Laying the foundation for building a Quantum Networking Benchmark suite using Quantum Network Applications

Evaluating the inclusion of the Clauser-Horne-Shimony-Holt game quantum network application

Thomas Maliappis[1]

Supervisor(s): Stephanie Wehner[1]**, Ravi Ashok Kumar Vattekkat**[1]

[1]EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 25, 2023

Name of the student: Thomas Maliappis
Final project course: CSE3000 Research Project
Thesis committee: Stephanie Wehner, Ravi Ashok Kumar Vattekkat, Andy Zaidman

1. Abstract

The rapid advancement of Quantum Network architectures necessitates a comprehensive and quantitative comparison to assess their effectiveness and performance. Unfortunately, there does not exist an implemented quantum network benchmark suite capable of determining the superior architecture. Hence, our study aims to establish the foundation for developing a benchmark suite by leveraging existing quantum network applications. However, the specific inclusion of quantum network applications in the suite remains to be determined. Therefore, to address this gap, our study will explore the potential inclusion of the Clauser-Horne-Shimony-Holt (CHSH) game based on its effectiveness in identifying errors within various properties of the quantum networking system. We use an exploratory research methodology involving experiments performed on simulated quantum networks utilizing SquidASM. Each experiment simulates multiple quantum networks, with a single property as the independent variable. For each value of the independent variable, we calculate both the success probability of the game and the number of successes per second. Subsequently, we employ the one-way ANOVA test to examine if there are significant variations in these performance metrics. Our results demonstrate that the CHSH game exhibits sensitivity to all properties affecting the quality of entanglement between nodes, execution time, and the error probability of both single-qubit gates and measure operations. Additionally, we compare the success probabilities based on different input combinations using the Root Mean Squared metric to uncover any underlying patterns within the data. As a result, we discovered a procedure for quantifying the difference between the error probabilities of measurements of zero and one. Based on the outcomes of our study, we consider the CHSH game to be a suitable addition to the benchmark suite if the testing requirements of the suite align with the qualities offered by the application. We anticipate that these results will aid the development of the benchmark suite and advance the understanding of quantum network architectures and their evaluation.

Keywords - Quantum Networks, Benchmarking, Clauser-Horne-Shimony-Holt (CHSH)

2. Introduction

In the field of computer networks, quantum networking is gaining traction as it has the potential to provide unbreakable encryption and enhanced privacy. However, this technology has not fully matured. Regularly, researchers are actively working on multiple hardware architectures and software modules, raising the question of which one is superior.

This research aims to lay the foundation for building a Quantum Network Benchmark Suite (QNBS). A benchmark suite consists of standardized tests used to objectively evaluate the performance and effectiveness of different hardware and software components on realistic workloads. In our case, the benchmark suite will assess the host network by executing a set of existing Quantum Network Applications (QNAs) to provide a set of performance metrics. These metrics are quantitative assessments of a quantum network system's performance. Hence, by executing a benchmark suite across various quantum network architectures, we can conduct a comparative analysis based on the performance metrics generated by the suite. As a result, a QNBS facilitates the identification of superior quantum network architectures, thereby streamlining the process of determining the most effective solution.

2.1. Research Question

This paper will determine how informative the Clauser-Horne-Shimony-Holt (CHSH) game [2] QNA is as a benchmark for quantum network systems. This investigation holds significant importance since the optimal QNAs for inclusion in the QNBS are presently unidentified. To comprehensively address this question, we have divided it into the following sub-questions:

1. How sensitive is the CHSH game in recognizing errors in the properties of the total quantum networking system?
2. Can we use the CHSH game to make quantitative predictions about a specific property?
3. Should the CHSH game be included in the QNBS?

Upon successfully addressing all the preceding sub-questions, this research holds the potential to mark one of the first inclusions in the QNBS and contribute valuable insights to guide the adoption or avoidance of similar applications, depending upon the research outcomes. Hence, at the very minimum, addressing these questions will provide invaluable knowledge to effectively guide the decision-making process of adding other QNAs.

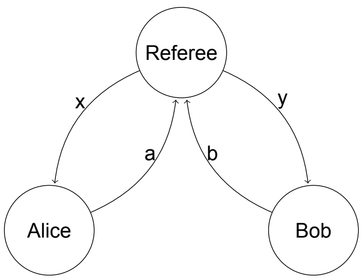
2.2. Background on the CHSH game quantum network application

The QNA considered in this paper is known as the Clauser-Horne-Shimony-Holt (CHSH) game [2]. This application is known as the gamification of Bell's Theorem [1] which states that a theory of local hidden properties that deterministically define the behavior of entangled particles cannot reflect reality. Based on this, Clauser and colleagues proposed "a decisive test between quantum mechanics and local hidden-variable theories" [2] focusing on the statistical correlation of an entangled particle's measurement outcome.

The structure of the CHSH game is shown in Figure 2.1a. In detail, it consists of two players (Alice and Bob) and a referee. The game starts with the referee providing a bit to each player, bit x to Alice, and bit y to Bob. Alice responds with bit a and Bob with bit b. The game is won if Equation 2.1 is satisfied:

$$x * y = a \oplus b \tag{2.1}$$

The players are not allowed to communicate once the game starts. Hence, they must agree on a winning strategy beforehand. The first Tsirelson bound [10] defines the upper bound of the success rate of any classical strategy at 75%. Figure 2.1b shows an example where both players always return zero.



(a) The network structure of the CHSH game. Suppose that Alice and Bob are two players that are not allowed to communicate. First, the Referee sends a single bit to each player x, y and then each player responds with a single bit a, b. A game is won if $x * y = a \oplus b$.

bit x	bit y	condition	bit a	bit b	success
0	0	$a = b$	0	0	TRUE
0	1	$a = b$	0	0	TRUE
1	0	$a = b$	0	0	TRUE
1	1	$a \neq b$	0	0	FALSE

(b) All possible cases of the zero strategy. Both players agree to always send back to the referee the bit value zero. Depending on the input pair x, y the success equation can be simplified to a simple condition between a and b. Three out of four input pair cases succeed so the success probability is 75%. This is the maximum value possible from a classical strategy.

