

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
UNIVERSITY SAAD DAHLEB OF BLIDA 1
FACULTY OF SCIENCES
PHYSICS DEPARTMENT



MASTER THESIS

Theoretical Physics

DEALING WITH QUANTUM COMPUTER READOUT
NOISE THROUGH HIGH ENERGY PHYSICS
UNFOLDING METHODS

Prepared by :

Hacene Rabah BENAÏSSA

Imene OUADAH

Defended on September 20, 2021 before the jury composed of :

Dr. K. A. BOUTELDJA	MAA	USDB1	President
Dr. A. MOUZALI	MCB	USDB1	Examiner
Dr. S. A. YAHIAOUI	MCB	USDB1	Examiner
Dr. A. YANALLAH	MCB	USDB1	Supervisor
Dr. N. BOUAYED	MCA	USDB1	Co-Supervisor

BLIDA1, 2020/2021

DEDICATION

إلى أمي وأبي الغاليين لو ملأت كل الدفاتر و الأوراق شكرا و تقديرا لكما لن أوفي ولو القليل من حقكما. الكون على اتساعه لا يضاهي سعة قلبكما. شكرا لكما على زرع الطموح في نفسي منذ الصغر، شكرا لكما على عطائكما و دعائكما الغير المنقطع. رضاكما وابتسامتكما نعمة لا تقدر بثمن.

إلى اخوتي سندي في الحياة لا أتذكر كم مرة طلبت منكم شراء الأوراق البيضاء لي. لم تترددوا يوما على الخروج لإحضارها. شكرا لكم على دعمكم و تشجيعاتكم لي.

إلى أختي الصغيرة التي لا تفارقني أينما ذهبت. اعلم انه كان من الصعب عليك أن تتركيني أعمل لساعات طوال دون قضاء الوقت معا كما اعتدنا سابقا. شكرا لك على كل زيارتك المشجعة لي. يمكننا أخيرا اللعب معا بعد التخرج.

أهدي لكم جميعا هذا العمل الذي استمتعت كثيرا بتحضيره. شكرا لكم لوجودكم في حياتي.

Imene OUADAH

To my mother and my father who have been so generous to me.
To my family, who provided me with a suitable space-time to achieve this work.
To over 100 teachers who taught me,

from أ، ب، ت، ... to $\alpha, \beta, \gamma, \dots$, and from $1 + 1$ to $|1\rangle \otimes |1\rangle$.

All this brought me right here... and **beyond!**

Hacene Rabah BENAÏSSA

ACKNOWLEDGMENTS

First and foremost we would like to express our deepest gratitude to our teachers and supervisors Dr. A. YANALLAH and Dr. N. BOUAYED for their helpful advice, insightful suggestions, and for their deep belief in our work. We really appreciated the learning process in the world of quantum computing. We can't thank them enough for offering us this wonderful opportunity. From the first year in the university, they were always a source of inspiration to us with their hard-working and teaching style. We have learned so much from them over the past years. We will forever be thankful for their patience and kindness and time to open our minds to new possibilities.

We have special thanks to IBM for giving the public free access to its quantum computers and to Dr. B. NACHMAN for sharing with us some Data of his work; it was very helpful when we started our learning process.

We would like to extend our sincere thanks to the jury members. We are very pleased that Dr. K. A. BOUTELDJA is the jury president of our thesis and we are very fortunate that the expert in this field Dr. A. MOUZALI accepted to be a jury member in our master thesis. We would like to offer our special thanks to our teacher and jury member Dr. S. A. YAHIAOUI for his interesting lectures and discussions and for caring about each individual student. Sincere appreciations to all the jury members of this master thesis for accepting reading this manuscript and for their valuable remarks and observations.

We also would like to thanks all our teachers in the department of physics and all who is working hard to provide better education to the future generations.

ABSTRACT

Quantum computers have the potential to solve problems that are intractable to classical computers, nevertheless they have high error rates. One significant kind of errors is known as “*Readout Errors*”. Current methods, as the *matrix inversion* and *least-squares*, are used to unfold (correct) readout errors. But these methods present many problems like oscillatory behavior and unphysical outcomes.

In 2020 BENJAMIN NACHMAN et al. suggested a technique currently used in HEP, to correct detector effects. This method is known as the “*Iterative Bayesian Unfolding*” (IBU), and they have proven its effectiveness in mitigating readout errors, avoiding problems of the mentioned methods.

Therefore, the main objective of our thesis is to mitigate readout noise of quantum computers, using this powerful unfolding method. For this purpose we generated a *uniform distribution* in the YORKTOWN IBM Q MACHINE, for 5 Qubits, in order to unfold it by IBU after being distorted by noise. Then we repeated the same experiment with a *Gaussian distribution*. Very satisfactory results and consistent with those of B. NACHMAN et al., were obtained. After that, we took a second purpose to explore unfolding in a larger qubit system, where we succeed to unfold a *uniform distribution* for 7 Qubits, distorted by noise from the MELBOURNE IBM Q MACHINE. In this case, the IBU method showed much better results than other techniques.

Keywords : Quantum computing, Qubit, Noise, Readout errors, Unfolding, High energy physics, Iterative Bayesian Unfolding, IBM Yorktown quantum computer, IBM Melbourne quantum computer.

ملخص

تمتلك الحواسيب الكمومية القدرة على حل المشكلات المستعصية على الحواسيب الكلاسيكية ، غير أن لديها معدلات أخطاء عالية. يعرف أحد أهم أنواع هذه الأخطاء بـ "أخطاء القراءة". يتم استخدام طرائق حالية ، مثل طريقة *matrix inversion* و *least-squares* ، لبسط (تصحيح) أخطاء القراءة هذه ، لكن هذه الأساليب تقدم العديد من المشاكل مثل السلوك التذبذبي والنتائج الغير فيزيائية.

في عام 2020 اقترح بنجامين ناشمان وآخرون تقنيةً مستخدمة حالياً في HEP ، لتصحيح تأثيرات الكاشف. تعرف هذه الطريقة باسم *Iterative* "IBU" (*Bayesian Unfolding*) ، وقد أثبتوا فعاليتها في التخفيف من أخطاء القراءة ، وتجنب مشاكل الطرق المذكورة.

لذلك ، فإن الهدف الرئيسي من أطروحتنا هو التخفيف من ضوضاء القراءة الصادرة من أجهزة الكمبيوتر الكمومية ، باستخدام هذه الطريقة القوية للبسط. لهذا الغرض ، أنشأنا توزيعاً منتظماً في آلة YORKTOWN IBM Q ، لـ 5 Qubits ، من أجل بسطها بواسطة IBU بعد تشويهاها بالضوضاء ، ثم كررنا نفس التجربة بتوزيع غاوسي. تم الحصول على نتائج مرضية للغاية ومتسقة مع نتائج ب. ناشمان وآخرون. بعد ذلك ، اتخذنا هدفاً ثانياً لاستكشاف البسط في نظام كيوبتي أكبر ، حيث بسطنا توزيعاً منتظماً لـ 7 Qubits ، مشوه بالضجيج الصادر عن آلة MELBOURNE IBM Q. في هذه الحالة ، أظهرت طريقة IBU نتائج أفضل بكثير من التي أظهرتها التقنيات الأخرى.

الكلمات المفتاحية: الحاسوب الكومومي ، الكيوبت ، الضوضاء ، أخطاء القراءة ، البسط ، فيزياء الطاقة العالية ، IBU ، حاسوب YORKTOWN IBM الكمي ، حاسوب MELBOURNE IBM الكمي.

List of Figures	vii
Introduction	1
1 Overview of Quantum Computing and Quantum Information	3
1.1 Superposition and Entanglement	4
1.1.1 Quantum Superposition Principle	4
1.1.2 Quantum Entanglement	5
1.1.2.1 Bell States	8
1.2 The Quantum Bit (Qubit)	10
1.2.1 Classical Bit	10
1.2.2 Single Qubit	12
1.2.3 Multiple Qubits	15
1.2.4 Physical Implementation of Qubit	18
1.3 Quantum Gates and Quantum Circuits	20
1.3.1 Classical Gates	20
1.3.2 Quantum Gates	22
1.3.3 Quantum Circuit Diagrams	33
1.4 Quantum Algorithms	42
1.4.1 Deutsch-Jozsa Algorithm	42
1.4.2 Shor's Algorithm	45
1.5 Error Detection and Correction	48
1.5.1 Classical Error Correction	49
1.5.2 Quantum Error Correction	50
1.5.2.1 Single Bit-Flip Error Correction Code	51
1.5.2.2 Single Sign-Flip Error Correction Code	53
1.5.2.3 The 9-Qubit Error Correction Shor Code	53
2 Unfolding for High Energy Physics	57
2.1 Inverse Problem and Unfolding Methods	58
2.2 Unfolding with Matrix Inversion	59
2.3 Iterative Bayesian Unfolding Method	60

3	Unfolding Quantum Computer Readout Noise	63
3.1	Readout Errors	63
3.2	Constructing the Calibration Matrix \mathcal{R}	65
3.2.1	One Qubit Calibration Matrix	66
3.2.2	Two Qubit Calibration Matrix	67
3.3	Unfolding Methods for QIS	69
3.3.1	Unfolding with Matrix Inversion Method	69
3.3.2	Unfolding with IBU	72
3.4	Simulation of the Unfolding Process with Python	76
3.4.1	Noise Generation with Python	76
3.4.2	Constructing R	80
3.4.3	Unfolding Methods	81
3.5	Unfolding Readout Errors from the Yorktown IBM Q Machine	86
3.5.1	Construction of the Response Matrix from the Yorktown IBM Q Machine	87
3.5.2	Unfolding a Uniform Distribution	92
3.5.3	Unfolding a Gaussian Distribution	101
3.6	Unfolding Readout Errors from The Melbourne IBM Q Machine	118
	Conclusion	123
	Appendices	
A	Boolean Algebra	127
B	Boolean Lows	128
C	Qubit Design	130
D	Python and Qiskit toolkit	134
E	Glossary	137
	References	139

List of Figures

1.1	The Double-Slit experiment.	5
1.2	A binary bit is true if a voltage is present and false if the voltage is absent. . . .	11
1.3	Geometrical representation of the 2D complex state vector $ \psi\rangle$, mapped by the Bloch vector onto real 3D space called the Bloch sphere , of <i>radius</i> $r = 1$, <i>colatitude angle</i> θ , and <i>azimuth angle</i> ϕ . Each qubit state is represented by a point. The state $ 0\rangle$ is at the north pole, and $ 1\rangle$ at the south pole.	13
1.4	The two extremum states of the qubit. In the left: the Bloch vector of the state $ \psi\rangle = 0\rangle$. In the right: the Bloch vector of the state $ \psi\rangle = 1\rangle$, for $(\theta = 0, \pi; \forall\phi)$, respectively.	14
1.5	Representation of a superposition balance state $ +\rangle$ for $(\theta = \frac{\pi}{2}, \phi = 0, r = 1)$. . .	14
1.6	The Q-sphere : a visualization of the entangled 2-qubit state $ \Phi^-\rangle$	17
1.7	Layout of IBM's four superconducting Qubit device	19
1.8	A visualization of the resulting states from the \mathbb{I} action on the states $ 0\rangle$ and $ 1\rangle$, respectively.	23
1.9	A visualization of the resulting states from the \mathbb{X} action on the states $ 0\rangle$ and $ 1\rangle$, respectively.	24
1.10	A visualization of the resulting states from the \mathbb{Y} action on the states $ 0\rangle$ and $ 1\rangle$, respectively.	25
1.11	A visualization of the resulting states from the \mathbb{Z} action on the states $ +\rangle$ and $ -\rangle$, respectively.	26
1.12	A visualization of the resulting states from the \mathbb{H} action on the states $ 0\rangle$ and $ 1\rangle$, respectively.	26
1.13	A visualization of the resulting states from the \mathbb{S} action on the states $ +\rangle$ and $ -\rangle$, respectively. (For the \mathbb{S}^\dagger gate, the figures will be inversed).	27
1.14	A visualization of the resulting states from the \mathbb{T} action on the states $ +\rangle$ and $ -\rangle$, respectively.	28
1.15	A visualization of the resulting states from the \mathbb{T}^\dagger action on the states $ +\rangle$ and $ -\rangle$, respectively.	29
1.16	A visualization of the resulting states from the $\sqrt{\mathbb{X}}$ action on the states $ 0\rangle$ and $ 1\rangle$, respectively.	30
1.17	A visualization of all 4 possible resulting states from the SWAP action on the 2-qubit state $ \alpha\beta\rangle$; $\alpha, \beta = 0, 1$	31
1.18	A visualization of all 4 possible resulting states from the \mathbb{CX} action on the 2-qubit state $ \alpha\beta\rangle$; $\alpha, \beta = 0, 1$	32
1.19	A visualization of all 8 possible resulting states from the \mathbb{CCX} action on the 3-qubit state $ \alpha\beta\gamma\rangle$; $\alpha, \beta, \gamma = 0, 1$	33

1.20	The interactive interface of IBM quantum composer.	33
1.21	Types of quantum gates.	34
1.22	The quantum circuit that generates the Bell state $ \Phi^+\rangle$	37
1.23	Two equivalent quantum circuits that generate the Bell state $ \Phi^-\rangle$	38
1.24	A visualization of the Bell states: $ \Phi^+\rangle$ (on the left) and $ \Phi^-\rangle$ (on the right).	39
1.25	The quantum circuit that generates the Bell state $ \Psi^+\rangle$	40
1.26	Two equivalent quantum circuits that generate the Bell state $ \Psi^-\rangle$	40
1.27	A visualization of the Bell states: $ \Psi^+\rangle$ (on the left) and $ \Psi^-\rangle$ (on the right).	40
1.28	A variety of controlled gates (from the left to the right): \mathbb{RCH} , \mathbb{CS} , \mathbb{CSWAP} , \mathbb{CZ} , $\mathbb{C}\sqrt{\mathbb{X}}$	41
1.29	The Toffoli gate (on the left) and its implementation (on the right).	41
1.30	Generic circuit for the Deutsch-Jozsa algorithm	42
1.31	Deutsch-Jozsa circuit example for a single input (n=1).	44
1.32	The classical error correction process.	49
1.33	Single bit flip error correction circuit.	52
1.34	Single sign flip error correction circuit.	53
1.35	SHOR error correction circuit.	54
2.1	Representation of an example of Smeared and True histogram [23].	58
2.2	Unfolding using matrix inversion: The true distribution (upper left), smeared distribution and points are data drawn from this distribution (upper right), unfolded data from matrix inversion with true distribution (Bottom left), error bars for the unfolded data (Bottom right) [25].	59
3.1	Measuring a 3-Qubit system of state $ 000\rangle$	64
3.2	The Frequency of measuring each basis state from the IBM Q 5 YORKTOWN MACHINE, for 8000 shots.	64
3.3	Measuring a qubit in the state $ 0\rangle$	66
3.4	Measured and unfolded counts via the matrix inversion method.	72
3.5	Measured and predicted bin counts via MI and IBU for 10^3 iterations.	75
3.6	Histogram representation of simulated true data and measured data.	79
3.7	Histogram representation of simulated true data and measured data.	81
3.8	Unfolding results of a Gaussian distribution using three methods: Matrix inversion (\hat{t}_{matrix}), ignis (\hat{t}_{ignis}), and IBU (\hat{t}_{IBU}) with 10 iterations and a uniform prior truth spectrum.	85
3.9	Response matrix for 5 qubits from the IBM Q Yorktown machine.	92
3.10	A circuit that generates a uniform distribution based on 5 qubits.	94
3.11	Unfolding results of a uniform distribution using three methods: Matrix inversion, ignis, and IBU with one iteration and a uniform prior truth spectrum.	101
3.12	A circuit that generates a Gaussian distribution based on 5 qubits.	104

3.13	Unfolding results of a Gaussian distribution using three methods: Matrix inversion, ignis, and IBU with one iteration and a prior truth spectrum $t_0=t_s$	111
3.14	Unfolding results of a Gaussian distribution using three methods: Matrix inversion, ignis, and IBU with one iteration and a prior truth spectrum $t_0=t_trig$	112
3.15	A comparison of unfolding results for a Gaussian distribution using IBU with three prior truth distributions: t_unif , t_s , t_trig and with 1 iteration.	113
3.16	A comparison of unfolding results for a Gaussian distribution using IBU with three prior truth distributions: t_unif , t_s , t_trig and with 10 iterations.	114
3.17	Unfolding results for a uniform distribution using IBU with 1,2,3,10 iterations and a uniform prior truth spectrum.	116
3.18	Unfolding results for a Gaussian distribution using IBU with 1,2,3,10 iterations and a prior truth spectrum $t_0=t_trig$	117
3.19	Response matrix for 7 qubits from the IBM Q Melbourne machine.	119
3.20	A circuit that generates a uniform distribution based on 7 qubits.	120
3.21	Unfolding results of a uniform distribution based on seven-qubits using 3 methods: Matrix inversion, ignis, and IBU with one iteration and a uniform prior truth spectrum.	121
C.1	Harmonic and anharmonic systems and their suitability as qubits [34].	131
C.2	The three basic Josephson-junction qubit circuits and their potential-energy landscapes, with the two lowest energy levels marked in red. (a) Charge qubit. (b) Flux qubit. (c) Phase qubit. For simplicity, the capacitance C_J is only shown in panel (a) , although it is also present in the circuits in panels (b) and (c) [34].	132

Introduction

We are in a new scientific revolution. A mystery box is open every day to enlighten a new era of science. Interdisciplinary research with collaboration and a spirit of a lifelong learner are the keys to gain a more well-developed perspective and solving problems whose solutions transcend the scope of a single discipline.

Forty years on, the theoretical physicist Richard Feynman published a paper: *simulating physics with computers* [1]. He suggested that to simulate the quantum behavior of a physical system, we need to use quantum computers. One of the significant advantages of quantum computing is the ability to massively parallel computations, using quantum superposition. In the last years, simulations of chemistry and condensed materials have developed as one of the dominant applications of quantum computing, offering an exponential speedup of solutions. However, quantum computers are noisy in consequence the results of calculations will have errors. One of the most important classes of errors are the readout errors.

For years several methods have been proposed in order to minimize these readout errors and to approach as close as possible to the true value after the correction of the measurement. This problem is known as the inverse problem and a well-known method is the Matrix inversion. This method leads to unphysical entries (negative values) and has oscillatory behavior. For this, we need a new method that will give us better results to reduce readout errors.

In 2020, Ben. Nachman et al. published a paper where he explained how they use a well-known method used in HEP and applied to quantum computing. The study focuses on a technique called “iterative Bayesian unfolding” (IBU). In their study, they simulated a quantum computer to compare the performance of three different error-correction: Matrix inversion, ignis, and IBU. The iterative Bayesian unfolding method have been proven very successful compared to the two other methods.

In this thesis, we want to explore more about this method by unfolding errors from a real quantum computer. That means we want to use measured data from a quantum computer and correct them with the three unfolding methods.

It is interesting to see the effect of noise in a real quantum computer on our data. For this, we will use two distributions as true data. The first one will be the uniform distribution, and the second one will be the Gaussian distribution. After the measurement, we will see how a noisy quantum computer will affect our data. Finally, we will use the unfolding methods to correct our data, and then, we will compare the three methods.

In summary, the plan of this thesis to reach our objective is as follows:

- **Chapter 1**

In the first chapter, we will present an overview of quantum computing and quantum information. We start with fundamental concepts of quantum mechanics: superposition and entanglement. Then, we will explain what is the difference between quantum and classical bit. Next, we will see the most famous quantum gates and examples of quantum circuits. After that, we briefly review examples of quantum Algorithms. We finish this chapter by explaining the concepts of error detection and correction.

- **Chapter 2**

This chapter is an introduction to unfolding for high-energy physics. First, we explain the concept of the inverse problem and unfolding methods. Then, we introduce two methods, the first one is the matrix inversion and the second one is the iterative Bayesian unfolding. The IBU method is what interests us the most.

- **Chapter 3**

After having all the needed tools, the last chapter is about applying unfolding methods from High Energy Physics to quantum information science (QIS). We start by explaining what is a readout error and the methodology to construct a calibration matrix of a real quantum computer. Then, we will see how to use unfolding methods from HEP to Quantum computing. Next, we will see how to simulate the unfolding process with python. Finally, we apply unfolding techniques to mitigate the readout noise of the 5 Yorktown IBM Q machine with 5 qubits and the 16 Melbourne IBM Q machine with 7 qubits.

Finally, we draw our conclusion and, additional details are submitted in appendices.

1

Overview of Quantum Computing and Quantum Information

Contents

1.1 Superposition and Entanglement	4
1.1.1 Quantum Superposition Principle	4
1.1.2 Quantum Entanglement	5
1.1.2.1 Bell States	8
1.2 The Quantum Bit (Qubit)	10
1.2.1 Classical Bit	10
1.2.2 Single Qubit	12
1.2.3 Multiple Qubits	15
1.2.4 Physical Implementation of Qubit	18
1.3 Quantum Gates and Quantum Circuits	20
1.3.1 Classical Gates	20
1.3.2 Quantum Gates	22
1.3.3 Quantum Circuit Diagrams	33
1.4 Quantum Algorithms	42
1.4.1 Deutsch-Jozsa Algorithm	42
1.4.2 Shor's Algorithm	45
1.5 Error Detection and Correction	48
1.5.1 Classical Error Correction	49
1.5.2 Quantum Error Correction	50
1.5.2.1 Single Bit-Flip Error Correction Code	51
1.5.2.2 Single Sign-Flip Error Correction Code	53
1.5.2.3 The 9-Qubit Error Correction Shor Code	53

Quantum information science QIS is a new field that combines *quantum physics* with *classical information science*. The foundations of *quantum information theory* were established in the late 1980s by CHARLES BENNETT and others [2].

In this chapter, we introduce the basic elements of QIS: *qubit*, *quantum logic gates*, *quantum circuit*, *quantum algorithms*, *quantum error correction*. But before tackling any of these topics, we first provide a glimpse of their classical analogs.

This chapter represents a reference to all that will follow because it provides us with all the necessary tools to achieve the goal of this thesis.

The quantum mechanics concepts on which QIS stands are *superposition* and *entanglement*, we, therefore, begin the chapter by recalling them.

1.1 Superposition and Entanglement

In this section, we recall the two concepts of quantum mechanics *superposition* and *entanglement*, then, we present *Bell states*.

1.1.1 Quantum Superposition Principle

The principle of superposition of states is a fundamental principle in quantum mechanics, which represents a property that every quantum object has. If a quantum object can be either in a state ψ_1 or in another state ψ_2 , then the superposition principle announce that this object is partially in ψ_1 and partially in ψ_2 , and the general description of its state must be viewed as the result of a *superposition* of the two original states

$$\Psi = c_1\psi_1 + c_2\psi_2 \quad (1.1)$$

where c_1, c_2 are complex coefficients.

The new state Ψ is completely defined by the original states ψ_1 and ψ_2 and has *intermediate properties* which are closer to the state with the greatest *probability amplitude* (c_1, c_2) [3].

We can cite illustrative examples of the superposition of states such as:

- An electron can be in a state of *spin-up* $|\uparrow\rangle$ and *spin-down* $|\downarrow\rangle$, *simultaneously*.

$$|\Phi\rangle_S = a|\uparrow\rangle + b|\downarrow\rangle$$

- An atom that can be both *excited* $|\phi^{(1)}\rangle$ or *not-excited* $|\phi^{(0)}\rangle$.

$$|\Phi\rangle_E = a|\phi^{(0)}\rangle + b|\phi^{(1)}\rangle$$

- A photon which can be in a superposition of *linear polarization* states; polarized *horizontally* $|\leftrightarrow\rangle$ and *vertically* $|\updownarrow\rangle$, at the same time.

$$|\Phi\rangle_P = a|\leftrightarrow\rangle + b|\updownarrow\rangle$$

We can generalize this concept to a number n ($n \in \mathbb{N}$) of states $\{\psi_1, \psi_2, \dots, \psi_n\}$ that can a system occupy:

$$\Psi = \sum_n c_n \psi_n \quad ; \quad c_n \in \mathbb{C} \quad (1.2)$$

It expresses the general state Ψ as a *linear combination* of the *discrete* original states ψ_n .

Although the *superposition principle* exists in *classical mechanics*, it is fundamentally different from any classical property, DIRAC wrote about the *non-classical nature* of superposition:

“the *superposition that occurs in quantum mechanics is of an essentially different nature from any occurring in the classical theory, as is shown by the fact that the quantum superposition principle*” [3].

The simplest experiment demonstrating this principle is the “*double-slit experiment*”, where we emit from a source a beam of *monochromatic light* to pass through an intermediate screen with two *identical slits*, that must be very *thin*, to finally hits a *detection screen*, Figure 1.1.

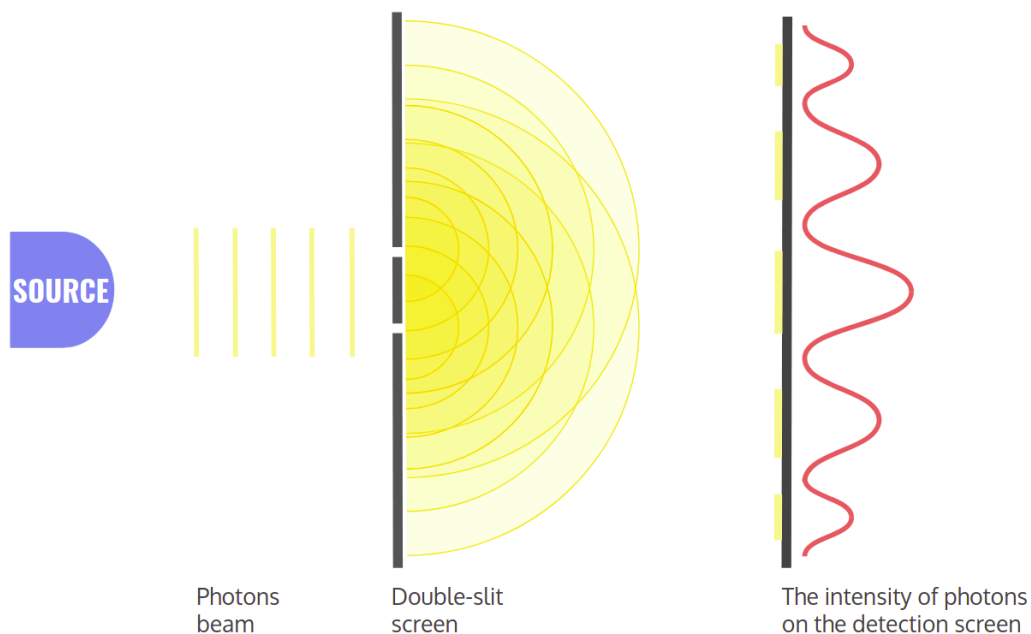


Figure 1.1: The Double-Slit experiment.

When crossing the two slits, the beam splits into two components and the two components are then caused to *interfere*. We need the *corpuscular* aspect of light, so consider a single photon γ . Arriving at the first screen, the photon will be in a *superposition* of the two paths; the photon is in a general state given by the *sum* of the two states associated with the two paths. This superposition leaves marks in the form of *interference fringes* in the *detection screen*, when we measure the position of the photons.

The principle of superposition gives rise to an important concept, called “*entanglement*”.

1.1.2 Quantum Entanglement

The entanglement phenomenon is a direct consequence of the superposition principle applied to composite systems, that occurs when a pair or group of particles is generated, and which are considered as an inseparable whole; where each constituent cannot be fully described without considering the other.

Quantum systems can become entangled through various types of interactions between subatomic particles. Experimentally there are various methods; we can cite a well-known one, named “*the spontaneous parametric down-conversion*” (SPDC) which is an optical process that converts one photon of higher energy into a pair of entangled photons of lower energy. In nature, a pair of entangled particles can be created for example by the decaying phenomena as the “*pair production process*” [4]:

$$\gamma \longrightarrow e^+ + e^-$$

where the pair (e^+ , e^-) are *entangled*; they form a *single system*.

For the mathematical aspect, consider a system formed by two particles, where the first particle is in the state $\psi \in \mathcal{H}_1$ and the second particle is in the state $\phi \in \mathcal{H}_2$, where \mathcal{H}_1 and \mathcal{H}_2 are respectively their *Hilbert-spaces*. To represent the **general** state of the *two-particle system*, we define the “*Tensor product*” of the two particles states $\psi \otimes \phi$, which is a vector in a new *Hilbert-space* written $\mathcal{H}_1 \otimes \mathcal{H}_2$ that will carry the description of the quantum states of this composite system. This state is expressed as a **linear superposition** of the product of the two-particle states $\psi_i \otimes \phi_j$:

$$\Upsilon = \sum_{ij} \alpha_{ij} \psi_i \otimes \phi_j \quad ; \Upsilon \in \mathcal{H}_1 \otimes \mathcal{H}_2 \quad (1.3)$$

where α_{ij} are the corresponding scalar components.

If Υ is an **entangled** state of the two particles, then it cannot be dissociated into separate states of each of them. Conversely, if one can write it as a single term $\psi_* \otimes \phi_*$ for some $\psi_* \in \mathcal{H}_1$ and $\phi_* \in \mathcal{H}_2$; which means that the description of the particles state in Υ can be written independently (each particle in its state), hence, in the state Υ the particles are **not entangled** [5]. Otherwise, Υ is not entangled if and only if $\alpha_{ij} = \alpha_i \alpha_j$, which leads to:

$$\Upsilon = \Psi \otimes \Phi \quad (1.4)$$

where $\Psi = \sum_i \alpha_i \psi_i$ and $\Phi = \sum_j \alpha_j \phi_j$

And vice versa, if $\alpha_{ij} \neq \alpha_i \alpha_j$, then the separation will be impossible.

We can illustrate this reasoning by taking *two-dimensional Hilbert-spaces* \mathcal{H}_1 and \mathcal{H}_2 . Let (e_1, e_2) the basis for \mathcal{H}_1 and (s_1, s_2) basis for \mathcal{H}_2 . Then, the general state is written as:

$$\Upsilon_G = g_{11} e_1 \otimes s_1 + g_{12} e_1 \otimes s_2 + g_{21} e_2 \otimes s_1 + g_{22} e_2 \otimes s_2 \quad (1.5)$$

This state is encoded by a matrix of coefficients:

$$\mathcal{G} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \quad (1.6)$$

If we have understood the reasoning above, the state is *not entangled* if there exist arbitrary constants $\alpha_1, \alpha_2, \beta_1, \beta_2$ such that:

$$\Upsilon_{\mathcal{G}} = (\alpha_1 e_1 + \alpha_2 e_2) \otimes (\beta_1 s_1 + \beta_2 s_2) \quad (1.7)$$

By expanding the expression (1.7) than identifying the constants with (1.5), we find the four equalities below:

$$g_{11} = \alpha_1 \beta_1 \quad ; \quad g_{12} = \alpha_1 \beta_2 \quad ; \quad g_{21} = \alpha_2 \beta_1 \quad ; \quad g_{22} = \alpha_2 \beta_2 \quad (1.8)$$

Which therefore leads to the following condition:

$$\mathcal{G} = \begin{pmatrix} \alpha_1 \beta_1 & \alpha_1 \beta_2 \\ \alpha_2 \beta_1 & \alpha_2 \beta_2 \end{pmatrix} \longrightarrow \det(\mathcal{G}) = 0 \quad (1.9)$$

So we conclude that $\Upsilon_{\mathcal{G}}$ is **not entangled** if and only if $\det(\mathcal{G}) = 0$. This result is available just for *two-dimensional* vector spaces [5].

We will introduce at next the most **important** FEATURE of the ENTANGLEMENT PHENOMENON. For this let's take the example of the *neutral meson* η_0 decay, which decays into an *entangled pair* of a *muon* μ^- and an *anti-muon* μ^+ .

$$\eta_0 \longrightarrow \mu^- + \mu^+ \quad (1.10)$$

In this decay process, the various conservation laws are respected (total energy, momenta, spin, ...). As known the η_0 is **spinless** and which decays to two particles with *spin-(1/2)*, Since the conservation laws require that the total *initial* and *final* spins must be *zero*. Thus if we found after a measurement that the 1st particle is *spin-up* on some axis, then the 2nd particle is found to be *spin-down* when measured on the same axis [5].

We can check this mathematically by firstly describing the entangled pair (μ^+, μ^-) state of total spin zero as follow:

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|+\rangle \otimes |-\rangle - |-\rangle \otimes |+\rangle) \quad (1.11)$$

where $|+\rangle \equiv |\uparrow\rangle$ and $|-\rangle \equiv |\downarrow\rangle$ are respectively, the *spin-up* and *spin-down* states. Also, $S_n^T |\psi\rangle = 0$, where S_n^T is the *total n-component of spin angular momentum*, given by $S_n^T = S_n \otimes \mathbb{1} + \mathbb{1} \otimes S_n$ and n an arbitrary direction.

If we measure $|\psi\rangle$ along the basis $\{|+\rangle, |-\rangle\}$, then the probability \mathbb{P} of $|\psi\rangle$ to be in the

states $|\pm\pm\rangle$, $|\pm\mp\rangle$ and $|\psi\rangle$ after measurement are respectively:

$$\mathbb{P}_{\pm\pm} = |\langle\pm\pm|\psi\rangle|^2 = 0 \quad ; \quad |\psi\rangle \longrightarrow |\psi\rangle_{\pm\pm} = 0 \quad (1.12)$$

$$\mathbb{P}_{\pm\mp} = |\langle\pm\mp|\psi\rangle|^2 = \frac{1}{2} \quad ; \quad |\psi\rangle \longrightarrow |\psi\rangle_{\pm\mp} = \pm|\pm\mp\rangle \quad (1.13)$$

These results allow us to say that the two *entangled* particles have always *opposite* spins. Hence, the measurement of one of them will affect the other, regardless of how much further they are apart! Hence the ENTANGLED particles are PERFECTLY CORRELATED¹.

ALAIN ASPECT describes the difference between classical and quantum correlations in bipartite systems as follows:

“If we have a pair of identical twins we do not know what their blood type is before testing them, but if we determine the type of one, we know for sure that the other is the same type. We explain this by the fact that they were born with, and still carry along, the same specific chromosomes that determine their blood type. Two entangled photons are not two distinct systems carrying identical copies of the same parameter. A pair of entangled photons must instead be considered as a single, inseparable system, described by a global wave function that cannot be factorized into single-photon states” [7].

1.1.2.1 Bell States

Bell states are defined as a *set of maximally entangled basis states*. Named after the Irish physicist JOHN STEWART BELL [5].

We aim in what follows, to construct these states in the particular case of a *two-particles* system of *spin-(1/2)*, which are described in the *Hilbert-space* $(\mathcal{H}_1 \otimes \mathcal{H}_2)$, where \mathcal{H}_1 and \mathcal{H}_2 are both *two-dimensional Hilbert-spaces*. Let's suppose that these two particles are produced in the specific *normalized* and *entangled* state below:

$$|\Phi_0\rangle = \frac{1}{\sqrt{2}} (|+\rangle_1 |+\rangle_2 + |-\rangle_1 |-\rangle_2) \quad (1.14)$$

We can quickly verify these two properties, respectively:

$$\langle\Phi_0|\Phi_0\rangle = \frac{1}{2} \langle++|++\rangle + \frac{1}{2} \langle--|--\rangle = 1 \quad (1.15)$$

$$\mathcal{G} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \longrightarrow \det(\mathcal{G}) = \frac{1}{2} \quad (1.16)$$

¹ This ENTANGLEMENT FEATURE seems to violate the *local realism* which gave rise in 1935 to what has been known as THE EINSTEIN-PODOLSKY-ROSEN (EPR) PARADOX [6].

The dimensionality of $(\mathcal{H}_1 \otimes \mathcal{H}_2)$ can be calculated by the following relation:

$$\dim(\mathcal{H}_1 \otimes \mathcal{H}_2) = \dim(\mathcal{H}_1) \times \dim(\mathcal{H}_2)$$

Since $\dim(\mathcal{H}_1 \otimes \mathcal{H}_2) = 4$, we therefore need four *entangled* states to construct our *Bell basis*. If we consider the state $|\Phi_0\rangle$ as the first one, then the other three can be generated by applying the operators $(\mathbb{1} \otimes \sigma_k)$ on $|\Phi_0\rangle$ [5].

$$|\Phi_k\rangle = (\mathbb{1} \otimes \sigma_k) |\Phi_0\rangle \quad ; \quad k = 1, 2, 3 \quad (1.17)$$

where σ_k are the *Pauli matrices*, and their actions on $|\pm\rangle$ are as follows:

$$\sigma_1 |\pm\rangle = |\mp\rangle \quad ; \quad \sigma_2 |\pm\rangle = \pm i |\mp\rangle \quad ; \quad \sigma_3 |\pm\rangle = \pm |\pm\rangle \quad (1.18)$$

where $\sigma_1 = \sigma_x$, $\sigma_2 = \sigma_y$, $\sigma_3 = \sigma_z$.

Applying this, we get the full list of *Bell states*, which are explicitly written:

$$|\Phi_0\rangle = (\mathbb{1} \otimes \mathbb{1}) |\Phi_0\rangle = \frac{1}{\sqrt{2}} (|+\rangle_1 |+\rangle_2 + |-\rangle_1 |-\rangle_2) \quad (1.19)$$

$$|\Phi_1\rangle = (\mathbb{1} \otimes \sigma_1) |\Phi_0\rangle = \frac{1}{\sqrt{2}} (|+\rangle_1 |-\rangle_2 + |-\rangle_1 |+\rangle_2) \quad (1.20)$$

$$|\Phi_2\rangle = (\mathbb{1} \otimes \sigma_2) |\Phi_0\rangle = \frac{i}{\sqrt{2}} (|+\rangle_1 |-\rangle_2 - |-\rangle_1 |+\rangle_2) \quad (1.21)$$

$$|\Phi_3\rangle = (\mathbb{1} \otimes \sigma_3) |\Phi_0\rangle = \frac{1}{\sqrt{2}} (|+\rangle_1 |+\rangle_2 - |-\rangle_1 |-\rangle_2) \quad (1.22)$$

By a simple calculation it will be clear that they form an *orthonormal* basis

$$\langle \Phi_\mu | \Phi_\nu \rangle = \delta_{\mu\nu} \quad (1.23)$$

where

$$|\Phi_\mu\rangle = (\mathbb{1} \otimes \sigma_\mu) |\Phi_0\rangle \quad ; \quad \mu = 0, 1, 2, 3 \quad (1.24)$$

and $\sigma_0 = \mathbb{1}$.

Finally, we have an *orthonormal basis* of *entangled states* which defines the ***Bell basis***.

1.2 The Quantum Bit (Qubit)

In this section, we introduce the *basic information unit* of QIS the “qubit”, then we explore its *physical implementation*. But before that, let’s take a look at its *classical equivalent*, the “bit”.

1.2.1 Classical Bit

In our daily life, we use decimal numbers for almost everything: weight, height, phone number, etc. These numbers are written in **Base-ten**. For this, we use ten digits: from 0 to 9, all of the numbers use some combination of these ten digits; each number can be decomposed into successive powers of 10. For example:

$$2021 = 2 \times 10^3 + 0 \times 10^2 + 2 \times 10^1 + 1 \times 10^0. \quad (1.25)$$

The number 9 is the last digit in base-ten. To go beyond 9, we need to change the row. If the row of units is complete, we start the row of tens, and we put the units back to zero. We continue the counting until 99, when we run out of digits, we add another 0 and one on the left (100) and so on.

In decimal, at every change of powers of 10, one passes from the units (10^0) to the tens (10^1), then to the hundreds (10^2). It is the same thing for **base-2**, but they are the powers of 2.

For example,

$$(1011)_2 = 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 = (11)_{10}. \quad (1.26)$$

A computer works with electric current, and to communicate with an electronic device, we must send electric signals to it. A computer’s processor is made up of transistors; we can imagine these transistors as small switches that can be turned ON or OFF; they let the current pass or not like a light switch. The processor interprets and executes the instructions by turning millions of connected transistors ON and OFF.

In the language of a computer, ON and OFF are represented by the numbers 1 and 0. This language is called the **binary language**; the computer does not work with decimal numbers but uses the binary system. *A single binary value is called a “bit”.*

A bit (binary digit) is the most basic unit of information in classical computing. To store the value of the bit, we use a capacitor inside a memory device. The value is stored as either above or below a specific level of electric charge in a capacitor; it can hold only one of two values: 0 or 1.

