one-shot executions and reproduce and validate the results against the expected theoretical ones for each approach [1]. The implementation of the discrimination process strategies requires an optimization over the input states and the output measurements (POVM). We explore several one-shot circuit layouts to see how helpful these can be for a 2-channel discrimination and beyond [2], leading to find a limit when more than two channels are involved, where we obtain interesting results not treated in the literature.

This document is structured in 3 chapters: *Quantum Amplitude Damping Channel Discrimination* covers all main concepts acquired and required from a theoretical point of view, while *Discrimination Process Implementation* provides the experimental perspective. It ends with some *Conclusions, applications and possible next steps*. For a detailed Table of Content go to Annex A.

Quantum Amplitude Damping Channel Discrimination

Quantum Amplitude Damping Channel (ADC)

A quantum channel can be seen as a linear mapping between quantum states, as we can describe the channel as a superoperator that will transform an input quantum state to another output quantum state. Noise effects are common when working with quantum states. This happens when the state describing a quantum system, as a whole or just one part of it (for multi-partite systems), evolves or degrades from its initial value to a different state not expected, due to reasons beyond the controlled variables and conditions. In particular, noise effects can decohere quantum states and hence make them potentially useless for applications. This happens in all areas when manipulating a quantum system, for instance when using its state to send information or when modifying the state by interacting with other systems to do calculations. Therefore, it is crucial to know how this phenomenon happens and model it, to find ways to correct it and define the thresholds to consider when a quantum state is still useful. Quantum channels help us describe these noise scenarios, and specifically there are three channels often highlighted [3]-[6]: depolarizing channel, dephasing channel (a.k.a. bit-flip and phase-flip channel) and ADC.

Phenomenon and mathematical representations

An ADC describes the natural transition from an excited quantum state to another one with less energy, so the system dissipates energy. This relates, for instance, to spontaneous emission of a photon by an atom, where the photon released has an energy equal to the difference of energy levels (Figure 1), and to many other scenarios where the system's energy is attenuated over time due to the interaction with an environment, like seeking thermal equilibrium. We say that the system decays, i.e. goes from one energy state to another one less energetic.

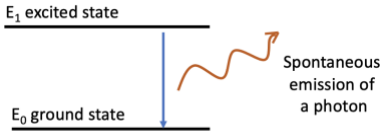


Figure 1 - Spontaneous emission

All these natural processes can be modeled as channels, which transform (or map) one input (initial) state to an output (final) state. This can be represented in several mathematical ways, all of them equivalent. The figure below, taken from [7], shows the operations required to move from one mathematical representation to the other.

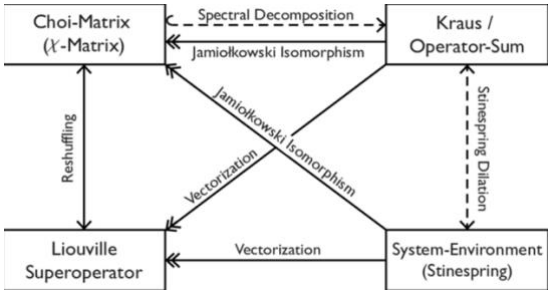


Figure 2 - Main mathematical representations for channels and how to convert them.

ADCs are well known for their effect of damping the amplitude, i.e., reducing the energy. The main parameters to describe the attenuation of the state going through the channel can be λ or η[a].

λ is defined between 0 and 1, representing respectively no action and full attenuation on the excited state. λ and η are related by the expression

$$\lambda = sin^2\,\eta$$

This means, η is defined between 0 and π/2. In our analysis to describe these ADCs, we will use both parameters.

The Stinespring and Kraus operators acting on the input state density matrix (ρ) will lead to a final state density matrix (ρ'), related by the expression:

$$\rho' = E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger \quad , \qquad \text{where } E_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\lambda} \end{bmatrix} \text{ and } E_1 = \begin{bmatrix} 0 & \sqrt{\lambda} \\ 0 & 0 \end{bmatrix}$$

Characteristics

These λ or η are key to describe the ADC, as they determine which kind of state will be found at the end. As λ gets closer to 1, the channel loses or dissipates more energy to the environment. Something we can't do is reverse this effect, as we lose quantum information when going through the channel. This energy dissipation to the environment also results in decoherence of the quantum state. As we can't control the environment, this information can't be recovered. This effect happens on the excited state, leaving the ground state intact. This results in an asymmetric action, which can be described also as non-unitariety or nonreversibility of the ADC [3].

Modeling an ADC

An ADC represented by η can be modeled in several ways. A beam splitter or a system reaching thermal equilibrium are shown at [4] 8.3.5. We use the circuit below with a RY gate with angle 2η and a CNOT, connected between q0 (qubit) and the environment.

[a] These Greek letters are an arbitrary selection as there is not a common standard to call these variables.

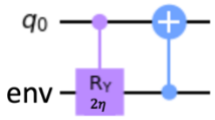


Figure 3 - Circuit reproducing an ADC characterized by η.

The equivalence of this circuit with the ADC is demonstrated in Annex C.

Metrics

Two tools will help to measure the effect of an ADC, to represent and compare the initial and final states, and how this relates with the channel parameter λ or η: Bloch Sphere and Fidelity.

Bloch Sphere Impact

To see how the ADC acts on the different states of a qubit, we pass 200 pure states ($\alpha|0\rangle + \beta|1\rangle$), which can approximate a Bloch Sphere, through 5 different ADC's with $\lambda \in \{0, 0.15, 0.5, 0.79, 1\}$.

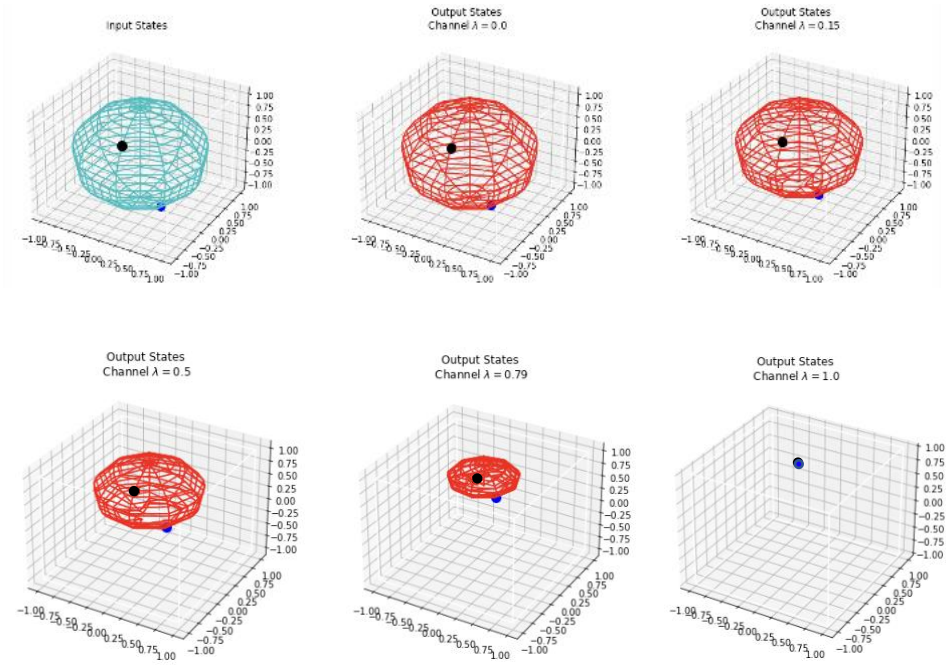


Figure 4 - Examples of the effect of an ADC with different attenuation factors.

The green sphere is the Bloch sphere, represented by all the initial states that go through the ADC's. All these states belong initially to the sphere's surface, which means that they are pure states with a Bloch vector of length 1.

The other "ellipsoids" (in red) represent the final states obtained after crossing one of the five channels. We observe that these channels compress and move up the input states towards the $|0\rangle$ state (ground state) distorting the sphere to an ellipsoid. In fact, this effect is larger as the λ value gets closer to 1, as shown in Figure 4.

- For $\lambda = 0$, the channel has no effect on the initial states. This situation is equivalent to a *lossless system* with no attenuation present.
- For $\lambda = 1$, we see all values converged at $|0\rangle$ state. Any state passing through this channel will lose all excited component ($|1\rangle$ state) and reach the $|0\rangle$ state. Therefore, any measurement in the computational basis gives $|0\rangle$ with 100% probability. We will be unable to recover any

information we put in our initial states or to go back. This scenario can well describe a *completely lossy system*.

- For λ values in-between, the states are compressed, and the center moves up, as the initial amplitudes for |0⟩ (α) are increased in the final states. These final states are mixed states (except those starting at α=1). These are *lossy systems*. Another observable effect is that the resulting figure does not resemble a sphere, because the values for coordinates x and y (horizontal plane) reduce slower than z (vertical axis). In fact, the reduction rate for each axis follows this relationship, [4] formula 8.122:

$$(X_{in}, Y_{in}, Z_{in}) \rightarrow (X_{in} \cos\eta, Y_{in} \cos\eta, Z_{in} \cos^2\eta)$$

We added two sample states to depict the effect of the channels on them. These states, initial and final ones, are represented by two dot points, black and blue. The reader can follow the attenuation effect via these two dots. We can see how the blue dot has a larger compression effect because it is closer to the |1> state than the black one. These two dots allow to see ADC preserves the phase, as no rotation is observed.

Fidelity analysis

Fidelity[b] is a measure of how close or related two states are. The fidelity, F, for two states represented by the density matrices ρ and σ, can be calculated with the following formula:

$$F(\rho,\sigma) = \left(tr\left|\sqrt{\rho}\sigma\sqrt{\rho}\right|\right)^2$$

The maximum value for fidelity is 1, when both states are so related that they are the same. The minimum value is 0, when both states are orthogonal.
The Figure 5 shows the fidelity calculated for a set of ADCs from λ = 0 to λ = 1 comparing the output state vs. the input state (left) and output state vs. the |0⟩ state (right).