Figure 2.1: Demonstration of how a CHSH game is played

The upper bound for the success probability of the quantum strategy is 85% as demonstrated by the first Tsirelson bound [10]. The optimal quantum strategy consists of the following steps. Firstly, the referee generates an entangled pair of qubits and sends one qubit to Alice and one to Bob. Subsequently, based on the bit received, each player measures their qubit after applying an operation as shown in Figure 2.2. Finally, both players respond with a single bit value that depends on the value measured from their qubit. Appendix A contains a rephrasing of well-known proof for $x = 0 \wedge y = 0$, showing a success probability of 85%. Similar proofs exist for all input bit combinations, each with a success probability of 85%.

Alice:	**Bob:**
– $x = 0$: apply the I gate	– $y = 0$: apply the Ry($\frac{\pi}{4}$) gate
– $x = 1$: apply the H gate	– $y = 1$: apply the Ry($\frac{-\pi}{4}$) gate

Figure 2.2: For the optimal quantum strategy an entangled pair of qubits is shared between the two players. Then each player performs an operation on their qubit before performing a measurement. This operations depends on the bit they received from the referee at the start of the game. This table shows what operation is applied by each player for all possible input bits x and y.

2.3. Related work

Before conducting any experiments, it was crucial to understand what research exists that focuses on benchmarking quantum networks. The results of an in-depth literature review, as presented in Section 3.1, revealed two relevant sources [8] [9]. However, it is noteworthy that none of these sources employ the CHSH game for benchmarking purposes. Instead, their focus primarily revolves around developing new protocols to assess the performance of quantum networks.

Furthermore, in parallel with this research, three similar studies have been conducted on three distinct QNAs. Although these QNAs differ significantly from the CHSH game, it is worth noting that scientific resources, such as referenced papers and statistical analysis methods, were shared and discussed among these research groups. This collaborative approach fostered a comprehensive understanding of the subject matter and facilitated a non-fraudulent exchange of information among researchers.

2.4. Conclusions

The conclusion reached by this study is that the inclusion of the CHSH ultimately depends on whether the benchmark suite's testing requirements align with the qualities offered by the application. Specifically, the CHSH game provides two performance metrics for system evaluation: the average success probability of a single game and the number of successes per second. Through these performance metrics, the CHSH application demonstrates sensitivity to compound properties, including the quality of entanglement between nodes, the execution time, and the error probability of both single-qubit gates and measure operations. Additionally, by comparing the success probabilities of different input combinations, we uncovered a method for calculating the difference between the probabilities of incorrectly measuring zero and one. Therefore, if this information is desirable in our benchmark suite, it is recommended to include the CHSH game.

2.5. Paper Overview

The paper is structured as follows. Chapter 3 presents the research methodology, detailing the literature review and experiments we conducted to address the research question. Building upon this, we showcase the results of these experiments in Chapter 4. Next, in Chapter 5, we provide a critical analysis of the ethical aspects of this research. Subsequently, Chapter 6 contains our reflection on the undertaken process and the derived results. Lastly, Chapter 7 summarizes the entire report, assessing the fulfillment of the research question and providing recommendations for future studies.

3. Methodology

This chapter will focus on the research methods we used during this study. Namely, Section 3.1 will present the results from a literature survey that we conducted to collect all relevant literature on quantum network benchmarking. Next, Section 3.2 will explain how SquidASM [11] was used in this research to perform experiments. Finally, Section 3.3 will provide an overview of the experimental setup and procedure.

3.1. Literature Review Results

Before conducting any experiments, we investigated any relevant work on quantum network benchmarking. The primary objective of this review is to identify any similar research and determine how our study is unique. For interested readers, Appendix D contains a more comprehensive explanation of the research plan we used to retrieve these results.

The literature review uncovered two relevant papers, summarized below. Firstly, Helsen et al.[8] propose a benchmarking protocol that estimates the quality of a path in a quantum network based on the standard randomized benchmarking protocol [7]. Secondly, Lee et al.[9] present a framework for quantifying the usefulness of a quantum network by extending IBM's Quantum Volume concept [4]. These two papers share a common approach of utilizing hypothesis-driven research to leverage existing knowledge from the field of quantum computing and apply it to the domain of quantum networking.

In contrast, this study will adopt an exploratory research approach to potentially repurpose existing quantum networking knowledge for benchmarking purposes. This alternative approach can potentially yield new theories and untapped ideas. Moreover, the scarcity of results from the literature review highlights that benchmarking quantum networks is an emerging concept, which greatly emphasizes the need for further research. Consequently, the research undertaken in this paper holds significant value.

3.2. How will SquidASM be used in this research

For this research, we used SquidASM [11] and its dependencies to perform experiments. These are the NetSquid quantum network simulator [3] and the Network Quantum Assembly (NetQASM) language [5]. This section explains SquidASM's relevant functionality for this project which is visualized in Figure 3.1.

SquidASM [11] is an open-source project that enables developers to create and simulate QNAs. In simple terms, to simulate a quantum network, SquidASM requires a QNA written in Python and a configuration file specifying all of the simulated network's properties. It then compiles this QNA to NetQASM and transfers this code to the hardware for execution. However, in our case, the execution was simulated using NetSquid.

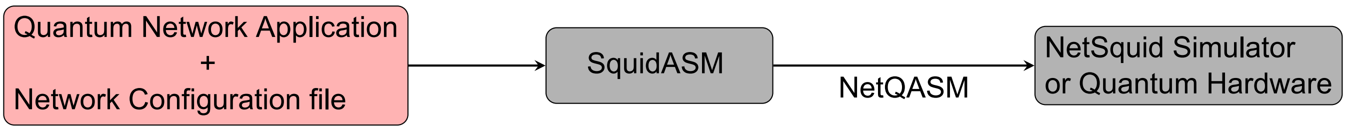


Figure 3.1: A simplification of the SquidASM pipeline. SquidASM converts a high level Quantum Network Application down to NetQASM [5] and it then simulates a quantum network running this application using a Network Configuration file and NetSquid [3].

We utilized SquidASM for this research in the following manner. Firstly, we adjusted this CHSH game implementation [1] to the SquidASM notation. Secondly, we used the Network Configuration files shown in Appendix C to simulate independent ideal quantum networks running the same application. In each experiment, the only difference between these networks was the value of the independent variable. Finally, SquidASM was used to set bit x and bit y at the beginning of each simulation and retrieve bit a, bit b, and the simulated time at the end.