Any number from our decimal system can be represented in binary; 0 in decimal is 0 in binary. Similarly, 1 in decimal is 1 in binary. 2 in decimal becomes 10 in binary².

The table below shows how decimal numbers convert to binary.

Decimal	0	1	2	3	4	5	6	7	8	...	15
Binary	0000	0001	0010	0011	0100	0101	0110	0111	1000	...	1111

²Create a new column and start afresh, using 1 and 0.

Bits are so small, it is uncommon to work with only one. Bits are usually put together in a series of eight-bit, to form a “byte”. A byte contains sufficient information to store a single *ASCII* character³, such as “O”: 01001111 and “B”: 01000010.

There is a unique binary number for each letter in upper and lower case. By knowing the binary number, we can write complete words with a series of 1s and 0s. For example:

- Benaissa: 01000010 01100101 01101110 01100001 01101001 01110011 01110011 01100001
- Ouadah: 01001111 01110101 01100001 01100100 01100001 01101000

Physical Implementation of Bit

Computers represent data as voltage or current signals at the hardware level. They could be analog, digital, or hybrid. Computers that use analog representations exclusively are uncommon. It is challenging to control physical signals with a high degree of accuracy.

In a digital computer, the input is converted and transmitted as electrical pulses and represented by two individual states, ON and OFF. The sequence of ON and OFF forms the electrical signals that the computer can understand.

In positive logic, the presence of a voltage is called: “1”, true or high state. If the voltage is absent then this is called the “0”, false or low state. The Figure 1.2 shows the output of a typical *complementary metal-oxide-semiconductor* (CMOS) circuit. The left side shows the condition with a true bit, and the right side shows a false. The output of each digital circuit consists of a P-type transistor “on top of” an N-type transistor.

In digital circuits, each transistor is essentially ON or OFF. If the transistor is ON, it is equivalent to a short circuit between its two output pins. If the transistor is OFF, it is equivalent to an open circuit between its outputs pins.

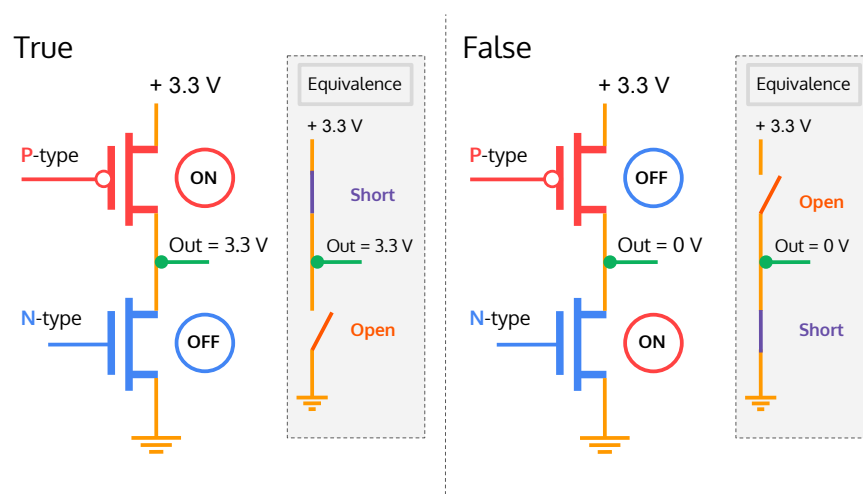


Figure 1.2: A binary bit is true if a voltage is present and false if the voltage is absent.

³Stands for "American Standard Code for Information Interchange" *ASCII* is a character encoding that uses numeric codes to represent characters.

1.2.2 Single Qubit

We saw that a classical bit is the elementary unit of information, which can be in one of two defined states 0 or 1. Similarly, a quantum bit or “**qubit**” is the elementary unit of quantum information in quantum computers, and which can be unlike the classical bit, in infinite states, but gives only an observed output of 0 or 1 [8].

A qubit is a *two-level* system that can be expressed using two orthonormal basis states:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad ; \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad (1.27)$$

where $|0\rangle$ is the qubit state which is certain to give 0 when measured, and $|1\rangle$ the qubit state that give 1 with certainty. The states $|0\rangle$ and $|1\rangle$ are called the “*computational basis states*”⁴.

In the quantum realm, the qubit is in a coherent **superposition** of these two states; hence the state describing it can be written as a linear combination of the vectors $|0\rangle, |1\rangle \in \mathbb{C}^2$ [8]:

$$|\psi\rangle = a|0\rangle + b|1\rangle \quad ; \quad a, b \in \mathbb{C} \quad (1.28)$$

When performing a measurement, the state $|\psi\rangle$ collapses irreversibly to either $|0\rangle$ or $|1\rangle$, with corresponding probabilities $|a|^2$ and $|b|^2$, respectively, where a and b represent the *probability amplitudes*. Furthermore, they achieve the relation $|a|^2 + |b|^2 = 1$, from the *normalization condition*, and which describes the *probability conservation*.

The fact that the qubit can be in infinite states, leads us to look for a representation, that is able to show the enormous *flexibility* of the qubit, in comparison with a classical bit.

Firstly, we can express the complex numbers (a, b) in *polar coordinates*; $a = |a|e^{i\phi_a}$ and $b = |b|e^{i\phi_b}$, then $|\psi\rangle$ becomes:

$$\begin{aligned} |\psi\rangle &= |a|e^{i\phi_a}|0\rangle + |b|e^{i\phi_b}|1\rangle \\ &= e^{i\phi_a}(|a||0\rangle + |b|e^{i\phi}|1\rangle) = e^{i\phi_a}|\chi\rangle \end{aligned}$$

where $\phi = \phi_b - \phi_a$.

If we calculate the probability of measuring $|\psi\rangle$ in the state $|\mathbf{q}\rangle$; $\mathbf{q} = 0, 1$:

$$\mathbb{P}_{\mathbf{q}} = |\langle \mathbf{q} | \psi \rangle|^2 = |e^{i\phi_a} \langle \mathbf{q} | \chi \rangle|^2 = |\langle \mathbf{q} | \chi \rangle|^2 \quad (1.29)$$

where $|e^{i\phi_k}| = 1$; $k = a, b$.

From this result, we can see that the probability for the state $|\chi\rangle$ is identical to that for the state $(e^{i\phi_a}|\chi\rangle)$. Hence, we conclude that the two states are equivalent $|\chi\rangle \equiv e^{i\phi_a}|\chi\rangle$, $\forall \phi_a \in \mathbb{R}$; we cannot differentiate between $|\chi\rangle$ and $e^{i\phi_a}|\chi\rangle$ when we measure.

The factor $e^{i\phi_a}$ is called the “**global phase**” and it’s not observable. What we can measure

⁴ We can assign $|0\rangle$ and $|1\rangle$, respectively, to the *spin-up* $|\uparrow\rangle$ and *spin-down* $|\downarrow\rangle$ states of an *electron*, or the *horizontal polarization* $|\leftrightarrow\rangle$ and the *vertical polarization* $|\updownarrow\rangle$ of a *photon*.

is the phase difference ($e^{i(\phi_b - \phi_a)}$) between the states $|0\rangle$ and $|1\rangle$, which is called the “**relative phase**” [8].

From this fact, we can choose the *complex* coefficient a to be *real*, then the *normalization condition* $|a|^2 + |b|^2 = 1$, can be identified with the *trigonometric identity* $\cos^2(x) + \sin^2(x) = 1$. Since $|a|^2$ and $|b|^2$ are *probabilities*, their values are confined between 0 and 1, i.e.

$$0 \leq (|a| = \sqrt{\mathbb{P}_0}; |b| = \sqrt{\mathbb{P}_1}) \leq 1 \implies 0 \leq (\cos(x); \sin(x)) \leq 1$$

So $0 \leq x \leq \pi/2$. For later reasons we take $x = \theta/2$, thus we can write:

$$|a| = \cos(\theta/2) \quad ; \quad |b| = \sin(\theta/2)$$

We can now describe the state of any qubit through the two variables ϕ and θ :

$$|\psi\rangle = \cos(\theta/2) |0\rangle + e^{i\phi} \sin(\theta/2) |1\rangle, \quad (1.30)$$

where $0 \leq \theta \leq \pi$, $0 \leq \phi < 2\pi$. If we interpret θ and ϕ as *spherical coordinates* (with $r = \langle \psi | \psi \rangle = 1$), then we can plot any single-qubit state on the surface of a unit sphere, known as the “*Bloch sphere*”. Each **point** on its surface represents a **qubit state** [8].

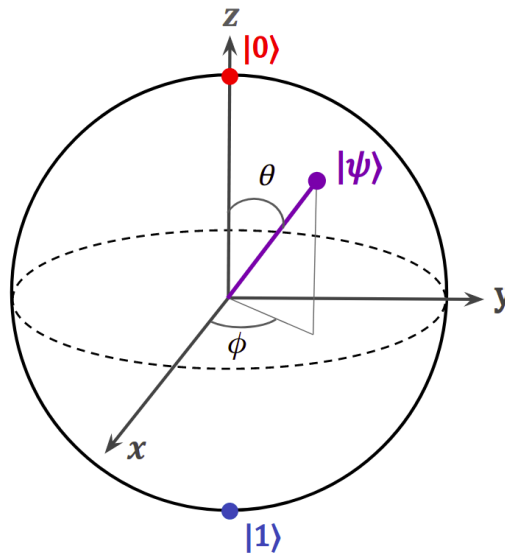


Figure 1.3: Geometrical representation of the 2D complex state vector $|\psi\rangle$, mapped by the **Bloch vector** onto real 3D space called the **Bloch sphere**, of radius $r = 1$, colatitude angle θ , and azimuth angle ϕ . Each qubit state is represented by a point. The state $|0\rangle$ is at the north pole, and $|1\rangle$ at the south pole.

The two extreme cases are when the qubit is completely in either the state $|0\rangle$ or $|1\rangle$, which can be obtained respectively, for $(\theta=0) \rightarrow |\psi\rangle = |0\rangle$ and $(\theta=\pi) \rightarrow |\psi\rangle = |1\rangle$, for any angle ϕ ; since the vectors $|0\rangle$ and $|1\rangle$ are perpendicular⁵ to the plane (x, y) , as shown in Figure 1.4:

⁵They are **orthogonal** in *Hilbert-space*. while, in the *Bloch sphere*, they are **perpendicular**, because the angles in the *Bloch sphere* are **double** those in the *Hilbert-space* (i.e., $\theta_B = \theta$; $\theta_H = \theta/2$).

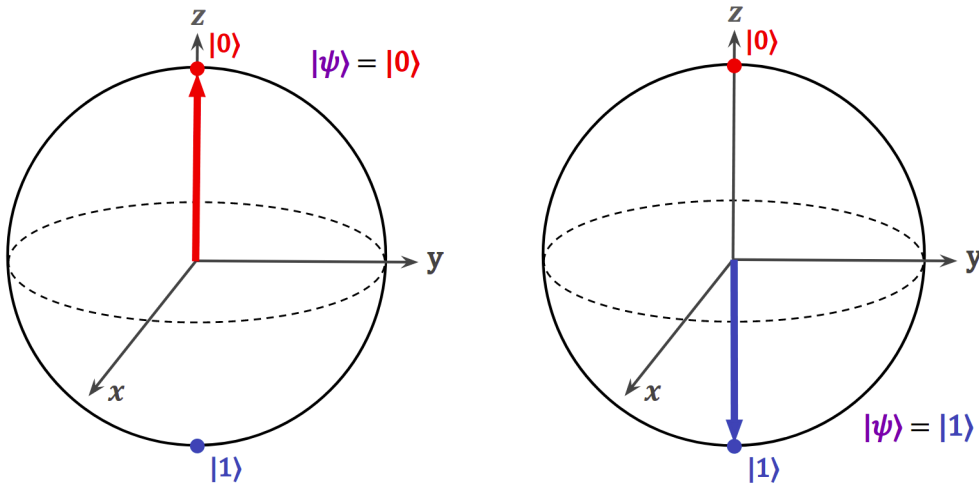


Figure 1.4: The two extremum states of the qubit. In the left: the Bloch vector of the state $|\psi\rangle = |0\rangle$. In the right: the Bloch vector of the state $|\psi\rangle = |1\rangle$, for $(\theta = 0, \pi; \forall\phi)$, respectively.

These two states are those in which a *classical bit* is confined. But, the qubit can be in any point of coordinates $(\theta, \phi, r = 1)$ on the *Bloch sphere*. This clearly shows the large difference between the qubit and its classical equivalent. We can take for example $(\pi/2, 0, 1)$ which gives from equation (1.30) the state $|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ (which is usually noted $|+\rangle$)⁶. This state is characterized by an equal probabilities $\mathbb{P}_{0,1} = \frac{1}{2}$, hence, the qubit is in perfect superpositions equilibrium of the two states $|0\rangle, |1\rangle$.

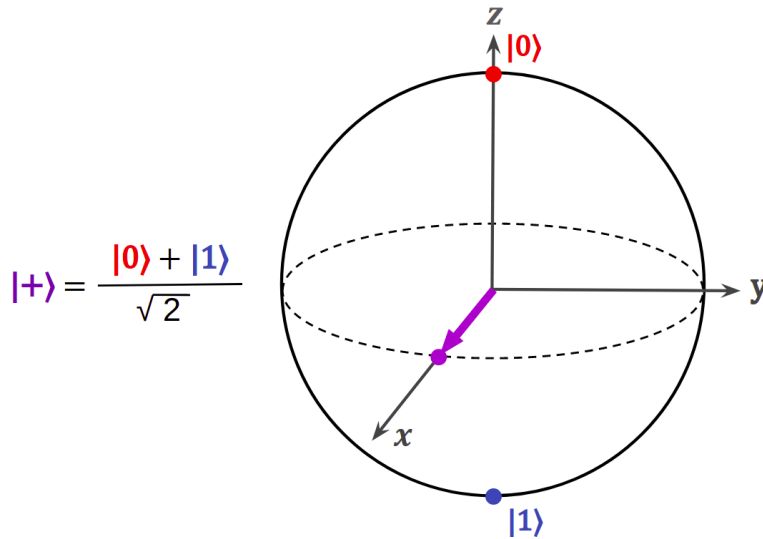


Figure 1.5: Representation of a superposition balance state $|+\rangle$ for $(\theta = \frac{\pi}{2}, \phi = 0, r = 1)$.

If we measure this specific qubit state or any other one, it will collapse to either $|0\rangle$ or $|1\rangle$. One may think: “if the qubit does eventually produce only a 0 or 1 when measured, then what can bring to us more than a classical bit?”. We say that the quantum properties of the qubit allow us to manipulate it behind the *classical limits*, in ways that can only be described through *quantum mechanics*; in other words, the modest 0 or 1 values, that we observe are

⁶To avoid any confusion, the state $|+\rangle$ represents an **equal superposition** state of a qubit, which has no relation to that of [subsection 1.1.2](#), which represents a **spin-up** state of a *spin-(1/2)* particle.

very *special*; by virtue of the *remarkable quantum* manipulations that allowed us to get them. Therefore, quantum computing is the art of exploiting these quantum properties to perform tasks that are impossible by the classical equivalent.

1.2.3 Multiple Qubits

A qubit by itself is not enough to represent the *useful information* needed. How much can a single LETTER tell us? Almost nothing. But if we have two or more letters, we can then form a meaningful WORD. By grouping words, we can syntax a SENTENCE, which is able to give us *significant information*. With those sentences you just read, we were able to provide *adequate information* to illustrate this fact. In classical information processing, one bit can give us just a 0 or 1, but by grouping bits together we can represent *useful information*. The same reasoning is going on quantum information processing, where the power of quantum computing appears when considering *multi-qubit* systems, in which the qubits interact with each other in a way that cannot be achieved by classical bits.

To understand how can **multi-qubit** systems make quantum computing more *powerful* than their *classical* equivalents, let's take a **2-qubit** system, where the two qubits are noted \mathbf{q}_1 and \mathbf{q}_2 , with their corresponding states $|\psi\rangle_1, |\psi\rangle_2$, and basis $\{|0\rangle_1, |1\rangle_1\} \in \mathcal{H}_1$ and $\{|0\rangle_2, |1\rangle_2\} \in \mathcal{H}_2$, respectively, where $\mathcal{H}_1, \mathcal{H}_2$ are *two-dimensional Hilbert spaces*. The two-qubit states are expressed:

$$|\psi\rangle_1 = a_1 |0\rangle_1 + b_1 |1\rangle_1 \quad ; \quad |\psi\rangle_2 = a_2 |0\rangle_2 + b_2 |1\rangle_2 \quad ; \quad a_1, a_2, b_1, b_2 \in \mathbb{C}$$

From the normalization condition we get $|a_1|^2 + |b_1|^2 = 1$; $|a_2|^2 + |b_2|^2 = 1$.

As seen in [subsection 1.1.2](#), composite systems are built by applying the “*tensor product*”, where the 2-qubit system is described by the state $|\psi\rangle_1 \otimes |\psi\rangle_2$, which lives in the *four-dimensional*⁷ complex vector space (*Hilbert-space*) $\mathcal{H}_1 \otimes \mathcal{H}_2$, with basis:

$$\{ |0\rangle_1 \otimes |0\rangle_2, |0\rangle_1 \otimes |1\rangle_2, |1\rangle_1 \otimes |0\rangle_2, |1\rangle_1 \otimes |1\rangle_2 \}$$

For brevity of notation we will omit the symbol “ \otimes ” between the kets $|0\rangle, |1\rangle$, and the subscripts 1, 2 on them:

$$\{ |00\rangle, |01\rangle, |10\rangle, |11\rangle \} \tag{1.31}$$

To explicit the form of these basis states, we compute their *tensor product*. Let's calculate the first one:

$$|00\rangle = |0\rangle \otimes |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 0 \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

⁷ Dimensions in a *tensor product* are **multiplied**: $\dim(\mathcal{H}_1 \otimes \mathcal{H}_2) = \dim(\mathcal{H}_1) \times \dim(\mathcal{H}_2)$

By analogous calculations, we obtain the explicit form of the four *orthonormal* basis states:

$$|00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} ; \quad |01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} ; \quad |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} ; \quad |11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad (1.32)$$

The general state of the 2-qubit system can be written as:

$$|\Upsilon\rangle = |\psi\rangle_1 \otimes |\psi\rangle_2 \quad (1.33)$$

$$= a_1 a_2 |00\rangle + a_1 b_2 |01\rangle + b_1 a_2 |10\rangle + b_1 b_2 |11\rangle, \quad (1.34)$$

Similar to the single-qubit state, the coefficients $(a_1 a_2, a_1 b_2, b_1 a_2, b_1 b_2)$ represent the *probability amplitudes*.

The state $|\Upsilon\rangle$ is of course normalized, which lead to the relation below:

$$\langle\Upsilon|\Upsilon\rangle = |a_1 a_2|^2 + |a_1 b_2|^2 + |b_1 a_2|^2 + |b_1 b_2|^2 = 1, \quad (1.35)$$

which can be seen as probability conservation.

When performing a measure, this state will collapse through projection to one of the states $|00\rangle, |01\rangle, |10\rangle, |11\rangle$, with respective probabilities

$$\mathbb{P}_{00} = |a_1 a_2|^2 ; \quad \mathbb{P}_{01} = |a_1 b_2|^2 ; \quad \mathbb{P}_{10} = |b_1 a_2|^2 ; \quad \mathbb{P}_{11} = |b_1 b_2|^2.$$

Unlike classical bits, the qubits have a *remarkable feature*, which is the **entanglement**; where the qubits are **tightly correlated** in a way that we need all of the constituents to have the *complete information* of the whole system.

We know from the [subsection 1.1.2](#) that a state is not *entangled* if we can separate it, like the example below:

$$|\Upsilon\rangle = \frac{1}{\sqrt{2}} |00\rangle + \frac{1}{\sqrt{2}} |10\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \otimes |0\rangle = |+\rangle \otimes |0\rangle,$$

where the state of the qubit \mathbf{q}_1 is $|\psi\rangle_1 = |+\rangle$ and \mathbf{q}_2 is $|\psi\rangle_2 = |0\rangle$.

Contrary to an *entangled* state, which is a *mixed* state and cannot be factorized into *pur* states:

$$|\Upsilon\rangle = \frac{1}{\sqrt{2}} |00\rangle + \frac{1}{\sqrt{2}} |11\rangle \quad (1.36)$$

If we measure the First qubit and find it to be in the state $|0\rangle$, then the second one will be in $|0\rangle$ with certainty.

$$|\Upsilon\rangle \xrightarrow{\text{measure}} |00\rangle$$

the same can be said for the state $|1\rangle$.

This state is one of four *entangled orthonormal* states known as the “**Bell states**”.

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \quad ; \quad |\Phi^-\rangle = \frac{1}{\sqrt{2}} (|00\rangle - |11\rangle) \quad (1.37)$$

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}} (|01\rangle + |10\rangle) \quad ; \quad |\Psi^-\rangle = \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle) \quad (1.38)$$

To visualize a 2-qubit system or more, we can think in the *Bloch sphere*, which was perfect to represent the state of 1-qubit system. But what for more than a single qubit?.

On the one hand, for a multi-qubit system, the entanglement applies to the qubits to be *correlated*; we cannot describe their states *separately*. On the other hand, the Bloch sphere gives a local view, where each qubit state is represented on its own Bloch sphere; therefore, we miss the important effect of *correlation* between the qubits when mapping them on *separate* Bloch spheres.

For this reason, we have to look for another model of visualization. The *Q-sphere* is the *primary visualization method*, that provides a *global viewpoint* of multi-qubit systems⁸.

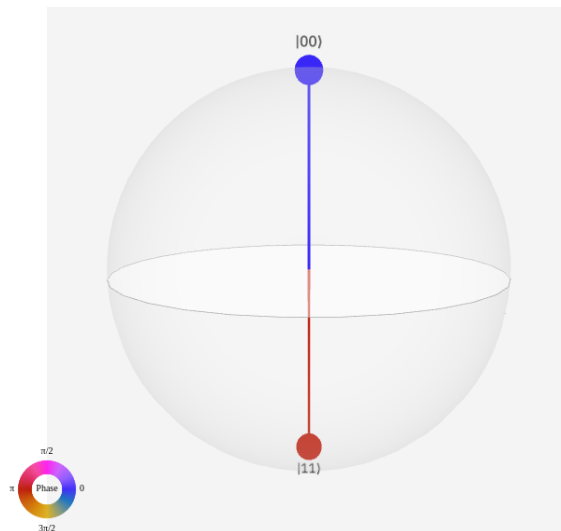


Figure 1.6: The Q-sphere : a visualization of the entangled 2-qubit state $|\Phi^-\rangle$.

- Each **computational basis state** is associated with a **point** on the surface of the sphere⁹, where the state with all **zeros** is at the *north pole* and that with all **ones** is at the *south pole*¹⁰.
- At each point there is a **node**, that represents the *probability amplitude*, and its **radius** is proportional to *the probability* \mathbb{P} of its basis states.
- The **color** of the *nodes* represents the *quantum phase* of each **computational basis state**.

⁸The Q-sphere is limited to a *small* number of qubits.

⁹The *spherical shape* has no importance, it's just a nice way to organize the *nodes*.

¹⁰The basis states with the same number of *zeros* or *ones* that lie on a shared *latitude* of the Q-sphere (e.g., $|011\rangle, |101\rangle, |110\rangle$).

The Q-sphere gives us *complete information* of the system state, at each given *manipulation* of the qubits.

We now move from the 2-qubit to the 3-qubit system in order to generalize these cases to an **n-qubit** system. Analogous to what we said above, the *Hilbert-space* associated with the 3-qubit system is written as $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$, with dimensionality $2 \times 2 \times 2 = 2^3 = 8$. The *computational basis* of this space is

$$\{ |000\rangle; |001\rangle; |010\rangle; |011\rangle; |100\rangle; |101\rangle; |110\rangle; |111\rangle \}, \quad (1.39)$$

each of these kets is an *eight-dimensional* column.

The general state of the 3-qubit system is expressed:

$$|\Upsilon\rangle = |\psi\rangle_1 \otimes |\psi\rangle_2 \otimes |\psi\rangle_3.$$

We can follow the same reasoning of separability to verify if it is entangled or not.

Now for the **generalization** to an **n-qubit** system. Let's have n qubit, where $n > 1$, with $n \in \mathbb{N}$, and their general state is constructed by the tensor product of the n single-qubit states; i.e.

$$|\psi\rangle_1 \otimes |\psi\rangle_2 \otimes \cdots \otimes |\psi\rangle_{n-1} \otimes |\psi\rangle_n.$$

This n -qubit state lives in the *Hilbert-space* $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_{n-1} \otimes \mathcal{H}_n$, which can briefly be written as $\mathcal{H} \otimes^n \mathcal{H} \equiv \mathcal{H}^{\otimes n}$. The *dimensionality* of the n -qubit space is $2 \times \dots \times 2 = 2^n$, which is the same as the n -qubit state. So the dimensionality of the state vector is increasing *exponentially* every time we add a qubit to the system [8]. The difficulty to simulate a *quantum computer* appears when considering a system of a large number of qubits. A modern laptop can easily simulate a general quantum state of 20 qubits, but simulating 100 qubits is too difficult for the largest **supercomputers**!

The **essence power** of *quantum computers* lies in the **superposition** feature of the qubits and which gives rise to the **entanglement** phenomenon, when considering multi-qubit systems.

1.2.4 Physical Implementation of Qubit

In quantum information theory, *the qubit is often described abstractly as a two-state (or two-level) quantum mechanical system*, in the beginning, we should not worry too much about the characteristics of these states or what their energies are. We assume that they are eigenstates of the Hamiltonian system with known eigenvalues. This approach allows us to concentrate on the fundamental properties of the system without getting lost in the details.

In 1994, the mathematician PETER SHOR presented two results [9] [10]: *the first, showing that a quantum computer would provide exponential speedup for number factorization. The second, showing that quantum error correction could overcome decoherence.* After that, the scientific community started to come up with ideas to build a quantum computer.

In this part, we introduce the notion of a **physical qubit** and distinguishes its relation with the **logical qubit** (described in subsection 1.2.2). We identify the essential criteria to realize quantum computing. Additional information on the technologies used to perform the physical implementation of qubits are provides in the Appendix C.

Physical and Logical Qubit

In quantum circuits and quantum algorithms, we use a **logical Qubit**. It specifies how a single Qubit should behave with quantum logical operations. We can use them indefinitely, they never lose state when they are not used, and we can apply as many gates to them as we wish.

A **physical Qubit** is a real Qubit produced by one of the implementation technologies and is physical hardware device. The Qubits used in quantum computers are constructed from microscopic degrees of freedom like the spin of electrons or nuclei, transition dipoles of atoms or ions in a vacuum. These degrees of freedom are naturally isolated from their environment and lead to slow decoherence [11].

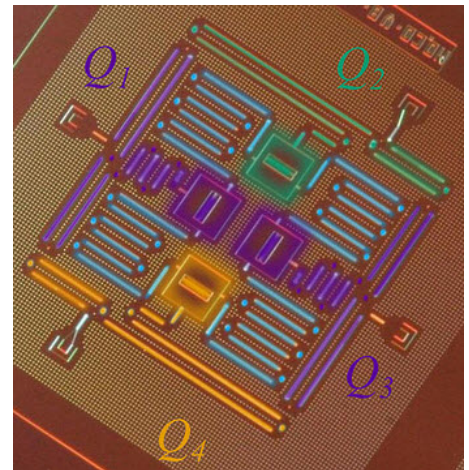


Figure 1.7: Layout of IBM's four superconducting Qubit device¹¹.

It is possible to create objects that act like logical qubits from physical, but it will require hundreds to thousands of physical qubits to make one logical Qubit. We use error-correcting schemes to have many physical qubits work together to create a virtual logical qubit. Hence, *A physical qubit is the hardware implementation of logical qubits* [8].

The quantum state of a physical qubit decays quickly and inevitably; for this, the main objective of researchers and engineers in quantum computing is to slow down this decay and correct the effects of the decoherence of qubits through the execution of circuits.

Criteria to Realize Quantum Computing:

In 2018, the theoretical physicist David DiVincenzo published a brief retrospective on creating and implementing quantum computers that respected some criteria. He laid out *five requirements for the implementation of quantum computation* [12].

The DiVincenzo Criteria that a quantum computer implementation must satisfy are as follows:

1. **A scalable physical system with well-characterized qubits.**

We should be able to add over time, qubits in our quantum system without any restriction. With a limited number of Qubits, we cannot make useful quantum calculations.

2. **The ability to initialize the state of the qubits to a simple fiducial state.**

It is common for algorithms, to start with a qubit in state $|0\rangle$ but sometimes we need to

¹¹<https://search.creativecommons.org/photos/30f43527-9f8d-4e77-859a-a85f0a0ada4c>

use $|1\rangle$ instead $|0\rangle$. For this, we must be able to initialize the qubit to a known initial state with a very high probability. This is called *high-fidelity state preparation*.

3. Long decoherence time

A quantum superposition state has a lifetime. This time is called the coherence time and depends on the environment of the qubit. Environmental disruptions cause a quantum superposition to dissipate and the final state will no longer be in a coherent superposition. If we have a long coherence time, but our gates take a long time to execute, that may be equivalent to a short coherence time but with fast gates. Long coherence, fast enough gates, and low error rates will be the key to success with *Noisy Intermediate-Scale Quantum computers* (NISQ).

4. A universal set of quantum gates

Without quantum logic gates, we cannot do quantum programming; that means, we cannot build quantum algorithms unless we have a sufficiently complete set of gates.

5. A qubit-specific measurement capability

When we measure a quantum system we force the state to take one of the basic orthogonal states. We must be able to force the qubit into one of the two orthogonal states $|0\rangle$ and $|1\rangle$. The error rate of this operation must be low enough to allow for useful quantum computation, if we get the wrong measurement answer fifty percent of the time, all the previous work on executing the circuit is lost. This is called *high fidelity readout*.

1.3 Quantum Gates and Quantum Circuits

Now that we understand what a qubit is, it's time to present the operators that manipulate its state called "*quantum logic gates*". Next, we introduce "*quantum circuits*" and see some examples. To understand this, we give a glimpse of "*classical gates*".

1.3.1 Classical Gates

In numerical systems, algebra describes the arithmetic operation on variables that can take any defined value. For Boolean algebra, it is the same thing except that the variables cannot take more than two values. Those variables are called logical variables. A logical variable can be presented in several ways, for example, 0 and 1 or TRUE and FALSE and it indicates the state of the system like voltage in a circuit. Boolean algebra has only three fundamental operators: NOT, AND, OR. We need to manipulate data using operators (like in mathematics); this operator is called "*gates*". Digital technologies are based on repeatedly applying the three operators (NOT, AND, OR) in various combinations.

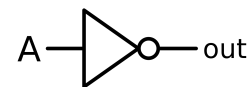
Logic Gates

A logic gate is an electronic circuit which is composed of transistors and resistors, and they are represented in schematic symbols. Each gate has one output that depends on its inputs; the output voltage depends on input voltage. A circuit or a digital system can be made using the three basic gates: AND, OR, NOT [13]. Other fundamental gates can be listed: NAND, NOR, XOR. These gates can be constructed using the three basic gates. To describe the output of a logical circuit, we need a truth table. We will see in this part the logic gates that are the most commonly used, their symbol, and their truth table¹².

1. NOT Gate

A NOT gate is a logic gate that inverts the digital input signal. If the input of a NOT gate is considered as A , then the output of the gate will be \bar{A} . When the input signal A is controlled by the NOT operation, the output O can be expressed as: $O = \bar{A} = A'$.

If the value of A is 1 then $\bar{A} = 0$ and in the opposite if the value of A is 0 then $\bar{A} = 1$.



A	Output
0	1
1	0

NOT gate symbol and truth table

2. AND Gate

An AND gate is a logic gate with two or more inputs and a single output, it can have any number of inputs, but the most common is 2 and 3. The AND gate operates on logical multiplication rules that mean the output of the gate, O , will be 0 unless all (both A and B) inputs are 1. The output can be expressed as $O = A.B$, which means: output O equals A AND B . The “.” sign stands for the AND operation.

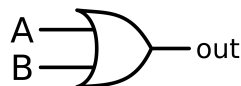


A	B	Output
0	0	0
0	1	0
1	0	0
1	1	1

AND gate symbol and truth table

3. OR Gate

An OR gate is a logic gate that performs logical OR operation. Just like an AND gate, an OR gate may have any input but only one output. When any of the OR gate inputs is 1, the output of the OR gate will be 1. However, the output of the OR gate is 0 only when all the input variables are 0. The output can be expressed as $O = A + B$, which means: output O equals A OR B . The “+” sign stands for the OR operation, and it is not an arithmetic addition.



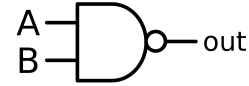
A	B	Output
0	0	0
0	1	1
1	0	1
1	1	1

OR gate symbol and truth table

¹²Truth tables list the output of a particular digital logic circuit for all the possible combinations of its inputs.

4. NAND Gate

The NAND gate is the inverse of the AND gate. Just like an AND gate, a NAND gate may have any number of inputs but only one output. The output of the gate, O , will be 1 when at least one input is 0 (A or B). The output can be expressed as $O = \overline{A \cdot B}$, which means: the output O equals $A \cdot B$ bar. The small circle at the output mean the inverted operation of output.

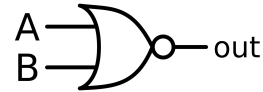


A	B	Output
0	0	1
0	1	1
1	0	1
1	1	0

NAND gate symbol and truth table

5. NOR Gate

A NOR gate is the inverse of the OR gate, and its circuit is produced by connecting an OR gate to a NOT gate. The NOR gate produces 1 only if all its inputs are 0, and 0 otherwise. The output can be expressed as $O = \overline{A + B}$, which means: the output O equals $A + B$ bar. This gate has a small circle on the output. This small circle represents the inversion operation.

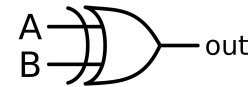


A	B	Output
0	0	1
0	1	0
1	0	0
1	1	0

NOR gate symbol and truth table

6. XOR Gate

A XOR (pronounced as Exclusive OR) is the logic gate that produces 1 in the output if and only if an odd number of inputs are 1. If both inputs are 1 or both inputs are 0, the output is 0. So a two-input XOR gate is quite different from the OR gate. XOR represents the inequality function, which means that the output is true if the inputs are not similar. The Boolean expression for the output from a two-input XOR gate is $O = A \cdot \overline{B} + \overline{A} \cdot B$



A	B	Output
0	0	0
0	1	1
1	0	1
1	1	0

XOR gate symbol and truth table

1.3.2 Quantum Gates

Now after we have seen classical gates we can have a better understanding of quantum gates. We know that classical logic gates manipulate the classical bit values 0 or 1. Likewise, quantum gates act on qubits and transform their states. For example by a rotation around the Bloch sphere. The main difference between classical and quantum gates is that quantum gates can manipulate multi-quantum states. Quantum gates are represented by unitary matrices (the inverse equals the conjugate transpose) and therefore, any quantum gate is reversible. When we apply a gate/operator \mathbb{G} on an arbitrary state $|\psi_{In}\rangle$, the input state will be transformed as follow: $|\psi_{Out}\rangle = \mathbb{G} |\psi_{In}\rangle$. We can reconstruct the input state from the output as $|\psi_{In}\rangle = \mathbb{G}^\dagger |\psi_{Out}\rangle$ [14].

In this part, we will see almost all the quantum gates provided by IBM quantum composer. Then, we will see some examples of quantum circuits that are very known in QIS.

Single-Qubit Gates

Identity Gate



The first gate is the IDENTITY GATE, it is represented by the identity matrix and defined for a single qubit as follow:

$$\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad ; \quad \mathbb{I}.\mathbb{I} = \mathbb{I} \quad (1.40)$$

When we apply this gate to the computational basis states $|0\rangle$ and $|1\rangle$, it leaves the states unchanged.

$$\mathbb{I}|0\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |0\rangle \quad (1.41)$$

$$\mathbb{I}|1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |1\rangle \quad (1.42)$$

This gate is useful for two reasons, one to prove that a gate is its own inverse. Second, it is useful to have a do-nothing operation, especially for the error models. This can be used as a simple version of decoherence of the qubit.

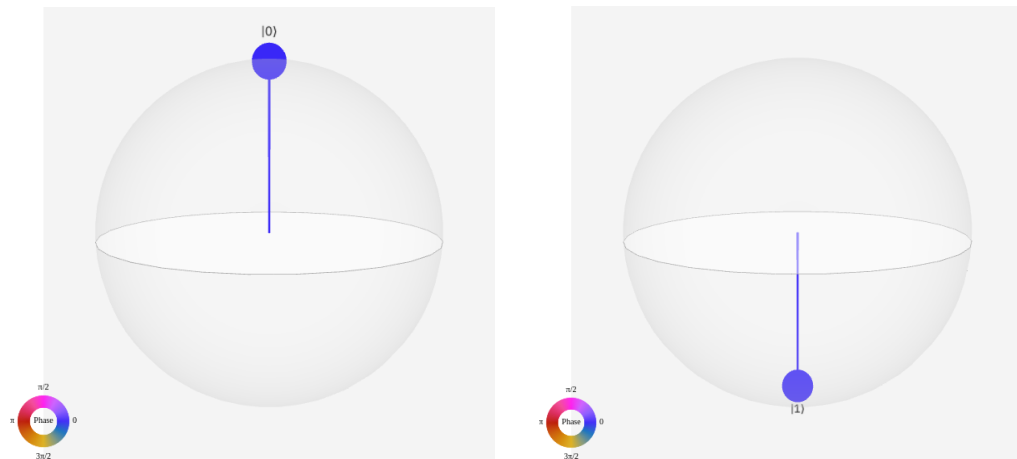


Figure 1.8: A visualization of the resulting states from the \mathbb{I} action on the states $|0\rangle$ and $|1\rangle$, respectively.

The X Gate



The NOT GATE is the quantum equivalent of the classical NOT gate, it is also known as the PAULI X gate. Its matrix representation is given by:

$$\mathbb{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad ; \quad \mathbb{X}.\mathbb{X} = \mathbb{I} \quad (1.43)$$

It acts in a single qubit by switching the **amplitudes** of the $|0\rangle$ and $|1\rangle$ basis.

$$\mathbb{X}|0\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |1\rangle \quad (1.44)$$

$$\mathbb{X}|1\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |0\rangle \quad (1.45)$$

The \mathbb{X} gate rotates the qubit state by π radians around the x -axis. It is sometimes called *bit-flip gate* as it maps $|0\rangle$ to $|1\rangle$ and $|1\rangle$ to $|0\rangle$.

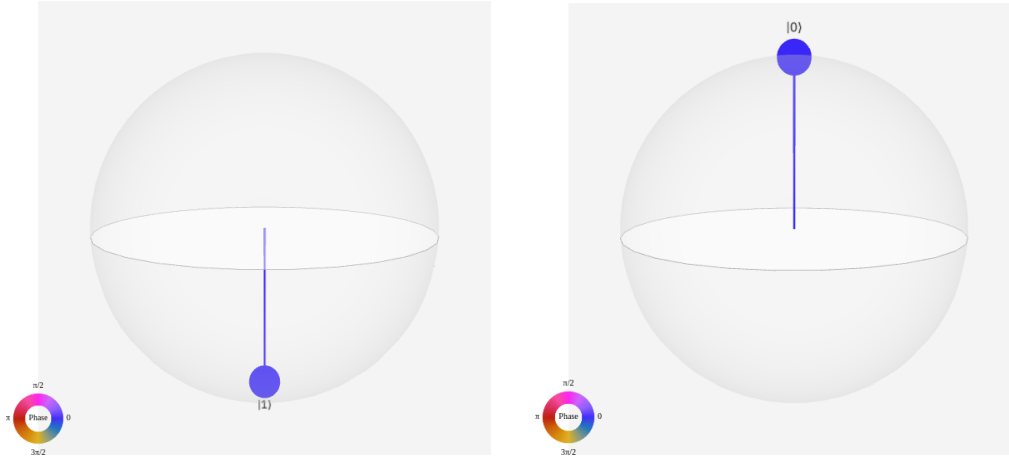


Figure 1.9: A visualization of the resulting states from the \mathbb{X} action on the states $|0\rangle$ and $|1\rangle$, respectively.