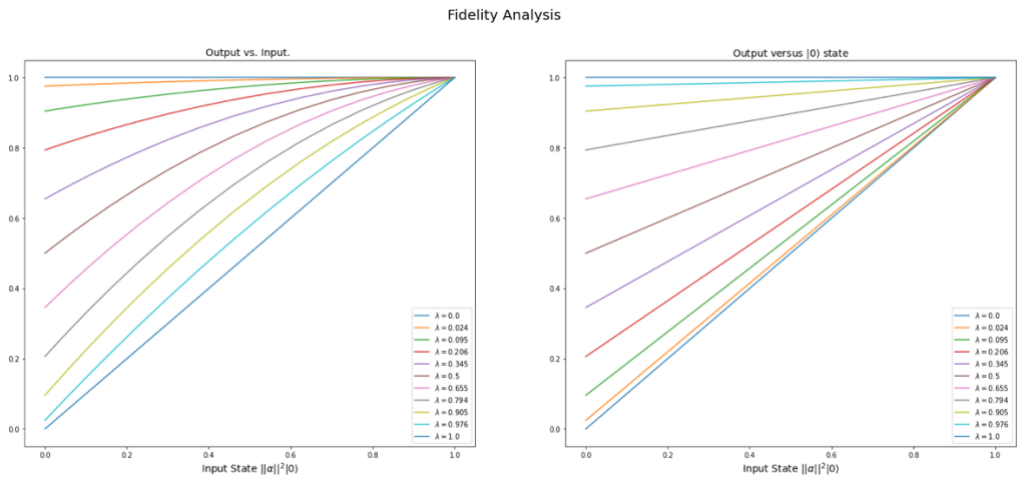


Figure 5 - Fidelity values between input and output states (left) for a set of attenuation factors (λ). Fidelity values between state |0⟩ and output states (right) for the same set of λ

On the left (Output vs Input) we see that fidelity for a lossless channel (λ = 0) is a flat line equal to 1, as expected, because the input crosses the channel without any alteration.
For a full lossy channel (λ = 1), fidelity is represented by a diagonal line, as it is 1 when the input state is |0⟩ and reach the 0 when the input is |1⟩. All other fidelities are the remaining lines in between, ordered from top to bottom as the λ values are higher.

[b] An alternative fidelity metric is used later: *square root fidelity* $F'(\rho,\sigma) = tr\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}$

On the right (Output vs |0⟩) we observe the opposite effect seen in left picture. In this case, fidelity for a lossless channel ($\lambda = 0$) is a diagonal line, as the probability for |0⟩ in input state (α^2) is preserved and contributing to the closeness. When α^2 gets to 1, the input is |0⟩ and we obtain F=1.
For a full lossy channel ($\lambda = 1$), fidelity is represented by a flat line equal to 1, because any input ends as a |0⟩. All other fidelities are the remaining lines in between, ordered from top to down as the λ values are lower.

Understanding how noise affects the experiment

All channel experiments are simulated considering ideal noiseless quantum devices. However, this is not the current real-life scenario. On the NISQ era, quantum devices have noise present in each experiment. In this section, we analyze how the experiment would be in noisy devices that do not implement error corrections yet.

The test consists of passing 70 input states to the same channel with $\lambda = 0.4$ on three different devices:

- an ideal noiseless device
- a mocked IBM Q device where you can simulate the noise
- a real IBM Q device

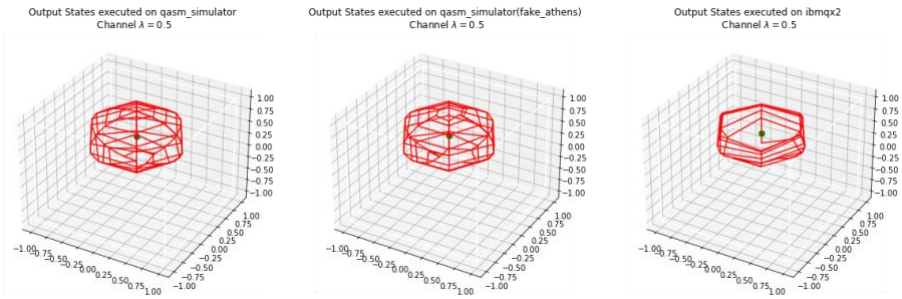


Figure 6 - Results of launching the same experiment using different quantum backend devices. From left to right: QASM ideal simulator, a mocked Athens IBM Q device with simulated noise, and an actual IBM Q device (ibmqx2). The experiment consists of sending 70 input states through an ADC with $\lambda = 0.4$

Comparing the output states' results, we see how the noise effect is translated into deformations, making it more difficult to distinguish the effect of attenuation factor from ADC and the noise from the real device. The deformation varies for each experiment. Sometimes there is nearly no noise present, and the result is much like the mocked device and the ideal one, but sometimes, as the one displayed, the noise is actively impacting the results, and those would not be usable for the project.

Discriminating ADCs

An ADC discrimination exercise can be described as the ability to determine which one of the given ADC's is present, among a discrete set of different and equiprobable ADCs. There exist several strategies to accomplish it, which try to maximize hits, i.e., the success probability. Basically, all of them seek an optimal initial state, such that the output at the end of any of the possible channels, after being prepared to be measured (POVM), provides measurement results that best help to discriminate the channels and hit the one present.

It has similitudes with the state discrimination problem because both search for an optimal POVM. However, discriminating channels is more complex, as the input state is also required in the optimization analysis. This leads to an increase in the number of variables to optimize depending on the strategies considered to solve the problem. Some of these strategies differ by the number of uses of the channel, from one-shot (only one use) or n-shots (we can repeat n times). As the number of uses increases more possibilities appear, because we can consider parallel or consecutive uses, use feedback from

measurement of previous shots to modify the next measurement, perform individual measurements or collective ones, and so on. In addition to that, we have to consider also different configurations of the input state, mainly in the number of qubits, the use of ancillas and entanglement, quickly scaling up the complexity. In fact, in its more general strategy like the adaptative one (consecutive use of the channel allowed), we can say that "*since the seminal work of Helstrom in the 70s, we know how to bound the error probability affecting the symmetric discrimination of two arbitrary quantum states. Remarkably, after about 40 years, a similar bound is still missing for the discrimination of two arbitrary quantum channels*" [8].

Our analysis is limited initially to two channels by using a one-shot scenario, which means we have only one try to discriminate between them. At this point we try to replicate the theoretical results available in [1] with an empirical approach. The circuit which models an ADC is the base of this analysis, and we show below how it evolves to replicate the results successfully based on a maximum likelihood estimate method.

After this, we increase the number of channels to discriminate to three, for which a theoretical solution of the maximum success probability is not known, and compare our results with theoretical bounds [2] to find some limits in the one-shot strategy and maximum likelihood estimate method. Surprisingly, we find that, for a general pure entanglement-assisted input and two-qubit projective measurement, the optimal strategy found by our analysis always guesses for the two most distinguishable channels, giving up on the third one.

We stress that the input states are pure and the measurements considered are projective, to avoid a further increase of optimization parameters. For states, we expect this to be optimal at least for two channels based on theoretical works [9]. For measurements, the binary discrimination problem is always solved by projective measurements [10], while for more than two hypotheses in general we need POVMs.

The following sections show the circuit used and the conditions considered to replicate successfully the results expected for each paper.

One-Shot Strategy: One Single Qubit as an input state

Our initial proposed circuit models the channel at Subsection II A [1]. It has two qubits: one for the input state, to be measured at the end, and another for the environment. Environment is always initialized with the $|0\rangle$ state.

The input state admits any possible pure quantum state,

$$|\Psi\rangle = \cos\theta\,|0\rangle + e^{i\phi}\sin\theta\,|1\rangle$$

The problem we want to analyze allows us to set $\phi = 0$, due to the symmetric action of ADC with respect to the z axis in the Bloch sphere. Taking this into account and by substituting $\sin^2\theta = x$, where $x \in [0, 1]$, we get the following expression for the input state:

$$|\Psi\rangle = \sqrt{1-x}\,|0\rangle + \sqrt{x}\,|1\rangle$$

The measurement should be as general as possible, ideally universal, to find the best configuration to maximize the hits for ADC discrimination. For that reason, we create a projective measurement with two gates, an x-rotation and y-rotation, and a measurement. Thus, we require two additional parameters, ϕ_{rx} and ϕ_{ry}, to define our measurement.

To illustrate it, Figure 7 shows an example of a one single qubit as an input state circuit:

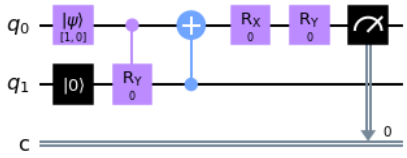


Figure 7 - Sample circuit used for a one single qubit ADC discrimination with values x=0, η=0, ϕ_{rx}=0 and ϕ_{ry}=0

Discriminating ADCs with this circuit becomes then an exercise of finding the optimal input state and measurement for any given two ADCs, characterized by η_0 and η_1, such that they bring the highest average success probability using a one-shot strategy. Therefore, we need to optimize three parameters (x, ϕ_{rx} and ϕ_{ry}) to find the optimal configuration and match the best probabilities provided in [1] to discriminate two ADCs.

One-Shot Strategy: Two Entangled Qubits as an input state

Another approach to discriminate two ADCs consists of using a different circuit layout that takes advantage of the quantum entanglement property. It specifies an input state with two entangled qubits to fit the proposed option at Subsection II B [1], *Side entanglement*, where one of the qubits is sent through the channel and the other acts as an ancillary qubit. At the end we measure both qubits.