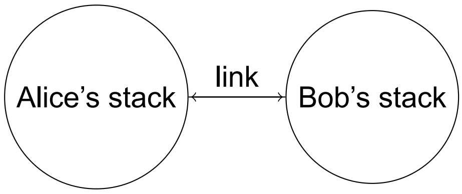
[1] https://github.com/QuTech-Delft/netqasm/tree/develop/netqasm/examples/apps/chsh

3.3. Experimental Procedure

In this Section, we will explain the experiments we have devised to determine the usefulness of the CHSH game as a benchmark for quantum networks. Initially, in Subsection 3.3.1, we will outline the structure of the experiments and describe the data collection process. Additionally, in Subsection 3.3.2, we will discuss post-experiment data processing. Lastly, in Subsection 3.3.3, we will present our speculations for the results and how we will answer the sub-questions introduced in Section 2.1.

3.3.1. Experiment structure

To execute the CHSH game, we first need to simulate a quantum network. Figure 3.2a illustrates the structure of the simulated quantum network used in all experiments. Since SquidASM only supports two-node networks, we eliminated the Referee node. Instead, the input bits x and y are predetermined before each experiment, while the bits a and b are obtained upon its completion. This simplification does not affect the success probability of the application. For every experiment, a specific stack and link type combination must be selected from Figure 3.2b.



(a) Structure of the Simulated network used in all experiments. Due to SquidASM's limitations the referee needed to be omitted. Instead bits x,y are passed to each stack before the start of the simulation and bits a,b are retrieved once it is done.

stack type	link type	ID
generic	magic state distributor	GD
generic	heralded	GH
nitrogen-vacancy	magic state distributor	NVD
nitrogen-vanancy	heralded	NVH

(b) All possible combinations of stack and link types that SquidASM supports. The Wehner magic state distributor link type is referred to as the depolarized link in SquidASM's [12] documentation.

Figure 3.2: The Structure of a simulated network used in each experiment and a table with all possible stack and link combinations that could be simulated.

The simulated quantum network is nearly perfect. Specifically, for each experiment, we employ a simulated quantum network, where we set all properties to their ideal values except for one property, which serves as the independent variable for the experiment. Appendix C contains all used ideal values. This setup enables us to evaluate the sensitivity of the CHSH game to variations in the independent variable.

Moreover, it is crucial to consider the data collected from each experiment. For a single CHSH game, the values of all input and output bits are recorded, along with the total simulated time in nanoseconds. For each independent variable value, we simulate eight thousand games that we divide into five batches. We do this to accurately determine the network's average success probability and its standard error of measurement. Hence, each batch comprises an equal number of all possible combinations of input values for bit x and bit y. Additionally, the range of values for the independent variable spans from fifteen to twenty equally spaced values. Finally, this data is collected and stored in a worksheet file for processing. Each file's name is formed by concatenating the identifier of the network shown in Figure 3.2b and the independent variable's name.

3.3.2. Experiment data processing

To determine the suitability of the CHSH game as a benchmarking tool, it becomes crucial to define metrics that assess the network's performance. Consequently, we have established the following performance metrics: the success probability of running the game and the number of successes per second. After each experiment, we calculate these performance metrics for each batch, averaging them to generate two plots, as depicted in Figure 3.3. These plots illustrate the variation of the performance metrics across different values of the independent variable. Furthermore, we display the standard error of measurement for each property value using error bars.

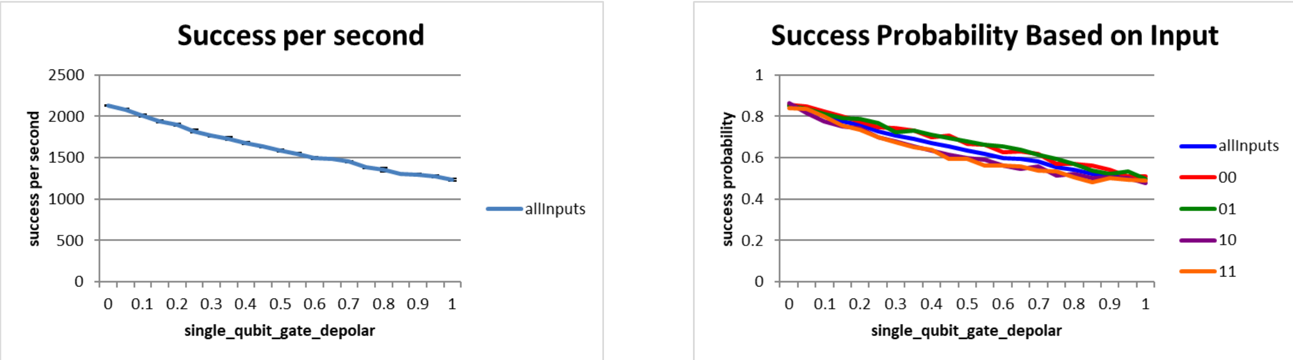


Figure 3.3: Example performance metrics plots for the network's $single\_qubit\_gate\_depolar$ property. For this experiment a magic state distributor link was used with generic Qdevices as stacks. The error bars are calculated for both metrics. In reality, a third plot exist that shows the error bars for the average success probability. However, to save space it was not included in this figure. Furthermore, we do not plot the error bars for all input combinations. The two bits shown in the right picture's legend state what is the value of the x bit and then the y bit.

However, simply visualizing the data is not sufficient. It is imperative to conduct a statistical analysis to gain meaningful insights into the variation of the performance metrics. Firstly, we require a statistical test to objectively determine if a performance metric is sensitive to the variance of the independent variable. Secondly, it would be beneficial to have a comparative method to quantify the deviation of the success probability plots based on different input combinations. The following paragraphs will explain why we have chosen the one-way ANOVA test [6] and Root Mean Square Error (RMSE) metric for these roles.

In this study, we employed the one-way ANOVA test [6] to analyze the success probability per game and the number of successes per second metrics. We chose this test as it allows us to detect if the variance between batches with different property values differs from that within batches of the same property value. Hence, if the resulting p-value is below the threshold of 0.05, it indicates significant variation among the property values. Therefore, we can objectively conclude that the performance metric is sensitive to detecting errors in that specific property.

Furthermore, we used the RMSE metric to calculate the deviations of the success probability plots across different input combinations. By comparing the RMSE values, we hope to uncover valuable patterns in the data that enable us to predict the value of a property. The RMSE was chosen based on two reasons. To begin with, it allows for straightforward interpretation as it expresses the deviations in the same unit of measurement as the dependent variable. Additionally, the RMSE applies a heavier penalty to larger errors. Consequently, higher RMSE values indicate significant variation or dispersion in the success probabilities among the various inputs.