The Y Gate



The Pauli Y gate rotates the qubit around the y -axis by π radians. It is represented by the following matrix:

$$\mathbb{Y} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} ; \quad \mathbb{Y} \cdot \mathbb{Y} = \mathbb{I} \quad (1.46)$$

When we apply this gate to the computational basis states $|0\rangle$ and $|1\rangle$, we get:

$$\mathbb{Y}|0\rangle = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ +i \end{pmatrix} = +i|1\rangle \quad (1.47)$$

$$\mathbb{Y}|1\rangle = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -i \\ 0 \end{pmatrix} = -i|0\rangle \quad (1.48)$$

This gate maps $|0\rangle$ to $i|1\rangle$ and $|1\rangle$ to $-i|0\rangle$.

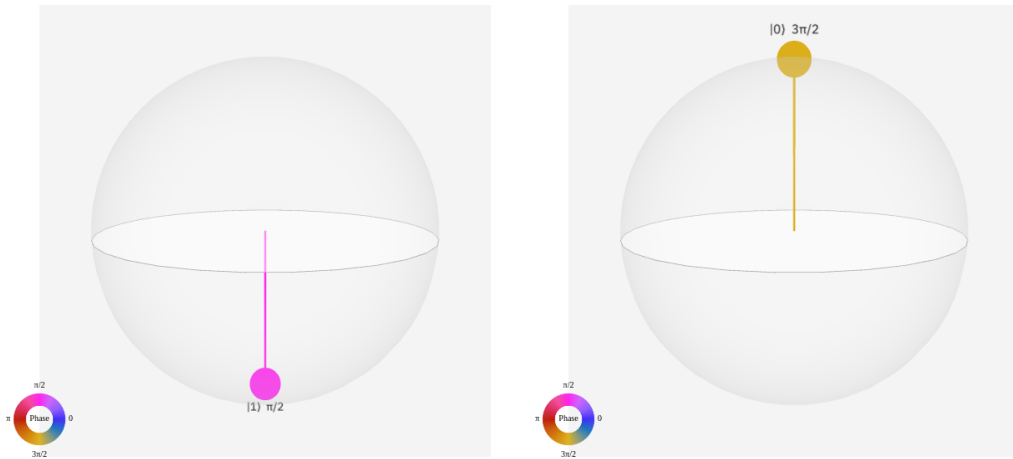


Figure 1.10: A visualization of the resulting states from the \mathbb{Y} action on the states $|0\rangle$ and $|1\rangle$, respectively.

The Z Gate



The PAULI Z GATE is represented by the *diagonal* matrix:

$$\mathbb{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad ; \quad \mathbb{Z} \cdot \mathbb{Z} = \mathbb{I} \quad (1.49)$$

Let's observe its action on the *computational* basis states $|0\rangle$ and $|1\rangle$.

$$\mathbb{Z} |0\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} +1 \\ 0 \end{pmatrix} = +|0\rangle$$

$$\mathbb{Z} |1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -|1\rangle$$

So, this gate acts as an \mathbb{I} on $|0\rangle$, and applies a phase of π on $|1\rangle$. Moreover, its action can be seen as a rotation of the qubit state by π around the *z-axis* on the *Bloch sphere*. Therefore, its action on the states $|+\rangle$ and $|-\rangle$ (recalling that $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$)

$$\mathbb{Z} |+\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} +\frac{1}{\sqrt{2}} \\ +\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} +\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = |-\rangle$$

$$\mathbb{Z} |-\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} +\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} +\frac{1}{\sqrt{2}} \\ +\frac{1}{\sqrt{2}} \end{pmatrix} = |+\rangle$$

The Z gate swaps these two states $|+\rangle \rightleftharpoons |-\rangle$. Since it reverses the sign of the 2nd amplitude, it is also called *sign-flip gate*.

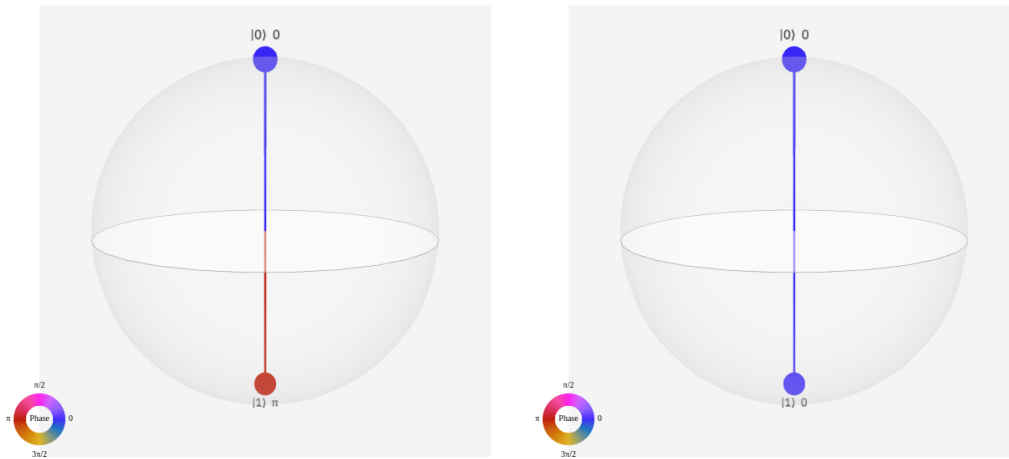


Figure 1.11: A visualization of the resulting states from the \mathbb{Z} action on the states $|+\rangle$ and $|-\rangle$, respectively.

Hadamard Gate



The HADAMARD GATE is represented by the matrix below:

$$\mathbb{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad ; \quad \mathbb{H} \cdot \mathbb{H} = \mathbb{I} \quad (1.50)$$

Applied to the two computational basis states $|0\rangle, |1\rangle$ gives:

$$\mathbb{H}|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = |+\rangle$$

$$\mathbb{H}|1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = |-\rangle$$

Therefore, the HADAMARD GATE puts the qubit in an **equal superposition**. The transformations $|0\rangle \rightarrow |+\rangle$ and $|1\rangle \rightarrow |-\rangle$ can be seen as a rotation around the *Bloch vector* of coordinates $(1, 0, 1)$ by an angle π . Without this gate, the qubit would still behave like a classical bit. Hence, implementing the HADAMARD GATE in any given classical circuit at any position will transform it into a quantum circuit.

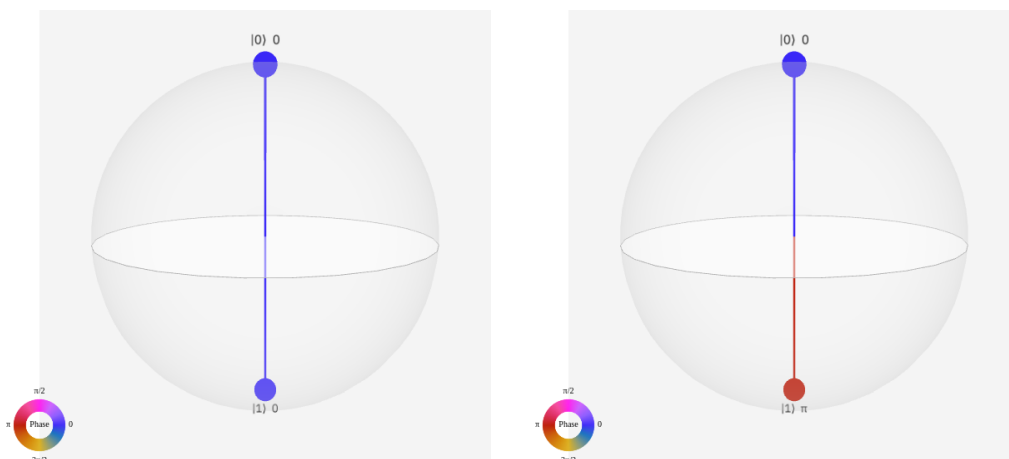


Figure 1.12: A visualization of the resulting states from the \mathbb{H} action on the states $|0\rangle$ and $|1\rangle$, respectively.

The S Gate



The S GATE is represented by the following *complex* matrix:

$$\mathbb{S} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \tag{1.51}$$

When applied to the basis states $|0\rangle$ and $|1\rangle$, one obtains

$$\mathbb{S} |0\rangle = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 |0\rangle$$

$$\mathbb{S} |1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ i \end{pmatrix} = i |1\rangle$$

Thus, this gate induces a phase of $(\frac{\pi}{2})$ just on $|1\rangle$, which is equivalent to a $(\frac{\pi}{2})$ rotation around the z -axis. If it is applied on the states $|+\rangle$ and $|-\rangle$, we obtain

$$\mathbb{S} |+\rangle = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{pmatrix} = |+i\rangle$$

$$\mathbb{S} |-\rangle = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-i}{\sqrt{2}} \end{pmatrix} = |-i\rangle$$

Let's calculate its square

$$\mathbb{S}^2 = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \mathbb{Z}$$

Therefore, the S GATE is not *hermitian*, and its *conjugate-transpose* defines the S-DAGGER or SDG GATE, given by the matrix:

$$\mathbb{S}^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{3\pi}{2}} \end{pmatrix} \quad ; \quad \mathbb{S}^\dagger = \mathbb{S}^3 \tag{1.52}$$

That applies a $(\frac{3\pi}{2})$ rotation around the z -axis.

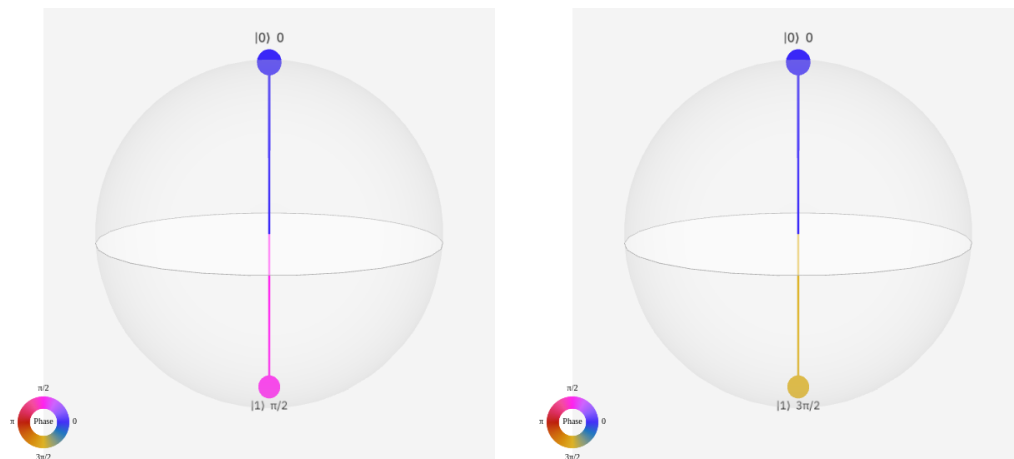


Figure 1.13: A visualization of the resulting states from the \mathbb{S} action on the states $|+\rangle$ and $|-\rangle$, respectively. (For the \mathbb{S}^\dagger gate, the figures will be inverted).

The T Gate



The only difference between the S GATE and the T GATE is the induced phase, where this gate applies a phase of $(\frac{\pi}{4})$, hence, we can write

$$\mathbb{T} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} \quad (1.53)$$

Thus, for its action on $|0\rangle$ and $|1\rangle$, the outputs are

$$\mathbb{T}|0\rangle = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |0\rangle$$

$$\mathbb{T}|1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ i \end{pmatrix} = e^{i\frac{\pi}{4}}|1\rangle$$

which is a rotation of $(\frac{\pi}{4})$ around the z -axis. Applied on the states $|+\rangle$ and $|-\rangle$, gives

$$\mathbb{T}|\pm\rangle = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} \begin{pmatrix} +\frac{1}{\sqrt{2}} \\ \pm\frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} (|0\rangle \pm e^{i\frac{\pi}{4}}|1\rangle)$$

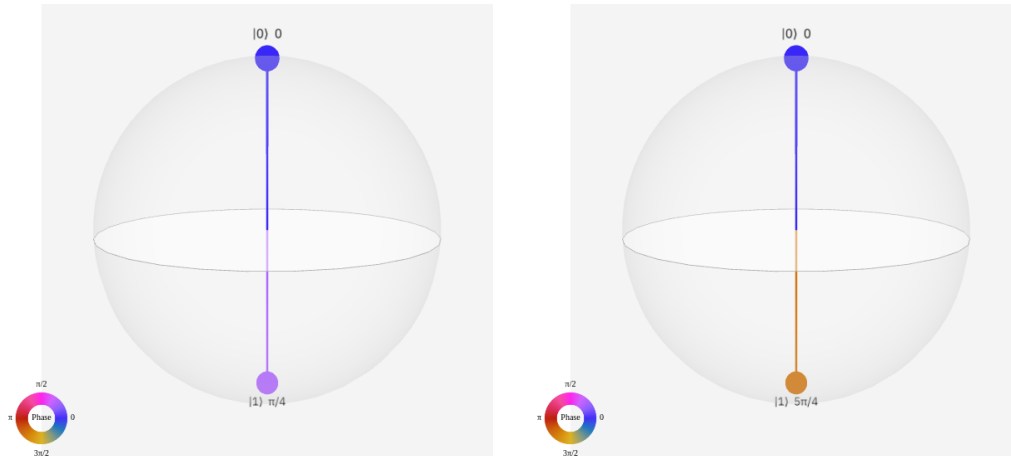


Figure 1.14: A visualization of the resulting states from the \mathbb{T} action on the states $|+\rangle$ and $|-\rangle$, respectively.

When calculating its square, we get

$$\mathbb{T}^2 = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{2}} \end{pmatrix} = \mathbb{S}$$

This means that the \mathbb{T} GATE is not *hermitian*, and its *conjugate-transpose* defines the T-DAGGER or TDG GATE, given by the *adjoint* of \mathbb{T}

$$\mathbb{T}^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & e^{-i\frac{\pi}{4}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{7\pi}{4}} \end{pmatrix} \quad ; \quad \mathbb{T}^\dagger = \mathbb{T}^7 \quad (1.54)$$

Which rotates the qubit state by $(\frac{7\pi}{4})$ around the z -axis on the *Bloch sphere*.

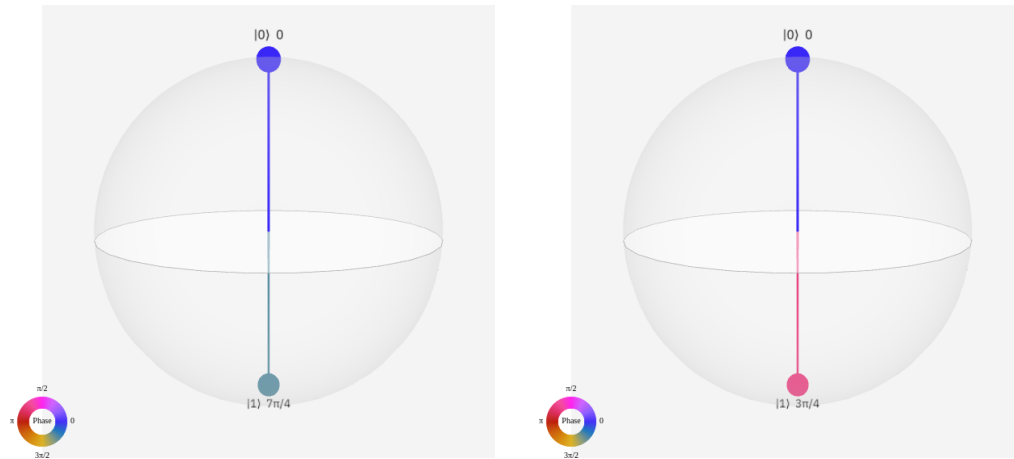


Figure 1.15: A visualization of the resulting states from the T^\dagger action on the states $|+\rangle$ and $|-\rangle$, respectively.

The P Gate



The generalization for a gate that applies any phase (θ), can be written as:

$$P_\theta = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix} \quad ; \quad 0 \leq \theta \leq 2\pi \tag{1.55}$$

This gate is known as the PHASE GATE, and which leads for specific values of θ to

$$P_{0,2\pi} = I \quad ; \quad P_\pi = Z \quad ; \quad P_{\frac{\pi}{2}} = S \quad ; \quad P_{\frac{3\pi}{2}} = S^\dagger \quad ; \quad P_{\frac{\pi}{4}} = T \quad ; \quad P_{\frac{7\pi}{4}} = T^\dagger$$

The PHASE GATE rotates the qubit state by θ around the z -axis on the Bloch sphere. Like any quantum gate, the matrix P is of course *unitary*

$$P \cdot P^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{-i\theta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

Square-Root NOT Gate



As the name suggests, it is the *square root* of the NOT gate, therefore its square gives the NOT gate

$$\sqrt{X} \cdot \sqrt{X} = X$$

It is represented by the matrix below :

$$\sqrt{\text{NOT}} \equiv \sqrt{X} = \frac{1}{2} \begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix} \tag{1.56}$$

$$\sqrt{X} |0\rangle = \frac{1}{\sqrt{2}} (e^{i\frac{\pi}{4}} |0\rangle + e^{i\frac{7\pi}{4}} |1\rangle)$$

$$\sqrt{X} |1\rangle = \frac{1}{\sqrt{2}} (e^{i\frac{7\pi}{4}} |0\rangle + e^{i\frac{\pi}{4}} |1\rangle)$$

It creates an **equal superposition** state, but the only difference with the HADAMARD GATE, is

that it applies a different *relative phase*. This gate is *natively* implemented in some HARDWARE.

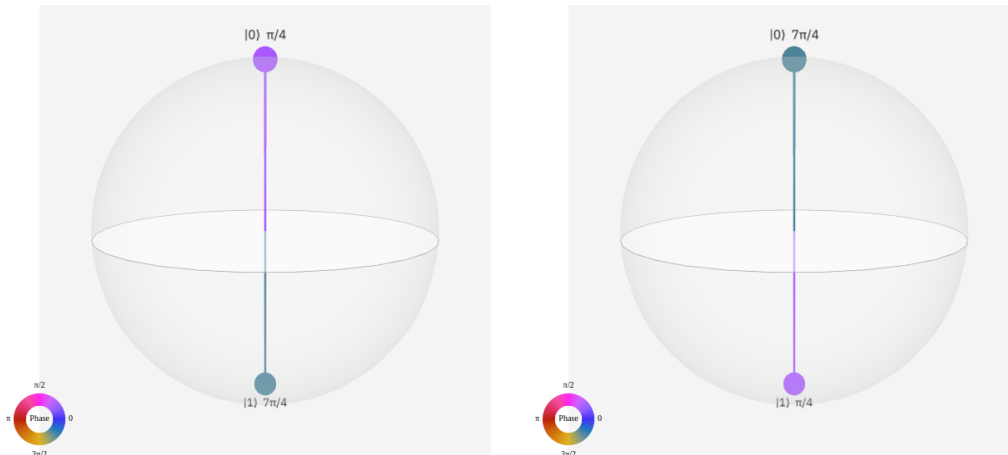


Figure 1.16: A visualization of the resulting states from the $\sqrt{\mathbb{X}}$ action on the states $|0\rangle$ and $|1\rangle$, respectively.

Square-Root NOT-Dagger Gate



This gate is the *inverse* of $\sqrt{\mathbb{X}}$ gate, so $\sqrt{\mathbb{X}^\dagger} \cdot \sqrt{\mathbb{X}} = \sqrt{\mathbb{X}} \cdot \sqrt{\mathbb{X}^\dagger} = \mathbb{I}$ and its square also gives the NOT gate $\sqrt{\mathbb{X}^\dagger} \cdot \sqrt{\mathbb{X}^\dagger} = \mathbb{X}$. Its matrix representation is given below:

$$\sqrt{\text{NOT}}^\dagger \equiv \sqrt{\mathbb{X}^\dagger} = \frac{1}{2} \begin{pmatrix} 1 - i & 1 + i \\ 1 + i & 1 - i \end{pmatrix} \quad (1.57)$$

$$\sqrt{\mathbb{X}^\dagger} |0\rangle = \frac{1}{\sqrt{2}} (e^{i\frac{7\pi}{4}} |0\rangle + e^{i\frac{\pi}{4}} |1\rangle)$$

$$\sqrt{\mathbb{X}^\dagger} |1\rangle = \frac{1}{\sqrt{2}} (e^{i\frac{\pi}{4}} |0\rangle + e^{i\frac{7\pi}{4}} |1\rangle)$$

We see that it also creates an **equal superposition** state, but with *relative phases* reversed with respect to $\sqrt{\mathbb{X}}$. Therefore, the Q-spheres of $\sqrt{\mathbb{X}^\dagger}$ are also reversed.

Two-Qubit Gates

SWAP Gate



The SWAP GATE is a two-qubit gate, so it is represented by a 4×4 *dimensional unitary matrix*, given by

$$\text{SWAP} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad ; \quad (\text{SWAP})^2 = \mathbb{I} \quad (1.58)$$

Let's observe its action on a two-qubit state

$$|\psi\rangle = |\phi\rangle_0 \otimes |\phi\rangle_1 = c_{00} |00\rangle + c_{01} |01\rangle + c_{10} |10\rangle + c_{11} |11\rangle \quad ; \quad c_{kl} = a_k b_l \quad (k, l = 0, 1)$$

where $|\phi\rangle_0$ and $|\phi\rangle_1$ are the states of the qubits q_0 and q_1 , respectively.

$$\begin{aligned} (\text{SWAP})|\psi\rangle &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_0 b_0 \\ a_0 b_1 \\ a_1 b_0 \\ a_1 b_1 \end{pmatrix} = \begin{pmatrix} a_0 b_0 \\ a_1 b_0 \\ a_0 b_1 \\ a_1 b_1 \end{pmatrix} \\ &= c_{00}|00\rangle + c_{10}|01\rangle + c_{01}|10\rangle + c_{11}|11\rangle = |\phi\rangle_1 \otimes |\phi\rangle_0 \end{aligned}$$

So this gate *switches* the states of two qubits

$$(|\phi\rangle_0 \otimes |\phi\rangle_1) \xrightarrow{\text{SWAP}} (|\phi\rangle_1 \otimes |\phi\rangle_0)$$

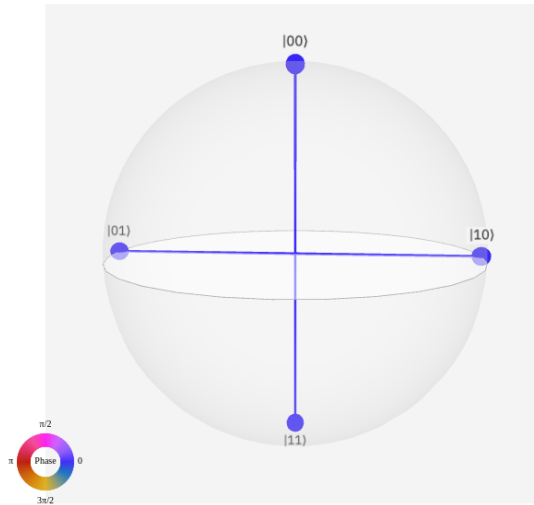


Figure 1.17: A visualization of all 4 possible resulting states from the SWAP action on the 2-qubit state $|\alpha\beta\rangle$; $\alpha, \beta = 0, 1$.

Controlled-NOT Gate



The CNOT or CX gate is a two-qubit gate, acting in the 2^2 dimensional *Hilbert-space*. The CNOT GATE is represented by the matrix of dimension 4×4 below:

$$\text{CNOT} \equiv \text{CX} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} ; \quad (\text{CX})^2 = \mathbb{I} \tag{1.59}$$

When the 2-qubit state $|\psi\rangle$ pass through this gate it becomes:

$$\text{CX}|\psi\rangle = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_0 b_0 \\ a_0 b_1 \\ a_1 b_0 \\ a_1 b_1 \end{pmatrix} = \begin{pmatrix} a_0 b_0 \\ a_0 b_1 \\ a_1 b_1 \\ a_1 b_0 \end{pmatrix}$$

We see that this gate swaps the two *amplitudes* of q_1 in the states $|10\rangle$ and $|11\rangle$ which is a NOT operation on q_1 . Moreover, it's applied only when q_0 is in $|1\rangle$. Therefore the CNOT gate is a **conditional** gate, where q_0 and q_1 are called **control** and **target**, respectively. If the control

qubit is in the state $|1\rangle$ the CNOT gate executes a NOT on the target qubit, otherwise it acts as an \mathbb{I} on the target.

If we turn the roles (*control* \leftrightarrow *target*), then this gate will become what we call a REVERSE CNOT gate and its matrix is given by :

$$\text{RCNOT} \equiv \text{RCX} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

The CNOT GATE is the gateway that **entangles** 2-qubits, only when the control qubit is in a **superposition**.

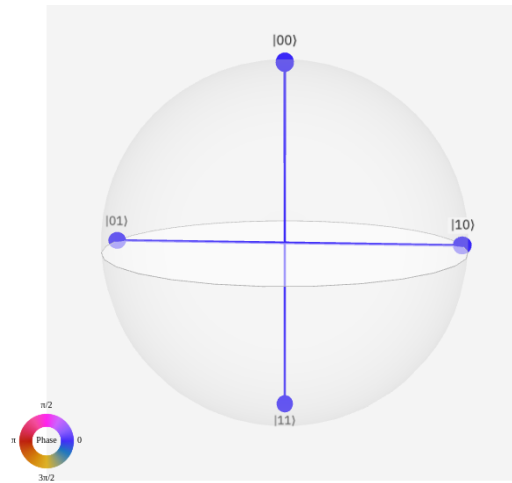


Figure 1.18: A visualization of all 4 possible resulting states from the CX action on the 2-qubit state $|\alpha\beta\rangle$; $\alpha, \beta = 0, 1$.

Three-Qubit Gates

Controlled-Controlled-NOT Gate



The CCNOT or CCX is a **conditional** three-qubit gate, with two *control* qubits and one *target*. It performs a NOT to the target if and only if the two control qubits are in $|1\rangle$, otherwise it acts as an \mathbb{I} on the target. It is represented by the matrix of dimension 8×8 below:

$$\text{CCNOT} \equiv \text{CCX} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}; \quad (\text{CCX})^2 = \mathbb{I} \quad (1.60)$$

The CCNOT is a *reversible* version of classical computing's AND and NAND gates. This is a

universal gate for *classical computation*, meaning that any program can be built from many instances of this gate.

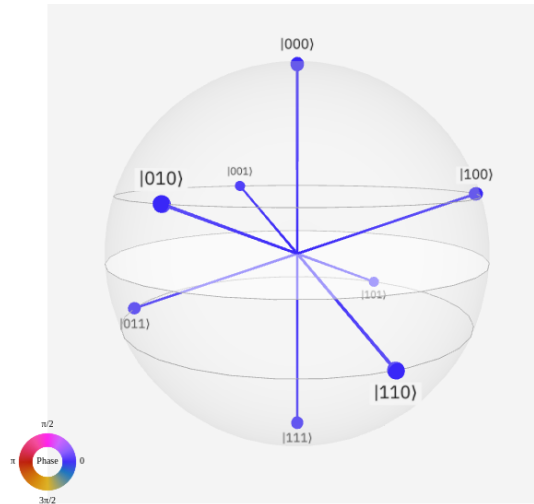


Figure 1.19: A visualization of all 8 possible resulting states from the CCX action on the 3-qubit state $|\alpha\beta\gamma\rangle$; $\alpha, \beta, \gamma = 0, 1$.

1.3.3 Quantum Circuit Diagrams

A quantum circuit is a sequence of gates applied to one qubit or more. The collection of qubits in a circuit is called a *quantum register*. So, to perform calculations in quantum computers, we need to manipulate qubits (inside a quantum register) by applying a series of gates. All qubits in the quantum register are initialized in the state $|0\rangle$ and by convention, we number the qubits in the register as $q_0, q_1, q_2, \dots, q_n$. To visualize quantum circuits, we need a *quantum circuit diagram*. In what follows, we will use the interactive interface of the IBM QUANTUM COMPOSER¹³ to represent these quantum circuits.

The following figure show an “empty” quantum circuit; we haven’t added any gates yet.

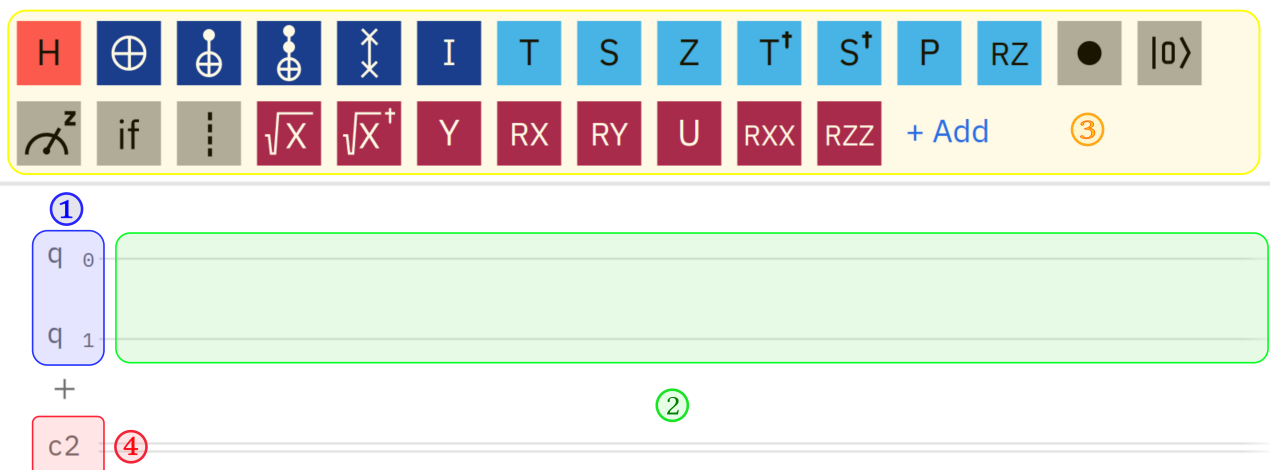


Figure 1.20: The interactive interface of IBM quantum composer.

¹³<https://quantum-computing.ibm.com/composer/files/new>

Let us describe each part of this empty circuit.

1. First, on the left-hand side, we see two qubits, the numbering starts from zero: q_0, q_1 . Each of these qubits is initialized to be in the $|0\rangle$ state.
2. In the right-hand side, we have two wires. In these wires we can place gates and measurements.
3. At the top, we see some of the quantum gates that we have already discussed.
4. On the bottom we see a C_2 , indicating that it is a classical register, a C_4 will indicate that the register contains four bits all initialized to zero, and a wire with a line through it will indicate that the register contains multiple bits. When we do the measurement a line will be drawn down to the classical register, indicating that the measurement is stored there.

Now let us see our quantum circuit with some gates on it.

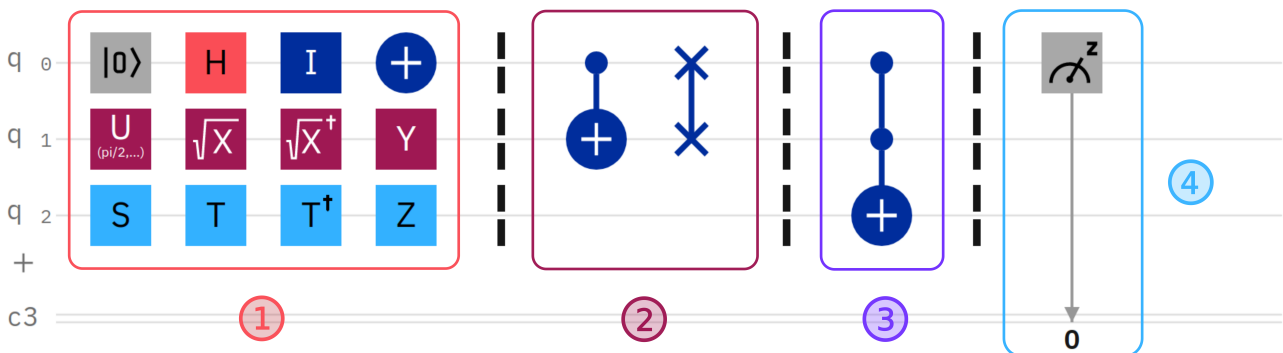


Figure 1.21: Types of quantum gates.

We have divided the circuit into four blocks.

- The first block includes single-qubit gates. These gates are applied from the left to the right, we apply first the closest gate to the qubit. On the other hand from the algebra perspective, we apply gates from the right to left. For example, the first line in the first block reads as $\text{XIIIH}|0\rangle$ which means we apply the HADAMARD first then the IDENTITY gate, and finally the NOT gate. The first operation from the left $|0\rangle$ returns a qubit to state $|0\rangle$, independently of its state before the operation was applied. It is not a reversible operation, so it is not a gate. The NOT gate is drawn graphically as (\oplus).
- The second block contains, two-qubit gates that span **two wires**. From the right to the left the CNOT and SWAP gate. The CNOT gate graphic contains the sign (\bullet) denoting the **control qubit** and the sign (\oplus) denoting the **target qubit**. If we draw it upside down, we will get the REVERSE CNOT.

- The third block contains, the CCNOT gate it operates on three qubits. If the first two qubits are $|1\rangle$ then it flips the third, otherwise, it does nothing. The only difference between CNOT and CCNOT gate is that there are two **control qubits** (••).
- Finally, the measurement operation returns a $|0\rangle$ or $|1\rangle$. The MEASUREMENT operation must always be at the **end** of any quantum circuit.

In quantum computing, all operations between *initialization* and *measurement* are **reversible**. The following table shows the gates and it inverses.

Gate	I	X	Y	Z	H	S	S [†]	T	T [†]
Reverse	I	X	Y	Z	H	S [†]	S	T [†]	T

Examples of Quantum Circuits

To explain how quantum circuits work, let's take two examples:

- The first example illustrate the SUCCESSION of operations:

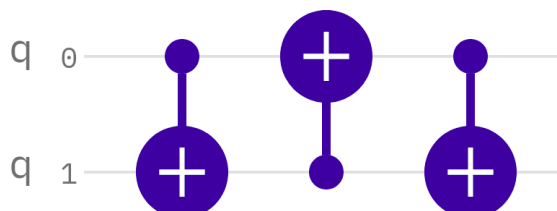


A qubit is initialized at the state $|\psi\rangle = |0\rangle$. Then each gate will act on its state and transform it according to its functionality. Mathematically we have:

$$\begin{aligned}
 |0\rangle \rightarrow ZTSH|0\rangle &= ZTSH|+\rangle = ZTS|+\rangle = ZT|+i\rangle \\
 &= \frac{1}{\sqrt{2}} Z(|0\rangle + ie^{i\frac{\pi}{4}}|1\rangle) = \frac{1}{\sqrt{2}} (|0\rangle + e^{i\frac{7\pi}{4}}|1\rangle)
 \end{aligned}$$

- The second example illustrate the EQUIVALENCE:

This circuit swaps two qubits. That means if we started in the $|01\rangle$ state, we would end up in the $|10\rangle$ state and vice versa. Inputs of $|00\rangle$ or $|11\rangle$ would be unchanged as swapping would have no effect. Note that for the first and third CNOTs, the control qubit is the first qubit and the target qubit is the second, while for the second CNOT, the control qubit is the second and the target qubit is the first.



Let's trace it out step by step:

Input	After first CNOT	After second CNOT	After the third CNOT (output)
00	00	00	00
01	01	11	10
10	11	01	01
11	10	10	11

If we calculate the matrix product, we will find certainly the SWAP matrix.

$$\text{CX.RCX.CX} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \text{SWAP}$$

Therefore

$$\text{SWAP} = \text{CX.RCX.CX}$$

Bell States Circuits

The central object of this part will be to construct **quantum circuits** that generate the **four Bell states**: $|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle, |\Psi^-\rangle$, which describe an entangled pair of qubits. We have seen previously simple manipulations on a single or two qubits system. But creating **entanglement** between two qubits, seems difficult at first glance, and one may think that we will use a varied assortment of **quantum gates**. But in fact, it is *extraordinary simple*.

Firstly, let's have two qubits, noted \mathbf{q}_0 and \mathbf{q}_1 , which are initialized in the state $|0\rangle$, so we have \mathbf{q}_0 in the state $|\psi\rangle_0 = |0\rangle_0$ and \mathbf{q}_1 in $|\psi\rangle_1 = |0\rangle_1$. The initial 2-qubit state is of course *separable* $|\psi\rangle_0 \otimes |\psi\rangle_1 = |0\rangle_0 \otimes |0\rangle_1 = |00\rangle$, and we want to construct the specific Bell state $|\Phi^+\rangle$, written

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \quad (1.61)$$

which is *inseparable*. But we can think of applying **operators/gates** to *separate* it into two distinct qubit states, and vis versa; by acting gates on the *initial separable* state, it will be transformed into an *inseparable* (entangled) state, i.e.

$$|\Phi^+\rangle = \hat{\mathbb{G}} |00\rangle \quad \rightarrow \quad |00\rangle = \hat{\mathbb{G}}^{-1} |\Phi^+\rangle \quad (1.62)$$

where $\hat{\mathbb{G}}$ is a *combination* of quantum gates, and $\hat{\mathbb{G}}^{-1}$ its *inverse*.

We saw previously a gate that *links* two qubits, through a *condition*, it is the CNOT GATE (noted $\text{CNOT} \equiv \text{CX}$), so it is a good candidate for the *linking* and the *separation*.

For the 1st example, we will try the *inverse* way; going from the *inseparable* state $|\Phi^+\rangle$ to the *initial separate* state $|00\rangle$, thus we apply the inverse of $\mathbb{C}\mathbb{X}$. But we know that it is its own *inverse* $\mathbb{C}\mathbb{X} = (\mathbb{C}\mathbb{X})^{-1}$.

$$(\mathbb{C}\mathbb{X})^{-1} |\Phi^+\rangle \equiv \mathbb{C}\mathbb{X} |\Phi^+\rangle = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} (|00\rangle + |10\rangle), \quad (1.63)$$

which is clearly separable

$$\mathbb{C}\mathbb{X} |\Phi^+\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \otimes |0\rangle = |+\rangle \otimes |0\rangle. \quad (1.64)$$

We can rewrite it in terms of the two *initial separate* states:

$$\begin{aligned} |\Phi^+\rangle &= \mathbb{C}\mathbb{X} (|+\rangle \otimes |0\rangle) \\ &= \mathbb{C}\mathbb{X} (\mathbb{H}|0\rangle \otimes \mathbb{I}|0\rangle) \\ &= \mathbb{C}\mathbb{X} (\mathbb{H} \otimes \mathbb{I}) (|0\rangle \otimes |0\rangle), \end{aligned}$$

where $\mathbb{H}|0\rangle = |+\rangle$.

So by acting the HADAMARD GATE \mathbb{H} , only on the \mathbf{q}_0 state $|0\rangle_0$, it becomes $|+\rangle_0$. Then by acting $\mathbb{C}\mathbb{X}$ on the new state $|+\rangle \otimes |0\rangle$, we obtain the *entangled* state $|\Phi^+\rangle$.

We can now apply this reasoning into a quantum circuit:

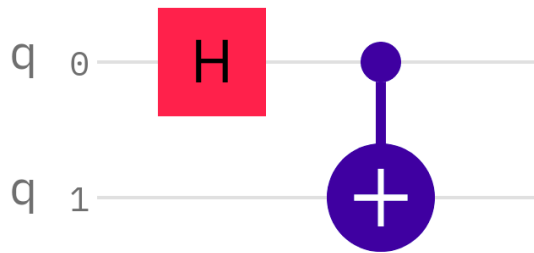


Figure 1.22: The quantum circuit that generates the Bell state $|\Phi^+\rangle$.

For the *interpretation*, we need some *philosophy*. We have a qubit \mathbf{q}_0 in the state $|0\rangle_0$ subjected to the action of the HADAMARD GATE, which will make it in a *superposition balance*. The second secret of the entanglement process lies in the CNOT GATE, which leaves \mathbf{q}_1 in $|0\rangle_1$ if \mathbf{q}_0 is in $|0\rangle_0$ and turns \mathbf{q}_1 to $|1\rangle_1$ if \mathbf{q}_0 is $|1\rangle_0$. But we know that \mathbf{q}_0 is **both** in $|0\rangle_0$ and $|1\rangle_0$; therefore, the CNOT GATE **simultaneously** turn and don't turn the \mathbf{q}_1 state!, which makes \mathbf{q}_0 also in a *superposition balance*; that is to say, the CNOT GATE transmits the *power of superposition* from \mathbf{q}_0 to \mathbf{q}_1 . Since \mathbf{q}_1 *depends directly* on \mathbf{q}_0 , they are therefore **linked**, which makes them **entangled**.