The input state $|\Psi\rangle$ uses y as the probability of the state that determines the entanglement, where $y \in [0, 1]$. When $y = 0$, a $|1\rangle$ is sent through the channel and $|0\rangle$ is assigned to the ancilla, and vice versa when $y = 1$.

$$|\Psi\rangle = \sqrt{1-y}\,|01\rangle + \sqrt{y}\,|10\rangle$$

This circuit measurement will also be a projective one with Rx, Ry rotations gates and a measurement for each input qubit, but we add a CNOT gate that connects both qubits, as they are entangled in their input state.

Figure 8 shows an example of a two entangled qubit used in simulations:

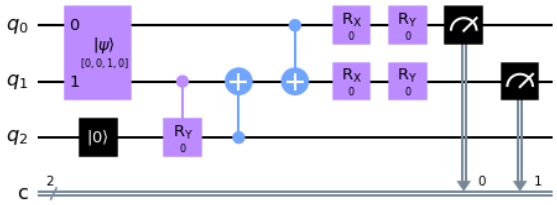


Figure 8 - Sample circuit for a two entangled qubit ADC discrimination

For this circuit layout, we need to optimize five parameters (y, $q_0\phi_{rx}$, $q_0\phi_{ry}$, $q_1\phi_{rx}$ and $q_1\phi_{ry}$) to find the optimal configuration that best discriminates two ADCs in [1].

One-Shot Strategy: Discriminate more than two channels

We also investigate if the one-shot strategy using a two-entangled qubit circuit layout can be used to discriminate more than two ADCs.

As an exercise, and without the constraints of following the input state proposed at [1], we also explore the possibilities of using the most universal two-qubit input state and measurement possible ([11],[12]) to investigate whether we could find better optimal configurations to discriminate each ADC.

Figure 9 depicts an example of universal two-qubit input and measurement circuit:

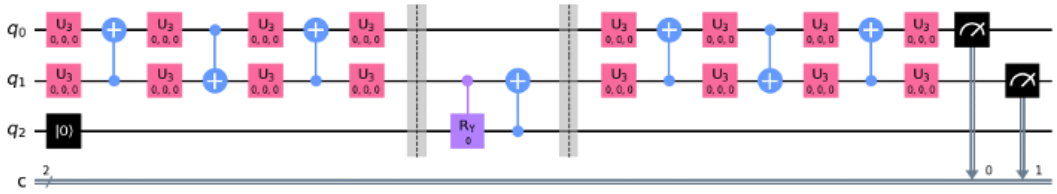


Figure 9 - Sample circuit for a universal two-qubit input and measurement ADC

Results show that this universal circuit does not improve the average success probability in terms of efficiency as it requires much more computational time, it requires forty-three additional parameters, and just one additional variable duplicates the space to explore by the optimization algorithms.

Therefore, we decide to use the same circuit from previous section, with the five parameters to optimize but, in this case, to discriminate three ADCs.

We did not consider POVM, which would be obtained by including additional ancillary systems in the measurement process, due to the further increase of the size for the parameter space this would imply.

Discrimination Process Implementation

This section describes how the results obtained from simulations reflect with accuracy the theoretical results from paper [1] and some exploration for discriminating more than two ADCs, comparing with the bounds of [2].

Additional detailed information about the process followed to obtain the results and further lines of work to improve the accuracy, can be found later in the Methodology section. Finally, Annex D contains an extended version of the simulation results and Annex E details the instructions on how to generate new results and try the source code library by yourself.

Discriminating 2 ADC with 1 qubit

Expected results

In paper [1] the success probability of discriminating 2 ADCs in a classical way, i.e., with a 1 input qubit going through the circuit, follows the formula

$$P_{succ} = \frac{1}{2}\left\{1 + (\cos\eta_1 - \cos\eta_0)\sqrt{x[1 - x(1-\gamma^2)]}\right\} \; where \; \gamma \equiv \gamma(\eta_1, \eta_0) \equiv \cos\eta_1 + \cos\eta_0$$

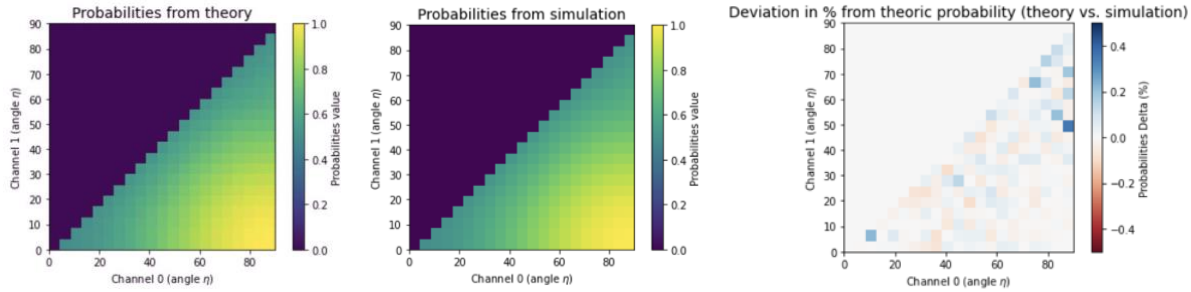
Plotting that theoretical success probability for different ADC pairs, see Figure 10 (a), we observe that it gets closer to 1 when the two channels are characterized by perpendicular angles, η_0=0° and η_1=90°. It means that they are orthogonal and then we can fully distinguish them. On the contrary, when channels are characterized by equal angles, for instance, η_0=20° and η_1=20°, we cannot discriminate them, and it would lead to a ½ probability over the diagonal. The trivial scenario of equal η hasn't been simulated.

From [1], we also know about the theoretically best input state to detect both ADC. This can be described in terms of the probability of |1> in the input (or the amplitude squared for |1>). Indeed, this component will be attenuated by the channel more and more as η gets closer to 90°.
Intuitively, the excited state as the input (|1>) provides a good reference to discriminate two channels, because the ground state (|0>) is not affected. However, as stated in [1], and seen in Figure 11 (a), optimal input state has also a small component of |0>, in the region where both ADCs have η approaching to 90°.

Results from simulation

We undertook a three-parameter optimization process: an input state probability (x) and 2 rotation gates (ϕ_{rx} and ϕ_{ry}) determining the projective measurement at the output of the unknown channel, to obtain the best average probability to discriminate two different ADCs.

After testing several optimization algorithms, the best results appear with the CRS one. Those are validated with an additional run using 10^6 shots by channel loading the optimal configuration determined by the algorithm. From the plotted results, see Figure 10 (b), it is difficult to appreciate any difference by a visual inspection between the theoretical and numerical results.



(a) (b) (c)

Figure 10 – (a) Theoretical maximum probability of success P(η_0, η_1, x) vs η_0 and η_1 (b) Resulting best average probability of success P(η_0, η_1, x) vs η_0 and η_1 from optimization with CRS algorithm and a validation process (c) Probability deviation in % from theoretical probability compared with the average success probability using 1 million times a circuit built with the best configuration resulting with a CRS optimization with 1,000 iterations and 10,000 shots

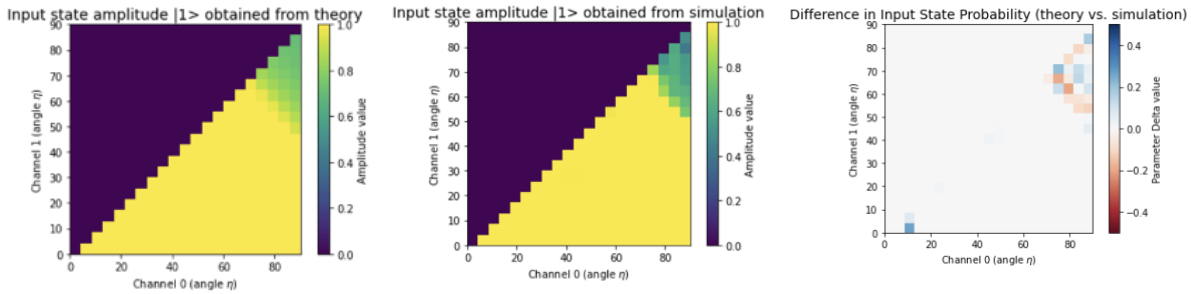
To appreciate differences, Figure 10 (c) shows the percentual deviation from theoretical probability result compared with best optimized configuration obtained for a sample of ADC pairs.

The maximum value is close to 0.4% for just one pair of η, while all other values are around +/-0.2%. In general, the accuracy is very close to the 3rd decimal of the theoretical probability.

From Figure 10 (c), and many other simulations, we can infer three different regions where optimization has a different result. Although algorithms have different behaviors to reach the minimum of the cost function, exploring the continuous parameters for *x*, ϕ_{rx} and ϕ_{ry}, we observe that:

- pairs close to be orthogonal (low-right corner) can be very precisely discriminated. The probability should be close to 1 and, as we can see them in white, the simulated results get to the expected values.
- pairs with η approaching to 90° (up-right corner) are apparently the more challenging in the simulation, and we can picture them in the more intense colors.
- pairs in other regions are statistically similar in tones.

Regarding the probability of having |1> in the input state, *x*, the obtained result is also close to the theoretical plot, as seen in Figure 11 (b), where most of the input values are 1, except from the top-right corner values. However, here the differences are visible at naked eye. This is the region in which ADC is characterized with high attenuation parameters. In this region, the results from simulations are aligned with the theory, showing values for *x* smaller than 1.



(a) (b) (c)

Figure 11 - (a) Theoretical amplitude values to best discriminate the given two ADCs. (b) Amplitude values obtained from the optimal simulated configuration from CRS algorithm. (c) Amplitude value difference between theory and optimal simulated configuration from CRS algorithm.

That difference for x between the theory and simulation hits the 2nd decimal for almost all regions, except from the top-right corner, where we obtain mainly an error in magnitude below +/- 0.2.