3.3.3. Expected Results

Before presenting our findings, we will mention our initial speculations. At the same time, we will address how we have used the tools mentioned in Subsection 3.3.2 to answer the main research question. In particular, we will focus on the sub-questions we introduced in Chapter 2.

Sub-questions:

1. How sensitive is the CHSH game in recognizing errors in the properties of the total quantum networking system?
2. Can we use the CHSH game to make quantitative predictions about a specific property?
3. Should the CHSH game be included in the QNBS?

To address the first sub-question, we present a table that outlines the specific performance metrics of the CHSH game, emphasizing if they exhibit sensitivity to variations in each property. Specifically, if the p-value of the one-way ANOVA test on the performance metric is less than 0.05, then it is sensitive to this property. However, we do not expect the CHSH game to be affected by all properties. Namely, our performance metrics should be insensitive to any property relating to quantum operations not present in the CHSH game. Examples include two-qubit gates and qubit initialization properties. Additionally, assuming that the total execution time of the application is less than the memory lifetimes of the qubits, then the memory lifetime properties should not affect our performance metrics.

Next, we investigated the second sub-question by evaluating the following hypothesis. We hypothesized that by comparing the success probability across different input combinations, we could discover a method for determining the value of a property. To perform this comparison, we utilized the RMSE metric to compare the deviation of two input combinations across different experiments. Subsequently, we examined the experiments with the highest RMSE values. If the prediction of a property value seemed feasible, we tried to form a mathematical proof.

Finally, we decided to answer the last sub-question based on the results obtained from the first two. Specifically, we expected to form a condition that states whether we should add the CHSH game to the benchmark suite, depending on the specific properties we aim to evaluate. However, while seeking to find the properties that our performance metrics are sensitive to, it is also crucial to consider that a scenario where both performance metrics are affected by all properties may not be desirable. In other words, it may be more beneficial for our performance metrics to be sensitive to only a few properties. In such cases, the CHSH game would possess the valuable characteristic of being a specialized test for that specific set of properties.

4. Results

This chapter summarizes the key findings from the statistical analysis conducted on all experiments. It is important to note that the property names mentioned in this chapter correspond to the ones utilized in SquidASM's configuration files, as illustrated in Appendix C. Firstly, we will present the results of the one-way ANOVA test to determine the performance metrics that exhibit sensitivity in detecting errors within specific properties. Subsequently, an analysis of the RMSE values reveals that it is possible to make quantitive predictions on two property values based on the deviation between the success probability across different input combinations.

Throughout this study, we have examined a total of twenty-three properties. Table 4.1 highlights the sensitivity of each performance metric to different properties. We define that a performance metric is sensitive to a property if the p-value of the one-way ANOVA test performed on that performance metric is less than 0.05. Specifically, properties marked in red are exclusively detectable through the success probability metric, while those in blue are exclusively detectable through the number of successes per second metric. Additionally, green properties are sensitive to both metrics.

It is crucial to emphasize that the results of the one-way ANOVA test presented in Table 4.1 do not accurately reflect reality. This discrepancy arises from the fact that we consistently simulate quantum networks without any delays for the red properties. Consequently, the number of successes per second metric always yields an infinite value. As a result, we could not use the one-way ANOVA test in this scenario. However, in practical situations, we suspect that both metrics would exhibit a strong correlation. Thus, in reality, the red properties would affect both metrics.

W.M.S.D link	Heralded link	Generic QDevice	Nv Qdevice
fidelity	length	init_time	electron_init_depolar_prob
t_cycle	p_loss_length	single_qubit_gate_time	electron_single_qubit_depolar_prob
prob_succ	p_loss_init	two_qubit_gate_time	prob_error_0
	dark_count_prob	measure_time	prob_error_1
	detector_eff	single_qubit_gate_depolar_prob	carbon_init_depolar_prob
	visibility	two_qubit_gate_depolar_prob	carbon_z_rot_depolar_prob
			ec_gate_depolar_prob
			measure

Table 4.1: This table presents the properties whose experiments yielded a p-value less than 0.05 for the one-way ANOVA test. Consequently, it highlights which performance metric is sensitive to each property. Red properties affect the success probability metric, blue properties affect the number of successes per second metric, and green properties affect both metrics. It is important to note that we could not perform the one-way ANOVA test on the red properties for the successes per second metric as all values of the metric were infinite. We have assigned all properties to the column of the component they belong to. W.M.S.D stands for Wehner Magic State Distributor.

In the following analysis, we observed a distinct pattern in the success probability plots regarding the probability of incorrectly measuring zero ($prob\_error\_0$) and one ($prob\_error\_1$) properties, as depicted in Figure 4.1. The plots indicate a substantial deviation in the success probability for inputs $x = 1 \wedge y = 1$ compared to the other input combinations. This observation gains further support through the RMSE values between the success probability of inputs $x = 1 \wedge y = 0$ and the success probability of inputs $x = 1 \wedge y = 1$ shown in Figure 4.2. Remarkably, the RMSE values for these properties are significantly higher than any subsequent property, underscoring the need for additional investigation.

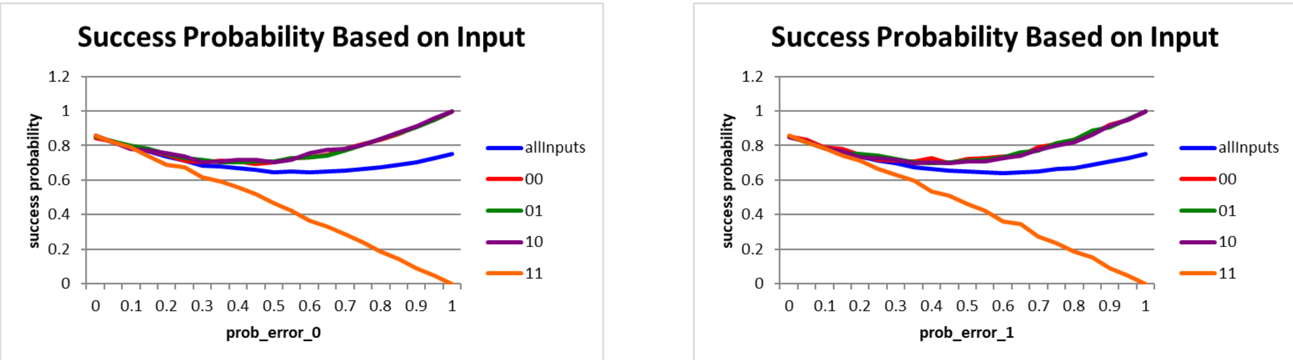


Figure 4.1: Success probability based on input plots for the $prob\_error\_0$ and $prob\_error\_1$ properties. The line for inputs $x = 1 \wedge y = 1$ deviates significantly from the rest of the inputs. This motivates the investigation of the mathematical relation between the two properties to determine any important patterns.