The steps followed to obtain $|\Phi^+\rangle$, begins by acting \mathbb{H} on $|0\rangle_0$. But what if, instead, we act \mathbb{H} on $|1\rangle_0$, what will be the resulting state?

For the 2nd example, we will take the *direct* way. Firstly, we can write $\mathbb{H}|1\rangle = \mathbb{H}(\mathbb{X}|0\rangle)$. Since the operation must act just on q_0 , therefore, our operator takes the form $(\mathbb{H}\mathbb{X} \otimes \mathbb{I})$.

$$|00\rangle \longrightarrow (\mathbb{H}\mathbb{X} \otimes \mathbb{I}) |00\rangle = \mathbb{H}\mathbb{X}|0\rangle \otimes \mathbb{I}|0\rangle = \mathbb{H}|1\rangle \otimes |0\rangle = |-\rangle \otimes |0\rangle \quad (1.65)$$

$$\mathbb{C}\mathbb{X}|-\rangle = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{-1}{\sqrt{2}} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{-1}{\sqrt{2}} \end{pmatrix} \quad (1.66)$$

which gives $|\Phi^-\rangle$

$$\mathbb{C}\mathbb{X}|-\rangle = \frac{1}{\sqrt{2}} (|00\rangle - |11\rangle) = |\Phi^-\rangle \quad (1.67)$$

Therefore, the gates that take us from $|00\rangle$ to $|\Phi^-\rangle$ are:

$$|\Phi^-\rangle = \mathbb{C}\mathbb{X} (\mathbb{H}\mathbb{X} \otimes \mathbb{I}) |00\rangle \quad (1.68)$$

If we draw the corresponding quantum circuit of this state, then the only difference compared to that of $|\Phi^+\rangle$, is the action of a NOT GATE before the HADAMARD GATE. This circuit is drawn on the left side of the figure below:

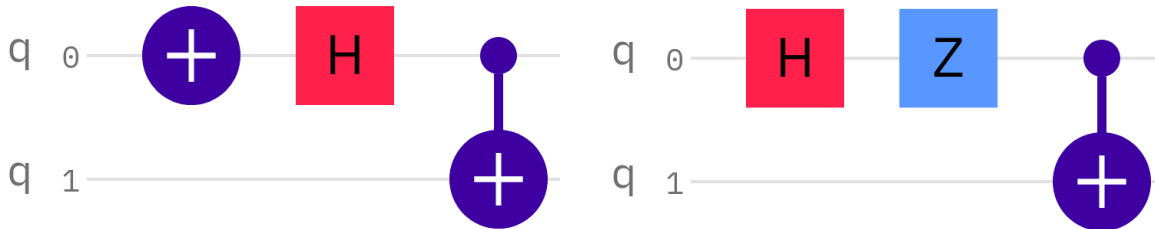


Figure 1.23: Two equivalent quantum circuits that generate the Bell state $|\Phi^-\rangle$.

We can mathematically prove that the state $|\Phi^-\rangle$, can be produced by the *equivalent* circuit, shown on the right above. Firstly, we recall that $\mathbb{Z}|\pm\rangle = |\mp\rangle$; hence, we can write $|-\rangle = \mathbb{Z}|+\rangle = \mathbb{Z}(\mathbb{H}|0\rangle)$ and we know that $|-\rangle = \mathbb{H}(\mathbb{X}|0\rangle)$, therefore

$$\mathbb{Z}\mathbb{H} = \mathbb{H}\mathbb{X}$$

Then we can obtain $|\Phi^-\rangle$, by applying the transformation below on $|00\rangle$:

$$|\Phi^-\rangle = \mathbb{C}\mathbb{X} (\mathbb{Z}\mathbb{H} \otimes \mathbb{I}) |00\rangle \quad (1.69)$$

This clearly explains the circuit on the right side.

We can plot the two states $|\Phi^\pm\rangle$ in the Q-sphere:

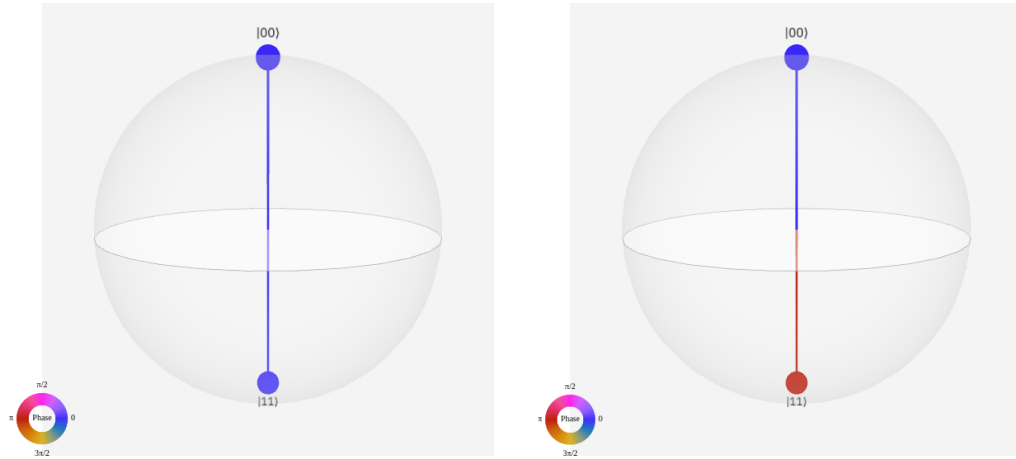


Figure 1.24: A visualization of the Bell states: $|\Phi^+\rangle$ (on the left) and $|\Phi^-\rangle$ (on the right).

In the Q-sphere, we represent the basis state with all **zeros** $|00\rangle$ at the *north pole* and that with all **ones** $|11\rangle$ at the *south pole*. On the left side, we see a representation of $|\Phi^+\rangle$, where the **phase** of $|0\rangle$ and $|1\rangle$ are both $= 0$. The right side concerns that of $|\Phi^-\rangle$, where the **phase** of $|0\rangle$ is 0 and that for $|1\rangle$ is π , which is reflected as a *red color* on the $|11\rangle$ node.

The 3rd and 4th examples relate to the states $|\Psi^\pm\rangle$. Recalling that we have found $|\Phi^\pm\rangle$ by acting CNOT on the states $|\pm 0\rangle$, i.e.

$$|\Phi^\pm\rangle = \text{CX} |\pm 0\rangle$$

What if we apply it on $|\pm 1\rangle$?

$$\text{CX} |\pm 1\rangle = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ 0 \\ \frac{\pm 1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{\pm 1}{\sqrt{2}} \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} (|01\rangle \pm |10\rangle) \quad (1.70)$$

which is clearly the states $|\Psi^\pm\rangle = \frac{1}{\sqrt{2}} (|01\rangle \pm |10\rangle)$.

We want to write it in terms of the initial states $|00\rangle$. We have two distinct states:

$$|\Psi^+\rangle = \text{CX} | + 1 \rangle \quad ; \quad |\Psi^-\rangle = \text{CX} | - 1 \rangle \quad (1.71)$$

$$|\Psi^+\rangle = \text{CX} (|+\rangle \otimes |1\rangle) = \text{CX} (\mathbb{H}|0\rangle \otimes \mathbb{X}|0\rangle) = \text{CX} (\mathbb{H} \otimes \mathbb{X}) |00\rangle \quad (1.72)$$

$$|\Psi^-\rangle = \text{CX} (|-\rangle \otimes |1\rangle) = \text{CX} (\mathbb{H}|1\rangle \otimes \mathbb{X}|0\rangle) = \text{CX} (\mathbb{HX} \otimes \mathbb{X}) |00\rangle \quad (1.73)$$

For the first case, we apply a HADAMARD GATE \mathbb{H} on \mathbf{q}_0 , and a NOT GATE \mathbb{X} on \mathbf{q}_1 . Then by acting CX on the resulting state $| + 1 \rangle$, we finally get the *entangled* state $|\Psi^+\rangle$. The corresponding circuit is drawn below:

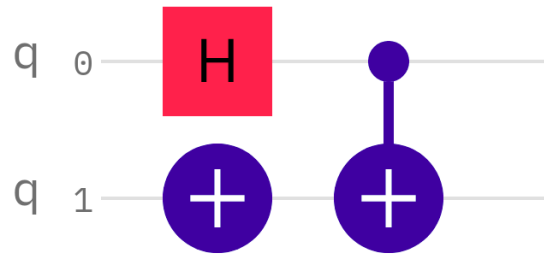


Figure 1.25: The quantum circuit that generates the Bell state $|\Psi^+\rangle$.

In the second case, it is the same process, except that we apply a NOT GATE before \mathbb{H} . Or like the 2nd example, where we apply a \mathbb{Z} GATE after \mathbb{H} , i.e.

$$|\Psi^-\rangle = \text{CX} (\mathbb{H}\mathbb{X} \otimes \mathbb{X}) |00\rangle = \text{CX} (\mathbb{Z}\mathbb{H} \otimes \mathbb{X}) |00\rangle$$

The corresponding quantum circuits are drawn in the figure below:

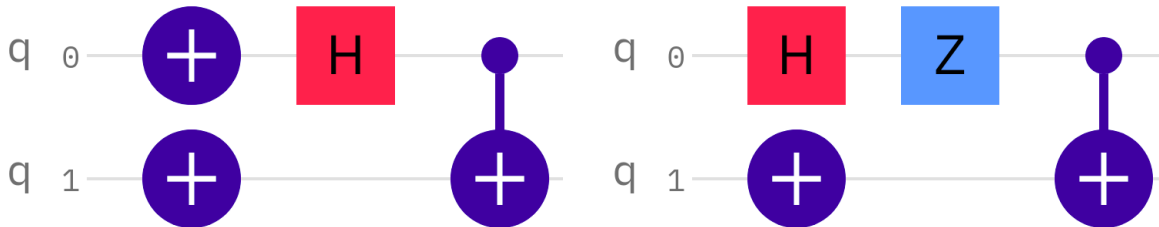


Figure 1.26: Two equivalent quantum circuits that generate the Bell state $|\Psi^-\rangle$.

We can represent the two states $|\Psi^\pm\rangle$ on the Q-sphere, in order to compare them with those of the states $|\Phi^\pm\rangle$.

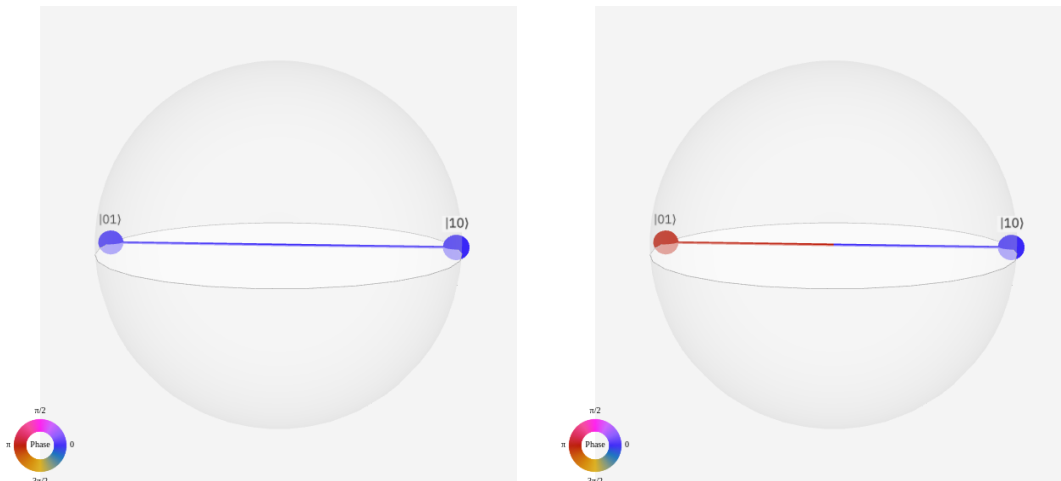


Figure 1.27: A visualization of the Bell states: $|\Psi^+\rangle$ (on the left) and $|\Psi^-\rangle$ (on the right).

If the basis state with all **zeros** $|00\rangle$ is represented at the *north pole* and that with all **ones** $|11\rangle$ at the *south pole*, then the basis states with an **equal number** of 0s and 1s (i.e., $|01\rangle, |10\rangle$) are represented at the **equator**, which explains our situation here. Another thing is the **quantum phase**, where in the left side, it is 0 for both $|0\rangle, |1\rangle$ and in the right side, it's 0

for $|0\rangle$ and π for $|1\rangle$, which explains the difference between the two colors.

Finally, we built the four quantum circuits that generate the four *Bell states*. As seen the two **necessary** gates to create **entanglement**, are the HADAMARD GATE that played the role of the “SOURCE” of the **superposition** property, and the CNOT GATE that played the role of a “KEY” that *close/entangle* or *open/disentangle* two qubit states.

We have seen in this section some *remarkable* quantum gates. But we are far from seeing the full picture; because there are many more gates than one can imagine, as the CH, CZ, CT, CS, CSWAP, . . . , which we can find in the IBM QUANTUM COMPOSER.

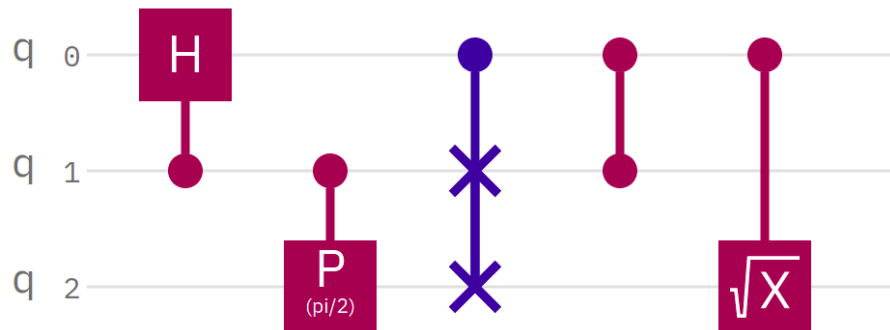


Figure 1.28: A variety of controlled gates (from the left to the right): $\mathbb{R}CH$, CS, CSWAP, CZ, $C\sqrt{X}$.

Unlike gates like $C\sqrt{X}^\dagger$, CCH, CXHT, . . . , and also gates of 4 and 5 qubits. But as shown in the SWAP gate example, we can create them through the existing basic gates. The figure¹⁴ below clearly explains this concept:

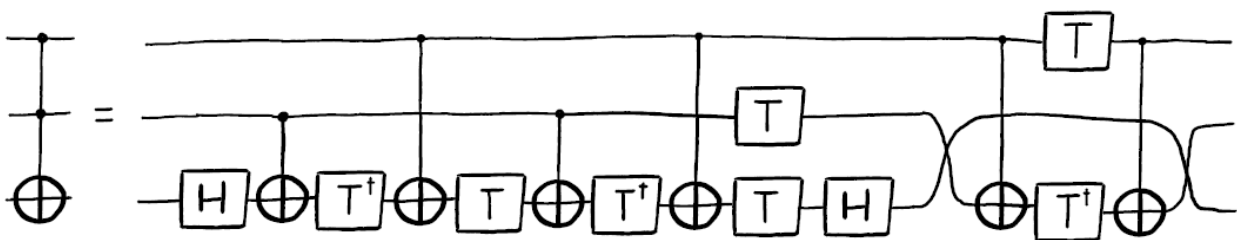


Figure 1.29: The Toffoli gate (on the left) and its implementation (on the right).

This figure shows the real face behind the CCNOT or as called the TOFFOLI gate. This *masterpiece* was invented by TOMMASO TOFFOLI in 1980¹⁵. Therefore any specific multi-qubit gate is valuable. The challenge lies in how to combine the basic gates in order to create a desired gate?

¹⁴<https://qiskit.org/textbook/ch-gates/more-circuit-identities.html>

¹⁵<https://medium.com/qiskit/ibm-quantum-challenge-2021-heres-what-to-expect-65a303753ffb>

1.4 Quantum Algorithms

A classical algorithm is a procedure or a finite sequence of instructions used in problem-solving. All of these instructions can be executed on classical computers. A quantum algorithm is also a step-by-step instruction but, it is performed in quantum computers. Quantum algorithms can solve some computational issues faster than their classical counterparts. For example, *integers factorization* is *exponentially* faster, and an *unordered search* is *quadratically* faster than any classical algorithm.

1.4.1 Deutsch-Jozsa Algorithm

The Deutsch-Jozsa algorithm was the first quantum algorithm that showed that quantum computing works better than the best classical algorithm and demonstrates exponential acceleration compared to a classical deterministic algorithm [15].

- **Deutsch-Jozsa Problem**

A hidden boolean function $f(x_1, \dots, x_n)$ takes as input a string of bits and returns 0 or 1, where x_n are 0 or 1. If the function returns all 0 or all 1 for any input, then it is a *constant function*. If the function returns 0 for exactly half of all inputs and 1 for the other half, then it is a *balanced function*. Our goal here is to determine if a given function is *balanced* or *constant*.

- **The Classical Solution:** If the task is performed with a classical algorithm on a classical computer, we can solve this problem after two calls to the function f . For example, if we find 0 as output for the first combination and 1 in the second combination, then the function f , is balanced. This is the best case. In the worst case, if we continue to see the same output for each input we try, we will have to check exactly half of all possible inputs plus one in order to be certain that f is constant.

- **The Quantum Solution:** We can solve this problem using quantum computing after only one call to the function f . This function is implemented as a quantum oracle, it maps the state from $|x\rangle|y\rangle$ to $|x\rangle|y \oplus f(x)\rangle$ ¹⁶.

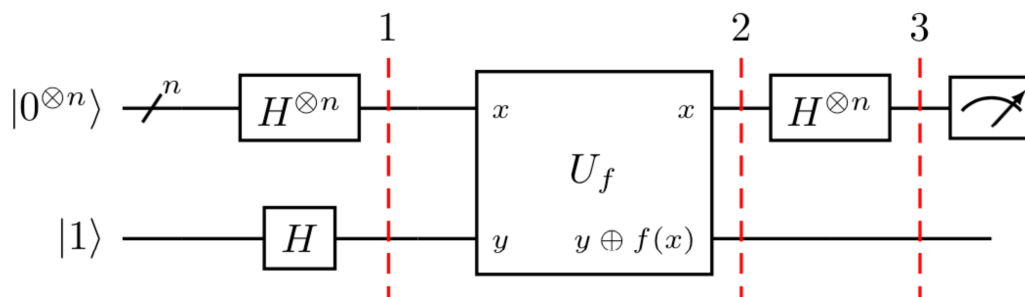


Figure 1.30: Generic circuit for the Deutsch-Jozsa algorithm¹⁷.

¹⁶<https://qiskit.org/textbook/ch-algorithms/deutsch-jozsa.html>

Now, let's go through the steps of the algorithm:

- Prepare two quantum registers. The first is an n -qubit register initialized to $|0\rangle$, and the second is a one-qubit register initialized to $|1\rangle$

$$|\psi_0\rangle = |0\rangle^{\otimes n} |1\rangle \quad (1.74)$$

- Apply a Hadamard gate to each qubit:

$$|\psi_1\rangle = (\mathbb{H}^{\otimes n} |0\rangle^{\otimes n})(\mathbb{H} |1\rangle) \quad (1.75)$$

$$\begin{aligned} &= (\mathbb{H} |0\rangle \otimes \mathbb{H} |0\rangle \otimes \dots \otimes \mathbb{H} |0\rangle) (\mathbb{H} |1\rangle) \\ &= \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes \dots \otimes \frac{|0\rangle + |1\rangle}{\sqrt{2}} \right) \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) \\ &= \left(\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \right) \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) \\ |\psi_1\rangle &= \frac{1}{\sqrt{2^{n+1}}} \sum_{x=0}^{2^n-1} |x\rangle (|0\rangle - |1\rangle) \end{aligned} \quad (1.76)$$

- Apply the *quantum oracle* $|x\rangle|y\rangle$ to $|x\rangle|y \oplus f(x)\rangle$:

$$\begin{aligned} |\psi_2\rangle &= \frac{1}{\sqrt{2^{n+1}}} \sum_{x=0}^{2^n-1} |x\rangle (|0 \oplus f(x)\rangle - |1 \oplus f(x)\rangle) \\ &= \frac{1}{\sqrt{2^{n+1}}} \sum_{x=0}^{2^n-1} (-1)^{f(x)} |x\rangle (|0\rangle - |1\rangle) \end{aligned} \quad (1.77)$$

Since for each x , $f(x)$ is either 0 or 1. At this point the second single qubit register may be ignored.

- Apply a Hadamard gate to each qubit in the first register:

$$\begin{aligned} |\psi_3\rangle &= \frac{1}{2^n} \sum_{x=0}^{2^n-1} (-1)^{f(x)} \left[\sum_{y=0}^{2^n-1} (-1)^{x \cdot y} |y\rangle \right] \\ &= \frac{1}{2^n} \sum_{y=0}^{2^n-1} \left[\sum_{x=0}^{2^n-1} (-1)^{f(x)} (-1)^{x \cdot y} \right] |y\rangle \end{aligned} \quad (1.78)$$

with $\mathbb{H} |x\rangle = \sum_{y=0}^{2^n-1} \frac{(-1)^{x \cdot y}}{\sqrt{2}} |y\rangle$ and $\mathbb{H}^{\otimes n} |x\rangle = \sum_{y=0}^{2^n-1} \frac{(-1)^{x \cdot y}}{\sqrt{2^n}} |y\rangle$

Where $x \cdot y = x_0 y_0 \oplus x_1 y_1 \oplus \dots \oplus x_{n-1} y_{n-1}$ is the sum of the *bitwise product*.

- Measure the first register. Notice that the probability of measuring $|0\rangle^{\otimes n} = \left| \frac{1}{2^n} \sum_{x=0}^{2^n-1} (-1)^{f(x)} \right|^2$, which evaluates to 1 if $f(x)$ is constant and 0 if $f(x)$ is balanced.

¹⁷<https://qiskit.org/textbook/ch-algorithms/deutsch-jozsa.html>

Example for $n = 1$

To better understand the Deutsch-Jozsa algorithm, let us take a simple example for a single input ($n = 1$). The quantum circuit that we need to use to solve the problem is very simple:

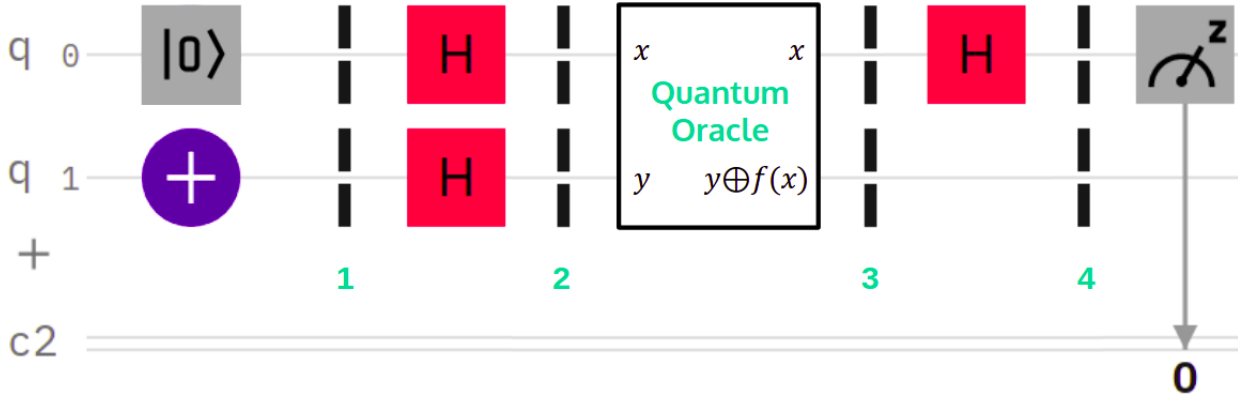


Figure 1.31: Deutsch-Jozsa circuit example for a single input ($n=1$).

The algorithm is divided by four steps. Let's look at what happens in each numbered state.

Step 1:

In this step, we prepare two quantum registers. The first is a one-qubit register initialized to $|0\rangle$, and the second is one-qubit register initialized to $|1\rangle$. The initial state is:

$$|\psi_0\rangle = |0\rangle |1\rangle \quad (1.79)$$

Step 2:

The next step is to apply the Hadamard gates, we get the following result:

$$\begin{aligned} |\psi_1\rangle &= \frac{(|0\rangle + |1\rangle)}{\sqrt{2}} \frac{(|0\rangle - |1\rangle)}{\sqrt{2}} \\ &= \frac{(|0\rangle + |1\rangle)(|0\rangle - |1\rangle)}{2} \\ &= \frac{|0\rangle(|0\rangle - |1\rangle)}{2} + \frac{|1\rangle(|0\rangle - |1\rangle)}{2} \end{aligned} \quad (1.80)$$

Step 3:

The next step is to apply the oracle function, the input (x, y) will become to $(x, y \oplus f(x))$ in the output. The state before the oracle function is

$$|\psi_1\rangle = \frac{|0\rangle(|0\rangle - |1\rangle)}{2} + \frac{|1\rangle(|0\rangle - |1\rangle)}{2} \quad (1.81)$$

Now let us see the effect of the oracle function. When we apply the oracle function, by linearity we obtain:

$$\frac{|0\rangle(|0 \oplus f(0)\rangle - |1 \oplus f(0)\rangle)}{2} + \frac{|1\rangle(|0 \oplus f(1)\rangle - |1 \oplus f(1)\rangle)}{2} \quad (1.82)$$

- When $f(0) = 0$, we get $|0 \oplus f(0)\rangle - |1 \oplus f(0)\rangle = |0\rangle - |1\rangle$.
- When $f(0) = 1$, we have $|0 \oplus f(0)\rangle - |1 \oplus f(0)\rangle = |0 \oplus 1\rangle - |1 \oplus 1\rangle = |1\rangle - |0\rangle = -(|0\rangle - |1\rangle)$.

- For the $f(1)$ we do the same thing, so our global state can be written as

$$|\psi_3\rangle = \frac{(-1)^{f(0)} |0\rangle (|0\rangle - |1\rangle)}{2} + \frac{(-1)^{f(1)} |1\rangle (|0\rangle - |1\rangle)}{2}. \quad (1.83)$$

This expression can be also written as

$$\frac{|0\rangle (|0\rangle - |1\rangle)}{2} + \frac{(-1)^{f(0)+f(1)} |1\rangle (|0\rangle - |1\rangle)}{2} \quad (1.84)$$

- If $f(0) = f(1)$, we will get

$$\frac{|0\rangle (|0\rangle - |1\rangle)}{2} + \frac{|1\rangle (|0\rangle - |1\rangle)}{2} = \frac{(|0\rangle + |1\rangle)(|0\rangle - |1\rangle)}{2} \quad (1.85)$$

Lets summarize what happen when $f(x)$ is constant or balanced.

When constant:

$$f(0) = 0 \text{ and } f(1) = 0 \rightarrow \frac{|0\rangle+|1\rangle}{\sqrt{2}} \text{ also } f(0) = 1 \text{ and } f(1) = 1 \rightarrow \frac{-|0\rangle-|1\rangle}{\sqrt{2}}$$

When balanced:

$$f(0) = 0 \text{ and } f(1) = 1 \rightarrow \frac{|0\rangle-|1\rangle}{\sqrt{2}} \text{ also } f(0) = 1 \text{ and } f(1) = 0 \rightarrow \frac{-|0\rangle+|1\rangle}{\sqrt{2}}$$

Step 4: The final step is to apply the Hadamard gate and the measure.

When we apply the Hadamard gate, we get the state $|1\rangle$ if it is balanced and $|0\rangle$ if it is constant. We can verifier this result by the measurement.

1.4.2 Shor's Algorithm

Shor's Factorization Algorithm is proposed by Peter Shor. It suggests that quantum mechanics allows the factorization to be performed in *polynomial time*, rather than *exponential time* achieved after using classical algorithms [15].

- The basic idea of Shor's algorithm:

The main problem to be solved by Shor's algorithm is given an integer N , we want to find its *prime factor*. In other words, we want to determine the two prime factors p_1 and p_2 which satisfy $p_1 \cdot p_2 = N$ for a given large number of N in polynomial time.

Shor algorithm is a *hybrid quantum-classical algorithm*. It begins with some classical processing, then uses the quantum computer to execute a particular part of the algorithm (a subroutine¹⁸). This subroutine is difficult to perform on classical computers, but it can be solved with a fast quantum algorithm using the *quantum Fourier transform* (QFT). The quantum computer essentially serves as a *co-processor* to the classical computer.

In the case of Shor's algorithm, the quantum subroutine is known as the period finding routine, which uses an important technique known as the quantum Fourier transform, or

¹⁸A sequence of computer instructions for performing a specified task that can be used repeatedly.

QFT, to create *interference* that gives us the period. We then use a classical technique known as *Euclid's algorithm* to find the prime factors of the number. It turns out that, if we can find the period r of a function $f(x) = a^x \bmod N$ as we increment the variable x , then we can find the factors of N with a is selected prime number inferior of N .

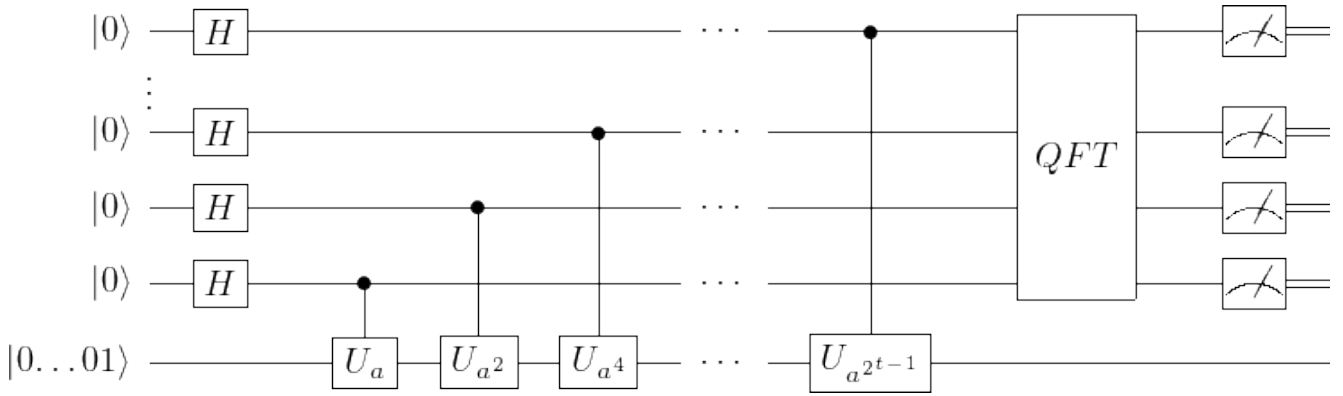
- Steps for Shor algorithm

The algorithm consists of two parts, a classical and quantum part. The classical part reduces the factorization to a problem of finding the period of the function. This is done classically using a classical computer. Quantum part which uses a quantum computer to find the period using the Quantum Fourier Transform.

Let's assume we're trying to factor a number N . The algorithm consists of the following steps:

1. Pick a random number a such that $a < N$.
2. Use Euclid's algorithm to check if is a factor of N ; if so, you're done.
3. Otherwise, prepare your quantum computer and a program for it to execute quantum period finding.
4. The first important quantum step is to calculate the superposition of a^x for all values of x , all done modulo N . $|x\rangle$ and the remainder both appear in our quantum register.
5. Measure the remainder (Technically, this is unnecessary).
6. Take the quantum Fourier transform (QFT) of $|x\rangle$.
7. Measure the register $|x\rangle$; call the result r , which should be a multiple of the period of the modular exponentiation function.
8. If r is even and greater than zero, calculate the numbers $a^{(r/2)} + 1$ and $a^{(r/2)} - 1$; if not, repeat the quantum portion of the algorithm to find a new r .
9. Use Euclid's algorithm to calculate the GCD (Great Common Divider) of each of those numbers and N , and we should have two factors of N !

The quantum circuit for Shor's algorithm is represented in the following figure [16]:



- The first t qubits, which form the first register, will hold the exponents to which we are raising a . The measurement at the end will yield (with high probability) some value y such that $\frac{y}{2^t}$ is close to a rational number $\frac{s}{r}$, where r is the period of $a \pmod{N}$. Thus increasing the number of these qubits, will simply increase the accuracy of this approximation.
- The last n qubits (initialized to be $|0 \dots 01\rangle$), which form the second register, will be where the powers $a^j \pmod{N}$ get stored. Hence we can take $n = \lceil \log_2 N \rceil$, the number of binary digits of N .
- First, we apply \mathbb{H} to each qubit in the input. The result will be that the first t qubits will each become $\mathbb{H}|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. The first t qubits taken together will be in the state

$$\frac{1}{2^{t/2}}(|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle) \otimes \dots \otimes (|0\rangle + |1\rangle) = \tag{1.86}$$

$$\frac{1}{2^{t/2}}(|0 \dots 00\rangle + |0 \dots 01\rangle + \dots + |1 \dots 11\rangle) = \frac{1}{2^{t/2}} \sum_{j=0}^{2^t-1} |j\rangle. \tag{1.87}$$

In other words, it will be in a *balanced superposition* of all the basis states.

- Then, we have a number of controlled U_x -gates. The gate U_x performs multiplication by $x \pmod{N}$. That mean, $U_x |y\rangle = |xy \pmod{N}\rangle$, for $0 \leq y < N$.

We can construct such gates efficiently since we can efficiently multiply two n digit numbers, and we can compute each $x = a^{2^k}$ simply by squaring $a \pmod{N}$.

- For example suppose $|j\rangle = |j_{t-1} \dots j_2 j_1 j_0\rangle$ is the content of the first register. Then the controlled U_x -gates will multiply the second register by a if $j_0 = 1$, then by a^2 if $j_1 = 1$, then by a^4 if $j_2 = 1$, and so forth. The result is that the second register will become $a^{\sum j_k 2^k} = a^j$.
- After these gates, the total state of the system will be a superposition of the form:

$$\frac{1}{2^{t/2}} \sum_{j=0}^{2^t-1} |j\rangle |a^j \pmod{N}\rangle. \tag{1.88}$$

We can write it as:

$$\frac{1}{2^{t/2}} \sum_{j,b} x_j^{(b)} |j\rangle |b\rangle, \quad (1.89)$$

where $x_j^{(b)} = 1$ if $b = a^j \pmod{N}$, otherwise $x_j^{(b)} = 0$. Thus, for any fixed b , the sequence $\{x_j^{(b)}\}_{j=0}^{2^t-1}$ will be periodic (in j) with period r .

- In this step, we input the first register into the QFT. This will send the state to

$$\frac{1}{2^{t/2}} \sum_{k,b} y_k^{(b)} |k\rangle |b\rangle, \quad (1.90)$$

where, for any fixed b , $\{y_k^{(b)}\}$ is the *discrete Fourier transform* of $\{x_j^{(b)}\}$. Since $\{x_j^{(b)}\}$ was periodic with period r , $|y_k^{(b)}|$ will be large when y_k is approximately a multiple of $\frac{2^t}{r}$.

Let $N = 2^t$. The quantum Fourier transform (QFT) is the linear transformation (acting on t qubits) that acts on the computational basis $|0\rangle, |1\rangle, \dots, |N-1\rangle$ by

$$|j\rangle \mapsto \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{2\pi i j k / N} |k\rangle. \quad (1.91)$$

This gives

$$\sum_{j=0}^{N-1} x_j |j\rangle \mapsto \sum_{k=0}^{N-1} y_k |k\rangle, \quad (1.92)$$

where $\{y_k\}_{k=0}^{N-1}$ is the discrete Fourier transform of $\{x_j\}_{j=0}^{N-1}$.

- Finally, we measure the first register. With high probability, this will give a value k such that k is approximately a multiple of $\frac{2^t}{r}$, that is, such that $\frac{k}{2^t}$ is close to a rational number with r as the denominator.

Once we have this value of k , we compute the continued fraction expansion of $\frac{k}{2^t}$ to approximate it by a *rational number* with a relatively small denominator r . Then we test to make sure that $a^r \equiv 1 \pmod{N}$.

1.5 Error Detection and Correction

In 1948, CLAUDE SHANNON proved that a reliable transmission of information over *noisy classical channels* is feasible. This discovery led to the search for *classical error correction* (CEC) codes. Then, in the mid-1990s PETER SHOR and ANDREW STEANE showed that *quantum error correction* (QEC) is possible [2]. So, in this section, we will give a glimpse of CEC and then we tackle QEC.

1.5.1 Classical Error Correction

Noise refers to any external and undesired disturbance superimposed on a useful signal, created by the operating environment and which leads to errors. In *data communication*, the four most important kind of noise are: *thermal noise*, *intermodulation noise*, *impulse noise* and *crosstalk*. If there are defects in the hardware, either original or acquired, this can cause errors as well.

To correct these errors, we use CODING THEORY which is critical for reliable *communication*, *information storage*, and *processing*. Software within the storage devices can often **detect** and **correct** errors quickly using as small extra data as possible. This is called “*fault-tolerance*” [8].

The aspects of error correction process can be sited in the 4 points below:

- An initial information to be **sent**, **stored**, or **processed**.
- **Encoding** the information in a larger system to protect it against noise.
- A **noise process** that alters information.
- **Decoding** the information, and correct it, which reduces the effects of noise.



Figure 1.32: The classical error correction process.

The trivial strategy for encoding information is to repeat it several times then the decoding is done by *majority decision*. This is known as a “*repetition code*”. For example, if we want to transmit the information 01101 through a *noisy channel* that corrupt the transmission in some places, we have to repeat it in order to reduce the error rate [8].

$$[01101 \quad 01101 \quad 01101] \rightarrow \langle \text{noisy channel} \rangle \rightarrow [01101 \quad 01100 \quad 01101]$$

An error occurred in one of The three copies, so we conclude that the original data is that occurs most often, which is **01101**. We see clearly the weak point of this code; if there are one more errors than we cannot correct, but we can avoid this by more redundancy. Generally, for a repetition of length n , the *error correction capacity* is $\frac{n-1}{2}$.

To detect an error there is another simple method is the use of an **even parity bit** [8]:

- If the data contains an **even** number of 1 bits, we precede it with a **0**, e.g. 10100 \rightarrow 010100.

- If the data contains an **odd** number of 1 bits, we precede it with a **1**, e.g. 10110 \rightarrow 110110.

In the two we have an **even** number of 1s, so if there is an **odd** number of 1s after the transmission, then at least one *bit flip* has occurred, anywhere. But if there is 2 or any even number of *bit flips*, then we will not know.

There are more powerful codes that protect information. The characteristics of a good code are:

1. The number of repetitions in the original message must be minimal.
2. Ease of encoding and decoding information.

The concepts of classical error detection (CED) and correction (CEC) introduced above are useless in the quantum case, because detecting errors in transmitted information requires measurement. But in the quantum case, this causes a collapsing of the transmitted qubit state, so we lost any information stored in its coefficients [17]. The philosophy of *quantum error correction* is to detect and correct information without measuring it!

1.5.2 Quantum Error Correction

Physical qubits are characterized by **decoherence**, which occurs when a state is affected by physical processes, and which cannot be reversed. For example, the effect of measures. In the current era of *quantum computing*, we are using a large number of noisy physical qubits to encode one **logical qubit**, through the process of quantum error correction (QEC) [8].

The decoherence is due to the **entanglement** of the qubit with its environment, which attempts to alter the state of the system. But as we will see, the encoding of logical qubits is maintained by constantly putting the physical qubits through a *highly entangling* circuit. PETER SHOR summarizes the fundamental idea of QEC as to “*fight entanglement with entanglement*” [2].