Conclusion

The theoretical results presented for probability and optimal input state in [1] to discriminate two ADC with one-shot and using a classical strategy (only a single qubit as input state) have been reproduced through a tuned optimization process based on the maximum likelihood for POVM results, with different accuracy. This accuracy, depending on the application required can be enough or could require an adjustment in the optimization parameters, which will also depend on the pair of attenuation factors involved, since we have observed three different categories of results.

Discriminating 2 ADC with 2 entangled qubits

Expected results

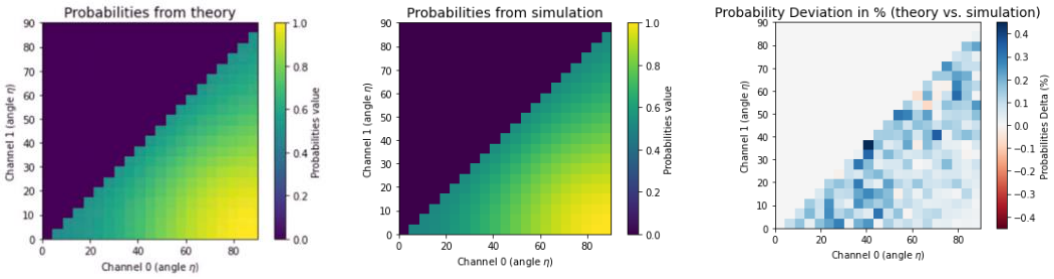
In this case, the proposed circuit consists of a two-entangled input qubit, which allows to discriminate the channels by taking advantage of the quantum entanglement property. Figure 12 (a) shows the theoretical maximum success probability resulting from the formula in [1]:

$$P_{succ} = \frac{1}{2} + \frac{1}{4}(\cos\eta_1 - \cos\eta_0)\left\{(1-y)\gamma + \sqrt{(1-y)[4y + (1-y)\gamma^2]}\right\}$$

The input state proposed is an entangled one, with the continuous variable y proposed as the probability for |10>. This helps to understand that the entanglement is used to differentiate the two ADCs when the values for y are different from zero. Using the formula provided in [1], we plot Figure 13 (a). We can observe that the y value increases as the attenuation angles get closer to 90° and similar to each other (a broader up-right corner than in previous section). So, we should have advantage in this region if we use entanglement in front of any classical strategy.

Results from simulation

Figure 12 shows the average success probability from the optimal configuration after optimizing the, currently five, parameters (y, $q_0\phi_{rx}$, $q_0\phi_{ry}$, $q_1\phi_{rx}$ and $q_1\phi_{ry}$) using DIRECT_L optimization algorithm, which provided the best results in this case.



(a) (b) (c)

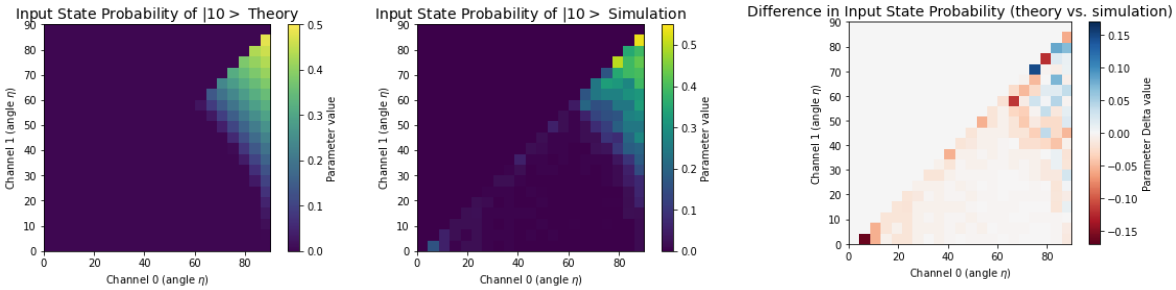
Figure 12 - Maximum probability of success $P(\eta_0, \eta_1, y^*)$ vs η_0 and η_1 (b) The resulting best average probability of success $P(\eta_0, \eta_1, y^*)$ vs η_0 and η_1 from optimization with DIRECT_L algorithm and a validation process (c) Probability deviation in % from theoretical probability compared with the average success probability using 1 million times a circuit built with the best configuration resulting with a DIRECT_L optimization with 4000 iterations and 100,000 shots.

As before, it is difficult to appreciate differences visually. To estimate how relatively close are both results, Figure 12 (c) presents the percentual deviation from the maximal theoretical probability compared with the validated average success probability using the best configuration for a sample of a given pair of ADCs, based on the maximum likelihood guess. As shown in that figure, the numerical results are similar to the theoretical in general, except for some pairs closer to 4%.

Again, in terms of accuracy, these results are like the previous classical strategy, but here we don't see the three regions we detected before. It is true that the low right corner keeps whiter than the rest, but nothing else can be said about the other parts.

There is another observation with the color to be made, though. Few red colors appear meaning that statistically, we achieve a slightly better result than expected. This might be an indicator that only in some cases we get to the best possible configuration, due to finite statistics. Note that we increased to 4.000 iterations and 100.000 plays for the optimization algorithm per ADC as we have two more parameters to optimize, while for the one qubit section before we used 1.000 iterations and 10.000 plays for just three parameters. These changes impact on the optimization calculation time.
The optimization process needs much more iterations to get the optimal parameter's configuration as we increase the number of parameters to optimize. Since we raised the parameters to optimize from three to five, the exploratory space to find optimal solutions increased 2^2 times, and therefore we also increased the number of iterations 2^2 times.

Figure 13 (b) shows *y* parameter, for input state |10> component, from the optimization process. Here we see similitudes to the expected results on the left picture, as we have a significant amount of *y* values different from zero for the larger attenuation angles (top-right corner). However, the optimization process was not able to hit exactly the optimal *y* value, as we can see some other values greater than zero, mainly close to the diagonal.



(a) (b) (c)

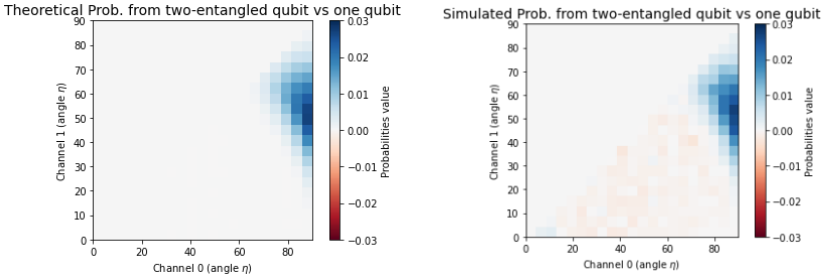
Figure 13 - (a) Optimal value y vs η_0 and η_1 (b) Best value y* vs $\eta 0$ and $\eta 1$ from the optimization with a DIRECT_L algorithm. (c) Amplitude value difference between theory and optimal simulated configuration from DIRECT_L algorithm.*

Figure 13 (c) depicts the calculated magnitudes differences between the expected and the simulated. We can observe that, in general, we obtained similar results as expected with maximum differences of +/-0.16, mainly through the diagonal, and in the area where the entanglement is required.

Circuit comparison: quantum entanglement advantage

From a theoretical standpoint, there is an actual advantage provided by the quantum entanglement. These formulas in [1] let us plot the difference between probabilities obtained by using the classical approach (one qubit as input state) and using entanglement (two qubits, one for signal and another as idler or ancilla). Figure 14 shows this comparison using the theoretical formulas and our numerical results.

We observe that entanglement can provide a small advantage, reaching a maximum improvement of 2,67 hundredths of probability. In addition to that, the figure shows that this advantage only happens in a reduced region, also as a result of the asymmetric action that the ADC perform on each of the two states |0> and |1>.



(a) (b)

Figure 14 – (a) Theoretical probability improvement using the two entangled qubits with respect to a one qubit circuit. (b)Same than previous from the simulation results

From the experimental side, see Figure 14 (b) we can observe that we get very similar results as expected, which means that we have confirmed the improvement provided due to the entanglement properties. Therefore, for further results, we are going to use the two-entangled qubits.

Conclusion

Results from [1] have been again reproduced with accuracy. Using an entangled input state provides an advantage in a reduced region of the possible ADC pairs. This small advantage comes after a more demanding optimization process, in terms of number of parameters in the input state and output measurement, as it requires a much higher effort in computational terms, to obtain similar results to the ones from the classical approach in the region where entanglement doesn't contribute. This additional effort comes by the increase of two additional parameters (corresponding to the ancilla gates), which increases the space of possible solutions to optimize and increases the number of operations for each circuit iteration.

Discriminating +2 ADC with 2 entangled qubits

At this point, with the experience gathered from simulations and circuits presented in previous sections, we moved to the idea exposed by the paper [2], which describes a discrimination of more than two ADCs.

In this case, we analyze the results of using the same two entangled qubits circuit, the same one-shot strategy and the maximum likelihood to get the optimal guesses, when discriminating three different channels. This process brings us to the conclusion that only 2 channels can be discriminated reliably with this method, which was previously unknown in the literature.

Expected results

No initial results were provided or theoretical results to compare with. The only formulas used come from [2] and previous articles cited therein, and are the upper and lower probability error bounds on the average error probability.

$$\frac{1}{2}\sum_{i\neq j}\pi_i\,\pi_j F_{i:j}^{2M} \;\leq\; p_{err} \;\leq\; \sum_{i\neq j}\sqrt{\pi_i\pi_j}\,F_{i:j}^{M}$$

where $F_{i:j} := F(\rho_i,\rho_j) = \| \sqrt{\rho_i}\sqrt{\rho_j} \|_1 = Tr\sqrt{\sqrt{\rho_i}\rho_j\sqrt{\rho_i}}$

The probability of having one of the three channels is represented by π_ι. In our case, we consider all three channel equiprobable, so any π=1/3. $F_{i:j}$ is the square root fidelity, so ρ_i and ρ_j are the density matrices for both end states compared after crossing channels i and j. And M is the number of shots used for the channel. As our strategy is one-shot, M=1.