Subsequently, we delved into the mathematical relationship between the two properties and the success probability. A comprehensive mathematical proof in Appendix E establishes the validity of Equation 4.1. This equation assumes an ideal quantum network with properties `prob_err_0` = e_0 and `prob_err_1` = e_1, where e_0 and e_1 range from zero to one. It states that the difference between the success probabilities for inputs $x = 1 \wedge y = 1$ and $x = 1 \wedge y = 0$ is equivalent to the squared difference between e_0 and e_1. However, the results from Figure 4.2 suggest that Equation 4.1 holds true even for non-ideal quantum networks, as all other properties generate a RMSE value close to zero between these input combinations.

filename	11_10_rms
NVD_prob_error_0.xlsx	0.4693469881
NVH_prob_error_1.xlsx	0.4683265802
NVH_prob_error_0.xlsx	0.4639727003
NVD_prob_error_1.xlsx	0.462203676
NVD_link_fidelity.xlsx	0.02008375321
GH_visibility.xlsx	0.01780349083

Figure 4.2: The six experiments with the highest root mean squared error values between the success probabilities of inputs $x = 1 \wedge y = 1$ and $x = 1 \wedge y = 0$. The high values of the `prob_error_0` and `prob_error_1` properties suggest that the deviation in the success probability, as depicted in Figure 4.1, is only present for these two properties. The small values assigned to all other properties may be attributed to the smaller size of the dataset available for each input combination.

$$\begin{aligned} &P(success|x = 1 \wedge y = 0) - P(success|x = 1 \wedge y = 1) \\ &\quad \text{(Both Equations were derived in Appendix E, labeled as E.2 and E.3)} \\ &\quad = 0.85(e0^2 + e1^2) - 0.70(e0 + e1) - 0.3e0e1 + 0.85 \\ &\quad\quad - (-0.15(e0^2 + e1^2) - 0.70(e0 + e1) + 1.7e0e1 + 0.85) \\ &\quad = (e0 - e1)^2 \end{aligned} \tag{4.1}$$

Moreover, for ideal quantum networks, Equation 4.2 can be used to calculate the exact value of $e0$ and $e1$. Assuming that we find $P(success|x = 1 \wedge y = 0) = P(success|x = 1 \wedge y = 1) = m$, then from Equation 4.1 we know that $e_0 = e_1 = e$. Hence, we can calculate e by substituting m into one of the two equations for the success probability. This method cannot be applied to non-ideal networks as the average success probability is affected by other properties.

$$\begin{aligned} &P(\text{success}|x = 1 \wedge y = 0) = m \\ &\quad \text{(Equations was derived in Appendix E, labeled as E.2)} \\ &\quad \Rightarrow 0.85(e_0^2 + e_1^2) - 0.7(e_0 + e_1) - 0.3e_0e_1 + 0.85 = m \\ &\quad \text{(Substitute } e = e_0 = e_1\text{)} \\ &\quad \Rightarrow 1.4e^2 - 1.4e + 0.85 = m \\ &\quad \text{(Therefore)} \\ &e = \frac{1.4 \pm \sqrt{1.4^2 - 4 * 1.4 * (0.85 - m)}}{2 * 1.4} \end{aligned} \tag{4.2}$$

5. Responsible Research

Initially, the research topic of quantum networking may appear harmless as it does not involve human subjects and therefore does not pose any risks to individuals or their rights. However, it is essential to ensure that generated code and data for this study adhere to the FAIR principles, which stand for findability, accessibility, interoperability, and reusability. In addition, all results should be easily verifiable. Therefore, in the following paragraphs we describe how we applied these principles to our study.

First and foremost, being able to find and access the generated data was our primary focus. To facilitate this, the dataset resulting from all experiments has been uploaded at 4TU.ResearchData [1], a public platform for storing and accessing data. The dataset includes metadata such as keywords and the field of study, enabling its discoverability by both humans and computers. Furthermore, a public repository has been established to host the code that was used in all experiments [2].

Next, we have used proper data management practices to ensure the interoperability and reusability of all generated data. Specifically, each experiment stores its data in a worksheet file using a row-by-row storage arrangement for processing. Hence, each row represents a distinct data point, while the columns represent different variables associated with the data. This organization enables future researchers to easily interface with these files using a worksheet program or any programming language. Furthermore, it is easier to reprocess to uncover undiscovered patterns.

Finally, we have guaranteed that our results are easily verifiable. All readers can reproduce the experimental setup by following the instructions in Appendix B. At the same time, the accompanied markdown file in the repository contains step-by-step instructions on running all experiments. In addition, all data processing is visible in the worksheet files showing all formulas and intermediate results. Lastly, the Python code developed for this project needed to be well-documented and of exceptional quality. Hence, all functions have a maximum length of twenty-five lines and include comprehensive descriptions. By doing so, readers will more easily understand the code, simplifying the verification of the study's results.

6. Discussion

During this research, we encountered certain limitations that necessitated our adjustment. This chapter will present these limitations and our corresponding responses. Additionally, we will explore potential alternative approaches that we could have pursued.

The primary limitation of this study arose from the utilization of SquidASM[11], the software employed for simulating quantum networks. As discussed in Subsection 3.3.1, SquidASM can simulate network structures with a maximum of two nodes and a single link. Consequently, it was necessary to modify the network structure of the CHSH game application, transitioning it from a three-node network to a two-node network. This adaptation involved removing the Referee node and establishing an EPR pair directly between the players. Additionally, before initiating the simulations, we assigned the input bits to each player and extracted the output bits from the nodes after the simulation. This simplification of the application enabled its use without encountering any drawbacks.

Furthermore, we encountered certain limitations, preventing us from assessing all possible properties. Firstly, we encountered difficulties with the qubit memory lifetime properties for Generic Qdevices, as they were not functioning as expected. Ideally, when the execution time of the application exceeds the memory lifetime of the qubits, a decrease in the success probability metric should be observed. However, contrary to this expectation, we found that this behavior was not occurring, suggesting the presence of a bug in the simulation of these properties. Consequently, we decided to skip evaluating all memory lifetime properties. Secondly, due to time constraints, we did not evaluate the gate execution times of the Nitrogen-Vacancy QDevice. Nonetheless, we anticipate these properties to behave similarly to their Generic Qdevice counterparts. Therefore, only the single-qubit gate execution times would affect the number of successes per second metric.