We have seen that to protect an information against noise requires creating multiple copies of it, in a way that decreases the probability of errors. But in quantum mechanics, the NO-CLONING THEOREM announces that a quantum state cannot be copied.

$$|\psi\rangle = a|0\rangle + b|1\rangle \not\rightarrow |\psi\rangle \otimes |\psi\rangle \otimes \cdots \otimes |\psi\rangle$$

For this, a formulation of QEC was suspected that is not possible. Therefore, we need another way to employ the reasoning of redundancy of classical error correction in QEC, and it turns out that this way is *entanglement*; indeed, instead of coping the state of a qubit multiple times, we can spread it onto a **highly entangled** state of multiple-qubits [8][2]. The encoding in QEC combines the concept of *redundancy* and *entanglement*, i.e.

$$|\psi\rangle = a|0\rangle + b|1\rangle \xrightarrow{\text{encoding}} |\psi\rangle = a|000\cdots\rangle + b|111\cdots\rangle$$

So we entangle one qubit carrying information, with $(n - 1)$ qubits of state $|0\rangle$, and create an n -qubit quantum state that is more resilient to errors.

The qubit state is prone to two types of quantum errors: **bit-flips** and **sign-flips** [18][19].

1.5.2.1 Single Bit-Flip Error Correction Code

Bit flips are the only kind of errors that occurs in the classical case and makes the interchange of $0 \rightleftharpoons 1$. This bit-flip can be traduced in the quantum case as the interchange of $|0\rangle \rightleftharpoons |1\rangle$; which means that a state $a|0\rangle + b|1\rangle$ becomes $a|1\rangle + b|0\rangle$, where the *amplitudes* are switched, which is a NOT action; the X GATE leads to a *bit-flip error*, and that can be corrected by applying another X GATE, since $\mathbb{X}.\mathbb{X} = \mathbb{I}$ [8].

In 1985 ASHER PERES proposed a method called “*3-qubit bit-flip code*” that corrects a single *bit-flip* error. For a given transmitted state $|\psi\rangle$ this method consists of the following steps:

1. **Encoding** $|\psi\rangle$ to an entangled 3-qubit state.
2. The *Noisy channel* cause just **bit-flip** errors.
3. **Decoding** the transmitted state.

For the 1st step, we have to *entangle* the qubit q_0 with both q_1 and q_2 , Therefore

$$|\psi 00\rangle \rightarrow (\mathbb{C}\mathbb{X}_{[0,2]})(\mathbb{C}\mathbb{X}_{[0,1]})(|\psi\rangle_0 \otimes |0\rangle_1 \otimes |0\rangle_2)$$

where

$$\mathbb{C}\mathbb{X}_{[0,1]} = (\mathbb{C}\mathbb{X} \otimes \mathbb{I}) = (|0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes \mathbb{X}) \otimes \mathbb{I}$$

$$\mathbb{C}\mathbb{X}_{[0,2]} = (|0\rangle\langle 0| \otimes \mathbb{I} \otimes \mathbb{I} + |1\rangle\langle 1| \otimes \mathbb{I} \otimes \mathbb{X})$$

hence

$$(\mathbb{C}\mathbb{X}_{[0,2]})(\mathbb{C}\mathbb{X} \otimes \mathbb{I})(|\psi 0\rangle_{01} \otimes |0\rangle_2) = (\mathbb{C}\mathbb{X}_{[0,2]})((a|00\rangle_{01} + b|11\rangle_{01}) \otimes |0\rangle_2) = a|000\rangle + b|111\rangle$$

This correspond to the 1st block in the circuit below:

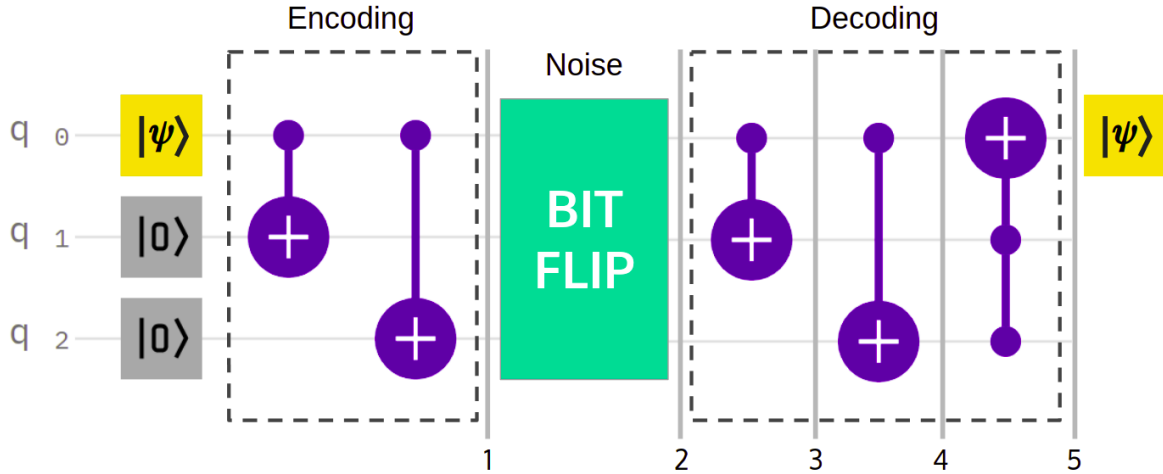


Figure 1.33: Single bit flip error correction circuit.

After encoding one **logical qubit** as 3 **physical qubits**, this one will pass through a **noisy environment** that may cause *bit-flips*. The decoding step consists of *disentangling* the *affected state* via the CNOT GATES. The role of TOFFOLI GATE is **detection** and **correction** of errors; where the 2 *control* parts work for *detection* and the *target* part is for *correction*.

$ \psi\rangle$	Nbr of Errors	Step 1	Step 2	Step 3	Step 4	Step 5	
0⟩	0 Errors	000⟩	000⟩	000⟩	000⟩	0 00⟩	
	1 Error	000⟩	100⟩	110⟩	111⟩	0 11⟩	
		000⟩	010⟩	010⟩	010⟩	0 10⟩	
		000⟩	001⟩	001⟩	001⟩	0 01⟩	
	2 Errors	000⟩	110⟩	100⟩	101⟩	1 01⟩	
		000⟩	101⟩	111⟩	110⟩	1 10⟩	
		000⟩	011⟩	011⟩	011⟩	1 11⟩	
	3 Errors	000⟩	111⟩	101⟩	100⟩	1 00⟩	
	1⟩	0 Errors	111⟩	111⟩	111⟩	111⟩	1 11⟩
		1 Error	111⟩	011⟩	011⟩	011⟩	1 11⟩
111⟩			101⟩	111⟩	110⟩	1 10⟩	
111⟩			110⟩	100⟩	101⟩	1 01⟩	
2 Errors		111⟩	001⟩	001⟩	001⟩	0 01⟩	
		111⟩	010⟩	010⟩	010⟩	0 10⟩	
		111⟩	100⟩	110⟩	111⟩	0 11⟩	
3 Errors		111⟩	000⟩	000⟩	000⟩	0 00⟩	

Here we can clearly understand what we have said before, about the fact that detection

and correction in QEC is done without any measurement. We see that the error correction capacity of this circuit does not exceed one bit flip. Furthermore, the correction concerns just q_0 , which is reasonable.

1.5.2.2 Single Sign-Flip Error Correction Code

Another kind of errors that can occur in quantum computers, it is the **sign-flip** or **phase-flip**, which is a π phase error that makes the interchange $(a|0\rangle + b|1\rangle) \rightleftharpoons (a|0\rangle - b|1\rangle)$, which describes the action of the Z GATE. Unlike bit-flips, this one exists only in the quantum case [8]. To construct the error correction circuit of sign-flips errors, we have to understand the identity obtained in subsection 1.3.3.

$$\mathbb{H}\mathbb{X} = \mathbb{Z}\mathbb{H} \longrightarrow \mathbb{X} = \mathbb{H}\mathbb{Z}\mathbb{H}$$

This means that the action of \mathbb{X} (*bit-flip*) on the *computational* basis $\{|0\rangle, |1\rangle\}$ is equivalent to the action of \mathbb{Z} (*sign-flip*) on the HADAMARD basis $\{|+\rangle, |-\rangle\}$; hence, to convert *bit-flips* to *sign-flips*, we have just to change the basis [8].

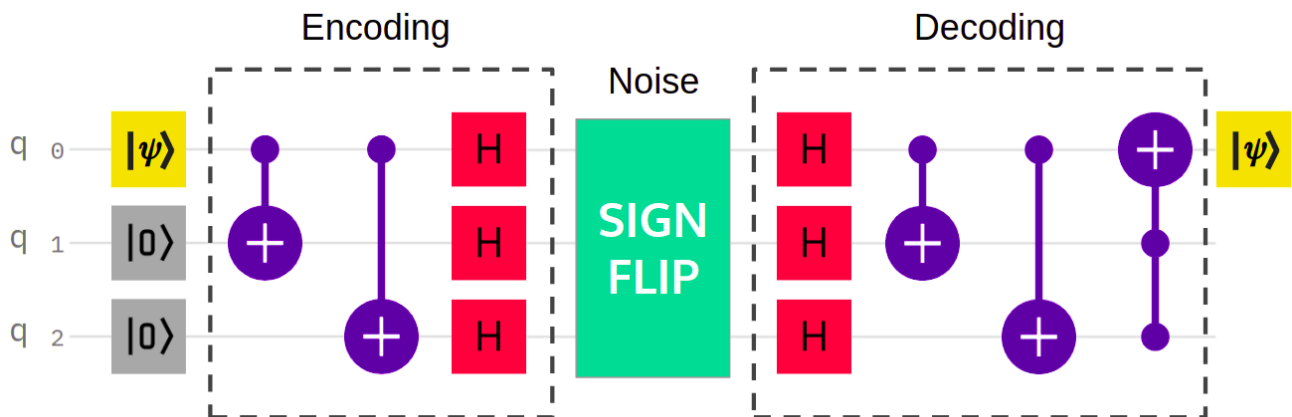


Figure 1.34: Single sign flip error correction circuit.

The encoding used in this case consist of the following mapping:

$$|\psi\rangle = a|0\rangle + b|1\rangle \rightarrow |\Psi\rangle_{bf} = a|000\rangle + b|111\rangle$$

$$|\Psi\rangle_{bf} \rightarrow (\mathbb{H} \otimes \mathbb{H} \otimes \mathbb{H})|\Psi\rangle_{bf} = |\Psi\rangle_{sf} = a|+++ \rangle + b|--- \rangle$$

As the *bit-flip* error correction circuit, this circuit corrects up to one *sign-flip* error in q_0 .

1.5.2.3 The 9-Qubit Error Correction Shor Code

In 1995 the mathematician PETER SHOR published work on an error correction circuit that can correct, a single *bit-flip* (\mathbb{X}), a single *sign-flip* (\mathbb{Z}) or both (\mathbb{Y}) [8][2]. This general error can be modeled as a *unitary* transform \mathbb{E} , that act on the qubit, and which expresses as

a *linear combination* of each error type.

$$\mathbb{E} = \alpha \mathbb{I} + \beta \mathbb{X} + \mu \mathbb{Z} + \nu \mathbb{Y} \quad ; \quad \alpha, \beta, \mu, \nu \in \mathbb{C}$$

The SHOR CODE, encodes the information of **one-qubit** onto a *highly entangled 9-qubit* state. The followed steps are:

1. The qubit q_0 state to be transmitted is encoded by 2-qubits (q_3 and q_6) via the circuit of **Sign-flip** error correction.
2. Each of the 3 qubits (q_0, q_3, q_6) is encoded via the **Bit-flip** error correction code, which gives 3 *bit-flip* code groups : $\{(q_0, q_1, q_2); (q_3, q_4, q_5); (q_6, q_7, q_8)\}$, respectively.
3. Each of these *bit-flip* groups receives its own error correction.
4. At the end, a global *sign-flip* correction is carried out.

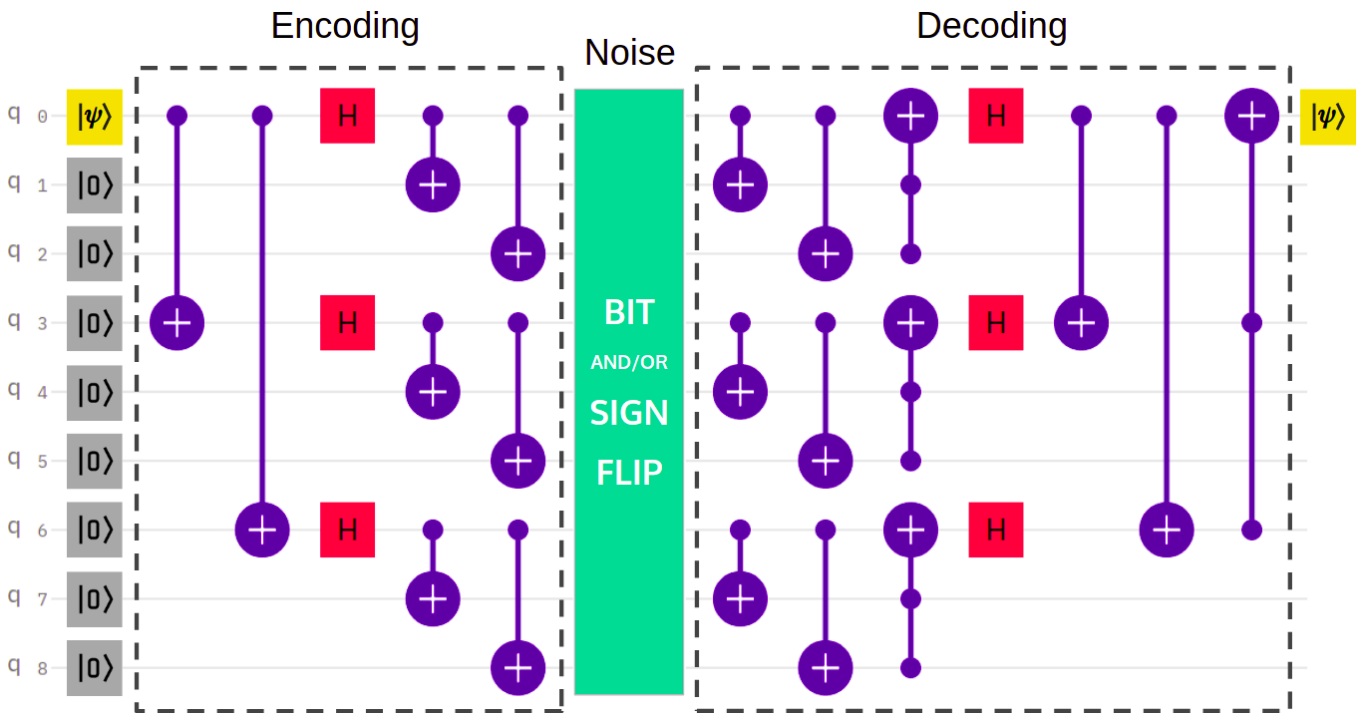


Figure 1.35: SHOR error correction circuit.

The SHOR CODE encoding requires the following transformation:

$$|0\rangle \rightarrow |0\rangle_{Shor} = \frac{1}{\sqrt{2^3}} (|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle) \otimes (|000\rangle + |111\rangle)$$

$$|1\rangle \rightarrow |1\rangle_{Shor} = \frac{1}{\sqrt{2^3}} (|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle) \otimes (|000\rangle - |111\rangle)$$

The encoded qubit state is written

$$|\psi\rangle = a|0\rangle + b|1\rangle \longrightarrow |\Psi\rangle_{Shor} = a|0\rangle_{Shor} + b|1\rangle_{Shor}$$

For this, the 9-QUBIT SHOR CODE is able to correct any *arbitrary* error.

After understanding the QEC basic error correction codes, the problem that appears is that to correct the errors of a given qubit, we need 2 *correction-qubits*, then we need 2 more for each *correction-qubit*, and so on. Therefore, other methods are used, like the “*surface codes*” based on GROUP THEORY and TOPOLOGY¹⁹.

Another important kind of errors occurs when reading out data from a quantum computer, they are known as “READOUT ERRORS” [20][21]. This type of error represents a *central* problem in quantum computing, because it intervenes when extracting an outcome after any quantum computation, and gives us results that are not entirely correct. Readout errors and their correction methods, will be developed in depth in [chapter 3](#).

¹⁹Lecture Notes 5: <https://qiskit.org/learn/intro-qc-qh>

2

Unfolding for High Energy Physics

Contents

2.1 Inverse Problem and Unfolding Methods	58
2.2 Unfolding with Matrix Inversion	59
2.3 Iterative Bayesian Unfolding Method	60

In the Large Hadron Collider (LHC), beams of particles are accelerated and collide to make debris of new particles; these particles fly out from the collision point in all directions. Detectors are situated where the beams can be crossed. When a particle interacts with the detector, an electrical signal is produced and sent to a readout system. They are digitized and read out by a computer and stored on a hard disk.

The measurements of the detector are grouped into events containing all information associated with one scattering process, and events are grouped into certain regions of phase-space, called bins. The measurement of distributions $f(x)$ of some kinematical quantity x is affected by the finite resolution of the particle detectors. If we have an ideal detector then we could measure this quantity in every event and could obtain the distribution by a histogram. But, with a real detector, determination of a distribution $f(x)$ is more difficult because of three factors [22]:

- **Finite resolution of the detector:** There is only a statistical relation between the true kinematical variable x and the measured quantity y . The measured quantity y is smeared out due to the limited measurement accuracy of the detector.
- **Limited acceptance of the detector:** Acceptance of a detector is the range of the kinematic parameters which particles are potentially detectable in it. The probability to observe a given event in a region is called acceptance.
- **Non-linear response:** The transformation from x to y can be caused by the non-linear response of a detector component. Instead of the quantity x we found a different quantity y in measurement.

To solve this problem and to estimate truth distribution (for example, cross-sections) from measurement, we use a statistical technique called unfolding methods.

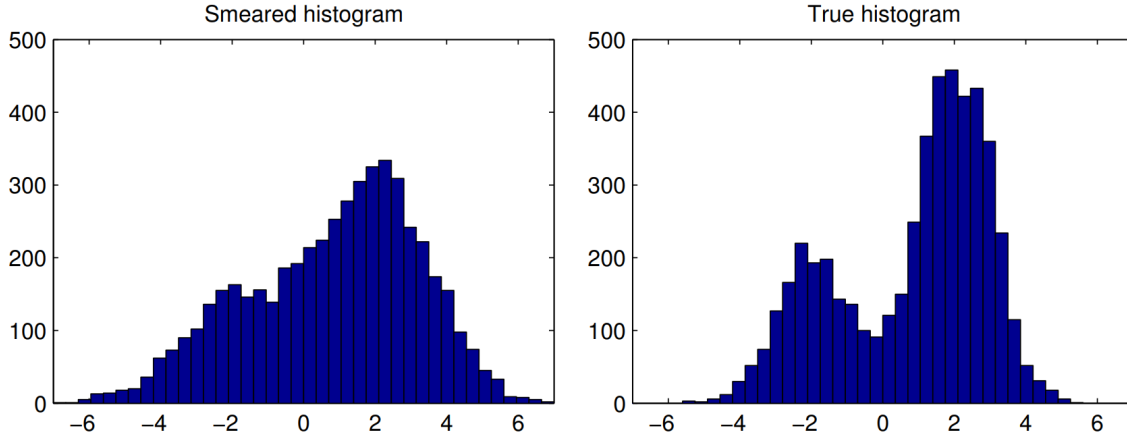


Figure 2.1: Representation of an example of Smeared and True histogram [23].

2.1 Inverse Problem and Unfolding Methods

Unfolding is the ensemble of statistical techniques used to estimate the true spectrum using the smeared observations. The distribution $f(x)$ of the true variable x is related to the measurement distribution $g(y)$ of the measured quantity y by a Fredholm integral equation of the first kind:

$$g(y) = \int_{\omega} R(y, x) f(x) dx \quad (2.1)$$

The problem to determine the distribution $f(x)$ from measured distribution $g(y)$ is an inverse problem and it is called unfolding. $R(y, x)$ represents the detector effect.

The equation is usually discretized when we want to represent the probability density functions $f(x)$ and $g(y)$ in histograms. We can construct a histogram of y with M bins.

The integral equation becomes:

$$m_i = \sum_{j=1}^M \mathcal{R}_{ij} t_j \quad i = 1, \dots, N. \quad (2.2)$$

The response matrix \mathcal{R} could be interpreted simply as a conditional probability:

$$\mathcal{R}_{ij} = \mathbb{P}(\text{measured in bin } i | \text{true value in bin } j) \quad (2.3)$$

where $t = (t_1, \dots, t_M)$ is the expectation values for the histogram of x and $m = (m_1, \dots, m_N)$ gives the expected number of events in bins of the observed variable y . The data are given as a vector of numbers $n = (n_1, \dots, n_N)$.

The efficiency ϵ_j is obtained by summing \mathcal{R}_{ij} :

$$\sum_i^N \mathcal{R}_{ij} = \mathbb{P}(\text{observed anywhere} | \text{true value in bin } j) = \epsilon_j \quad (2.4)$$

In the general form, the equation (2.2) should include a number of background events as follow:

$$m = \mathcal{R}t + \beta. \quad (2.5)$$

Unfolding is useful when we want to compare the measurement with other experiments, or to compare the measurement with future theories. It is most often used in measurement analyses.

2.2 Unfolding with Matrix Inversion

There are many unfolding methods and the most simple is the *matrix inversion*. This is possible only if the number of bins observed is equal to the number of bins on the truth level.

We assume that the equation $m = \mathcal{R}t + \beta$ can be inverted [24]:

$$t = \mathcal{R}^{-1}(m - \beta) \quad (2.6)$$

The matrix inversion returns an unbiased result, because it is a simple linear transformation of the result.

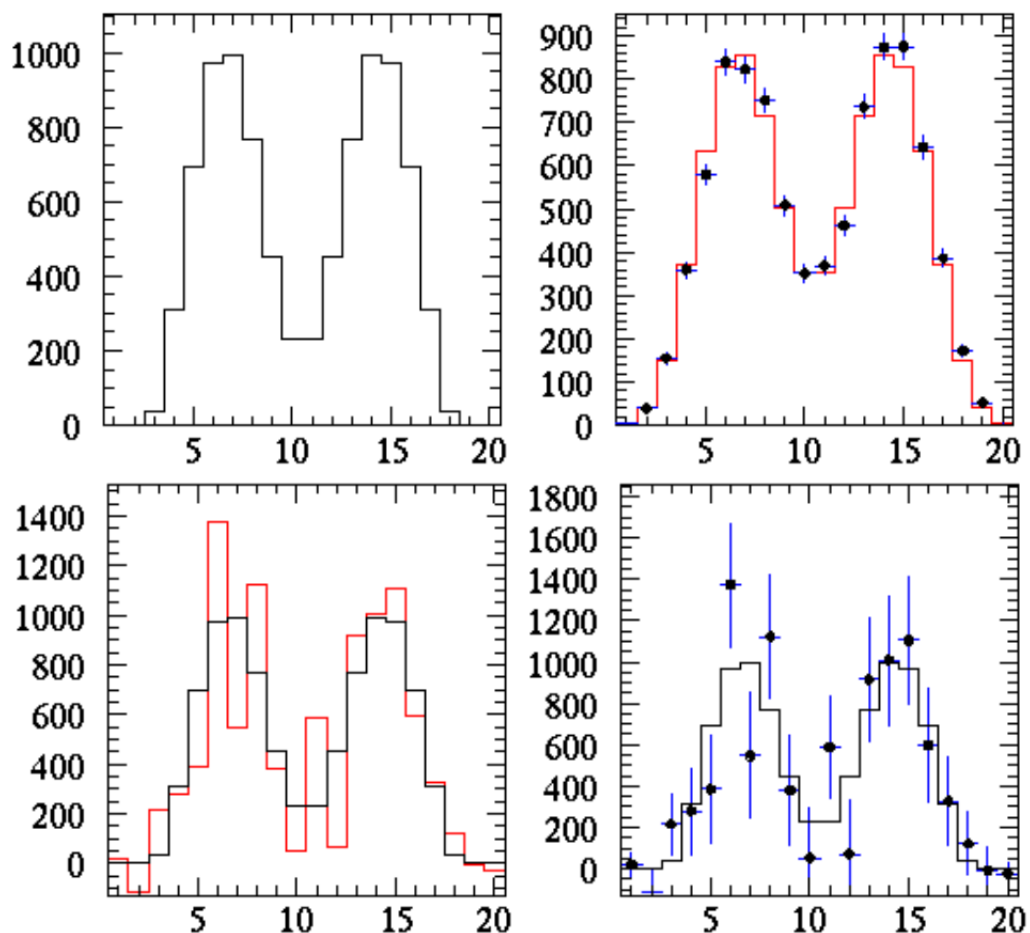


Figure 2.2: Unfolding using matrix inversion: The true distribution (upper left), smeared distribution and points are data drawn from this distribution (upper right), unfolded data from matrix inversion with true distribution (Bottom left), error bars for the unfolded data (Bottom right) [25].

2.3 Iterative Bayesian Unfolding Method

In 1994 GIULIO D'AGOSTINI proposed an iterative method based on *Bayes' theorem*, known in HEP as the *Iterative Bayesian Unfolding* (IBU). This technique gives approximations very close to that of the *true distributions*. The notable advantage of IBU over the *matrix inversion* is that the outcome is a *probability* (i.e., *normalization* and *non-negativeness*). This Bayes approach requires knowledge of the *initial distribution*, but D'AGOSTINI has overcome this problem by using an *iterative procedure* [26] [27] [28].

For an illustration, we start by stating the BAYES' THEOREM, using a picture of *causes* and *effects*, where multiple independent *causes* c_i ($i = 1, 2, 3, \dots, n_c$), which can produce a single *effect* (e_1). The *initial probability* of the causes $\mathbb{P}(c_i)$ is assumed to be known, as well as the *conditional probability* of the i^{th} cause to produce the effect $\mathbb{P}(e_1|c_i)$, then *Bayes formula* is written in the form [26]:

$$\mathbb{P}(c_i|e_1) = \frac{\mathbb{P}(e_1|c_i) \cdot \mathbb{P}(c_i)}{\sum_k^{n_c} \mathbb{P}(e_1|c_k) \cdot \mathbb{P}(c_k)}$$

where $\mathbb{P}(c_i|e_1)$ represents the probability that a single effect e_1 has been due to the i^{th} cause c_i . Moreover, many possible effects e_j ($j = 1, 2, \dots, n_e$) can be created from a given cause c_i , then the above formula can be generalized by the change $e_1 \rightarrow e_j$.

The causes correspond to the **true bin counts** before the *distortion* from *detector effects*, noted (t_i), and effects correspond to the **measured bin counts**, noted (m_j). Then the previous formula becomes

$$\mathbb{P}(t_i|m_j) = \frac{\mathbb{P}(m_j|t_i) \cdot \mathbb{P}(t_i)}{\sum_k^{n_t} \mathbb{P}(m_j|t_k) \cdot \mathbb{P}(t_k)} \quad (2.7)$$

From the *law of total probability*, we can write $\mathbb{P}(t_i)$ as

$$\mathbb{P}(t_i) = \sum_j \mathbb{P}(t_i \cap m_j) = \sum_j \mathbb{P}(t_i|m_j) \cdot \mathbb{P}(m_j)$$

where we used the definition of *conditional probability* : $\mathbb{P}(a|b) \stackrel{\text{def}}{=} \frac{\mathbb{P}(a \cap b)}{\mathbb{P}(b)}$; $\mathbb{P}(b) \neq 0$.

Therefore, if we replace the expression of $\mathbb{P}(t_i|m_j)$ in the one above, we will have

$$\mathbb{P}(t_i) = \sum_j \left(\frac{\mathbb{P}(m_j|t_i) \cdot \mathbb{P}(t_i)}{\sum_k^{n_t} \mathbb{P}(m_j|t_k) \cdot \mathbb{P}(t_k)} \right) \times \mathbb{P}(m_j) \quad (2.8)$$

The probabilities $\mathbb{P}(m_j|t_i)$ must be calculated or estimated with MONTE CARLO methods [26]. The desired solution $\mathbb{P}(t_i)$, is also **unknown**, so we have to use an estimate of it, therefore we can

write

$$\mathbb{P}^{n+1}(t_i) = \sum_j \left(\frac{\mathbb{P}(m_j|t_i) \cdot \mathbb{P}^n(t_i)}{\sum_k \mathbb{P}(m_j|t_k) \cdot \mathbb{P}^n(t_k)} \right) \times \mathbb{P}(m_j) \quad ; \quad n \in \mathbb{N} \quad (2.9)$$

This leads to an *iterative procedure*, where the choice commonly used to estimate the initial probability $\mathbb{P}^0(t_i)$ is the **uniform distribution**, so $\mathbb{P}^0(t_i) = 1/n_t$ [26] [28][20].

Recalling that the *response matrix* is written as $\mathcal{R}_{kl} = \mathbb{P}(m_k|t_l)$, and with $t_l = \mathbb{P}(t_l)$, $m_l = \mathbb{P}(m_l)$, then we can write :

$$t_i^{n+1} = \sum_j \left(\frac{\mathcal{R}_{ji} t_i^n}{\sum_k \mathcal{R}_{jk} t_k^n} \right) \times m_j \quad (2.10)$$

where n is the number of iterations. The unfolding procedure starts by choosing a prior *truth spectrum* $t_i^0 = \mathbb{P}^0(t_i)$ from the best knowledge of the process under study, otherwise, one can take $\mathbb{P}^0(t_i) = 1/n_t$. The number of iterations needed to converge depends on the desired precision. It is important to highlight that despite of the fact that the outcome is a probability, the unfolded outcome can have *negative entries*, because the *measured bin counts* may have *negative entries*, which results from *background subtraction*, so the result is *positive* only when $m \geq 0$ [20].

D'AGOSTINI mentioned in his paper 10 advantages of the IBU, in comparison with other unfolding methods [26]:

1. It is theoretically well-grounded.
2. Can be applied to *multidimensional* problems.
3. It can use cells of different sizes for the distribution of the true and the *experimental values*.
4. The domain of the definition of the *experimental values* can differ from that of the *true values*.
5. It can take into account *any kind of smearing* and *migration* from the *true values* to the *measured ones*.
6. In terms of its ability to reproduce the *true distribution*, it gives the *best* results if one makes a realistic guess about the distribution that the true values follow. But, in case of total ignorance, satisfactory results are obtained even starting from a *uniform distribution*.
7. Can take different sources of *background* into account.
8. Does not require *matrix inversion*.
9. Provides the *correlation matrix* of the results.

10. Can be implemented in a *short, simple* and *fast* program, which deals directly with distributions and not with *individual events*.

These advantages made the IBU a widely used method in HEP, as well as *statistical astronomy* in rectification and deconvolution problems [27]. The next step will be to implement it in *quantum information science* (QIS).

3

Unfolding Quantum Computer Readout Noise

Contents

3.1	Readout Errors	63
3.2	Constructing the Calibration Matrix \mathcal{R}	65
3.2.1	One Qubit Calibration Matrix	66
3.2.2	Two Qubit Calibration Matrix	67
3.3	Unfolding Methods for QIS	69
3.3.1	Unfolding with Matrix Inversion Method	69
3.3.2	Unfolding with IBU	72
3.4	Simulation of the Unfolding Process with Python	76
3.4.1	Noise Generation with Python	76
3.4.2	Constructing \mathcal{R}	80
3.4.3	Unfolding Methods	81
3.5	Unfolding Readout Errors from the Yorktown IBM Q Machine	86
3.5.1	Construction of the Response Matrix from the Yorktown IBM Q Machine	87
3.5.2	Unfolding a Uniform Distribution	92
3.5.3	Unfolding a Gaussian Distribution	101
3.6	Unfolding Readout Errors from The Melbourne IBM Q Machine	118

Quantum computers use the principle of superposition and entanglement to solve complex problems. But today’s quantum computers are noisy. Their are two sources of this noise “*qubit decoherence*” and “*gate infidelity*”. One important type of errors that arise from *qubit decoherence* are known as “**Readout Errors**”. These errors are similar to those caused by detector effects in HEP.

In this chapter we will introduce the so-called “*readout errors*”, where their effects can be represented by a matrix called “*calibration matrix*”. Then we try to use the unfolding methods from HEP to QIS. We also provide a simulation of an *unfolding process*, Finally, we mitigate readout noise through HEP unfolding methods in order to achieve the objective of this thesis.

3.1 Readout Errors

After performing any quantum computation we have to get an outcome, so we have to measure our system of qubits, which cause a collapse of the system state to either 0 or 1. But we know that the measurement operation will take a certain time. On the other hand, qubits have a *coherence time*, and which is *negligible* compared to the *measurement times*; this means that

during a measurement a qubit in the state $|0\rangle$ can decay to $|1\rangle$, and vice versa (i.e., a *bit-flip error*). In addition, the probability distributions of measured *physical quantities* corresponding to the states $|0\rangle$ and $|1\rangle$ have an *overlapping support* and there is a low probability of measuring the opposite value. This leads to biased outcomes, which therefore change our interpretation of these results [20][21]. We can investigate this fact directly, through the circuit below:

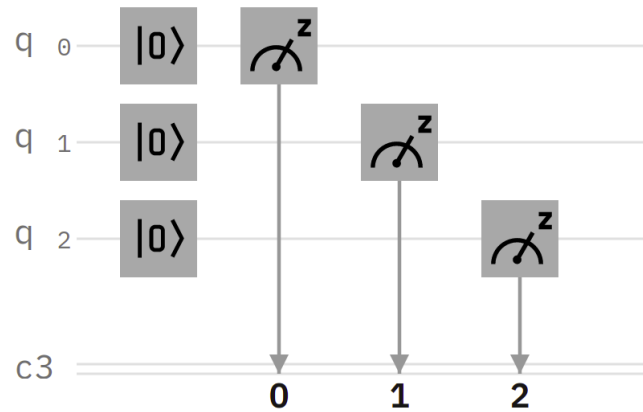


Figure 3.1: Measuring a 3-Qubit system of state $|000\rangle$.

In this circuit we have a system of 3-qubits initialized in the state $|000\rangle$, when performing a measurement, we expect to find them 100% in this state. But for the reasoning mentioned above, we will get the following values:

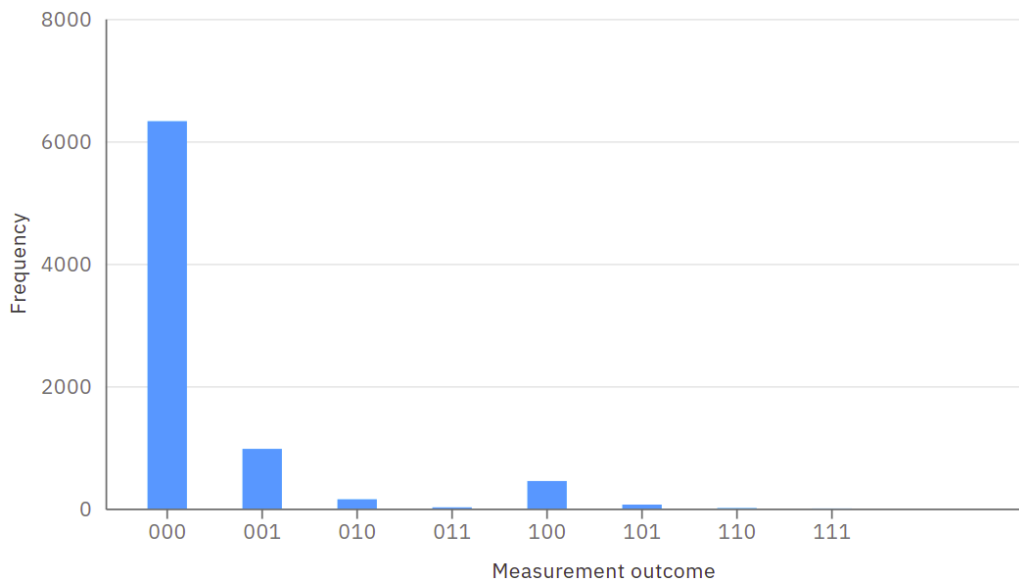


Figure 3.2: The Frequency of measuring each basis state from the IBM Q 5 YORKTOWN MACHINE, for 8000 shots.

We executed the circuit above in the IBM Q 5 YORKTOWN MACHINE that runs it 8000 times (the number of shots is proportional to the precision of the output probability distribution). We see in this histogram the frequency \mathcal{F} of measuring each basis state, where \mathcal{F} is proportional

to the *probability*, which achieves the following relation

$$\mathbb{P} = \frac{\mathcal{F}}{\mathcal{N}_{shot}}$$

where \mathcal{N}_{shot} is the total number of shots¹.

Due to the effects of noise, instead of having a probability of 1 to measure the state $|000\rangle$, we get $\mathbb{P}_{000} \simeq 0.7909$. This means that *bit-flips* have occurred during the measurement. For a *single bit-flip* on the state $|000\rangle$, we can get $|001\rangle$, $|010\rangle$, $|100\rangle$, with respective probabilities²

$$\mathbb{P}_{001} \simeq 0.1219, \quad \mathbb{P}_{010} \simeq 0.019, \quad \mathbb{P}_{100} \simeq 0.0563$$

If *two bit-flips* occurs, we get $|011\rangle$, $|101\rangle$, $|110\rangle$, with respective probabilities:

$$\mathbb{P}_{011} \simeq 0.0025, \quad \mathbb{P}_{101} \simeq 0.0081, \quad \mathbb{P}_{110} \simeq 0.0001$$

If *three bit-flips* occurs, we get $|111\rangle$, with the corresponding probability:

$$\mathbb{P}_{111} \simeq 25 \times 10^{-5}$$

The probability that more than a single *bit-flip* will occur is very low. For simple manipulations it is not significant, but for more complex manipulations (where we are in total ignorance of the results) it represents a challenge for interpreting results of experiments.

The methods of correcting readout errors will be discussed in [section 3.3](#). But before that, let's explore how to construct the central object of the unfolding process, the so-called "*calibration matrix*".

3.2 Constructing the Calibration Matrix \mathcal{R}

The calibration matrix is the essential tool in the mitigation of readout errors, which transforms true data into noisy data, so it is equivalent to the response matrix in HEP. This matrix is specific for each quantum computer; it depends directly on the hardware. Furthermore, for a given hardware, its elements changes every time due to *calibration drift*. The method often used to measure this matrix consist of constructing a set of quantum circuits called "*calibration circuits*", where for an n-qubit system, we must construct 2^n calibration circuit that generates all the corresponding computational basis states of the system, where this configurations are produced via the application of the X GATE [20][21].

To illustrate how the construction of the calibration matrix works we will take two instances. In the reste of this section, we will use the 5 YORKTOWN IBM Q MACHINE, and a number of 8000 of total shots.

¹The number of shots defines how many times a quantum circuit is executed in a quantum computer.

²The probabilities are not equal because the qubits are not **identical**; they have different *decoherence times*.

3.2.1 One Qubit Calibration Matrix

Let's have a system of one qubit which is initialized in the state $|0\rangle$, then we perform a measurement on it, as shown in the circuit below:

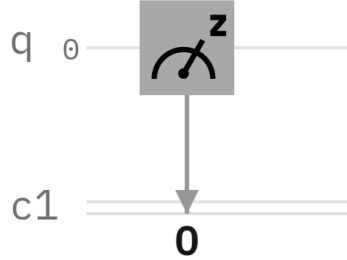
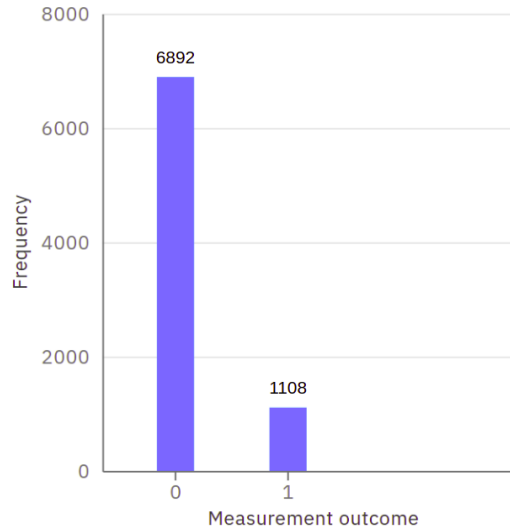


Figure 3.3: Measuring a qubit in the state $|0\rangle$.

After running this circuit in real hardware, we get the following outcomes:



As we said in the previous section, this qubit can decay to the state $|1\rangle$, during a measurement.