Results from simulation

Our simulation gave us error probability for ADC trios as well as the error probability bounds. The plots below (Figure 15) show the average error probability to discriminate three channels and the upper and lower bounds.

To simulate and better represent the results, it was required to define specific trios of ADC characterized by η. The results presented here set two fixed ADCs, for instance, η_0: 29°and η_1: 59°, while varying the third one from 0° to 90°. The values for η_0 and η_1 are shown with additional dotted vertical lines.

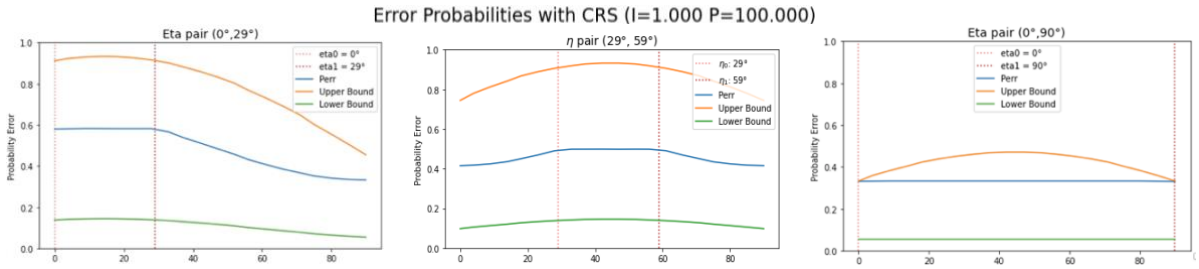


Figure 15 - Error probabilities from the best configurations to discriminate three amplitude damping channels compared with the theoretical upper and lower error probability bounds.

From Figure 15 plots, we can get the following observations:

- probability of error (blue line) between η_0 and η_1 is always constant, while outside of this range it declines.
- probability of error (blue line) is always higher than 0.3
- the bounds (orange and green lines) seem to be generous bounds, as they are far away from each other.

To better understand how the simulation selects the correct channel, how probabilities behave depending on the trio members, and why it is constant for any value of η_2 between η_0 and η_1, we analyze the average success probability and its breakdown for each channel.

From Figure 16, we can state that the simulation guesses always bet for the two farthest channels, meaning those two channels that have their η angles with larger distance.

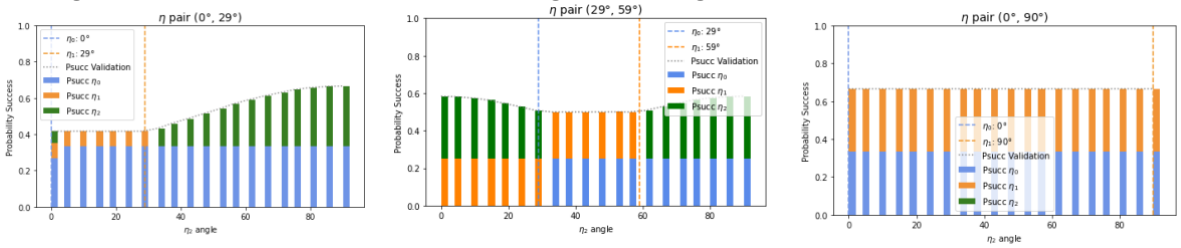


Figure 16 - Success Probabilities breakdown for the best selected η to discriminate the three channels.

Each column represents the success probability for a trio of ADCs. This probability can be broken down into 3 parts of 1/3 for each channel. The contribution from each channel, the success from our guess, is represented in colors (blue, orange, green) for each of them (η_0, η_1, η_2).

Instantly, we can observe that each guess for any trio of ADC, has only the contribution of just two ADCs, since only two colors are presented for each column. The bet on just two of the three ADCs explains why the error probability is always larger than 1/3, as there is one ADC missed all the time.

Looking carefully, we see also that the two ADC bet by the guess are always those with parameters η more distant. This helps to understand too, why the probability of error is constant for values of η_2 between η_0 and η_1, because in this situation η_0 and η_1 are the more distant η.

Finally, with the results obtained, we want to confirm if we can use the theoretical success probability obtained from [1] to compare it with the one resulting from the simulation. Scaling it to the maximum theoretical probability to 2/3, because we have three channels instead of two, we compare these theoretical results with the ones coming from the simulation (Figure 17).

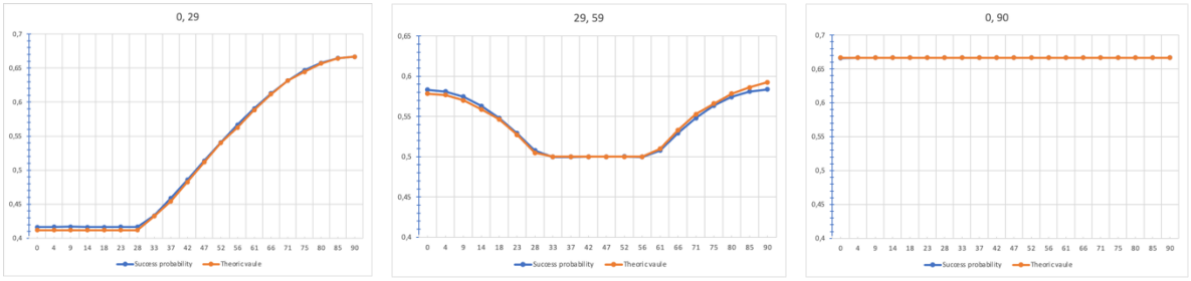


Figure 17 - Average success probability from the simulation compared with the maximum theoretical one scaled to 2/3.

We observe that both probabilities align very closely, which means that this set up belongs to the strategy that always discriminate between two channels, those that are the farthest.

Why only two?

Looking into the details of the numerical results obtained, the probabilities for each possible result measured at the end of each channel must fall in any of the four possible buckets: 00, 01, 10 and 11. There is no way to rotate the final state vector for each of the possible channels in such a way it goes into a different bucket for each one. The ADC acts only on the excited state |1> crossing the channel, while leaving |0> untouched. All three ADCs will act in a similar way, just with differences on the intensity, depending on η. So, the channel with the value of η intermediate, will just act on |1> in a "moderated way" compared with the other two channels, which will be impossible to distinguish, based on one-shot, maximum likelihood guess and just two qubits, independently on the number of gates and rotations at the end our measurement does.

Conclusion:

Our initial idea was that the proposed optimized measurement with 4 possible outcomes (00, 01, 10, 11) using the two-entangled qubit circuit, would be enough to accommodate the results. The experimental results concluded that we were wrong.

There is a limit of 2 ADC in a discrimination process based on a one-shot strategy, one-ancilla for input entanglement, a projective measurement on the two qubits and using the maximum likelihood as guess function. Therefore, to discriminate three or more ADCs it might be required a new circuit proposal (perhaps adding more qubits to perform a non-projective POVM) or a new strategy to allow more than one shot.

Methodology and Process Description

To build the circuits and obtain the results described in the previous sections, we coded it in Python language using the Qiskit [13] open-source SDK framework from IBM with additional extra libraries to meet the ADC requirements.

The overall process was structured in three phases: Optimization, Validation and Representation.

Optimization

In this first phase, we obtain the optimal configurations, i.e., input state and rotation angles values, that maximizes the success probability to discriminate ADCs.

For doing that, we generate a limited set of groups of η avoiding repetitions. Groups of two when discriminating between 2 ADCs or triplets between 3 ADCs.

For each of these pairs or triplets, we iterate the circuit to find the optimal configuration that maximizes the success rate. This is achieved by using the global Non-Linear Optimization algorithms included in Qiskit Aqua [14] to minimize a cost function, crafted to fulfill the statistical method of the maximum likelihood estimate. For each parameter setup, the corresponding circuit is run for multiple shots with different channels from those to be discriminated, estimating the conditional probabilities of obtaining a certain outcome given a certain channel. These estimates are used to make a maximum-likelihood guess, i.e., selecting for each outcome the most probable channel.

Listing this phase key aspects, we can find that:

- the number of parameters involved are directly related to the number of circuit gates
- regarding the total number of algorithm iterations and the number of shots for each iteration to calculate the cost function, we discovered that:
    - Using a small number of iterations reduces the chances to explore the space of possible solutions defined by the continuous variable ranges, whereas, too many iterations will rapidly scale up the required computation time to get the results. However, it is needed to increase the number of iterations to also increase accuracy in the results.
    - Few shots lead to wrong configuration easily, just because the cost function is based on measurements, which are purely stochastic. Using a high number of shots will provide a better statistical result, reducing bias. Increasing shots too much will not contribute to improve accuracy, as it helps mainly to reduce statistical noise. Again, increasing the number of shots will rapidly scale the computation time.
    - As the number of iterations increases, shots should also increase.
- It is also relevant to select the appropriate optimization algorithm. We observed significant deviation from expected results depending on the algorithm. Specifically, we obtained better results for CRS and DIRECT_L and DIRECT_L_RAND than the evolutive ones. Regarding the cost function, despite not being continuous, it seems to have a unique minimum reachable in a monotonous way. This discovery would justify why the first optimizers (CRS, DIRECT_L and DIRECT_L_RAND) do a better job, for the same number of iterations, than the evolutive ones (ESCH, ISRES).

Validation

On the next phase, we create the corresponding circuits to get the statistical guesses and probabilities from the optimal configuration values obtained previously.

Basically, the optimal values for input state and gates are used to define our circuit for each pair or triplet. We run these circuits with a large number of shots (10^6 shots), so the values obtained for the guesses and probabilities are statistically refined.

This step could be avoided if the number of plays in the optimization phase is high enough, but unfortunately the optimization is very time consuming and the results with the required accuracy would take too long. Typically the optimization process takes many hours or several days, while the validation usually takes in the order of several minutes.