Finally, we must acknowledge the limitations of our data processing approach. During this research, we employed the RMSE metric to assess if we can use the CHSH game to make quantitive predictions

[1] Dataset can be found at: https://doi.org/10.4121/68ef97d7-8aeb-4fdc-92e3-a0947a53400c.v1

[2] Repository used for the study: https://gitlab.com/tmaliappis28/research-project

for a quantum network's property. However, our results may not provide a comprehensive analysis. By preserving our dataset for future research, we invite the use of alternative data processing methods to uncover new patterns.

7. Conclusions

We conclude this paper by addressing the main research question and sub-questions introduced in Chapter 2. For convenience, we restate them below. Specifically, we will form our answers based on the results obtained from Chapter 4 and compare them to our speculations found in Subsection 3.3.3. Furthermore, we will provide recommendations for the future.

Sub-questions:

1. How sensitive is the CHSH game in recognizing errors in the properties of the total quantum networking system?
2. Can we use the CHSH game to make quantitative predictions about a specific property?
3. Should the CHSH game be included in the QNBS?

To address the first sub-question, we have defined the success probability and the successes per second performance metrics. Then we performed experiments using SquidASM on simulated quantum networks with a single property as the independent variable. For each experiment, we used the one-way ANOVA test to determine if each performance metric is sensitive to the variations of a property. Our results show that our performance metrics are affected by sixteen out of twenty-three tested properties. Specifically, this includes all properties that affect the execution time of the application, the quality of entanglement between nodes, and the error probability of single-qubit gate and measuring operations.

For the second sub-question, we demonstrated a method to calculate the difference between the measuring error properties. Initially, we support our claim by constructing a mathematical proof for an ideal quantum network. In detail, our proof states that the square difference between the two measuring errors is equal to the difference in the success probabilities of inputs $x = 1 \wedge y = 0$ and $x = 1 \wedge y = 1$. Moreover, we observed that the RMSE of these success probabilities for all other experiments is nearly zero. Therefore, we concluded that this relationship possibly applies to non-ideal quantum networks.

For the final sub-question, we decided to form a condition that states whether we should add the CHSH game to the benchmark suite, depending on the specific properties we aim to evaluate. Namely, we should consider what each performance metric offers. Firstly, the success probability metric quantifies the fidelity of the quantum link and the error probabilities associated with single-qubit gates in the system. Secondly, the successes per second metric encompasses the aforementioned properties and the system's execution time for the application. Thirdly, we have seen that we can predict the relative difference between the measuring error properties. Therefore, if these qualities are desirable in our benchmark suite, then we should include the CHSH game.

Lastly, it is crucial to note that our study has the potential for further expansion. By observing the deviations in the success probability based on different inputs, we have uncovered a method for quantitative prediction. Consequently, we hypothesize that conducting more comparisons will enable the prediction of additional property values. Such deviations also exist within the visibility and dark count probability experiments. Unfortunately, due to the time constraints imposed on this project, we did not conduct an in-depth exploration of these deviations. However, in the future, employing additional data processing techniques can facilitate a deeper understanding and exploration of these patterns.

Reference Documents

[1] Bell, J. S. (1964). On the Einstein Podolsky Rosen Paradox. *Physics Physique Физика*, *1*(3), 195–200.

[2] Clauser, J. F., & Horne, M. A. (1974). Proposed experiment to test local hidden-variable theories. *Physical Review Letters*, *23*(15), 880–884. https://doi.org/10.1103/PhysRevLett.23.880

[3] Coopmans, T., Knegjens, R., Dahlberg, A., Maier, D., Nijsten, L., Oliveira Filho, J., Papendrecht, M., Rabbie, J., Rozpedek, F., Skrzypczyk, M., et al. (2021). Net-squid, a network simulator for quantum information using discrete events. *Communications Physics*, *4*(1), 164.

[4] Cross, A. W., Bishop, L. S., Smolin, J. A., & Gambetta, J. M. (2017). Quantum volume. *Nature*, *549*(7671), 195–202.

[5] Dahlberg, A., van der Vecht, B., Delle Donne, C., Skrzypczyk, M., Raa, I., Kozlowski, W., & Wehner, S. (2022). Netqasm–a low-level instruction set architecture for hybrid quantum–classical programs in a quantum internet. *Quantum Science and Technology*, *7*(3), 035023.

[6] Girden, E. R. (1992). *ANOVA: Repeated measures*. 84. Sage.

[7] Helsen, J., Roth, I., Onorati, E., Werner, A., & Eisert, J. (2022). General framework for randomized benchmarking. *PRX Quantum*, *3*(2). https://doi.org/10.1103/prxquantum.3.020357

[8] Helsen, J., & Wehner, S. (2021). A benchmarking procedure for quantum networks.

[9] Lee, Y., Dai, W., Towsley, D., & Englund, D. (2022). Quantum network utility: A framework for benchmarking quantum networks.

[10] Tsirelson, B. S. (1980). Quantum generalizations of bell's inequality. *Letters in Mathematical Physics*, *4*(2-3), 93–100.

[11] van der Vecht, B., Wehner, S., & Dahlberg, A. (2021). Squidasm github repo. https://github.com/QuTech-Delft/squidasm.

[12] van der Vecht, B., Wehner, S., & van Hooft, M. (2021). Squidasm docs. https://squidasm.readthedocs.io/en/latest/index.html.

A. Proof for the Upper bound for the CHSH inequality

The Tsirelson bound [10] was used to show that the CHSH inequality [2] has an upper bound of 85%. Many proofs exist but here we will rephrase the one derived from Wikipedia [1].

Proof. Suppose, that Alice and Bob each possess one qubit of the following 2-qubit entangled state: $|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ and let $x = y = 0$. By replacing the input bits in the equation $x * y = a \oplus b$ it is evident that the game will be won if and only if $a \oplus b = 0$ or $a = b$. It will be shown using a proof by case analysis that P($a = b$ | $|\Phi\rangle$) = P($a = b = 0$ | $|\Phi\rangle$) + P($a = b = 1$ | $|\Phi\rangle$) = 0.85.