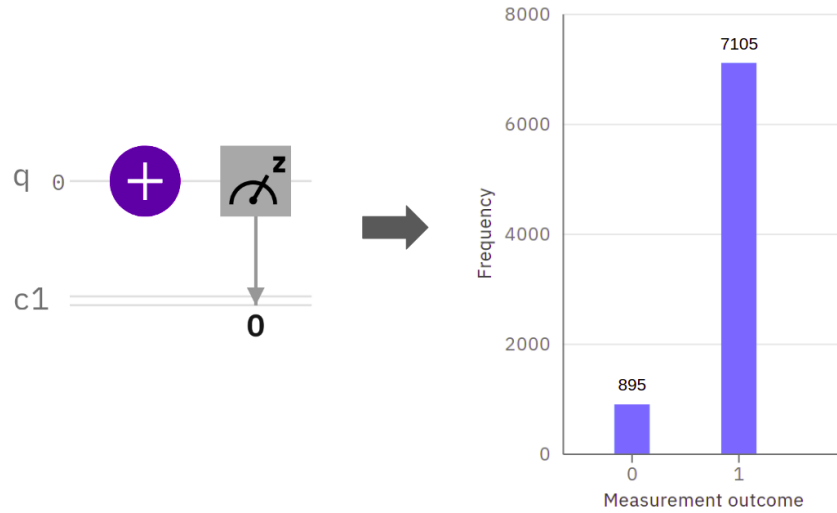
The calibration matrix \mathcal{R} is interpreted as a *conditional probability*, where the coefficient \mathcal{R}_{ij} represents the probability of measuring the value i under the condition that the true value is j

$$\mathcal{R}_{ij} = \mathbb{P}(\text{measured is } : i | \text{true is } : j) = \mathbb{P}(m = i | t = j)$$

Therefore, the above results can be translated into a column vector

$$\mathcal{R}_{i0} = \begin{pmatrix} \mathcal{R}_{00} \\ \mathcal{R}_{10} \end{pmatrix} = \begin{pmatrix} \mathbb{P}(0|0) \\ \mathbb{P}(1|0) \end{pmatrix} = \begin{pmatrix} \frac{6892}{8000} \\ \frac{1108}{8000} \end{pmatrix} = \begin{pmatrix} 0.861500 \\ 0.138500 \end{pmatrix}$$

The second step consists of building the circuit generating the other basis state $|1\rangle$, and which can be done by applying an X gate, before measuring the one-qubit system, as shown below:



These results represent the second column of the calibration matrix

$$\mathcal{R}_{i1} = \begin{pmatrix} \mathcal{R}_{01} \\ \mathcal{R}_{11} \end{pmatrix} = \begin{pmatrix} \mathbb{P}(0|1) \\ \mathbb{P}(1|1) \end{pmatrix} = \begin{pmatrix} \frac{895}{8000} \\ \frac{7105}{8000} \end{pmatrix} = \begin{pmatrix} 0.111875 \\ 0.888125 \end{pmatrix}$$

Finally, with these two columns, we can now construct the 2×2 calibration matrix of a single qubit system:

$$\mathcal{R}_{1Q} = \begin{pmatrix} \mathcal{R}_{00} & \mathcal{R}_{01} \\ \mathcal{R}_{10} & \mathcal{R}_{11} \end{pmatrix} = \begin{pmatrix} 0.86150 & 0.111875 \\ 0.13850 & 0.888125 \end{pmatrix} \quad (3.1)$$

If there was no noise, then we will have : $\mathbb{P}(0|0) = \mathbb{P}(1|1) = 1$ and $\mathbb{P}(1|0) = \mathbb{P}(0|1) = 0$, which leads to an identity matrix.

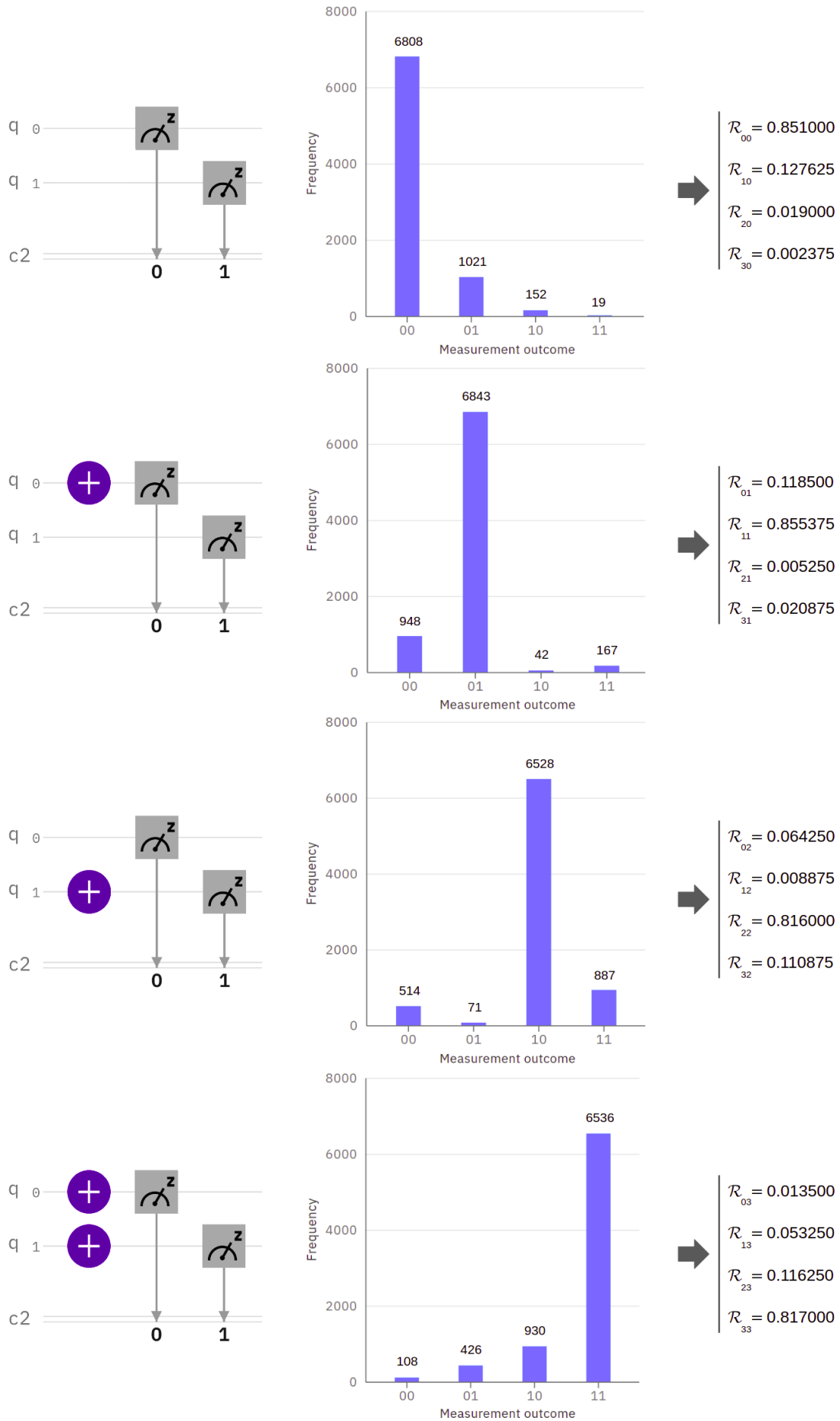
$$\mathcal{R}_{1Q} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{1}$$

3.2.2 Two Qubit Calibration Matrix

Let's have two qubits that are initialized in the state $|00\rangle$, when we measure them, instead of having a probability of 1 for the state $|00\rangle$, this probability will be distributed to the 2^2 possible computational basis states ($|00\rangle, |01\rangle, |10\rangle, |11\rangle$). To construct the calibration matrix we have to build 2^2 calibration circuits, where we use the X gate to prepare these 4 qubit configurations. After the readout, we get the corresponding column of each calibration circuits.

It is important to note that in the IBM QUANTUM COMPOSER, the convention used of the qubit order is that qubits start from the left to the right, which is the opposite of our convention, i.e., $|q_0 q_1\rangle \mapsto |q_1 q_0\rangle_{\text{IBM}}$. For brevity of notation, we will write the *binary* numbers i, j of the matrix elements \mathcal{R}_{ij} , in the *decimal* base, (i.e., $(00)_2 = (0)_{10}$, $(01)_2 = (1)_{10}$, $(10)_2 = (2)_{10}$, $(11)_2 = (3)_{10}$).

The figure below shows the calibration circuits and their measurement outcomes which correspond respectively to the computational basis states $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$:



$$\mathcal{R}_{2Q} = \begin{pmatrix} \mathbf{0.851000} & 0.118500 & 0.064250 & 0.013500 \\ 0.127625 & \mathbf{0.855375} & 0.008875 & 0.053250 \\ 0.019000 & 0.005250 & \mathbf{0.816000} & 0.116250 \\ 0.002375 & 0.020875 & 0.110875 & \mathbf{0.817000} \end{pmatrix} \quad (3.2)$$

Finally, we get the $2^2 \times 2^2$ calibration matrix, where we see that the correct results (located in the diagonal) have the greatest probability. As the number of qubits is fixed, the calibration matrix is always *square*. Moreover, it is reduced to an identity matrix $\mathcal{R}_{ij} = \delta_{ij}$, in the absence of noise.

In the next section, we will discuss how to correct readout errors (“*Unfolding*”), and show how it is done through the *calibration matrix*.

3.3 Unfolding Methods for QIS

Unfolding is a method of determining *approximately* the *true distribution* from the *measured disturbed distribution* (after distortions from readout noise). Several methods are used to achieve this approximation [29], and which are mainly based on knowledge of the calibration matrix (a.k.a., correction matrix). The most basic and commonly used method to correct readout errors is matrix inversion. In 2020, physicist BENJAMIN NACHMAN et al. investigated various unfolding methods currently used in HEP, in order to connect them with current methods used in the field of *quantum information science* (QIS), because current QIS methods have pathologies that can be avoided with HEP techniques. The *iterative Bayesian unfolding* method, introduced in [chapter 2](#), is shown to avoid *pathologies* from the *matrix Inversion* method [20].

By way of illustration, we will apply in what follows these two unfolding methods (MI, IBU), in a specific case. As we have built the YORKTOWN device-specific response matrix, we will continue to use it.³

3.3.1 Unfolding with Matrix Inversion Method

The Matrix Inversion is the most naive unfolding method, it assumes that the true bin counts (t) and its corresponding measured counts (m) are related by a calibration matrix \mathcal{R} as $m = \mathcal{R}t$. The additional assumption of this method is that \mathcal{R} is *regular*. Hence, it can be inverted to

³It is clear that each device has its specific response matrix; it makes no sense to unfold readout errors of the SANTIAGO device (for example) using the YORKTOWN device response matrix.

give us an *approximation* of the true counts, denoted by \hat{t}_{MI} .

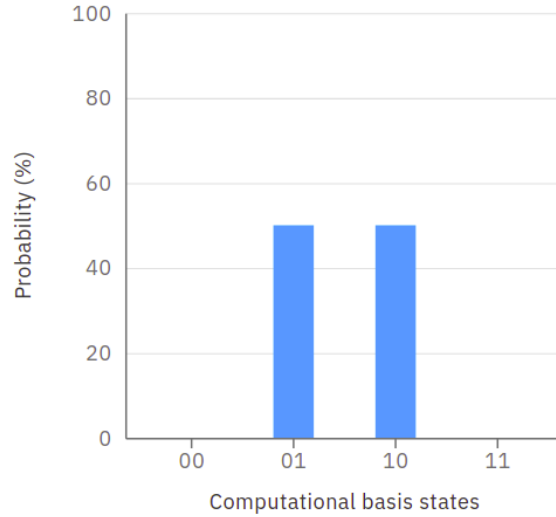
$$\hat{t}_{\text{MI}} = \mathcal{R}^{-1} m \quad (3.3)$$

To illustrate the unfolding procedure through this method, let's take the following example, where we want to mitigate the readout errors of two-qubit system generated in the specific Bell state $|\Psi^+\rangle$, written as:

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}} |01\rangle + \frac{1}{\sqrt{2}} |10\rangle$$

The probability of finding the two qubits in different states (i.e., $|01\rangle$, $|10\rangle$), and that of finding them in the same states (i.e., $|00\rangle$, $|11\rangle$), are given respectively:

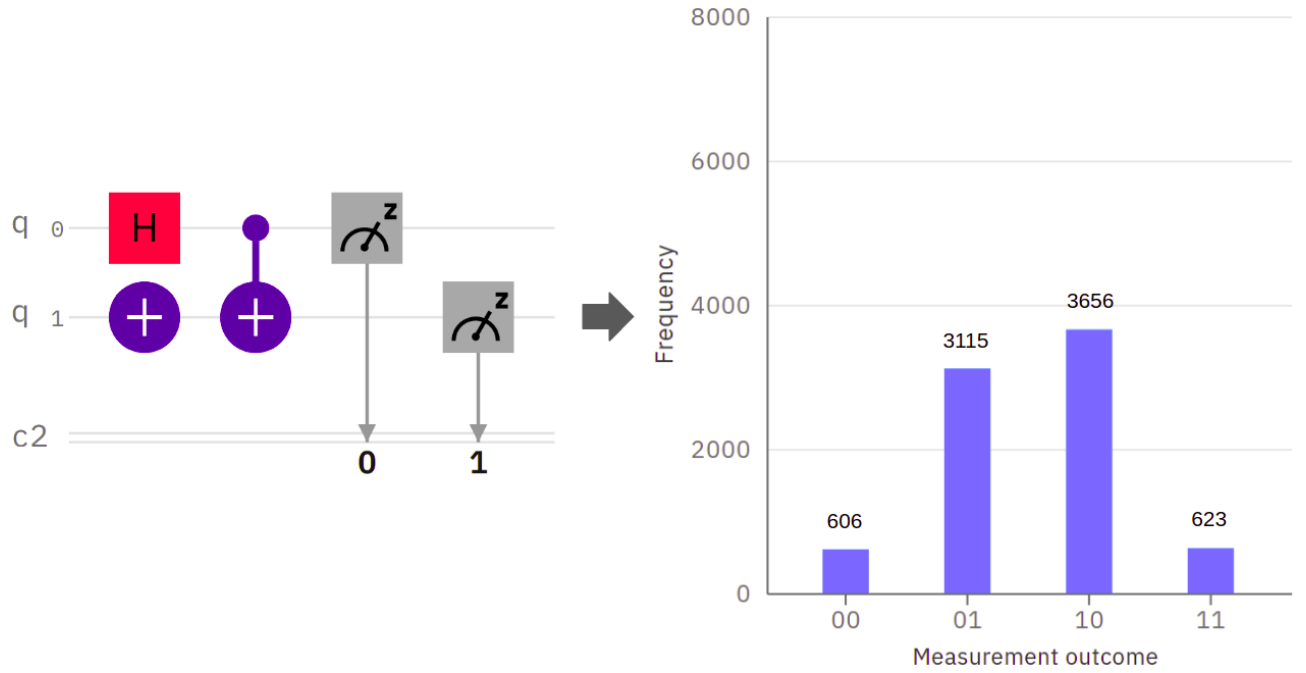
$$\mathbb{P}_{01} = \mathbb{P}_{10} = 0.5 \quad ; \quad \mathbb{P}_{00} = \mathbb{P}_{11} = 0$$



Hence, the vector representing the true bin counts can be expressed as:

$$t = \begin{pmatrix} 0 \\ 0.5 \\ 0.5 \\ 0 \end{pmatrix}$$

After building the circuit of $|\Psi^+\rangle$ and running it on the 5 YORKTOWN IBM Q MACHINE for 8000 shots, the following results are obtained:



As before, these measured bin counts can be translated into a column vector. But for a better observation of the changes, we express its element in terms of *frequency* instead of *probability*, the same thing for the true bin counts.

$$m = \begin{pmatrix} 606 \\ 3115 \\ 3656 \\ 623 \end{pmatrix} ; \quad t = \begin{pmatrix} 0 \\ 4000 \\ 4000 \\ 0 \end{pmatrix}$$

We have a two-qubit system, so, we are going to use \mathcal{R}_{2Q} , and all we have to do is simply to invert this matrix, in order to multiply it by the measured vector m . But before this, we have to check if \mathcal{R}_{2Q} is regular by calculating its determinant: $\det(\mathcal{R}_{2Q}) \simeq 0.46448 \neq 0$, it is therefore not singular, and its inverse is given below

$$\mathcal{R}_{2Q}^{-1} = \begin{pmatrix} +1.20206606 & -0.16605952 & -0.09341977 & +0.00425315 \\ -0.17936727 & +1.19572556 & +0.01152758 & -0.07661078 \\ -0.02752246 & +0.00046624 & +1.25179634 & -0.17769229 \\ +0.00482367 & -0.03013228 & -0.16990414 & +1.25004992 \end{pmatrix} \quad (3.4)$$

We can now calculate the approximation of the true bin counts t .

$$\hat{t}_{MI} = \mathcal{R}_{2Q}^{-1} m = \begin{pmatrix} -0127.716 \\ +3610.404 \\ +4450.638 \\ +0066.672 \end{pmatrix} \quad (3.5)$$

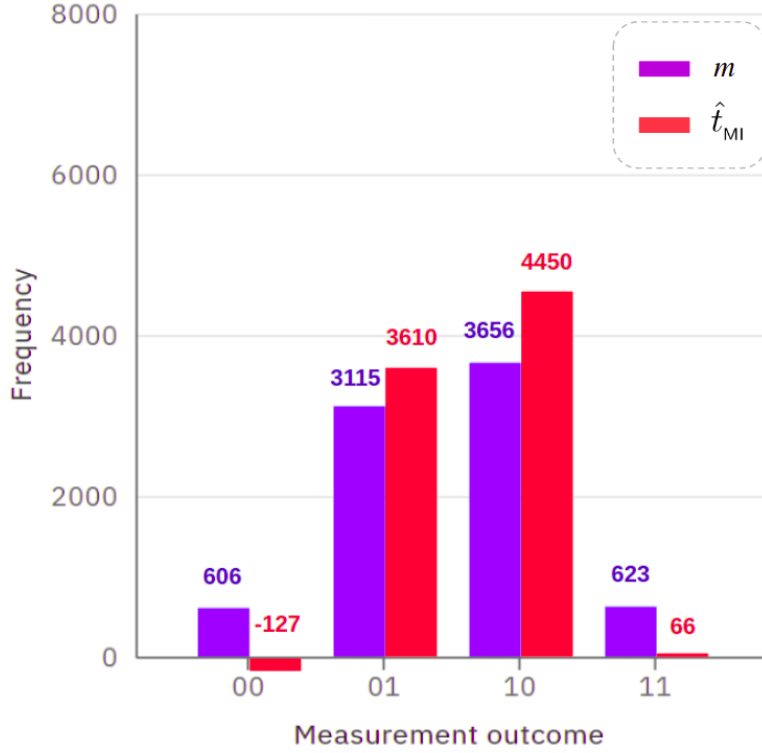


Figure 3.4: Measured and unfolded counts via the matrix inversion method.

After unfolding the readout noise, we notice that the frequencies of the computational basis state $|01\rangle$, $|10\rangle$ have increased, unlike the basis state $|00\rangle$, $|11\rangle$, where their frequencies has decreased, which is suitable. But, we also notice that for the state $|00\rangle$, the mitigation causes it to decrease to a negative frequency, which is not physical. But that happens just in some cases; for example, if one gets after a measurement $m = [750, 3100, 3550, 600]^T$, then the predicted counts will be

$$\hat{t}_{IBU} = [57.676, 3567.180, 4318.065, 57.077]^T$$

which is free from *unphysical outcomes*. But the fact remains that the opposite can be found.

3.3.2 Unfolding with IBU

The iterative Bayesian unfolding IBU, which was introduced in [chapter 2](#), it is an iterative technique, where the *approximate* (predicted) true counts vector \hat{t}_{IBU} is updated with each iteration in order to converge to the real counts t , recalling that the initial probability $t_i^0 = \mathbb{P}^0(\text{true value is } i)$, is often chosen as a *uniform distribution*. The IBU formula is written in the form

$$t_i^{n+1} = \sum_j \left(\frac{\mathcal{R}_{ji} t_i^n}{\sum_k \mathcal{R}_{jk} t_k^n} \right) m_j \quad ; \quad n \in \mathbb{N} \quad (3.6)$$

where n is the number of iterations.

We will take the same example of $|\Psi^+\rangle$, with the same outcomes in order to compare the two methods. But before that, let's take, for simplicity, an instance of a one qubit

system prepared in the state

$$|+\rangle = \mathbb{H} |0\rangle = \frac{1}{\sqrt{2}} |0\rangle + \frac{1}{\sqrt{2}} |1\rangle \quad (3.7)$$

The probabilities that the qubit is in the states ($|0\rangle$, $|1\rangle$) are given respectively:

$$\mathbb{P}_0 = \mathbb{P}_1 = 0.5$$

After running the corresponding circuit of $|+\rangle$ in the same device and with the same number of shots (8000), we get $m = [4151, 3849]^T$, while the true values are $t = [4000, 4000]^T$. As we have a one qubit system, we will use for the unfolding the matrix

$$\mathcal{R}_{1Q} = \begin{pmatrix} 0.86150 & 0.111875 \\ 0.13850 & 0.888125 \end{pmatrix}$$

The IBU formula gives us the *approximate components* of the true vector t , these components construct the predicted vector \hat{t}_{IBU} , which is a *two-dimensional* column vector in our case. By expanding the IBU equation for $(i, j, k = 0, 1)$, we obtain:

$$t_i^{n+1} = \sum_{j=0}^1 \left(\frac{\mathcal{R}_{ji} t_j^n}{\mathcal{R}_{j0} t_0^n + \mathcal{R}_{j1} t_1^n} \right) m_j \quad (3.8)$$

$$t_0^{n+1} = \left(\frac{\mathcal{R}_{00} t_0^n}{\mathcal{R}_{00} t_0^n + \mathcal{R}_{01} t_1^n} \right) m_0 + \left(\frac{\mathcal{R}_{10} t_0^n}{\mathcal{R}_{10} t_0^n + \mathcal{R}_{11} t_1^n} \right) m_1$$

$$t_1^{n+1} = \left(\frac{\mathcal{R}_{01} t_1^n}{\mathcal{R}_{00} t_0^n + \mathcal{R}_{01} t_1^n} \right) m_0 + \left(\frac{\mathcal{R}_{11} t_1^n}{\mathcal{R}_{10} t_0^n + \mathcal{R}_{11} t_1^n} \right) m_1$$

We can schematize this iteration procedure as follows:

$$\hat{t}_{\text{IBU}}^0 = \begin{pmatrix} t_0^0 \\ t_1^0 \end{pmatrix} \longrightarrow \hat{t}_{\text{IBU}}^1 = \begin{pmatrix} t_0^1 \\ t_1^1 \end{pmatrix} \longrightarrow \hat{t}_{\text{IBU}}^2 = \begin{pmatrix} t_0^2 \\ t_1^2 \end{pmatrix} \longrightarrow \dots \longrightarrow \hat{t}_{\text{IBU}}^{n+1} = \begin{pmatrix} t_0^{n+1} \\ t_1^{n+1} \end{pmatrix} \approx t$$

Since we express the elements of vectors t and m in terms of frequency, the *range* of the *uniform distribution* will be extended from $[0, 1]$ to $[0, 8000]$, and any value of $t_i^0 \in [0, 8000]$ hold for the iteration procedure, so we can choose

$$t_0^0 = 1 \quad ; \quad t_1^0 = 1$$

For $n = 0$:

$$t_0^1 = 4193.165 \quad ; \quad t_1^1 = 3806.834$$

In the 1st iteration these *new* values will replace the *old* ones $(t_0^0, t_1^0) \rightarrow (t_0^1, t_1^1)$. So, we have:

For $n = 1$:

$$t_0^2 = 4277.457 \quad ; \quad t_1^2 = 3722.542$$

For $n = 2$:

$$t_0^3 = 4314.417 \quad ; \quad t_1^3 = 3685.582$$

and so on

$$\vdots \quad \vdots \quad \vdots \quad ; \quad \vdots \quad \vdots \quad \vdots$$

For $n = 9$:

$$t_0^{10} = 4343.410 \quad ; \quad t_1^{10} = 3656.589$$

We notice that the values closest to the truth $t = [4000, 4000]^T$ are obtained for just **1** iteration, and the subsequent iterations diverge from t , where the values of t_0^n are increasing, while those of t_1^n are decreasing, which is the opposite of what we want.

We can now understand the unfolding example of the Bell state $|\Psi^+\rangle$ via the IBU technique. In this case, the vectors m , t and \hat{t}_{IBU} are *four-dimensional*, so we are going to use an *algorithm*, written with `python` language. In doing so, we found the following results for 1, 5, 10, 100, 1000 iterations, where we take the prior truth vector \hat{t}_{IBU}^0 as a uniform distribution.

$$\begin{aligned} \hat{t}_{\text{IBU}}^0 &= [001.000, 0001.000, 0001.000, 0001.000]^T \\ \hat{t}_{\text{IBU}}^1 &= [947.003, 2651.760, 3255.226, 1146.008]^T \\ \hat{t}_{\text{IBU}}^5 &= [187.168, 3366.011, 4151.338, 0295.481]^T \\ \hat{t}_{\text{IBU}}^{10^1} &= [065.849, 3468.337, 4302.088, 0163.724]^T \\ \hat{t}_{\text{IBU}}^{10^2} &= [000.001, 3524.170, 4400.027, 0075.801]^T \\ \hat{t}_{\text{IBU}}^{10^3} &= [10^{-50}, 3524.179, 4400.088, 0075.731]^T \end{aligned}$$

Note that this is the opposite of the previous example, where the \hat{t}_{IBU}^1 is not the closest vector to the truth t . But the iterations converges it more and more to t . Also, note that the amount of improvement decreases with each iteration. For our situation, the values stabilize for $n \geq 10^3$, i.e., subsequent iterations will give the same outcomes. Hence, $\hat{t}_{\text{IBU}}^{10^3}$ is the *maximum approximation* of t .

$$t = [000.000, 4000.000, 4000.000, 000.000]^T$$

Finally, we can represent the predicted values in a histogram in order to compare the two methods MI and IBU.

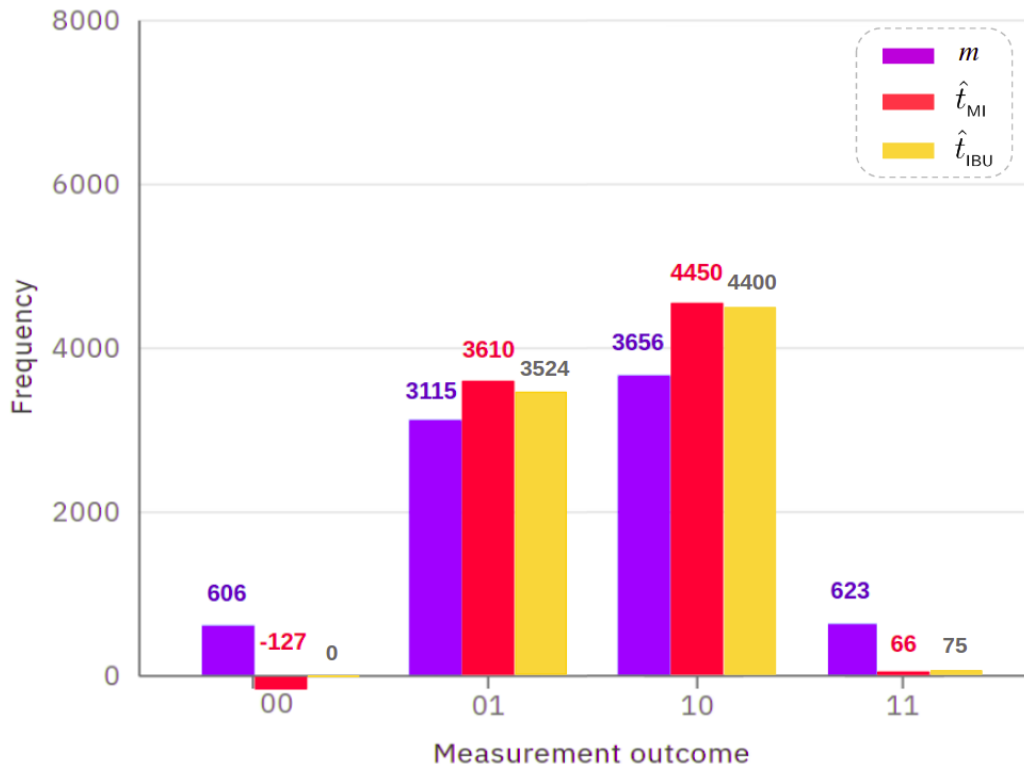


Figure 3.5: Measured and predicted bin counts via MI and IBU for 10^3 iterations.

We notice that in this specific example of $|\Psi^+\rangle$, the predicted values from matrix inversion method \hat{t}_{MI} and that from IBU \hat{t}_{IBU} , are very close. Furthermore, no negative outcomes for IBU, if we look at state $|00\rangle$, for IBU its error attenuation has reached 100%. The reason of the *non-negativeness* of the IBU outcomes, as we have mentioned before, is that the results are probabilities, which are always positive.

The difference in the behavior of IBU between the 1st and 2nd example comes from the choice of \hat{t}_{IBU}^0 , where in the 1st one, both \hat{t}_{IBU}^0 and t were uniform distributions. But for the 2nd example t was not uniform, unlike \hat{t}_{IBU}^0 . If we took \hat{t}_{IBU}^0 similar to t , e.g. $[0, 1, 1, 0]^T$, then we get:

$$\hat{t}_{\text{IBU}}^1 = [0, 3598.040, 4401.959, 0]^T$$

$$\hat{t}_{\text{IBU}}^{10} = [0, 3533.524, 4466.475, 0]^T$$

which shows the same behavior as in the 1st example. We see that it gives better results than before, for one iteration.

3.4 Simulation of the Unfolding Process with Python

In this section, we will explain how to simulate the unfolding process with python. To perform unfolding, we need the calibration matrix. So, in this simulation, we will assume that we know this matrix exactly and we will use it in the three unfolding methods: Matrix Inversion, IBU, and Ignis from [Qiskit](#).

It is interesting to see the effect of noise in data, so we will generate some noise using a pseudo-random generator provided by python tools, this will conduct us to simulate the measured data. Then, we will see how to use an arbitrary response matrix in the unfolding methods. The matrix will be constructed with python code, but we know precisely its values. Finally, we will compare the output of the tree unfolding methods with the measured and the true data in the histograms.

3.4.1 Noise Generation with Python

The purpose of this part is to produce noisy data using python tools. The first step consists of generating *pseudo-random numbers* for true data which follow the **normal law**, named `true_data`. In a second step, we generate pseudo-random numbers that follow the uniform law, named as `randnoise`. We will modify the `randnoise` by using conditions on it using true data the outcome is the reconstructed noisy data. To visualize the result, we will represent it in a histogram.

Python packages

Before starting, it is necessary to import some packages. We will need only two packages. The first is `numpy` which provides mathematical computations on arrays and matrices. The second is `matplotlib` for data visualization.

```
1 import numpy as np
2 import matplotlib.pyplot as plt
```

True Data

It is easier to visualize a little number of data so for the first example we will use $N = 100$. The true data is generated using the function `np.random.randn(N_data)*sigma`, the variable follow the normal law with `sigma = 3` the result will be clipped to the interval $[-10, 10]$ using the function `np.clip()` and will be named as `true_data`. The final step is to round the data to the nearest integer numbers by the function `round()`.

```
1 N_data = 100
2 sigma = 3
3 true_data = np.clip(np.random.randn(N_data)*sigma,-10.,10.)
4 true_data_int = [round(x) for x in true_data]
```

Now we can visualize the result by the print command as follow:

```
1 print("true_data =\n",true_data)
```

The output for the true_data:

```
1 true_data =
2 [ 5.36588542  1.30952955  0.2894924  -5.59047811 -0.83216461 -1.06427694
3 -0.24822444 -1.88100203 -0.13145451 -1.43165409 -3.94159426  2.65386714
4  2.64395413  5.12871919  0.15010093 -1.21403224 -1.63607984 -4.63943195
5  2.9471023  -3.30320289 -3.55513958 -0.6169497  4.45844507  0.7101488
6 -3.07135542 -2.1389796  1.8757349  -0.48154009 -2.30650905 -0.69009217
7  2.2351688  5.92833235 -3.73236999 -1.87925073 -2.41129828 -7.25724952
8 -2.77137607 -3.07162728  3.37193388 -0.3957427  -4.86985634  1.94002636
9 -1.06881228 -5.22942311 -1.78994893 -1.76578314 -2.62164689  0.08914145
10 -6.7447733  -0.80328559  3.03955033  2.55839352  3.3245625  3.35817197
11  4.4626294  -3.35490205  2.53750022 -5.58266859 -1.80865531 -5.74341613
12  3.14444254  4.00121346 -0.59224404  5.32393509 -2.02418253  0.4518506
13  0.45883711 -3.19258582  1.31383983  5.81693538 -3.07479262  2.69801534
14 -0.46352056  5.30888191  1.45136504  2.0286492  1.92948984  0.74726012
15 -4.18729051  4.17498872 -4.11200704  0.71568958  1.84223126 -2.51373682
16  0.43518964  3.50364686 -0.07231341 -2.66597225 -8.74721326 -2.91552151
17 -1.77323622 -1.5492521  -2.87998854  1.1318857  -1.72412526 -0.328363
18  2.0372148  -2.56631151 -0.90061822  6.47444803]
```

The output for the true_data_int:

```
1 true_data_int =
2 [5, 1, 0, -6, -1, -1, 0, -2, 0, -1, -4, 3, 3, 5, 0, -1, -2, -5, 3, -3, -4, -1,
3 4, 1, -3, -2, 2, 0, -2, -1, 2, 6, -4, -2, -2, -7, -3, -3, 3, 0, -5, 2, -1, -5,
4 -2, -2, -3, 0, -7, -1, 3, 3, 3, 3, 4, -3, 3, -6, -2, -6, 3, 4, -1, 5, -2, 0, 0,
5 -3, 1, 6, -3, 3, 0, 5, 1, 2, 2, 1, -4, 4, -4, 1, 2, -3, 0, 4, 0, -3, -9, -3, -2,
6 -2, -3, 1, -2, 0, 2, -3, -1, 6]
```

Noisy Data

In this part, we want to construct noisy data from true data, this will represent our measured data. Later, in the unfolding part, we will see how to correct this noisy data. So, now let's see how to generate some noise starting from the true data with python tools.

First, we generate N pseudo-random numbers (which represent the noise) that follows the uniform law. The seed is equal to 5 and the values of this data are in the interval $[0, 1]$.

```
1 np.random.seed(5)
2 randnoise = np.random.uniform(0.,1.,N_data)
```

The output for the randnoise:

```

1 randnoise =
2 [2.74780505e-01 6.52223125e-01 9.56449511e-01 4.35520556e-01
3 7.01325051e-02 5.77314878e-02 8.28710188e-02 9.59707187e-01
4 5.40760836e-01 8.37462433e-01 1.70033544e-01 2.60345073e-01
5 6.91977512e-01 8.95570328e-01 3.40688484e-01 6.46731980e-02
6 8.64119669e-01 2.90872446e-01 7.41082406e-01 1.58033655e-01
7 6.94963435e-01 8.41419619e-01 7.27152079e-01 3.59107525e-01
8 7.26689751e-01 1.39467124e-01 3.13819115e-01 4.19582757e-01
9 8.77212039e-01 1.53740209e-01 8.80124790e-01 7.98964319e-01
10 9.71624297e-01 3.67702983e-01 2.04939769e-01 2.40570320e-01
11 8.27862801e-01 9.65228149e-01 6.98809998e-01 4.82497042e-01
12 2.87049765e-01 8.33687884e-01 8.72179508e-01 9.21315918e-02
13 2.15949471e-01 8.31761090e-01 8.48303897e-01 3.14652999e-01
14 2.79294597e-01 4.30815022e-01 5.39446500e-01 9.55668150e-02
15 8.36912139e-01 5.34734870e-01 7.74967815e-01 2.30836266e-01
16 9.65293351e-01 7.51027307e-01 3.43093864e-01 9.48527647e-01
17 7.00511779e-01 8.40561085e-01 4.54973059e-02 5.56415411e-02
18 7.42737274e-01 3.04686433e-01 5.16784366e-01 1.56262424e-01
19 9.77952410e-01 5.02751048e-01 8.29001078e-01 7.40377963e-02
20 4.78915452e-01 6.22794804e-02 8.84241431e-01 4.45810179e-01
21 6.85499183e-02 7.64962823e-02 5.38792658e-01 7.55664045e-02
22 1.83772318e-01 4.36357084e-01 4.97782831e-01 5.83311915e-01
23 6.20512681e-01 3.72811500e-01 6.18736583e-01 1.57244658e-01
24 2.75508475e-01 7.98718271e-01 1.53089301e-01 2.23322973e-01
25 2.42978180e-01 4.79507305e-01 7.45522042e-04 3.03113605e-02
26 4.61548152e-01 1.62520685e-01 6.79501804e-01 7.95204587e-01]

```

The second step is to build new data using conditions on the real data and the random noise. That means, each value in the coordinate i will be added or subtracted by 1, this will depend on the condition between the random noise and the true value. The new reconstructed data is the measured data.

```

1 reco_data_int = []
2 for i in range(len(true_data_int)):
3     reco_data_int+=[true_data_int[i]]
4     if (true_data_int[i]==-10 and randnoise[i]<0.2):
5         reco_data_int[i]+=1
6     elif (true_data_int[i]==10 and randnoise[i]<0.3):
7         reco_data_int[i]-=1
8     elif (abs(true_data_int[i]) < 10 and randnoise[i]<0.2):
9         reco_data_int[i]-=1
10    elif (abs(true_data_int[i]) < 10 and randnoise[i]<0.8):
11        reco_data_int[i]+=1
12    pass
13 pass

```

Before seeing the result let us reproduce the Reconstructed data by hand for the five

values using the preview conditions.

i	0	1	2	3	4
True Data	5	1	0	-6	-1
Random Noise	2.74780505e-01	6.52223125e-01	9.56449511e-01	4.35520556e-01	7.01325051e-02
Reconstructed Data	5+1=6	1+1=2	0 (pass)	-6+1=-5	-1-1= -2

The output for the Reconstructed Noisy data:

```

1 reco_data_int =
2 [6, 2, 0, -5, -2, -2, -1, -2, 1, -1, -5, 4, 4, 5, 1, -2, -2, -4, 4, -4, -3, -1,
3 5, 2, -2, -3, 3, 1, -2, -2, 2, 7, -4, -1, -1, -6, -3, -3, 4, 1, -4, 2, -1, -6,
4 -1, -2, -3, 1, -6, 0, 4, 2, 3, 4, 5, -2, 3, -5, -1, -6, 4, 4, -2, 4, -1, 1, 1,
5 -4, 1, 7, -3, 2, 1, 4, 1, 3, 1, 0, -3, 3, -5, 2, 3, -2, 1, 5, 1, -4, -8, -2, -3,
6 -1, -2, 2, -3, -1, 3, -4, 0, 7]

```

The Plot

To visualize the effect of noise and the true data together, we will represent the result in two histograms. The true data t and the measured data m are discretized with 15 uniform unit width bins spanning the interval $[-10, 10]$.