Representation

This last phase gathers all information resulting from previous phases to plot the optimization and validation results and the comparison against theoretical ones. This requires not only the empirical data from previous phases, but also the calculations of the theoretical values. So, several formulas from papers [1] and [2] are implemented here.

Conclusions, applications and possible next steps

Main conclusions

We have experimentally replicated the results described for one-shot strategy in [1] with accuracy, following numerical simulation and optimization, together with a maximum likelihood guessing

strategy. In doing that, we had to put in place an optimization process that involved several parameters directly related with the quantity of gates taking part in the implementation for the input state and output measurement.

We managed to replicate the theoretical success probability of discriminating ADCs and the optimal input state and output measurement parameters, in both approaches: the classical (represented by a one-qubit circuit) and the one taking advantage of quantum entanglement.

We had to tune the optimization algorithms to make results from both approaches comparable in precision. The number of optimizer iterations and shots had to be higher when more parameters are considered, or when the exploratory space of possible configurations increases. It is worth mentioning that we had to introduce a fast validation phase, using the optimization results with a previous reasonable number of shots, and testing with a huge number of shots, to get curated probability results, and reduce the bias introduced by the optimization phase.

We found a limit of 2 as the maximal number of ADC able to discriminate on a one-shot strategy using input states with entanglement and only one ancilla.

Applications

There are multiple motivations to study the Quantum Channel Discrimination (QCD) that come from several areas and have their own applications. The following list does not pretend to be exhaustive but mentions the ones we considered most relevant.

Quantum Illumination

In quantum information processing, Quantum Illumination offers a paradigm that increases the signal-to-noise ratios by using nonclassical states, allowing performance enhancements also in presence of quantum decoherence. These systems, as described in [15]-[16] are comprised by a source that generates entangled signal and idler beams, the interaction that the signal has with a target, if present it will reflect the signal, which also suffers potential sources of decoherence, and finally a receiver that makes the joint measurement, if the signal returns, with the idler beam, which was stored in a memory for this moment.

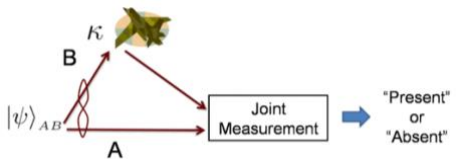


Figure 18 - Schematic diagram of quantum target detection of low-reflectivity object. Taken from [16]

In this case, QCD strategies using entanglement show a better measurement and a clear advantage compared to the classical approach. This is equivalent to the advantage seen when using an ancilla qubit (here the idler beam) and the qubit traveling the circuit (the signal beam returned).

Quantum Reading

Quantum Reading is a mechanism that exploits entangled states to read (retrieve) information from memories more efficiently than classical strategies.

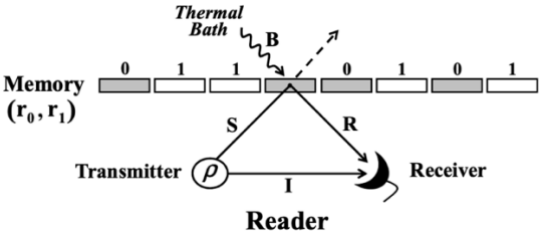


Figure 19 - Model of memory taken from [17]

In this case, as shown in Figure 19, a reader has the transmitter, which generates an entangled pair S (signal) and I (idler). The signal S probes the memory and interacts with it. The memory model used in [18] is made of beam splitter mirrors, which reflects one part (R) returning to a reader. The reader makes a joint measurement of R and I. The information is encoded in the memory by different reflectivities in beam splitters (r_0=0 and r_1=1). As stated in the first section, a beam splitter models also an ADC, and therefore equivalent to considering each of the two beam-splitters reflectivities as two different lossy channels [19]. The relationship of making this optical memory efficient and the QCD problem is direct.

Anomaly Detection

Detecting an anomaly of a device, which forms part of a network of N equivalent devices can be very complex, but it has been shown that if these devices fulfill some conditions, it is possible to identify the faulty device with optimal probability doing a global quantum measurement for entangled input probes with N ancillae [20].

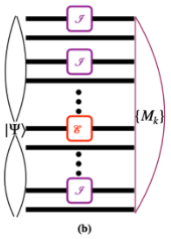


Figure 20 - Position error identification for channels. All channels are programmed to perform the identity except for a faulty device [20] Fig1. (b)

Sensing and Metrology

Quantum entangled states provide advantages versus classical ones, inclusive in decoherent environments breaking entanglement [21]. This helps to know the noise that is going to occur in a system, to protect against it and improve the signal-to-noise ratio as seen in [16].

Related with previous

Other areas that will improve as a direct consequence are:

- Secure Communication: Quantum Illumination can resist the eavesdropper and therefore help Quantum Communication and Secret Key generation.
- Biological Imaging: As sensing improves, better discrimination and categorization for signals can be done, leading to more accurate images.
- Other quantum imaging applications: some mentioned in [22]: Ghost imaging, sub shot noise imaging, proposals for beating Rayleigh diffraction limit by exploiting quantum states of light, entangled imaging, (i.e., the generation of two correlated images by exploiting twin beams entanglement), noiseless amplification of images, quantum imaging with undetected photons and more.

How to follow

We suggest three lines of action depending on the reader's interest from this point:

- To keep the path to discriminate more than two channels, we suggest the exploration of more than two qubits in order to help the optimization process to use more than 2^2 outcomes.
- To increase the probabilities obtained, we encourage to work with other strategies where the channel can be used more than once and check how adaptive strategies help to achieve the goal.
- To help others to work in the QCD problem you can use the code created and freely available (see Annex E) and expand it to go from ADC to, for instance, superchannels and quantum combs.

References and bibliography

[1] M. Rexiti, S. Mancini, Discriminating qubit amplitude damping channels (2021) arXiv:2009.01000
[2] L. Banchi, Q. Zhuang, S. Pirandola, Quantum-enhanced barcode decoding and pattern recognition (2020) arXiv:2010.03594
[3] J Preskill, lecture notes on quantum information theory, chapter 3. http://theory.caltech.edu/~preskill/ph219/chap3_15.pdf
[4] M. A. Nielsen & I. L. Chuang, Quantum computation and quantum information – pp. 373-386 (Cambridge University Press, Cambridge, 2000).
[5] J. Jones, & D. Jaksch, Quantum Information, Computation and Communication - pp. 146-150 (Cambridge University Press, Cambridge, 2012)
[6] D. McMahon, Quantum Computing Explained – pp. 251-278 (John Wiley & Sons, Inc., Hoboken, New Jersey, 2007)
[7] C.J. Wood, J.D. Biamonte, D.G. Cory, Tensor networks and graphical calculus for open quantum systems - Figure 1 - (2015) arXiv:1111.6950
[8] S. Pirandola, R. Laurenza, C. Lupo, & J. L. Pereira, Fundamental limits to quantum channel discrimination – https://arxiv.org/abs/1803.02834
[9] U. Herzog, Minimum-error discrimination between a pure and a mixed two-qubit state – https://arxiv.org/pdf/quant-ph/0307038.pdf - 2003
[10] C. W. Helstrom, Quantum Detection and Estimation Theory (New York: Academic, 1976).
[11] IBM, Proving Universality, Learn Quantum Computation using Qiskit – https://qiskit.org/textbook/ch-gates/proving-universality.html
[12] IBM, Simulating Molecules using VQE, Learn Quantum Computation using Qiskit – https://qiskit.org/textbook/ch-applications/vqe-molecules.html#Simple-Variational-Forms
[13] IBM, Open-Source Quantum Development – https://qiskit.org/
[14] IBM, Optimizers (qiskit.aqua.components.optimizers), Qiskit Documentation – https://qiskit.org/documentation/apidoc/qiskit.aqua.components.optimizers.html#module-qiskit.aqua.components.optimizers
[15] S. Lloyd, Quantum Illumination – https://arxiv.org/pdf/0803.2022.pdf - 2008
[16] Z. Zhang, S. Mouradian, F.N. C. Wong, & J. H. Shapiro, Entanglement-Enhanced Sensing in a Lossy and Noisy Environment – https://arxiv.org/pdf/1411.5969.pdf - 2014
[17] S. Zhang, J. Guo, W. Bao, J-H Shi, C. Jin, X. Zou, G. Guo, (2014). Quantum illumination with photon-subtracted continuous-variable entanglement. Physical Review A. 89. 10.1103/PhysRevA.89.062309.https://www.researchgate.net/publication/277564495_Quantum_illumination_with_photon-subtracted_continuous-variable_entanglement – FIG.1
[18] – G. Spedalieri, C. Lupo, S. Mancini, S.L. Braunstein, & S. Pirandola – Quantum reading under a local energy constraint - https://arxiv.org/pdf/1204.3448.pdf - 2015
[19] – G. Ortolano, E. Losero, S. Pirandola, M. Genovese, I. Ruo-Berchera, Experimental quantum reading with photon counting - https://advances.sciencemag.org/content/7/4/eabc7796 - 2021 - 2021
[20] - M. Skotiniotis, R. Hotz, J. Calsamiglia, and R. MuñozTapia, Identification of malfunctioning quantum devices, https://arxiv.org/pdf/1808.02729.pdf - 2018
[21] – Broken quantum links still work. Nature 499, 129 (2013). https://doi.org/10.1038/499129a
[22] – M. Genovese - Real applications of quantum imaging - https://arxiv.org/pdf/1601.06066.pdf - (2016)

Annex A – Table of Contents

Annex B – List of Figures

Annex C – Equivalence between ADC and circuit

Let's solve the exercise 8.20 in [4]. Show that the following circuit in models the amplitude damping quantum operation, with $\sin^2\eta = \lambda$.