When Alice receives bit $x = 0$, she will measure her qubit in the basis $|0\rangle, |1\rangle$ and respond with 0 if the measurement outcome is $|0\rangle$, and 1 if it is $|1\rangle$.

$$Ry(\frac{\pi}{4}) = \begin{bmatrix} \cos\frac{\pi}{8} & -\sin\frac{\pi}{8} \\ \sin\frac{\pi}{8} & \cos\frac{\pi}{8} \end{bmatrix}$$

When Bob receives bit $y = 0$, he will apply the Ry($\frac{\pi}{4}$) gate and measure his qubit. Hence, the measurement will occur in the basis $|a_0\rangle, |a_1\rangle$ where $|a_0\rangle = (\cos\frac{\pi}{8})|0\rangle + (\sin\frac{\pi}{8})|1\rangle$ and $|a_1\rangle = (-\sin\frac{\pi}{8})|0\rangle + (\cos\frac{\pi}{8})|1\rangle$. He then responds with 0 if the result is $|a_0\rangle$, and 1 if it is $|a_1\rangle$.

$$P(a = b = 0||\Phi\rangle) = |(\langle 0| \otimes \langle a_0|)|\Phi\rangle|^2 = \frac{1}{2}\cos^2(\frac{\pi}{8}) = \frac{0.85}{2} \tag{A.1}$$

$$P(a = b = 1||\Phi\rangle) = |(\langle 1| \otimes \langle a_1|)|\Phi\rangle|^2 = \frac{1}{2}\cos^2(\frac{\pi}{8}) = \frac{0.85}{2} \tag{A.2}$$

Therefore, P($a = b$ | $|\Phi\rangle$) = P($a = b = 0$ | $|\Phi\rangle$) + P($a = b = 1$ | $|\Phi\rangle$) = 0.85.

□

For the other three input pairs, a similar derivation can show that the success probability is also 0.85. Therefore, the overall success probability of the quantum strategy is 85%.

B. Manual for duplicating the experimental setup

1. The host computer must be running Ubuntu 20.04.6 LTS (Focal Fossa) [1]
2. Create an account at the NetSquid forum [2]
3. Install the following list of dependencies in order:
   - sudo apt install make
   - sudo apt install pip
   - pip install netqasm
   - pip3 install pytest
   - pip install openpyxl
   - pip install xlsxwriter
   - pip3 install –extra-index-url https://pypi.netsquid.org netsquid
4. Install SquidASM [3]
5. Clone this repository https://gitlab.com/tmaliappis28/research-project

[1] https://en.wikipedia.org/wiki/CHSH_inequality#Optimal_quantum_strategy
[1] https://www.releases.ubuntu.com/focal/
[2] https://forum.netsquid.org/ucp.php?mode=register
[3] https://github.com/QuTech-Delft/squidasm

C. Perfect stack & link configuration files

```
# 1) Fidelity of entanglement (or link fidelity) between the EPR pair qubits a measure of how close the entangled pair is to the required entangled pair.
fidelity: 1

# 2) Time in nanoseconds for an attempt to generated entanglement
t_cycle: 0

# 3) Chance for each attempt at entanglement to succeed (non-zero)
prob_success: 1
```

Figure C.1: The depolarise link perfect config file used to simulate a quantum network without any noise or delays.

```
# 1) total length [km] of heralded connection
length: 1.0
# 2) attenuation coefficient [dB/km] of fiber on either side.
p_loss_length: 0.25
# 3) probability that photons are lost when entering connection on either side.
p_loss_init: 0.0
# 4) speed of light [km/s] in fiber on either side.
speed_of_light: 200_000
# 5) dark-count probability per detection
dark_count_probability: 0
# 6) probability that the presence of a photon leads to a detection event
detector_efficiency: 1.0
# 7) Hong-Ou-Mandel visibility of photons that are being interfered
visibility: 1.0
# 8) whether photon-number-resolving detectors used for the Bell-state measurement
num_resolving: False
```

Figure C.2: The heralded link configuration file used to simulate a perfect quantum networking with a single heralded link. The heralded link uses the double click model as developed and described by: https://arxiv.org/abs/2207.10579"""

```
# Memory lifetimes or coherence over time (state decay constants)
T1: 10_000_000_000 # longitudinal relaxation time (how long the state decays)
T2: 1_000_000_000 # transverse relaxation (dephasing time)

#### Gate execution times
init_time: 0 # Qubit initialization time
single_qubit_gate_time: 0 # Single qubit gate execution time
two_qubit_gate_time: 0 # Two qubit gate execution time
measure_time: 0 # Qubit measurement time

#### Gate nose model
# Probability of error in each single qubit gate operation
single_qubit_gate_depolar_prob: 0
# Probability of error in each two qubit gate operation
two_qubit_gate_depolar_prob: 0
```

Figure C.3: Part of the perfect Generic Qdevice configuration file showing the ideal parameter values used.

```
# initialization error of the electron spin
electron_init_depolar_prob: 0
# error of the single-qubit gate
electron_single_qubit_depolar_prob: 0
# Chance of 0 being measured as 1
prob_error_0: 0
# Chance of 1 being measured as 0
prob_error_1: 0
# initialization error of the carbon nuclear spin
carbon_init_depolar_prob: 0
# error of the Z-rotation gate on the carbon nuclear spin
carbon_z_rot_depolar_prob: 0
# error of the native NV two-qubit gate
ec_gate_depolar_prob: 0
# coherence times
electron_T1: 10_000_000_000
electron_T2: 1_000_000_000
carbon_T1: 10_000_000_000
carbon_T2: 1_000_000_000
# gate execution times
carbon_init: 0
carbon_rot_x: 0
carbon_rot_y: 0
carbon_rot_z: 0
electron_init: 0
electron_rot_x: 0
electron_rot_y: 0
electron_rot_z: 0
ec_controlled_dir_x: 0
ec_controlled_dir_y: 0
measure: 0
```

Figure C.4: Part of the perfect NV-Qdevice configuration file showing the ideal parameter values used.

D. Literature Review Research Plan

This Appendix outlines the process of investigating relevant work on quantum network benchmarking. It contains a comprehensive explanation of the targeted information sources and the construction of the search query we used to retrieve the results shown in Section 3.1.

For this literature review, the main types of document required are journal articles and conference papers. This is because they can provide scientific information and current research results relevant to the field of quantum networking. Consequently, the following resources were searched. Firstly, two resources were chosen relating to quantum computing from the TU delft WorldCat website [1]. These were Nature quantum Information [2]and PRX quantum [3]. Secondly, two multidisciplinary resources of scientific information were chosen. These were, Google Scholar [4]and arXiv [5]. By utilizing domain-specific and multidisciplinary resources, this investigation broadened its scope.