The code that we use to draw the histograms is as follow:

```

1 m,b,_ = plt.hist(reco_data_int,bins=np.linspace(-10.5,10.5,16),facecolor="orange",
2                 edgecolor='none',label='m')
3 t,b,_ = plt.hist(true_data_int,bins=np.linspace(-10.5,10.5,16),facecolor="none",
4                 edgecolor='black',label="t")
5 plt.legend(frameon=False)
6 plt.xlabel('Value')
7 plt.ylabel('Counts')

```

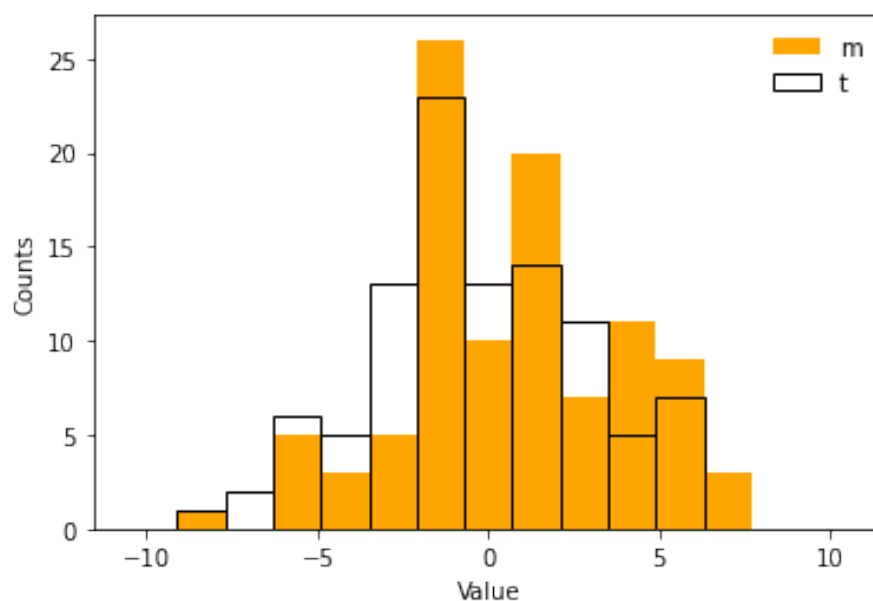


Figure 3.6: Histogram representation of simulated true data and measured data.

We can see that the noise changed the real values to the point where it is not possible to see the pattern of the distortion in the histogram. So, we will need to solve this inverse problem by applying the unfolding technique. For the next example, we will show only the histograms⁴. But first, let us see how to construct the response matrix with python.

3.4.2 Constructing R

We assume that we know our calibration matrix and will use the previews information about true and measured data. The dimension of m is equal to the dimension of t and they are related by the equation $m = Rt$.

The elements of the matrix are constructed with the following python code:

```

1 R = []
2 for i in range(len(m)):
3     R += [[]]
4     for j in range(len(t)):
5         if (i==0):
6             if (j==0):
7                 R[i] += [0.75]
8             elif (j==1):
9                 R[i] += [0.25]
10            else:
11                R[i] += [0.]
12        elif (i==len(t)-1):
13            if (j==len(t)-1):
14                R[i] += [0.75]
15            elif (j==len(t)-2):
16                R[i] += [0.25]
17            else:
18                R[i] += [0.]
19        else:
20            if (abs(i-j)==0):
21                R[i] += [0.5]
22            elif (abs(i-j)==1):
23                R[i] += [0.25]
24            else:
25                R[i] += [0.]

```

The index i takes values from 0 to $\text{dim}(m)$ and the index j takes values from 0 to $\text{dim}(t)$. In our case, the calibration matrix will have dimension 15×15 .

The output of the calibration matrix:

```

1 [[0.75, 0.25, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0],
2  [0.25, 0.5, 0.25, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0],
3  [0.0, 0.25, 0.5, 0.25, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0],
4  [0.0, 0.0, 0.25, 0.5, 0.25, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0],

```

⁴Because it is not practical to see data in dozens of pages.


```

5 [0.0, 0.0, 0.0, 0.25, 0.5, 0.25, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0],
6 [0.0, 0.0, 0.0, 0.0, 0.25, 0.5, 0.25, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0],
7 [0.0, 0.0, 0.0, 0.0, 0.0, 0.25, 0.5, 0.25, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0],
8 [0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.25, 0.5, 0.25, 0.0, 0.0, 0.0, 0.0, 0.0],
9 [0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.25, 0.5, 0.25, 0.0, 0.0, 0.0, 0.0],
10 [0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.25, 0.5, 0.25, 0.0, 0.0, 0.0],
11 [0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.25, 0.5, 0.25, 0.0, 0.0],
12 [0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.25, 0.5, 0.25, 0.0],
13 [0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.25, 0.5, 0.25],
14 [0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.25, 0.5, 0.25],
15 [0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.25, 0.75]]

```

We can also introduce each value of a calibration matrix by hand like the following example:

```

1 R = [[0.75, 0.25, 0.00, 0.00],
2       [0.25, 0.50, 0.25, 0.00],
3       [0.00, 0.25, 0.50, 0.25],
4       [0.00, 0.00, 0.25, 0.75]]

```

3.4.3 Unfolding Methods

In this part, we will use three unfolding methods to correct the measured data. Each unfolding method can be defined as a function in `python`. Here, we will change the number of data `N_data` (from 100 to 10000) and the number of bins (from 15 to 21) and also the measured data (Reconstructed Data). All the other parameters from the previews code will remain unchanged. The new histograms of measured and true data are represented as follow:

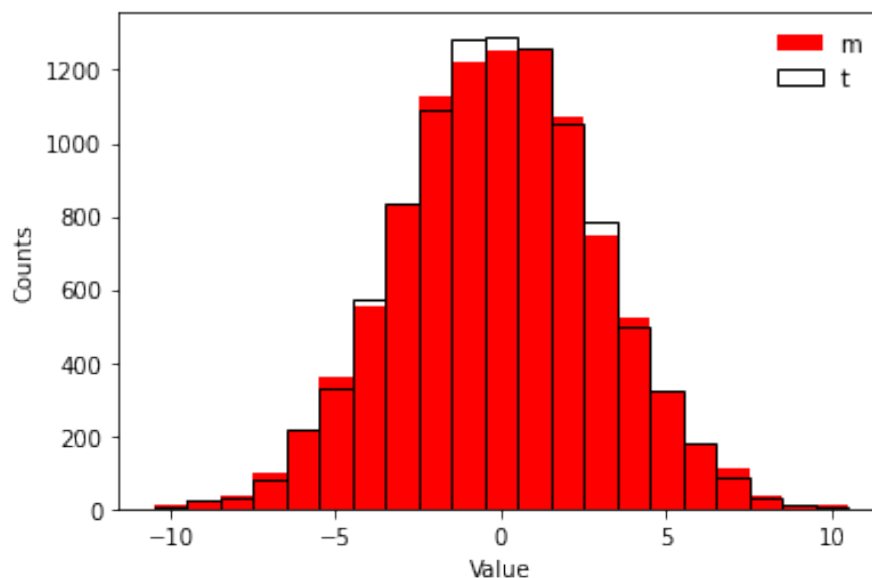


Figure 3.7: Histogram representation of simulated true data and measured data.

Matrix Inversion

One of the methods to unfold the true distribution from experimental data is to calculate the inverse of the response matrix, then we will get the true distribution using the following equation $\hat{t} = R^{-1}m$.

To calculate the inverse of a matrix in `python` we use `np.linalg.inv()` function. The multiplication between the inverse of the matrix and the measured distribution vector is performed by the function: `np.matmul()`. The `np.ravel` return a contiguous flattened 1-D array, containing all the elements of the input.

```

1 def MatrixInversion(y_mes,Rin):
2     return np.ravel(np.matmul(np.linalg.inv(np.matrix(Rin)),y_mes))

```

Ignis Function

The ignis function uses the Sequential Least Squares Programming (SLSQP) method. This code can be found in `qiskit-ignis`⁵.

```

1 def fun(x,y_mes,Rin):
2     mat_dot_x = np.ravel(np.matmul(Rin,x))
3     return sum((y_mes - mat_dot_x)**2)
4
5 def Ignis(y_mes,Rin):
6     x0 = np.random.rand(len(y_mes))
7     x0 = x0 / sum(x0)
8     nshots = sum(y_mes)
9     cons = ({'type': 'eq', 'fun': lambda x: nshots - sum(x)})
10    bnds = tuple((0, nshots) for x in x0)
11    res = minimize(fun, x0, method='SLSQP',constraints=cons,
12                 bounds=bnds, tol=1e-6,args=(y_mes,Rin))
13    return res.x

```

IBU Function

The iterative Bayesian unfolding function is defined in this code as IBU. Its input parameters are the measured data `y_mes` and the `t0` is an initial truth spectrum of the process, in our case, it is an array of ones. The third parameter is the response matrix, and finally `n` is the number of iterations. The lines from 3 to 9 are for initialization. The IBU method start in line ten. The code for the IBU function is as follow:

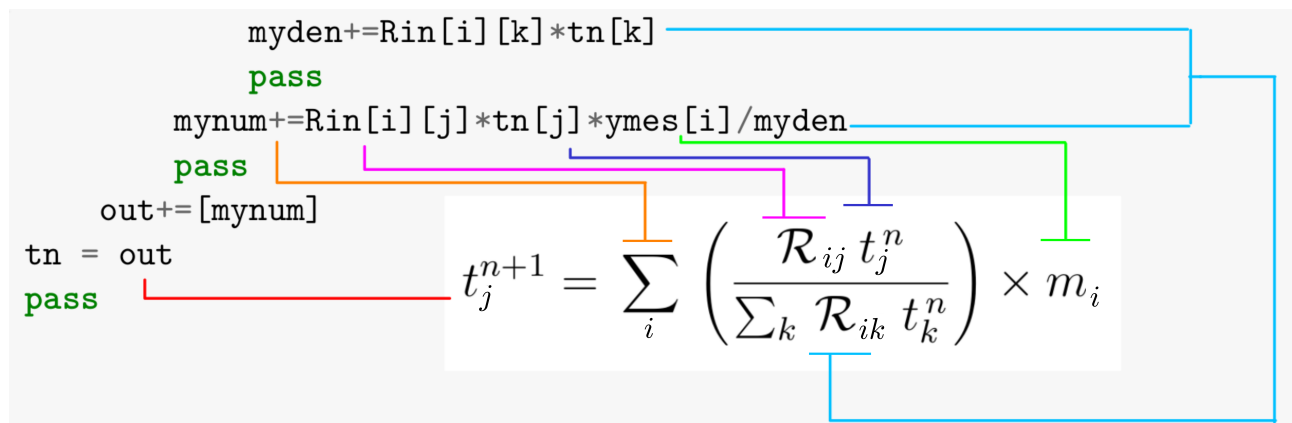
⁵https://github.com/Qiskit/qiskit-ignis/blob/master/qiskit/ignis/mitigation/measurement_filters.py

```

1 def IBU(ymes,t0,Rin,n):
2
3     tn = t0
4     for q in range(n):
5         out = []
6         for j in range(len(t0)):
7             mynum = 0.
8             for i in range(len(ymes)):
9                 myden = 0.
10                for k in range(len(t0)):
11                    myden+=Rin[i][k]*tn[k]
12                    pass
13                mynum+=Rin[i][j]*tn[j]*ymes[i]/myden
14                pass
15            out+=[mynum]
16        tn = out
17        pass
18
19    return tn

```

We recall the expression of this unfolding method:



Now, let us unfold by including the measured data and the calibration matrix to our functions. The matrix inversion and ignis have two parameters \mathbf{m} and \mathbf{R} . For IBU it has four parameters: first the measured data \mathbf{m} , the initial truth data $\mathbf{t0}$ that we choose it as a vector of elements 1 (*uniform distribution*) and which have the same dimension as the measured data, the response matrix, and finally the number of iterations $n=10$.

```

1 MI    = MatrixInversion(m, R)
2 ignis = Ignis(m, R)
3 ibu   = IBU(m, np.ones(len(m)), R, 10)

```

To see the result of the three unfolding methods, we use the `print` function.

Matrix Inversion

```

1 print(MI)
2
3 [2.36190476e+01 -3.48571429e+01  1.50095238e+02 -1.21333333e+02
4  4.80571429e+02  1.90476190e-01  9.67047619e+02  2.89714286e+02
5  1.77752381e+03  6.71238095e+02  1.77600000e+03  7.88761905e+02
6  1.65447619e+03  1.82285714e+02  9.60952381e+02 -2.01904762e+01
7  3.83428571e+02 -2.26666667e+01  1.01904762e+02 -3.71428571e+01
8  2.83809524e+01]
```

Ignis

```

1 print(ignis)
2
3 [4.6175683800  21.291404430  59.2658346200  0.          334.52132545
4  165.440539820  787.74328185  478.197504920  1584.36645844  864.83510512
5  1585.82646343  971.91899933  1481.55800967  342.012482900  816.99820065
6  105.680456820  277.58084116  61.4881026600  40.7416841000  0.
7  15.9157361800]
```

IBU

```

1 print(ibu)
2
3 [5.259372656831409, 25.388591380196175, 33.557337461054985,
4  67.96709437048914, 208.8410238062819, 356.97817161684395,
5  521.2370635355416, 822.9011004066444, 1193.1006184446076,
6  1246.8670968072922, 1232.6156342943655, 1329.7638967204725,
7  1095.5138161773816, 719.1611893853731, 514.3011592136613,
8  299.016654127236, 181.16794027605107, 100.54732163607031,
9  23.88223196059284, 9.002561105297106, 12.930124617715322]
```

The True Data

```

1 bins=np.linspace(-10.5,10.5,22)
2 np.histogram(true_data,bins)
3
4 (array([ 5, 23, 32, 78, 219, 329, 571, 837, 1089, 1287, 1292,
5         1260, 1052, 783, 500, 324, 182, 83, 32, 14, 8]),
6  array([-10.5, -9.5, -8.5, -7.5, -6.5, -5.5, -4.5, -3.5, -2.5,
7         -1.5, -0.5, 0.5, 1.5, 2.5, 3.5, 4.5, 5.5, 6.5,
8         7.5, 8.5, 9.5, 10.5]))
```

The plot of the truth and measured data with the result of the three unfolding methods.

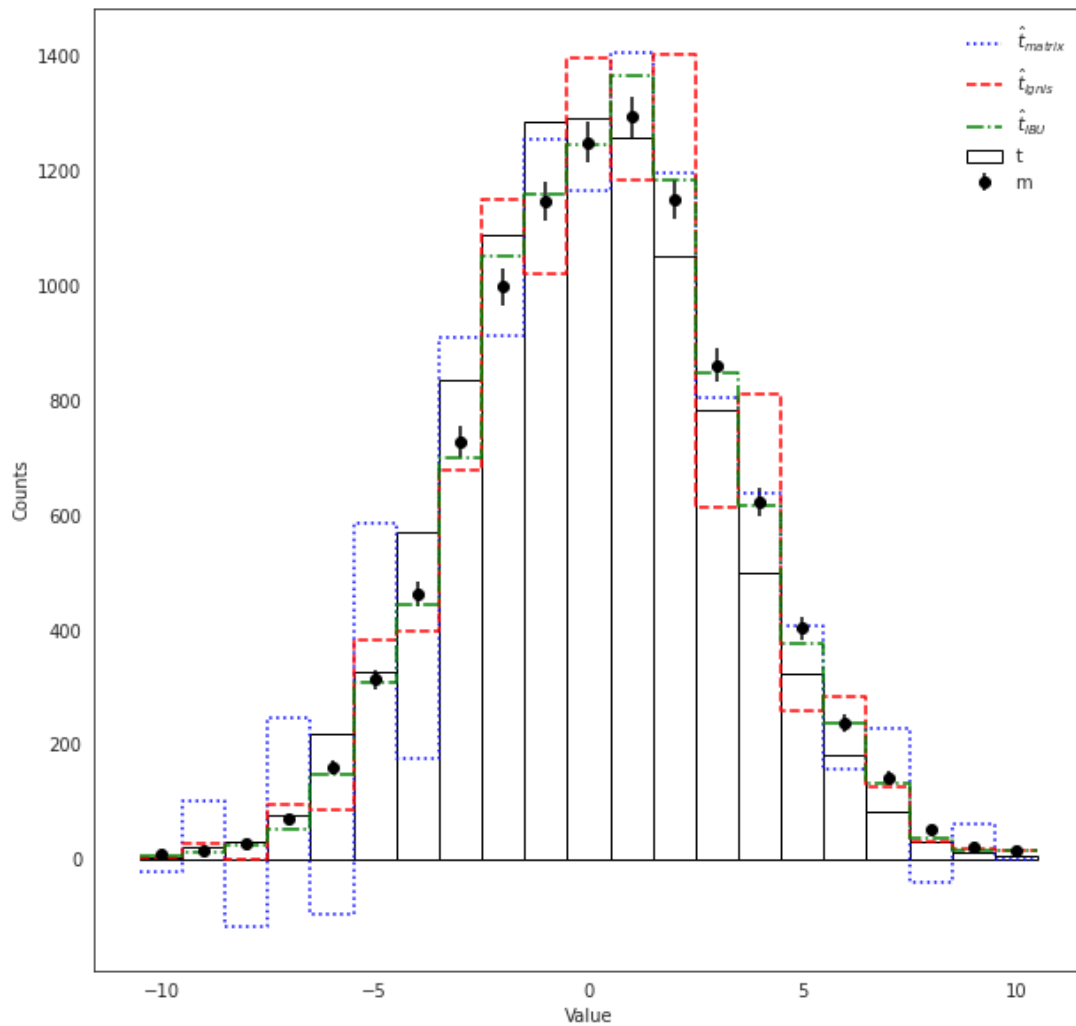


Figure 3.8: Unfolding results of a Gaussian distribution using three methods: Matrix inversion (\hat{t}_{matrix}), ignis (\hat{t}_{ignis}), and IBU (\hat{t}_{IBU}) with 10 iterations and a uniform prior truth spectrum.

We can clearly see that the IBU results are close to the real spectrum compared to matrix inversion and ignis. We can also see that the results of ignis and matrix inversion show large oscillations. We also notice that matrix inversion give negative outcomes, while IBU does not encounter this problems. Finally, we can deduce that the IBU leads us to better results compared to the other two methods in this simulation.

3.5 Unfolding Readout Errors from the Yorktown IBM Q Machine

The central objective of the article [20] published in 2020 by BENJAMIN NACHMAN et al., is to provide a more powerful technique to mitigating readout errors in the field of QIS. This technique is the *iterative Bayesian unfolding* IBU. They studied three techniques: **matrix inversion**, **ignis**, and **IBU** in order to compare them. This study is based on **Qiskit Simulator**⁶, to mitigate readout errors for 5 qubits generated in a gaussian distribution. The unfolding was done through a **realistic** response matrix (from the JOHANNESBURG IBM Q Machine), and also for an **artificial** response matrix, and they demonstrated the advantages of the IBU over other methods.

But it is very interesting to explore the behavior of unfolding techniques, through **real noise** coming from a **real quantum computer**. The *International Business Machines corporation* (IBM) provide us with 9 devices that we can use. In the following, we will be using the YORKTOWN IBM Q machine (a.k.a. IBMQX2). The reason we chose it is because it has the highest value of readout errors (average 5.826×10^{-2}).

We will start our study with a *uniform distribution* as the true distribution, then we will tackle the *Gaussian distribution*, both are based on 5 qubits. The number of shots taken in all experiments is fixed at 8192⁷.

Another interesting question we are trying to answer is how do these unfolding methods behave for a system beyond 5-qubits. We will use for this purpose the 16 MELBOURNE IBMQ machine, which is a 15-Qubit quantum computer, we will study the unfolding a 7-qubit system, in the specific case of a *uniform distribution*. In this situation, we are dealing with readout noise described by a $2^7 \times 2^7$ dimensional *response matrix*, which represents a challenge for any *unfolding technique*.

Our work can be schematized in 4 steps:

1. Constructing the response matrix \mathcal{R} .
2. Generating the desired distribution to be the truth.
3. Distortion of the true distribution by readout noise.
4. Unfolding readout errors via MI, Ignis, IBU.

⁶A local simulation is provided by Aer. In the article [20] they used the so-called “qasm_simulator”.

⁷This is the maximum number allowed.

3.5.1 Construction of the Response Matrix from the Yorktown IBM Q Machine

In this part, we will explain how to extract a **calibration matrix** from a quantum computer. In our work, we will use the IBM 5 YORKTOWN system. This system contains 5 qubits, and we will use all the five qubits to build this matrix. So the calibration matrix we want to build has dimension $2^5 \times 2^5$. The process will be done in IBM QUANTUM LAB⁸. The response matrix will be applied later to unfold distorted distributions by readout noise.

Now, let us explore how to construct the calibration matrix of the IBM Yorktown system step by step.

Step 1: Import the packages

The first step as always is to import the packages that we will need in our code. When we open a new file in the IBM QUANTUM LAB, a list of packages will be available in the notebook.

```
1 import numpy as np
2 # Importing standard Qiskit libraries
3 from qiskit import QuantumCircuit, transpile, Aer, IBMQ
4 from qiskit.tools.jupyter import *
5 from qiskit.visualization import *
6 from ibm_quantum_widgets import *
7 # Loading your IBM Quantum account(s)
8 IBMQ.load_account()
```

We will add additional packages as we need more qiskit modules to construct the calibration matrix. The imported modules are qiskit ignis, and the measurement calibration for the complete measurement calibration. The execute module is to execute the circuit, matplotlib and seaborn to visualize the results.

```
1 import qiskit
2 from qiskit import QuantumRegister
3 # Execute a list of circuits or pulse schedules on a backend
4 from qiskit import execute
5 # Measurement calibration modules.
6 from qiskit.ignis.mitigation.measurement import complete_meas_cal
7 from qiskit.ignis.mitigation.measurement import CompleteMeasFitter
8 # For data visualization
9 import matplotlib.pyplot as plt
10 import seaborn as sns
```

⁸<https://quantum-computing.ibm.com>

Now let us see how to construct the *calibration circuits* with 5-qubits for the ibmq Yorktown system.

Step 2: Construct the calibration circuits

The first step after importing the packages is to access to quantum devices by a provider. All IBMQ providers are specified by a [hub](#), group, and project. To access a given provider we should use the `get_provider()` function of the IBMQ account. In our case, our provider gives us access to the public IBM Quantum devices available to all IBM quantum users.

```
1 provider = IBMQ.get_provider(hub='ibm-q')
2 # List of all available providers
3 IBMQ.providers()
```

To see all available quantum computers, we use the following function:

```
1 provider.backends(simulator = False)
```

We get the whole list of systems without simulators:

```
1 [<IBMQBackend('ibmqx2') from IBMQ(hub='ibm-q', group='open', project='main')>,
2  <IBMQBackend('ibmq_16_melbourne') from IBMQ(hub='ibm-q', group='open',
3  ↪ project='main')>,
4  <IBMQBackend('ibmq_armonk') from IBMQ(hub='ibm-q', group='open', project='main')>,
5  <IBMQBackend('ibmq_athens') from IBMQ(hub='ibm-q', group='open', project='main')>,
6  <IBMQBackend('ibmq_santiago') from IBMQ(hub='ibm-q', group='open',
7  ↪ project='main')>,
8  <IBMQBackend('ibmq_lima') from IBMQ(hub='ibm-q', group='open', project='main')>,
9  <IBMQBackend('ibmq_belem') from IBMQ(hub='ibm-q', group='open', project='main')>,
10 <IBMQBackend('ibmq_quito') from IBMQ(hub='ibm-q', group='open', project='main')>,
11 <IBMQBackend('ibmq_manila') from IBMQ(hub='ibm-q', group='open', project='main')>]
```

Now we will generate a list of measurement calibration circuits for the full Hilbert space. Each circuit creates a basis state. For 5 qubits, we get 2^5 calibration circuits.

```
1 qr = qiskit.QuantumRegister(5)
2 qubit_list = [0,1,2,3,4]
3 meas_calibs, state_labels = complete_meas_cal(qubit_list, qr, circlabel='mcal')
```

ibmq_5_yorktown system (ibmqx2) contains five qubits, and we want to use all the qubits to build our calibration matrix. So, we need five quantum registers. The qubit list is a list of qubits to perform the measurement on, and the last function return a list of objects containing the calibration circuits, and a list of the calibration state labels.

Step 3: Calculate the Calibration Matrix

Now, we can execute the calibration circuits, we choose the `ibmq_5_yorktown` backend to run our circuit, and we specify the maximum number of shots⁹.

```
1 backend = provider.get_backend('ibmq_5_yorktown')
2 job = qiskit.execute(meas_calibs, backend, shots=8192)
```

To get the result from the `job`, we use the `result()` function:

```
1 cal_results=job.result()
2 counts=cal_results.get_counts()
```

After we run the calibration circuits and obtain the results `cal_results`, we can construct the calibration matrix and it will be ordered according to the `state_labels`.

```
1 meas_fitter = CompleteMeasFitter(cal_results, state_labels, circlabel='mcal')
2 R= meas_fitter.cal_matrix
3 print(R)
```

Finally, our response matrix for the IBM Yorktown quantum computer is as follows:

```
1 array([[7.23632812e-01, 1.22558594e-01, 3.18603516e-02, ...,
2         2.44140625e-04, 0.00000000e+00, 0.00000000e+00],
3        [7.58056641e-02, 7.46582031e-01, 3.05175781e-03, ...,
4         1.46484375e-03, 0.00000000e+00, 1.22070312e-04],
5        [3.51562500e-02, 6.22558594e-03, 7.75512695e-01, ...,
6         0.00000000e+00, 9.76562500e-04, 0.00000000e+00],
7        ...,
8        [0.00000000e+00, 0.00000000e+00, 0.00000000e+00, ...,
9         6.40136719e-01, 4.63867188e-03, 3.78417969e-02],
10       [0.00000000e+00, 0.00000000e+00, 0.00000000e+00, ...,
11        3.54003906e-03, 6.21093750e-01, 8.39843750e-02],
12       [0.00000000e+00, 0.00000000e+00, 0.00000000e+00, ...,
13        1.48925781e-02, 8.98437500e-02, 6.16088867e-01]])
```

To better visualize the result we will use the code provided with the article to represent the matrix in the form of a heatmap.

Step 4: Heatmap Representation of the Response Matrix

First, we need some functions that will help us to represent the heatmap.

The following code allows us to construct the state labels for the heatmap. Let us print the output to better understand what this function is doing.

⁹How many times an algorithm is run to get a probability distribution of results.

```

1 nqbits = 5
2 def mybin(input,q,r=1):
3     if (r==0):
4         return bin(input).split('b')[1].zfill(q)
5     else:
6         return bin(input).split('b')[1].zfill(q)[::-1]
7 xvals = []
8 xlabs = []
9 for i in range(2**nqbits):
10    xvals+=[i+0.5]
11    xlabs+=[r+'$|'+mybin(i,nqbits)+r+'\rangle$']
12    pass

```

The `bin()` function returns the binary version of a specified integer. The result will always start with the prefix `0b`. For example, if `input=7` the result will be as follow:

```
1 '0b111'
```

We do not need the prefix `0b` so we will divide the result and we will retain the second part in position 1 (position 0 is for the first part `0b`)

```
1 bin(7).split('b')[1]
```

And the result is as follows:

```
1 '111'
```

And the `zfill` will fill the string with zeros until it is (`q`) characters long, for example, `zfill(5)` will give us:

```
1 bin(7).split('b')[1].zfill(5)
```

That will give us:

```
1 '00111'
```

Finally, let us see how the `mybin()` function works. For `input=7`, `q=5`, and `r=1` and `r=0`:

```

1 mybin(7,5,0)
2 mybin(7,5,1)

```

We get the following result:

```
1 '11100'
2 '00111'
```

The following code allows us to generate labels for the heatmap:

```
1 response_normalized = np.zeros((2**nqbits,2**nqbits))
2 response_normalized_reversed = np.zeros((2**nqbits,2**nqbits))
3 response_labels = np.zeros((2**nqbits,2**nqbits))
4 for i in range(2**nqbits):
5     mysum = 0.
6     for j in range(2**nqbits):
7         mysum+=meas_fitter.cal_matrix[j,i]
8         pass
9     for j in range(2**nqbits):
10        response_labels[j][i] = '0'
11        if (mysum > 0):
12            response_normalized[j,i] = 100.*meas_fitter.cal_matrix[j,i]/mysum
13            response_normalized_reversed[i,j] = response_normalized[j,i]
14            response_labels[j][i] = '%1.0f' % (response_normalized[j,i])
15        pass
16    pass
17    pass
```

Finally, we used this code to represent the response matrix for the Yorktown quantum computer in heatmap representation.

```
1 f = plt.figure(figsize=(20, 20))
2 sns.set(font_scale=1.5)
3 ax = sns.heatmap(R,annot=response_labels,annot_kws={"size": 15}
4 ,cmap="rocket_r")
5
6 ax.figure.axes[-1].yaxis.label.set_size(18)
7 cbar = ax.collections[0].colorbar
8 cbar.set_label('Pr(Measured | True) [%]', labelpad=30)
9
10 plt.xticks(xvals,xlabs,rotation='vertical',fontsize=15)
11 plt.yticks(xvals,xlabs,rotation='horizontal',fontsize=15)
12 plt.xlabel('True',fontsize=18)
13 plt.ylabel('Measured',fontsize=18)
14
15 plt.subplots_adjust(bottom=0.18,left=0.18)
16 plt.savefig("York_Calibration.pdf",bbox_inches='tight')
```

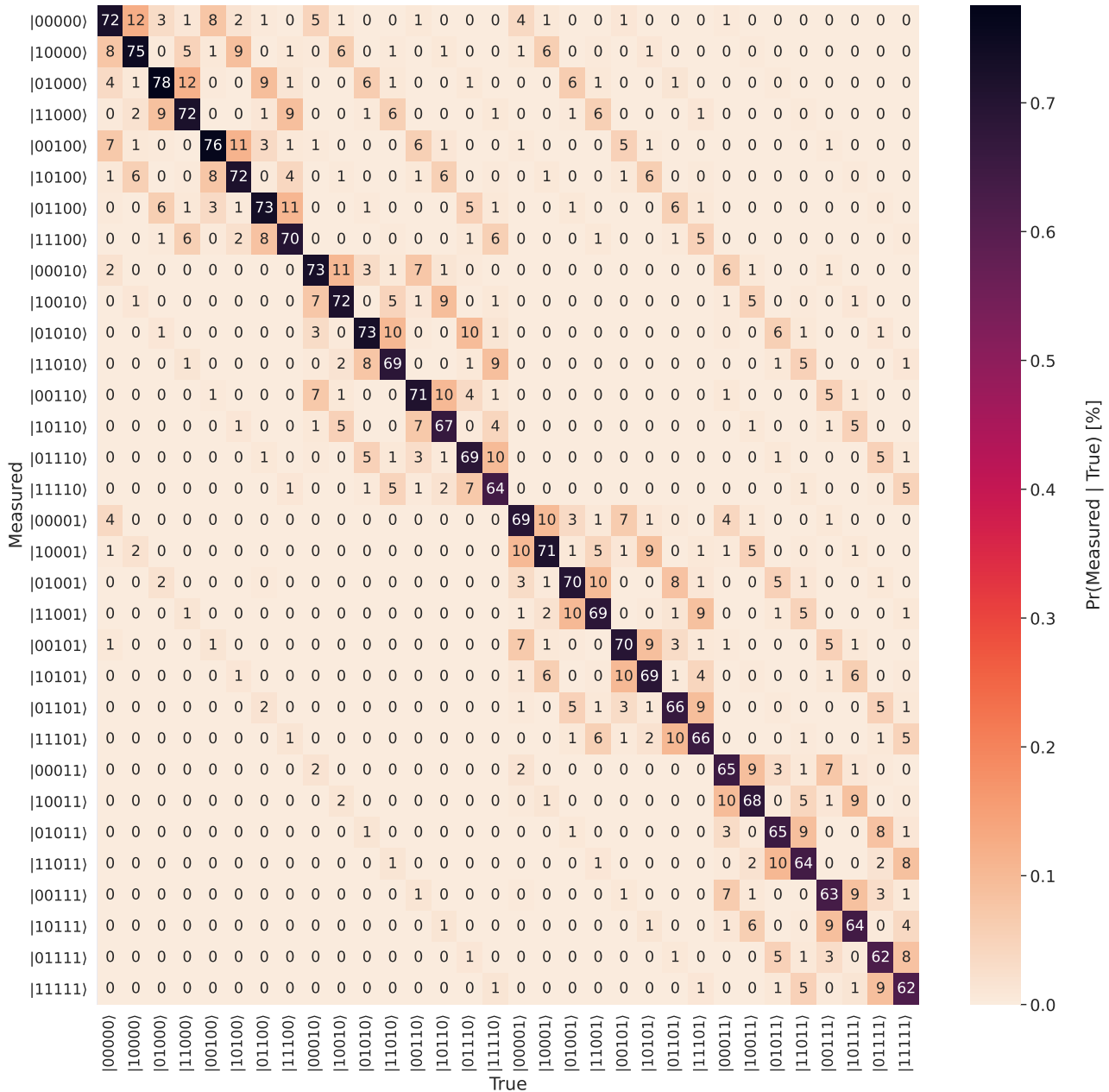


Figure 3.9: Response matrix for 5 qubits from the IBM Q Yorktown machine.

This is the result of our calibration matrix from the Yorktown machine for 5 qubits. The dimension of this matrix is $2^5 \times 2^5$. First, we can see that the probability is high in the diagonal component, when the measured and true states are the same. Secondly, we observe that the diagonal stripes with transition probabilities of about 4 – 11% are the result of the qubit flipping from 0 to 1 and vice versa. For a number of flips equal to 2 or more, the probability is almost zero in this transition matrix.

3.5.2 Unfolding a Uniform Distribution

In this part, we will use the unfolding methods on measured data that is obtained from the Yorktown quantum computer. First, we construct a quantum circuit, theoretically, this circuit will generate a uniform distribution (but quantum computer is noisy so it will destroy the

distribution). Then, the circuit will be launched in the IBM Yorktown computer, and after the data collection, we will use our calibration matrix. Finally we will use the code provided by Ben. Nachman et al., to unfold our data by the three techniques: Matrix inversion, Ignis and IBU.

Packages

First, we complete our needed packages, some of them are needed to build our quantum circuit and other packages are for data visualization.

```

1 from qiskit import QuantumCircuit, QuantumRegister
2 from qiskit.circuit import ClassicalRegister
3 #-----
4 # Packages for plotting
5 from matplotlib.gridspec import GridSpec
6 from mpl_toolkits.axes_grid1 import make_axes_locatable
7 import matplotlib.lines as mlines
8 #-----
9 from scipy.optimize import minimize

```

Generation of a uniform distribution with a quantum circuit

The quantum circuit that we want to construct is five HADAMARD gates in parallel. Initially, all the five qubits are in the $|0\rangle$ state. When we apply a HADAMARD gate on the first qubit it will give us an *equal superposition* state $|+\rangle = \frac{(|0\rangle+|1\rangle)}{\sqrt{2}}$. For five HADAMARD gates in parallel acting on $|0\rangle^{\otimes 5}$, we have the tensor product $\mathbb{H}^{\otimes 5} |0\rangle^{\otimes 5} = \frac{1}{\sqrt{2^5}} \sum_x |x\rangle$.

The code below is for creating this quantum circuit. We need five quantum registers and five classical registers. We first apply a HADAMARD gate in each quantum register. Then we measure each qubit and we store the result in the classical registers.

```

1 qreg_q = QuantumRegister(5, 'q')
2 creg_c = ClassicalRegister(5, 'c')
3 circuit = QuantumCircuit(qreg_q, creg_c)
4 circuit.h(qreg_q[0])
5 circuit.h(qreg_q[1])
6 circuit.h(qreg_q[2])
7 circuit.h(qreg_q[3])
8 circuit.h(qreg_q[4])
9 circuit.measure(qreg_q[0], creg_c[0])
10 circuit.measure(qreg_q[1], creg_c[1])
11 circuit.measure(qreg_q[2], creg_c[2])
12 circuit.measure(qreg_q[3], creg_c[3])
13 circuit.measure(qreg_q[4], creg_c[4])
14
15 circuit.draw(output='mpl')

```

The circuit diagram of five Hadamard gates in parallel is displayed as follow:

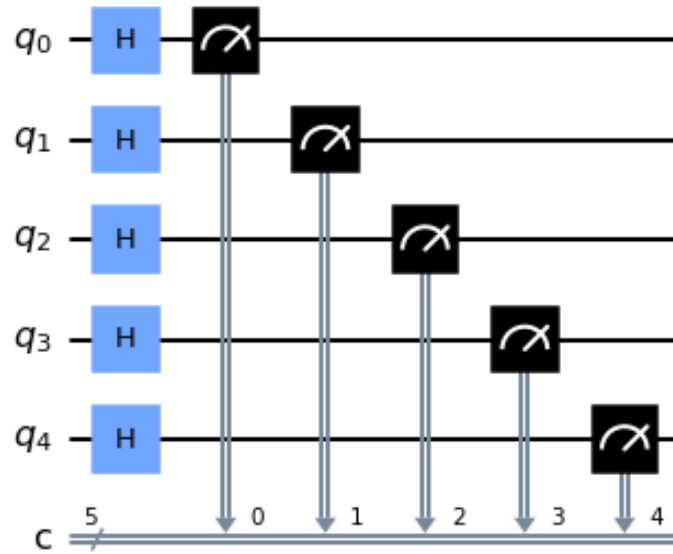


Figure 3.10: A circuit that generates a uniform distribution based on 5 qubits.

This circuit will create a uniform distribution, where the amplitude of each basis state will be $a = \frac{1}{\sqrt{2^5}} \approx 0.177$ and the equal probability is $\mathbb{P} = |a|^2 \approx 0.031$.

Executing the circuit in the IBM Yorktown QC

We will now run the circuit in the Yorktown quantum computer with 8192 shots. If the computer has no noise, we will have equal probabilities. That means each state among the $2^5 = 32$ states will have the same number of counts and it is equal to $8192/32 = 256$. In a real quantum computer, this is not the case because it is very noisy so the result will not be what we expect. This is why unfolding is so important to get closer to the true data. The following code allows us to execute the circuit in the IBM Yorktown system.

```

1 # We use the Yorktown quantum computer
2 backend = provider.get_backend('ibmqx2')
3 # Number of shots to run the program (experiment); maximum is 8192 shots.
4 shots = 8192
5 job = execute(circuit, backend, shots)

```

The result of the measurement

The results do not appear directly. We have to wait in a **chain** and then wait for the execution time (Sometimes it takes 5 minutes and sometimes it takes hours). The total time differs depending on the number of people on the waiting list and the speed of the quantum computer. We use the following code to collect the measurement.

```
1 m=job.result().get_counts(circuit)
```

When we print `m` we get our result for a uniform distribution. We can see that the counts are not equal because of readout noise from the quantum device, so they will not be uniformly distributed.

```
1 {'00000': 324, '00001': 220, '10000': 264, '10001': 240,
2  '10010': 265, '10011': 274, '10100': 281, '10101': 268,
3  '10110': 299, '10111': 289, '11000': 220, '11001': 211,
4  '11010': 224, '11011': 216, '11100': 242, '11101': 196,
5  '11110': 222, '11111': 221, '00010': 306, '00011': 262,
6  '00100': 309, '00101': 294, '00110': 342, '00111': 311,
7  '01000': 243, '01001': 188, '01010': 236, '01011': 215,
8  '01100': 283, '01101': 218, '01110': 267, '01111': 242}
```

Data Organization by label

The labeled measures are not arranged according to the binary system, so we will use the following code to arrange our measurement.

```
1 M = {val[0] : val[1] for val in sorted(m.items(), key = lambda x: (x[0], x[1]))}
```

After the arrangement we get:

```
1 {'00000': 324, '00001': 220, '00010': 306, '00011': 262,
2  '00100': 309, '00101': 294, '00110': 342, '00111': 311,
3  '01000': 243, '01001': 188, '01010': 236, '01011': 215,
4  '01100': 283, '01101': 218, '01110': 267, '01111': 242,
5  '10000': 264, '10001': 240, '10010': 265, '10011': 274,
6  '10100': 281, '10101': 268, '10110': 299, '10111': 289,
7  '11000': 220, '11001': 211, '11010': 224, '11011': 216,
8  '11100': 242, '11101': 196, '11110': 222, '11111': 221}
```

Counts extraction from the measurement

Now, we will extract the counts from the arranged data, but we will keep the same order. The following code will allow us to do this.

```
1 m_u = []
2 for key in M.keys()
3 m_u.append(M[key])
```

Finally, our measures are ready to use in the unfolding methods.

```

1 [324, 220, 306, 262, 309, 294, 342, 311, 243, 188, 236,
2  215, 283, 218, 267, 242, 264, 240, 265, 274, 281, 268,
3  299, 289, 220, 211, 224, 216, 242, 196, 222, 221]
```

Now, let us prepare our true data for a uniform distribution.