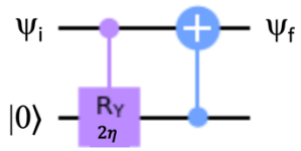


Figure 21 - Circuit modeling ADC

where $\psi_i = \alpha|0\rangle + \beta|1\rangle$

We can do this by calculating the state at the points 0, 1 and 2

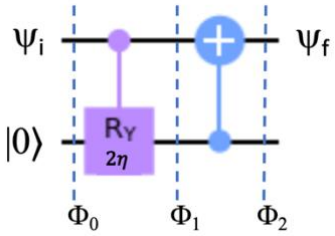


Figure 22 - Points to evaluate the quantum state.

The state Φ_0 is the tensor product of ψ_i with the environment: $|\Phi_0\rangle = \alpha\,|0\rangle\otimes|0\rangle + \beta\,|1\rangle\otimes|0\rangle$

The state Φ_1 is the result of apply the gate RY, which is a rotation in Y of 2η:

$$|\Phi_1\rangle = \alpha|0\rangle\otimes|0\rangle + \beta\,|1\rangle\otimes\{\cos\eta\,|0\rangle + \sin\eta\,|1\rangle\} = \alpha|0\rangle\otimes|0\rangle + \beta\cos\eta\,|1\rangle\otimes|0\rangle + \beta\sin\eta\,|1\rangle\otimes|1\rangle$$

The state Φ_2 is the result of apply the CNOT: $|\Phi_2\rangle = (\alpha\,|0\rangle + \beta\cos\eta\,|1\rangle)\otimes|0\rangle + \beta\sin\eta\,|0\rangle\otimes|1\rangle$

The final state ψ_i will result of applying the partial trace over the environment:

$$\mathrm{tr}_{env}\,|\Phi_2\rangle\langle\Phi_2| = (\alpha\,|0\rangle + \beta\cos\eta\,|1\rangle)\,(\alpha^*\,\langle 0| + \beta^*\cos\eta\,\langle 1|) + |\beta|^2\sin^2\eta|0\rangle\langle 0| =$$

$$= (|\alpha|^2 + |\beta|^2\sin^2\eta)\,|0\rangle\langle 0| + |\beta|^2\cos^2\eta\,|1\rangle\langle 1| + \alpha^*\beta\cos\eta\,|1\rangle\langle 0| + \alpha\beta^*\cos\eta\,|0\rangle\langle 1|$$

If we compare this result with the expected one from the ADC with Kraus operators E_0 and E_1, where $E_0 = \begin{bmatrix}1 & 0\\ 0 & \sqrt{1-\lambda}\end{bmatrix}$ *and* $E_1 = \begin{bmatrix}0 & \sqrt{\lambda}\\ 0 & 0\end{bmatrix}$

$$\Lambda_{AD}(|\Phi_0\rangle\langle\Phi_0|) = E_0\begin{pmatrix}|\alpha|^2 & \alpha^*\beta\\ \alpha\beta^* & |\beta|^2\end{pmatrix}E_0^\dagger + E_1\begin{pmatrix}|\alpha|^2 & \alpha^*\beta\\ \alpha\beta^* & |\beta|^2\end{pmatrix}E_1^\dagger = \begin{pmatrix}|\alpha|^2 + |\beta|^2\lambda & \alpha^*\beta\sqrt{1-\lambda}\\ \alpha\beta^*\sqrt{1-\lambda} & |\beta|^2(1-\lambda)\end{pmatrix} =$$

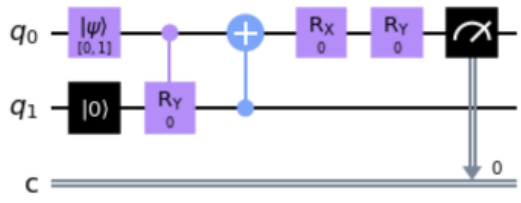
$$\begin{pmatrix}|\alpha|^2 + |\beta|^2\sin^2\eta & \alpha^*\beta\cos\eta\\ \alpha\beta^*\cos\eta & |\beta|^2\cos^2\eta\end{pmatrix}$$

We have the same result, so the circuit models the ADC channel properly.

Annex D – Exhaustive Results Obtained for Each Implementation

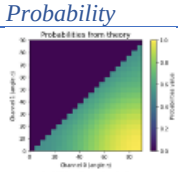
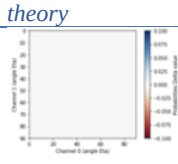
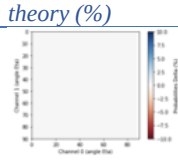
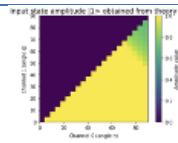
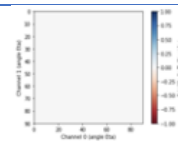
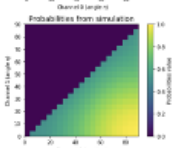
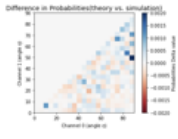
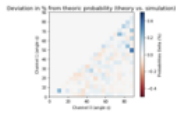
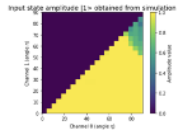
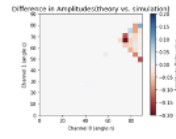
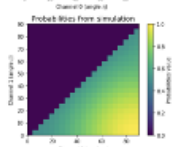
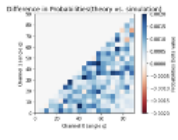
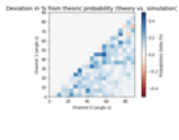
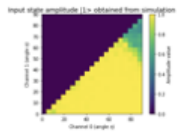
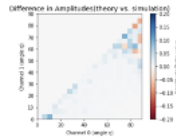
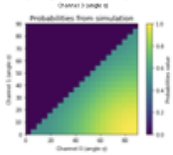
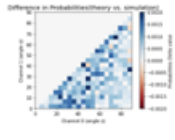
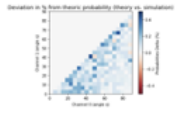
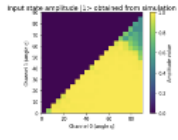
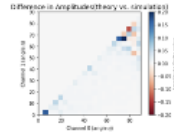
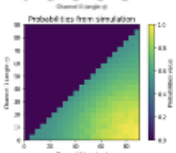
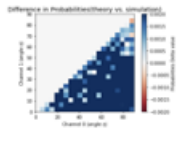
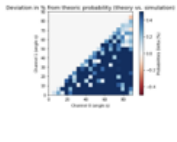
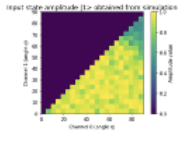
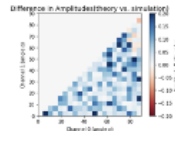
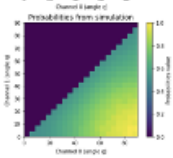
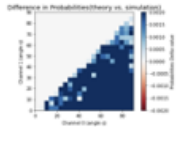
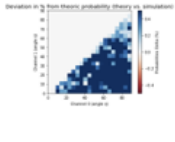
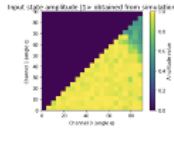
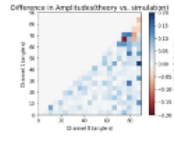
Discriminating 2 ADC with 1 Qubit Implementation Results

Circuit layout



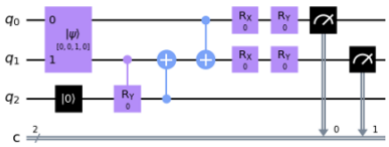
Results

Results from the specified optimization algorithm launched with 1.000 iterations and 100.000 shots per circuit, and validated the optimal configuration executed with 1.000.000 shots per circuit.

	Success Probability	*Delta prob vs. theory*	*Delta prob vs. theory (%)*	*X (prob of |1>)*	*Delta x vs. theory*
Theoretical					
CRS					
DIRECT_L					
DIRECT_L_RAND					
ESCH					
ISRES					

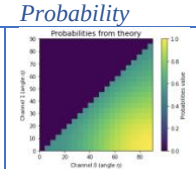
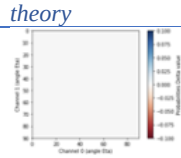
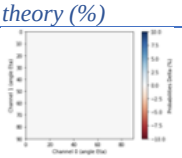
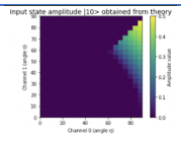
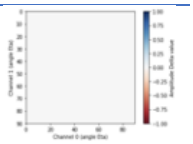
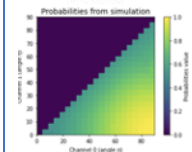
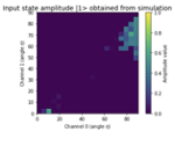
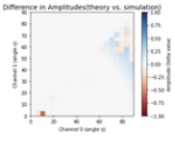
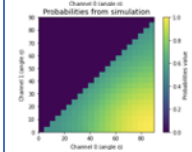
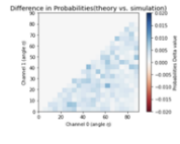
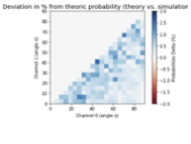
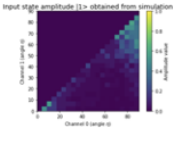
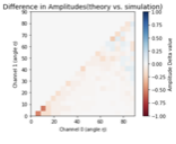
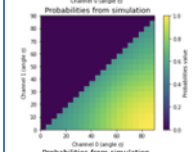
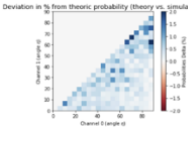
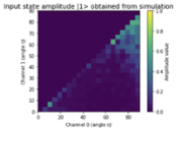
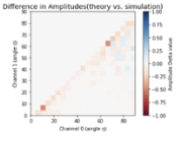
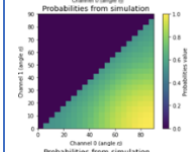
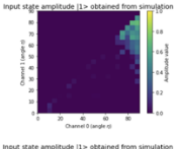
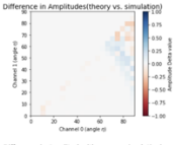
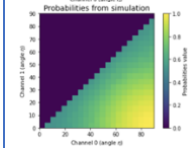
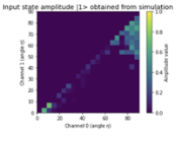
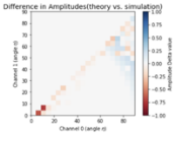
Discriminating 2 ADC with 2 Entangled Qubits Implementation Results