The general searchable query utilized for this literature review is depicted in Figure D.1. To summarize, the objective of this overview was to identify any research about benchmarking quantum networks.

[1] https://tudelft.on.worldcat.org/atoztitles/browse/journals
[2] https://www.nature.com/search?q=""
[3] https://journals.aps.org/search
[4] https://scholar.google.com/
[5] https://arxiv.org/search/advanced

The main concepts of interest in this inquiry are ”quantum networking” and ”benchmarking”. From these concepts, several synonyms were defined, and the following initial query was formulated.

Question: What research has been done on benchmarking quantum networks?
Key Concepts: quantum networking, benchmarking
Query: (“Quantum network*” OR ”Distributed quantum computing”) AND (“Benchmark suite” OR ”Testing suite” OR ”Metric collection” OR ”Assessment tool” OR ”Comparative analysis”)

Figure D.1: Initial Search Query that was used to find related work about Quantum Networking Benchmark suites. The question shown in the figure is broken down into key concepts which are then concatenated in a query with several synonyms.

It is worth noting that the query shown in Figure D.1 was modified for each resource during the search process. These adjustments were carried out to filter out irrelevant findings or to reduce the number of results to less than a hundred. The definitive compilation of queries employed for each resource can be found in Figure D.2. Simultaneously, results were filtered to ensure their timeliness, focusing on those published from 2018 onwards to gather up-to-date information. Among all the retrieved results, only the first fifty most relevant were subject to title scanning to confirm their relevance. Based on these results, the conclusions in Section 3.1 were drawn.

1. Npj quantum Information: ((quantum network* OR distributed quantum comput*) AND benchmark*)
   - Additional filters: (Article type = Research)
   - Number of Results: 110
2. PRX quantum: (quantum network* OR quantum computing*)
   - Additional filters: None
   - Number of Results: 26
3. Google Scholar: (”Quantum network*”) AND ((”Benchmark*”) AND (“suite” OR “tool”))
   - Additional filters: None
   - Number of Results: 305
4. arXiv: (”Quantum network*”) AND (”Benchmark*” OR “comparative analysis”)
   - Additional filters: None
   - Number of Results: 60

Figure D.2: List of Queries that were used in the literature review for each resource that was used. Only results from 2018 and onwards were considered and the fifty most relevant ones were subject to title scanning to confirm relevance. This list also specifies the number of results that the query retrieves as well as any additional filters that were used.

E. Success probability and error of measurement proof

Claim 1. *Suppose, the CHSH game is played with all of the following assumptions satisfied.*

Assumption 1. *Suppose, the CHSH game is played using the following 2-qubit entangled state:* $|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ *and let* x, y, a, b *be bits where* $x = 1 \land y = 0$. *It can be shown using a proof similar to the one shown in Appendix A that:*

- $P(a = 0 \land b = 0 \mid |\Phi\rangle) = P(a = 1 \land b = 1 \mid |\Phi\rangle) \approx \frac{0.15}{2}$
- $P(a = 0 \land b = 1 \mid |\Phi\rangle) = P(a = 1 \land b = 0 \mid |\Phi\rangle) \approx \frac{0.85}{2}$

Assumption 2. *Suppose, the CHSH game is played using the following 2-qubit entangled state:* $|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ *and let* x, y, a, b *where* $x = y = 1$. *It can be shown using a proof similar to the one shown in Appendix A that:*

- *P*($a = 0 \wedge b = 0$ | $|\Phi\rangle$) = *P*($a = 1 \wedge b = 1$ | $|\Phi\rangle$) $\approx \frac{0.85}{2}$
- *P*($a = 0 \wedge b = 1$ | $|\Phi\rangle$) = *P*($a = 1 \wedge b = 0$ | $|\Phi\rangle$) $\approx \frac{0.15}{2}$

Assumption 3. *Suppose, the quantum network uses nitrogen-vacancy stacks where all properties are set to their ideal value except the probability that a 0 is measured as 1 and the probability that 1 is measured as 0*

Abbreviations:

- *P*($a = 0 \wedge b = 0$ | $|\Phi\rangle$) = *P(00)*
- *P*($a = 0 \wedge b = 1$ | $|\Phi\rangle$) = *P(01)*
- *P*($a = 1 \wedge b = 0$ | $|\Phi\rangle$) = *P(10)*
- *P*($a = 1 \wedge b = 1$ | $|\Phi\rangle$) = *P(11)*
- *Probability 0 measured as 1 = e0*
- *Probability 1 measured as 0 = e1*

Then,

$$P(success|x = 1 \wedge y = 0) \geq P(success|x = 1 \wedge y = 1) \tag{E.1}$$

***Proof by Contradiction**.* Suppose that there exist values for $e0, e1$ such that E.1 does not hold. These values will be determined.

$$\begin{aligned} P(success|x = 1 \wedge y = 0) &= P(00)(1 - e0)^2 + P(01)(1 - e0)e1 + P(10)e1(1 - e0) + P(11)e1^2 \\ &+ P(00)e0^2 + P(01)e0(1 - e1) + P(10)(1 - e1)e0 + P(11)(1 - e1)^2 \\ &\text{(Substitute probabilities from Assumption 1 and simplify)} \\ &= 0.85(e0^2 + e1^2) - 0.70(e0 + e1) - 0.3e0e1 + 0.85 \end{aligned} \tag{E.2}$$

$$\begin{aligned} P(success|x = 1 \wedge y = 1) &= P(00)(1 - e0)e0 + P(01)(1 - e0)(1 - e1) + P(10)e1e0 + P(11)e1(1 - e1) \\ &+ P(00)(1 - e0)e0 + P(01)e0e1 + P(10)(1 - e1)(1 - e0) + P(11)(1 - e1)e1 \\ &\text{(Substitute probabilities from Assumption 2 and simplify)} \\ &= -0.15(e0^2 + e1^2) - 0.70(e0 + e1) + 1.7e0e1 + 0.85 \end{aligned} \tag{E.3}$$

$$E.2 - E.3 < 0 => (e0 - e1)^2 < 0 \tag{E.4}$$

Contradiction, the inequality shown in E.4 cannot be satisfied. Therefore, the proof's assumption that there exist values for $e0, e1$ for which E.1 does not hold is False. Thus, E.1 holds for all values of $e0, e1$. □