Circuit layout



Results

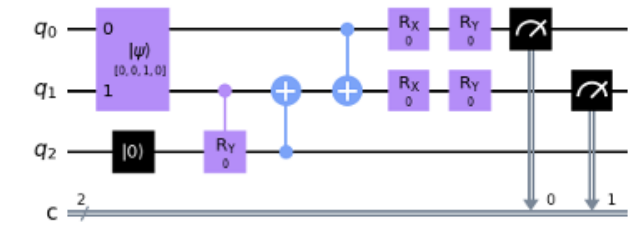
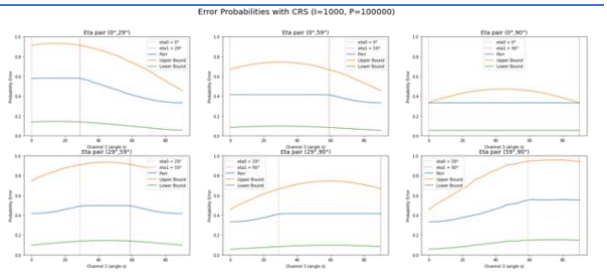
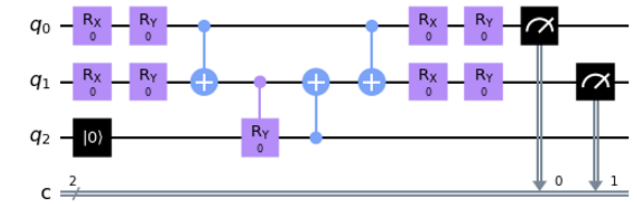
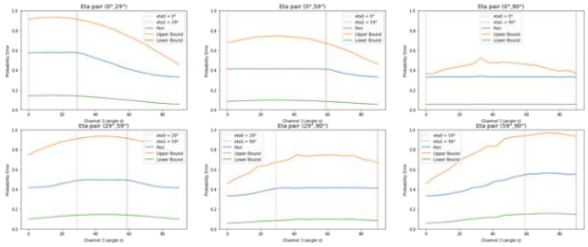
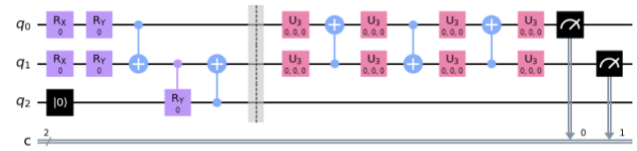
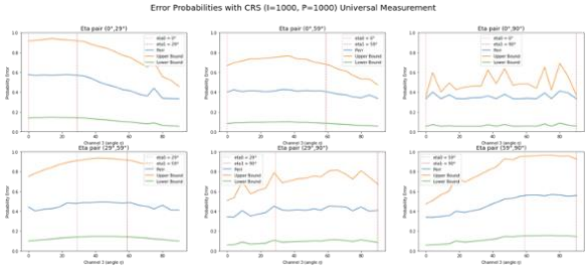
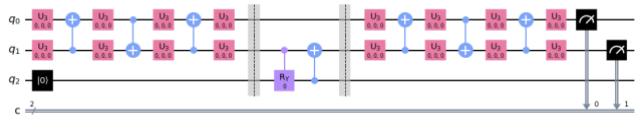
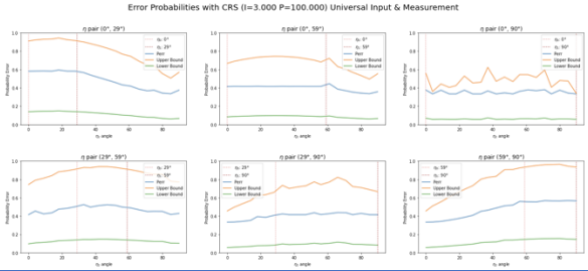
Results from the specified optimization algorithm launched with 500 iterations and 10.000 shots per circuit, and validated the optimal configuration executed with 1.000.000 shots per circuit.

	Success Probability	*Delta prob vs. theory*	*Delta prob vs. theory (%)*	*y (prob of |10>)*	*Delta y vs. theory*
Theoretical					
CRS		-	-		
DIRECT_L					
DIRECT_L_RAND					
ESCH		-	-		
ISRES		-	-		

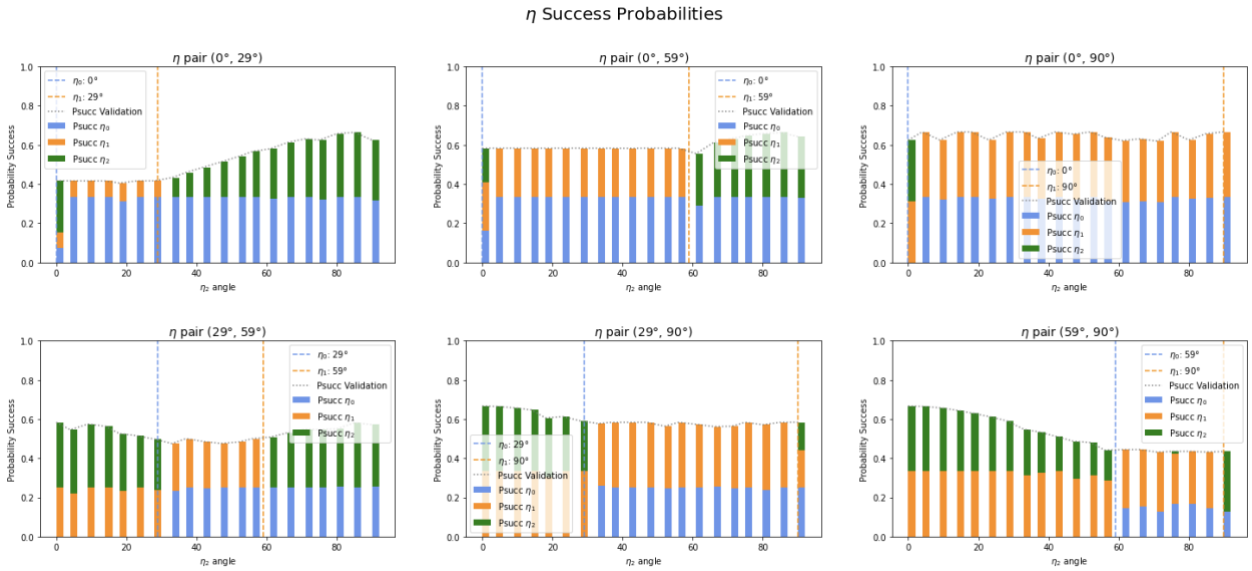
Discriminating +2 ADC with 2 Entangled Qubits Implementation Results

Average Error Probabilities with different Circuit layouts

Results launched with a CRS optimization algorithm with 1.000 iterations and 100.000 shots per circuit and the optimal configuration validated with 1.000.000 shots per circuit, except the last one, the most universal, that was executed with 3.000 iterations and 100.000 shots per circuit.

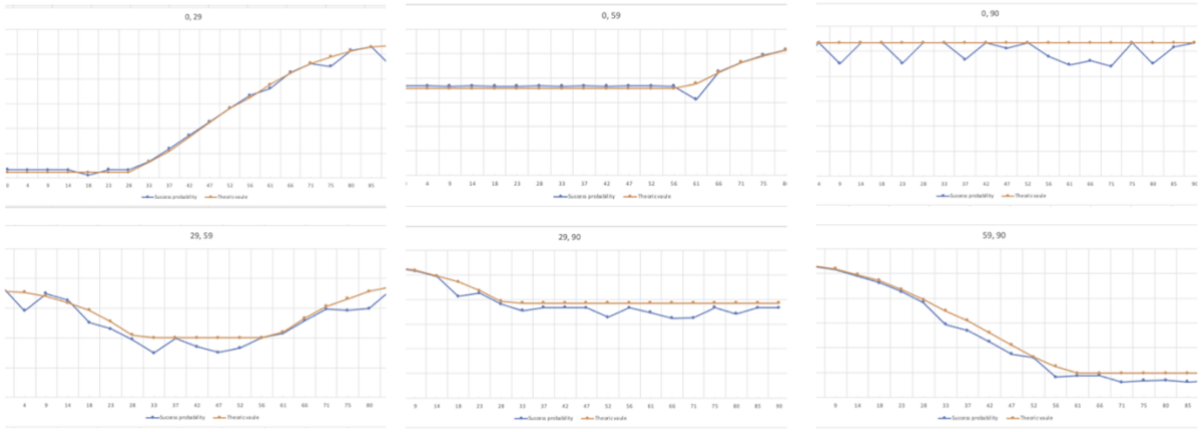
Circuit	Error Probabilities
	
	
	
	

Average Success Probabilities Breakdown



Maximal Theoretical Probability vs Best Simulated Average Success Probability

Theoretical probability is scaled by a 2/3 factor to be comparable with the exercise of discriminating three channels.



Annex E – QCD Project Library and notebooks source code

We have been working on a GitHub project repository and it is an open-source project based on Apache 2.0 license.

We have prepared a documentation on how to use the Quantum Channel Discrimination (QCD) project library in a Jupyter notebook: https://github.com/iamtxena/quantum-channel-discrimination/blob/main/docs/qcd-user-guide.ipynb

There is also available the QCD library source code at: https://github.com/iamtxena/quantum-channel-discrimination