The true count

The true data is a uniform dist. That means for a number of shots equal to 8192, every state will have $8192/2^5 = 256$ of counts.

```

1 t_u = [8192/32 for i in range(32)]
```

The result:

```

1 [256.0, 256.0, 256.0, 256.0, 256.0, 256.0, 256.0, 256.0, 256.0, 256.0, 256.0,
2  256.0, 256.0, 256.0, 256.0, 256.0, 256.0, 256.0, 256.0, 256.0, 256.0, 256.0,
3  256.0, 256.0, 256.0, 256.0, 256.0, 256.0, 256.0, 256.0, 256.0, 256.0]
```

If we want to represent the true counts in a histogram, we need to label each state:

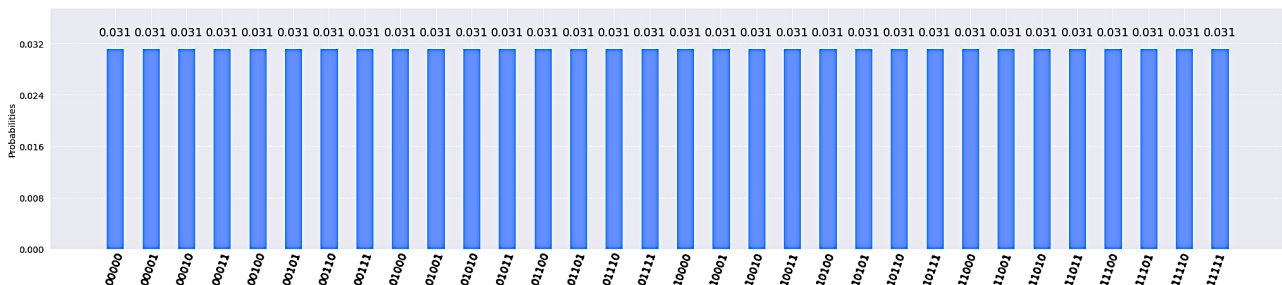
```

1 T_u = {
2  '00000': 256.0, '00001': 256.0, '00010': 256.0, '00011': 256.0, '00100': 256.0,
3  '00101': 256.0, '00110': 256.0, '00111': 256.0, '01000': 256.0, '01001': 256.0,
4  '01010': 256.0, '01011': 256.0, '01100': 256.0, '01101': 256.0, '01110': 256.0,
5  '01111': 256.0, '10000': 256.0, '10001': 256.0, '10010': 256.0, '10011': 256.0,
6  '10100': 256.0, '10101': 256.0, '10110': 256.0, '10111': 256.0, '11000': 256.0,
7  '11001': 256.0, '11010': 256.0, '11011': 256.0, '11100': 256.0, '11101': 256.0,
8  '11110': 256.0, '11111': 256.0}
```

We can see the representation of our data in the histogram by the `plot_histogram()` function:

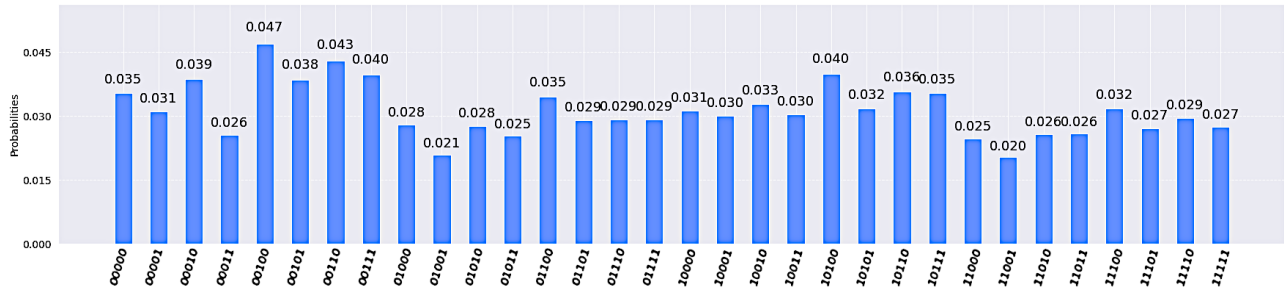
```

1 plot_histogram(T_u, figsize=(35,7))
```



The measured data are represented also in the following histogram:

```
1 plot_histogram(M,figsize=(30,6))
```



We can see clearly the effect of noise from the quantum computers by comparing true and measured data.

Unfolding functions

Now, we will use the unfolding functions provided by Ben. Nachman et al., to unfold our measured data from a real quantum computer. This part of the code is already explained in the python simulation section. The matrix inversion and the ignis function have two arguments (m_u , R). The IBU have four arguments (m_u , t_0 , R , n). At this point, we have all the ingredients for the unfolding methods. We can choose any t_0 and any n . But we will fix the number of iterations to $n=1$. The unfolding methods are defined as follow:

```
1 def MatrixInversion(y mes,R in):
2     return np.ravel(np.matmul(np.linalg.inv(np.matrix(R in)),y mes))
3
4 def fun(x,y mes,R in):
5     mat_dot_x = np.ravel(np.matmul(R in,x))
6     return sum((y mes - mat_dot_x)**2)
7
8 def Ignis(y mes,R in):
9
10    x0 = np.random.rand(len(y mes))
11    x0 = x0 / sum(x0)
12    nshots = sum(y mes)
13    cons = ({'type': 'eq', 'fun': lambda x: nshots - sum(x)})
14    bnds = tuple((0, nshots) for x in x0)
15    res = minimize(fun, x0, method='SLSQP',constraints=cons, bounds=bnds, tol=1e-6,
16                args=(y mes,R in))
17    return res.x
18
19 def IBU(y mes,t0,R in,n):
20
21    tn = t0
22    for q in range(n):
23        out = []
```

```

24     for j in range(len(t0)):
25         mynum = 0.
26         for i in range(len(yes)):
27             myden = 0.
28             for k in range(len(t0)):
29                 myden+=Rin[i][k]*tn[k]
30                 pass
31             mynum+=Rin[i][j]*tn[j]*yes[i]/myden
32             pass
33         out+=[mynum]
34     tn = out
35     pass
36     return tn

```

Unfolding a uniform distribution

Now, let us unfold our measured distribution by the three unfolding methods. For matrix inversion and Ignis we need only `m_u` and the response matrix. For IBU, we additionally need `t0` and `n`. These is all our ingredients in one place:

The true data `t_u`:

```

1 array([256, 256, 256, 256, 256, 256, 256, 256, 256, 256, 256, 256,
2        256, 256, 256, 256, 256, 256, 256, 256, 256, 256, 256, 256,
3        256, 256, 256, 256, 256, 256])

```

The measured data `m_u`:

```

1 array([324, 220, 306, 262, 309, 294, 342, 311, 243, 188, 236,
2        215, 283, 218, 267, 242, 264, 240, 265, 274, 281, 268,
3        299, 289, 220, 211, 224, 216, 242, 196, 222, 221])

```

The response matrix `R`:

```

1 array([[7.39257812e-01 1.07177734e-01 4.65087891e-02 ... 0.00000000e+00
2         0.00000000e+00 0.00000000e+00]
3        [7.34863281e-02 7.59277344e-01 4.39453125e-03 ... 1.09863281e-03
4         0.00000000e+00 2.44140625e-04]
5        [3.97949219e-02 7.32421875e-03 7.72827148e-01 ... 0.00000000e+00
6         8.54492188e-04 1.22070312e-04]
7        ...
8        [0.00000000e+00 0.00000000e+00 0.00000000e+00 ... 6.33178711e-01
9         5.24902344e-03 4.82177734e-02]
10       [1.22070312e-04 0.00000000e+00 0.00000000e+00 ... 5.24902344e-03
11        6.30249023e-01 8.23974609e-02]
12       [0.00000000e+00 0.00000000e+00 0.00000000e+00 ... 2.61230469e-02
13        9.57031250e-02 6.31103516e-01]])

```

We choose the prior truth distribution t_0 , as a uniform distribution.

```
1 array([1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
   ↪  1, 1, 1, 1, 1, 1, 1, 1])
```

Let us see our output of the three unfolding methods:

```
1 unfolded_Matrix_r = MatrixInversion(m_u, R)
2 unfolded_IBU_r = IBU(m_u, np.ones(len(m_u)), R, 1)
3 unfolded_ignis_r = Ignis(m_u, R)
```

Visual representation of unfolding result

In this part, we will compare the three unfolding methods by representing the results of unfolded data graphically.

We will need the following code to label our histograms.

```
1 nqubits_r = 5
2
3 mymapping_r = {}
4 mymapping_inverse_r = {}
5 for i in range(2**nqubits_r):
6     mymapping_r[i] = bin(i).split('b')[1]
7     mymapping_inverse_r[bin(i).split('b')[1]] = i
8     pass
9
10 xvals_r = []
11 xlabs_r = []
12 xlabs0_r = []
13 for i in range(2**nqubits_r):
14     xvals_r+= [i]
15     xlabs_r+= [r'$|'+mybin(i,nqubits_r)+r'\rangle$']
16     xlabs0_r+= ['']
17     pass
```

The following code allow us to represent the measured data, the true data and the unfolded data via the three unfolding methods.

```
1 def myre(x):
2     return np.concatenate([[x[0]],x])
3 bincenters_r = np.array(xvals_r)+0.5
4 sns.set_style("white")
5 f = plt.figure(figsize=(8.5, 8.5))
6 gs = GridSpec(2, 1, width_ratios=[1], height_ratios=[3.5, 1])
7 ax1 = plt.subplot(gs[0])
```

```

8  ax1.set_xticklabels( () )
9
10 mybincenters = np.linspace(-0.5,31.5,33)
11 plt.step(mybincenters,myre(t_u),color='black',label=r"t")
12 plt.errorbar(0.5*(mybincenters[1:]+mybincenters[:-1]),m_u,yerr=np.sqrt(m_u),
13 label='m',marker='o',linestyle='none',color='black')
14 plt.step(mybincenters,myre(unfolded_Matrix_r),color='blue',linestyle=':'
15 ,label=r"$\hat{t}_{matrix}$",linewidth=3.0)
16 plt.step(mybincenters,myre(unfolded_ignis_r),color='red',linestyle='--'
17 ,label=r"$\hat{t}_{ignis}$")
18 plt.step(mybincenters,myre(unfolded_IBU_r),color='green',linestyle='-.'
19 ,label=r"$\hat{t}_{IBU}$")
20
21 _=plt.xticks(xvals_r,xlabs0_r,rotation='vertical',fontsize=10)
22 plt.ylabel('Counts')
23 plt.ylim([0,2.*max(t_u)])
24 plt.legend(loc='upper left',fontsize=12,frameon=False)
25 _=plt.text(20., max(t_u)*1.8, "5 Yorktown Machine", fontsize=18)
26 _=plt.text(20., max(t_u)*1.7, "IBMQ Yorktown Readout Errors", fontsize=12)
27
28 ax2 = plt.subplot(gs[1])
29 ratio_ignis2_r = np.zeros(2**nqubits_r)
30 ratio_ignis_r = np.zeros(2**nqubits_r)
31 ratio_matrix_r = np.zeros(2**nqubits_r)
32 ratio_unfold_r = np.zeros(2**nqubits_r)
33 for i in range(len(ratio_unfold_r)):
34     ratio_unfold_r[i]=1.
35     ratio_matrix_r[i]=1.
36     ratio_ignis_r[i]=1.
37     ratio_ignis2_r[i]=1.
38     if (t_u[i] > 0):
39         ratio_unfold_r[i] = unfolded_IBU_r[i]/t_u[i]
40         ratio_matrix_r[i] = unfolded_Matrix_r[i]/t_u[i]
41         ratio_ignis_r[i] = unfolded_ignis_r[i]/t_u[i]
42     pass
43     pass
44
45 plt.ylim([0.8,1.2])
46 plt.step(mybincenters,myre(ratio_ignis_r),color='red',linestyle="--")
47 plt.step(mybincenters,myre(ratio_matrix_r),color='blue',linestyle=":",linewidth=3.0)
48 plt.step(mybincenters,myre(ratio_unfold_r),color='green')
49 plt.ylabel('Extracted / Truth')
50 xx2 = [1.,1.]
51 plt.plot([xvals_r[0],xvals_r[len(xvals_r)-1]],xx2,color='black',linestyle=':')
52 _=plt.xticks(xvals_r,xlabs_r,rotation='vertical',fontsize=12)
53 plt.savefig("Unfold_york.pdf",bbox_inches='tight')

```

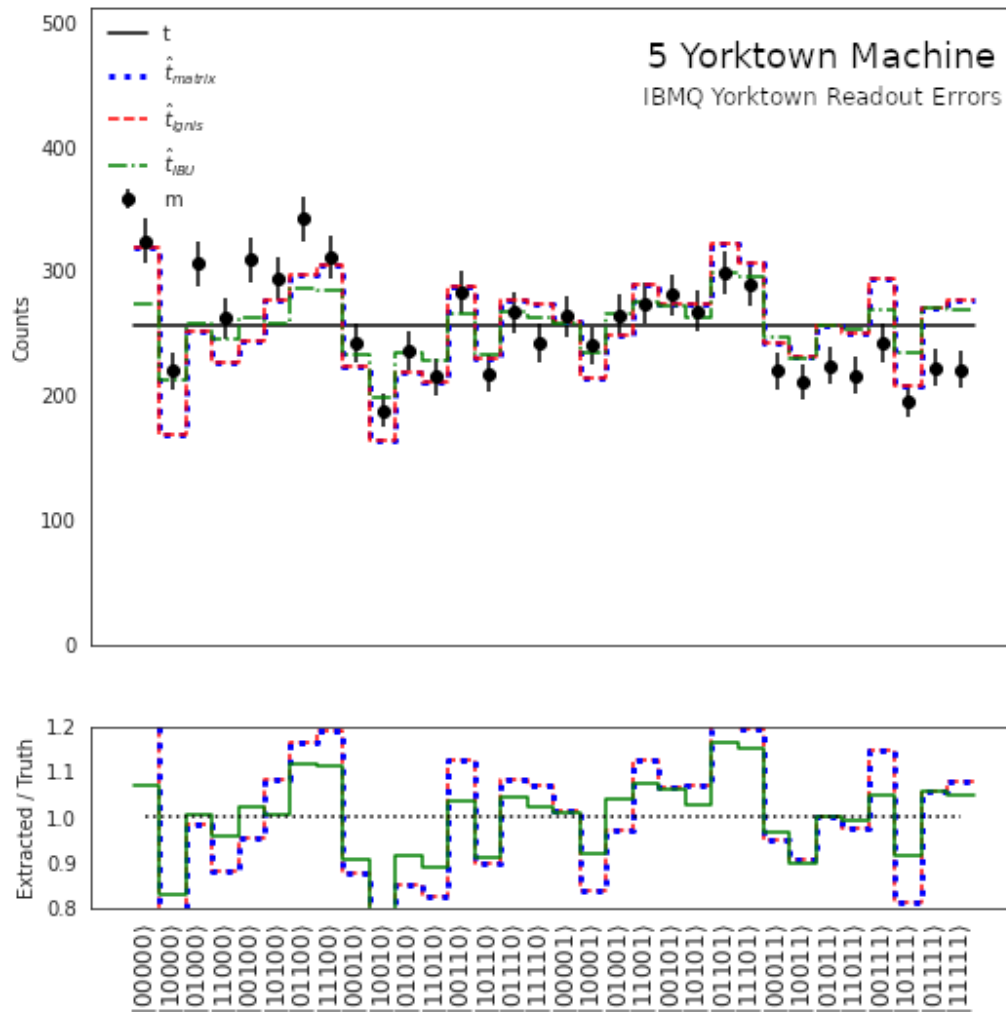


Figure 3.11: Unfolding results of a uniform distribution using three methods: Matrix inversion, ignis, and IBU with one iteration and a uniform prior truth spectrum.

Now, let us compare the three unfolding methods. First, we can see clearly that the values of the unfolded data with ignis and Matrix inversion are nearly identical. For IBU, the result is better than the two previous methods. where its values are near to the true data. The number of iterations that we used here is $n=1$ and the test data (t_0) is an array of 32 ones. We noticed that these parameters give better results for IBU.

3.5.3 Unfolding a Gaussian Distribution

In this part we will study unfolding in the case of a Gaussian distribution (the ground state of a harmonic oscillator), the steps followed are similar to that of a uniform distribution. But we will use a more developed tool to build the circuit corresponding to the Gaussian case. The general idea is that we can initialize our 5-qubit system in any desired state vector, specifying the *complex amplitudes* of the general 5-qubit state.

$$|\psi\rangle = \alpha_{00000} |00000\rangle + \dots + \alpha_{11111} |11111\rangle$$

We will explain this concept as we go along. But first, let's define our quantum circuit

with its corresponding quantum and classical registers.

```

1 nq = 5
2
3 qr = QuantumRegister(nq, 'qr')
4 cr = ClassicalRegister(nq, 'cr')
5 circuit_g = QuantumCircuit(qr, cr)

```

Since we want a Gaussian distribution, we define the function: $\text{HO}_{gs}(x) = e^{-\frac{x^2}{2}}$

```

1 def HO_groundstate(x):
2     return np.exp(-x**2/2)

```

which represents the *harmonic oscillator ground state*. The general form of the Gaussian distribution *probability density function* is written as

$$f(z) \sim e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2},$$

where μ is the mean, and σ is the standard deviation.

Instead of having z continuous in our specific case it is discrete, where $z \in [0, 31]$. We will plot the Gaussian in this interval, therefore, to center it we take $\mu = (31/2)$, and we can choose $\sigma = 8$. So we have

$$x = \frac{z - \mu}{\sigma} = \frac{z - (16 - 0.5)}{8} = 2 \left(\frac{z + 0.5 - 2^4}{2^4} \right).$$

We now define our desired vector, as the vector of elements $e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2}$, $z \in [0, 31]$:

```

1 desired_vector = []
2 for z in range(2**nq):
3     desired_vector += [HO_groundstate(2.*( z+ 0.5-2**(nq-1))/2**(nq-1))]
4     pass

```

The above algorithm creates the vector

$$\text{desired_vector} = \left[e^{-\frac{1}{2}\left(\frac{0-\mu}{\sigma}\right)^2}, e^{-\frac{1}{2}\left(\frac{1-\mu}{\sigma}\right)^2}, e^{-\frac{1}{2}\left(\frac{2-\mu}{\sigma}\right)^2}, \dots, e^{-\frac{1}{2}\left(\frac{31-\mu}{\sigma}\right)^2} \right]$$

```

1 print(desired_vector)
2
3 [0.15305573773635817, 0.19348058160704673, 0.24079047427224856, 0.295022
4 65617444284, 0.35586525214213843, 0.4226004432231888, 0.4940699735316383
5 4, 0.568671053718672, 0.6443887248251953, 0.7188675697091758, 0.78952156

```

```

6 96607879, 0.8536763613451477, 0.9087337563610701, 0.9523447998951764, 0.
7 9825754689579004, 0.9980487811074755, 0.9980487811074755, 0.982575468957
8 9004, 0.9523447998951764, 0.9087337563610701, 0.8536763613451477, 0.7895
9 215696607879, 0.7188675697091758, 0.6443887248251953, 0.568671053718672,
10 0.49406997353163834, 0.4226004432231888, 0.35586525214213843, 0.29502265
11 617444284, 0.24079047427224856, 0.19348058160704673, 0.1530557377363581]

```

Since these components represent *complex amplitudes*, they must check the *normalization condition*:

$$|\alpha_{00000}|^2 + |\alpha_{00001}|^2 + \dots + |\alpha_{11111}|^2 = 1$$

Thus, we need to normalize our desired vector (`desired_vector`/ $\sqrt{\sum_{k=0}^{31} v_k^2}$).

```

1 desired_vector /= np.sqrt(np.sum([v**2 for v in desired_vector]))
2
3 print(desired_vector)

```

Output

```

1 [0.04074024 0.05150049 0.06409340 0.07852887 0.09472390 0.11248742
2  0.13151111 0.15136836 0.17152283 0.19134754 0.21015416 0.22723083
3  0.24188595 0.25349430 0.26154108 0.26565975 0.26565975 0.26154108
4  0.25349430 0.24188595 0.22723083 0.21015416 0.19134754 0.17152283
5  0.15136836 0.13151111 0.11248742 0.09472390 0.07852887 0.06409340
6  0.05150049 0.04074024]

```

To configure the general 5-qubit state $|\psi\rangle$ in this desired vector, we use the `initialize` instruction, which requires a vector of **complex amplitudes** (`desired_vector`) and a **list of qubits**, which can be done by the loop below:

```

1 qrs = []
2 for i in range(nq):
3     qrs+=qr[i]
4     pass
5 print(qrs)

```

Output

```

1 [Qubit(QuantumRegister(5, 'qr'), 0), Qubit(QuantumRegister(5, 'qr'), 1),
2  Qubit(QuantumRegister(5, 'qr'), 2), Qubit(QuantumRegister(5, 'qr'), 3),
3  Qubit(QuantumRegister(5, 'qr'), 4)]

```

We have therefore defined a list of 5 qubits $[q_0, q_1, q_2, q_3, q_4]$. We can now initialize them in a *Gaussian distribution*, through the following input:

```
1 circuit_g.initialize(desired_vector, qrs)
```

Once the initialization is done, the last thing to do is to measure the system. Here we are using the `measure()` function, with the parameters `(qr,cr)`, which represent a shorthand that allow us to implement a *measurement operation* on each qubit.

```
1 circuit_g.measure(qr, cr)
```

Since we are curious to see this circuit, we visualize it.

```
1 circuit_g.draw('latex')
```

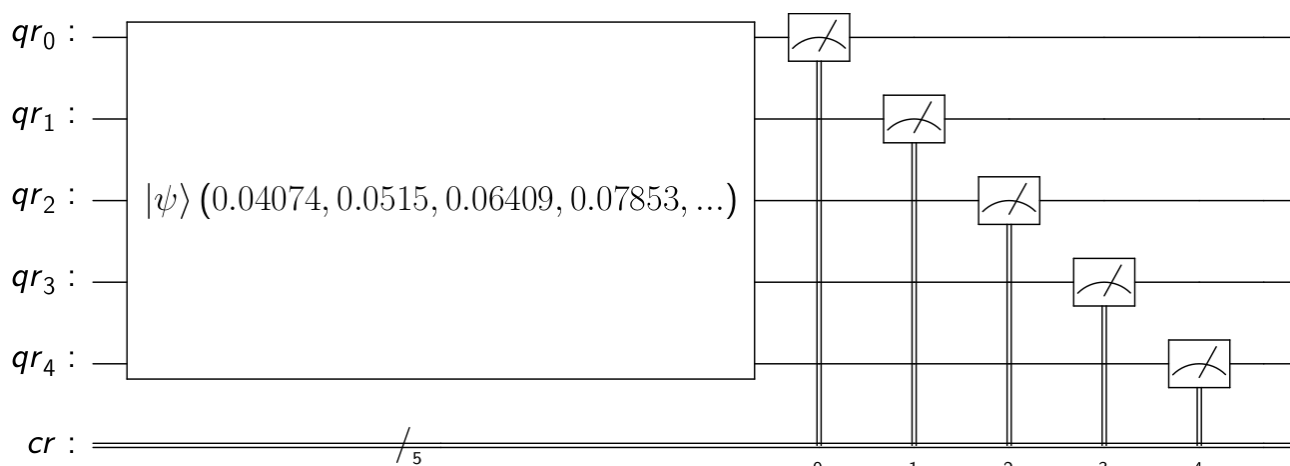


Figure 3.12: A circuit that generates a Gaussian distribution based on 5 qubits.

Note that `initialize` contains a `reset` instruction, which is not *unitary*, therefore it is an **Instruction** and not a **Gate**.

Simulators are characterized by negligible readout noise, therefore, we can use one of them to get the *true Gaussian distribution*, denoted by `t_g`. Here we are using the simulator called “`qasm_simulator`” provided by Aer. To use it we enter the same commands that allow us to use a real quantum computer, with the only difference instead of the command `provider = IBMQ.load_account()`, we use Aer.¹⁰

```
1 simulator = Aer.get_backend('qasm_simulator')
2 result = execute(circuit_g, simulator, shots=8192).result()
3 count_s = result.get_counts(circuit_g)
4 print(count_s)
```

¹⁰For simulators provided by IBM Q (non-local), we use the `provider = IBMQ.load_account()`.

Output

```

1 {'00000': 12, '00001': 18, '10000': 603, '10001': 563, '10010': 540, '10011
2 ': 470, '10100': 449, '10101': 369, '10110': 324, '10111': 236, '11000': 19
3 0, '11001': 150, '11010': 96, '11011': 67, '11100': 50, '11101': 25, '111
4 10': 21, '11111': 11, '00010': 31, '00011': 51, '00100': 75, '00101': 103
5 , '00110': 141, '00111': 195, '01000': 240, '01001': 288, '01010': 342,
6 '01011': 429, '01100': 449, '01101': 527, '01110': 566, '01111': 561}

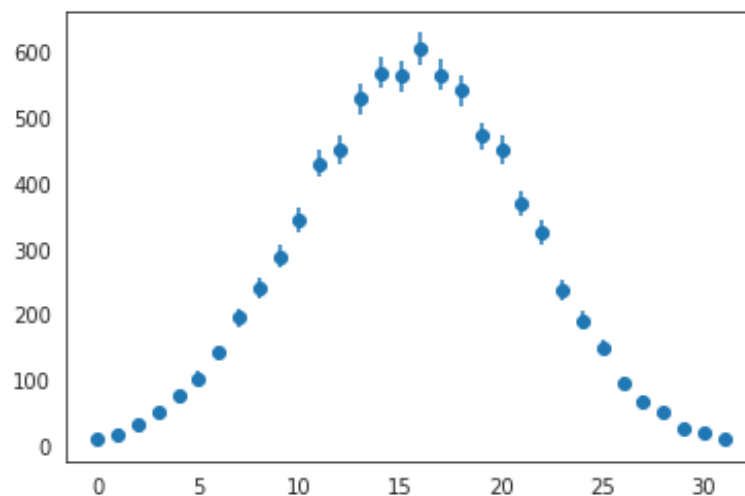
```

To plot this distribution, we write the following code:

```

1 nq=5
2 xvals = []
3 yvals = []
4 for j in range(2**nq):
5     xvals+= [j]
6     yvals+= [count_s[mybin(j,nq)[::-1]]]
7
8 sns.set_style("white")
9 plt.errorbar(xvals,yvals,yerr=np.sqrt(yvals),linestyle='none',marker='o')

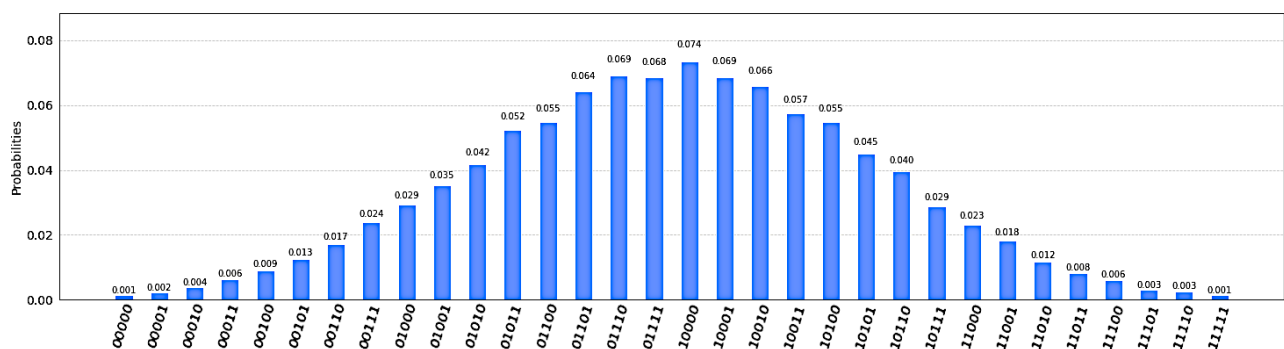
```



```

1 plot_histogram(count_s,figsize=(25,6))

```



We have now confirmed that we really have a Gaussian distribution.

The readout noise from the Yorktown machine will distort this perfect Gaussian and gives us a disturbed distribution, denoted by `m_g`. To realize this, we execute our circuit (`circuit_g`) in the “`ibmq_5_yorktown`” machine.

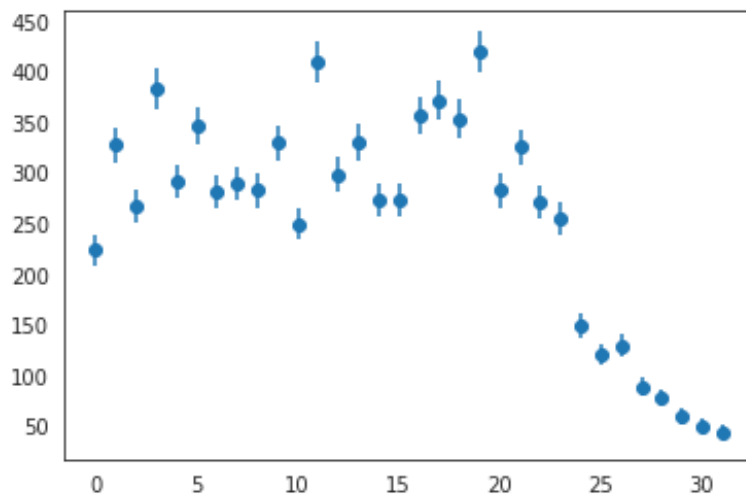
```
1 backend = provider.get_backend('ibmq_5_yorktown') #or:'ibmqx2'
2 result = execute(circuit_g, backend,shots=8192).result()
3 count_y = result.get_counts(circuit_g)
4 print(count_y)
```

Output

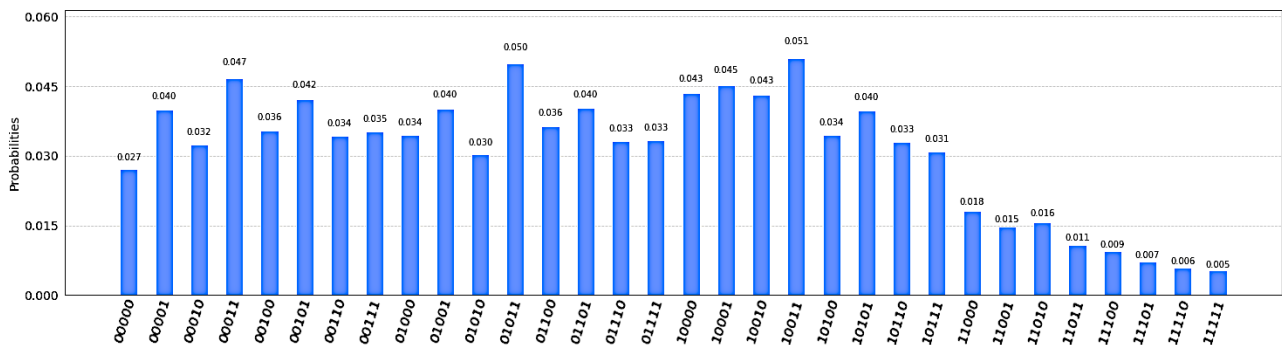
```
1 {'00000': 223, '00001': 327, '10000': 356, '10001': 371, '10010': 353, '1
2 10011': 419, '10100': 282, '10101': 325, '10110': 270, '10111': 254, '11
3 000': 149, '11001': 120, '11010': 129, '11011': 88, '11100': 77, '11101'
4 : 59, '11110': 49, '11111': 43, '00010': 266, '00011': 383, '00100': 291
5 , '00101': 346, '00110': 281, '00111': 289, '01000': 282, '01001': 329,
6 '01010': 249, '01011': 409, '01100': 298, '01101': 330, '01110': 272, '0
7 1111': 273}
```

Same inputs for visualization:

```
1 xvals = []
2 yvals = []
3 for j in range(2**nq):
4     xvals+= [j]
5     yvals+= [count_y[mybin(j,nq)[::-1]]]
6
7 sns.set_style("white")
8 plt.errorbar(xvals,yvals,yerr=np.sqrt(yvals),linestyle='none',marker='o')
```



```
1 plot_histogram(count_y,figsize=(25,6))
```



We notice that the readout noise has distorted the Gaussian (true) distribution to the point that we cannot recognize its original shape.

Now we have just to extract the *frequencies* from (count_s) and (count_y), so we use the same code seen for the uniform case, in order to finally obtain the vectors `t_g`, and `m_g`.

```
1 true_g= {val[0] : val[1] for val in sorted(count_s.items(), key = lambda
2                                     x: (x[0], x[1]))}
3 t_g = []
4 for key in true_g.keys() :
5     t_g.append(true_g[key])
6     pass
7 print(t_g)
8
9 [12, 18, 31, 51, 75, 103, 141, 195, 240, 288, 342, 429, 449, 527, 566,
10 561, 603, 563, 540, 470, 449, 369, 324, 236, 190, 150, 96, 67, 50, 25,
11 21, 11]
```

We now extract the measured distribution vector `m_g`:

```
1 meas_g= {val[0] : val[1] for val in sorted(count_y.items(), key = lambda
2                                     x: (x[0], x[1]))}
3 m_g = []
4 for key in meas_g.keys() :
5     m_g.append(meas_g[key])
6     pass
7 print(m_g)
8
9 [223, 327, 266, 383, 291, 346, 281, 289, 282, 329, 249, 409, 298, 330, 2
10 72, 273, 356, 371, 353, 419, 282, 325, 270, 254, 149, 120, 129, 88, 77,
11 59, 49, 43]
```

Here we are at the crucial part of this experience; the reconstruction of truth “Unfolding”. For this objective, we need the *Yorktown response matrix* and the result obtained from the

Yorktown readout noise. The two unfolding methods `MatrixInversion` and `Ignis` (commonly used in QIS), give us the following outputs:

```

1 MatrixInversion(m_g,R)
2
3 array([151.96727863, 278.69220815, 183.01138799, 361.83956380,
4         231.16886853, 318.11018472, 228.28325324, 261.47569561,
5         262.90470574, 336.10572204, 204.19246932, 511.82630002,
6         303.77496880, 395.06103499, 303.84402141, 307.09007413,
7         383.34934568, 372.73466745, 360.07688363, 510.94824455,
8         271.18915214, 358.70672324, 295.37171705, 251.92702464,
9         176.83096465, 128.69790459, 163.81956139, 87.15086908,
10        75.30630507, 42.56811204, 38.78854417, 35.18624349])

```

```

1 Ignis(m_g,R)
2
3 array([151.96736728, 278.69210044, 183.01138211, 361.83954193,
4         231.16889604, 318.11019459, 228.28326267, 261.47568405,
5         262.90461210, 336.10575357, 204.19247446, 511.82626943,
6         303.77501047, 395.06101942, 303.84403513, 307.09006300,
7         383.34934156, 372.73459805, 360.07668674, 510.94840325,
8         271.18909030, 358.70681406, 295.37191834, 251.92685430,
9         176.83112506, 128.69782248, 163.81966319, 87.15077385,
10        75.30624717, 42.56814870, 38.78852282, 35.18632344])

```

We notice that the unfolding results for `MatrixInversion` and `Ignis` are very close, despite the fact that these are two different methods.

The IBU from HEP have 4 parameters (`m_g`, `t0`, `R`, `n`), the only parameters we are free to choose are the *number of iterations* (`n`) and the *prior truth distribution* (`t0`). We recall the 6th advantage of the IBU mentioned by D’agostini in his article [26], where he states that satisfactory results are obtained if we start from a *uniform distribution*. But it gives the best results if one makes a realistic guess about the true distribution. Therefore, we can choose `t0` for example `[0, 1, 2, ..., 2, 1, 0]`, which is of course a 2^5 dimensional vector. In `python` it can be produced as follows:

```

1 c = []
2 for s in range(16):
3     c.append(s)
4     pass
5 t_s=[ *c , *c[::-1] ]
6 print(t_s)

```

Output

```
1 [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 15, 14, 13, 12,
2 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0]
```

In the algorithm above, we are adding the elements `s` to the empty vector `c` by the function `append()`. The object `[::-1]` reverse the elements of `c`, we can then concatenate the two vectors.

The second parameter to set, is the number of iterations, where we just take `n=1`.

```
1 print(IBU(m_g,t_s,R,1))
2
3 [0.0, 87.29209365373235, 115.43210028814272, 210.70351275475588
4 , 208.88814802549277, 285.4138302647153, 238.21228525158762, 30
5 0.7623880678843, 279.92435625976367, 353.78385668972953, 285.58
6 200778083545, 433.46993222855866, 351.6549239367971, 424.667582
7 80482253, 346.4533319644723, 406.80177950302266, 454.7858110448
8 58, 474.61896151747544, 430.9510312411226, 481.0646239820014, 3
9 29.50013439662126, 344.7721592177831, 294.3503960979849, 278.53
10 393415499374, 184.08290907346077, 151.61192190751058, 153.01420
11 394758765, 121.38914195438865, 78.25374972502975, 52.2064547630
12 8817, 33.82243750178114, 0.0]
```

These values are clearly different from the two previous methods. We have now to visualize the three unfolding results in order to get a more accurate observation of their behavior and to compare them. We use the same algorithm as the uniform case, where the only changes to be made are the true and measured distribution `t_g`, `m_g`, and the prior truth distribution `t0 = ts`.

```
1 unfolded_Matrix = MatrixInversion(m_g, R)
2 unfolded_IBU = IBU(m_g, t_s, R, 1)
3 unfolded_ignis = Ignis(m_g, R)
4
5
6 bincenters_r = np.array(xvals_r)+0.5
7 sns.set_style("white")
8 f = plt.figure(figsize=(8.5, 8.5))
9 gs = GridSpec(2, 1, width_ratios=[1], height_ratios=[3.5, 1])
10 ax1 = plt.subplot(gs[0])
11 ax1.set_xticklabels( ( ) )
12
13 def myre(x):
14     return np.concatenate([[x[0]],x])
15
16 mybincenters = np.linspace(-0.5,31.5,33)
```

```

17
18 plt.step(mybincenters,myre(t_g),color='black',label=r"t")
19 plt.errorbar(0.5*(mybincenters[1:]+mybincenters[:-1]),m_g,yerr=np.sqrt(m
20             _g),label='m',marker='o',linestyle='none',color='black')
21 plt.step(mybincenters,myre(unfolded_Matrix),color='blue',linestyle=':',
22             label=r"$\hat{t}_{matrix}$",linewidth=3.0)
23 plt.step(mybincenters,myre(unfolded_ignis),color='red',linestyle='--',
24             label=r"$\hat{t}_{ignis}$")
25 plt.step(mybincenters,myre(unfolded_IBU),color='green',linestyle='-.',
26             label=r"$\hat{t}_{IBU}$")
27
28 _=plt.xticks(xvals_r,xlabs0_r,rotation='vertical',fontsize=10)
29 plt.ylabel('Counts')
30 plt.ylim([0,2.*max(t_g)])
31 plt.legend(loc='upper left',fontsize=12,frameon=False)
32 _=plt.text(20., max(t_g)*1.8, "5 Yorktown Machine", fontsize=18)
33 _=plt.text(20., max(t_g)*1.7, "IBMQ Yorktown Readout Errors", fontsize=12)
34
35 nqubits = 5
36
37 ax2 = plt.subplot(gs[1])
38 ratio_ignis2 = np.zeros(2**nqubits)
39 ratio_ignis = np.zeros(2**nqubits)
40 ratio_matrix = np.zeros(2**nqubits)
41 ratio_unfold = np.zeros(2**nqubits)
42 for i in range(len(ratio_unfold)):
43     ratio_unfold[i]=1.
44     ratio_matrix[i]=1.
45     ratio_ignis[i] =1.
46     ratio_ignis2[i]=1.
47     if (t_g[i] > 0):
48         ratio_unfold[i] = unfolded_IBU[i]/t_g[i]
49         ratio_matrix[i] = unfolded_Matrix[i]/t_g[i]
50         ratio_ignis[i] = unfolded_ignis[i]/t_g[i]
51     pass
52     pass
53
54 plt.ylim([0.8,1.2])
55 plt.step(mybincenters,myre(ratio_ignis),color='red',linestyle="--")
56 plt.step(mybincenters,myre(ratio_matrix),color='blue',linestyle=":",
57             ,linewidth=3.0)
58 plt.step(mybincenters,myre(ratio_unfold),color='green')
59 plt.ylabel('Extracted / Truth')
60 xx2 = [1.,1.]
61 plt.plot([xvals_r[0],xvals_r[len(xvals_r)-1]],xx2,color='black',lines
62             tyle=':')
63 _=plt.xticks(xvals_r,xlabs_r,rotation='vertical',fontsize=12)

```

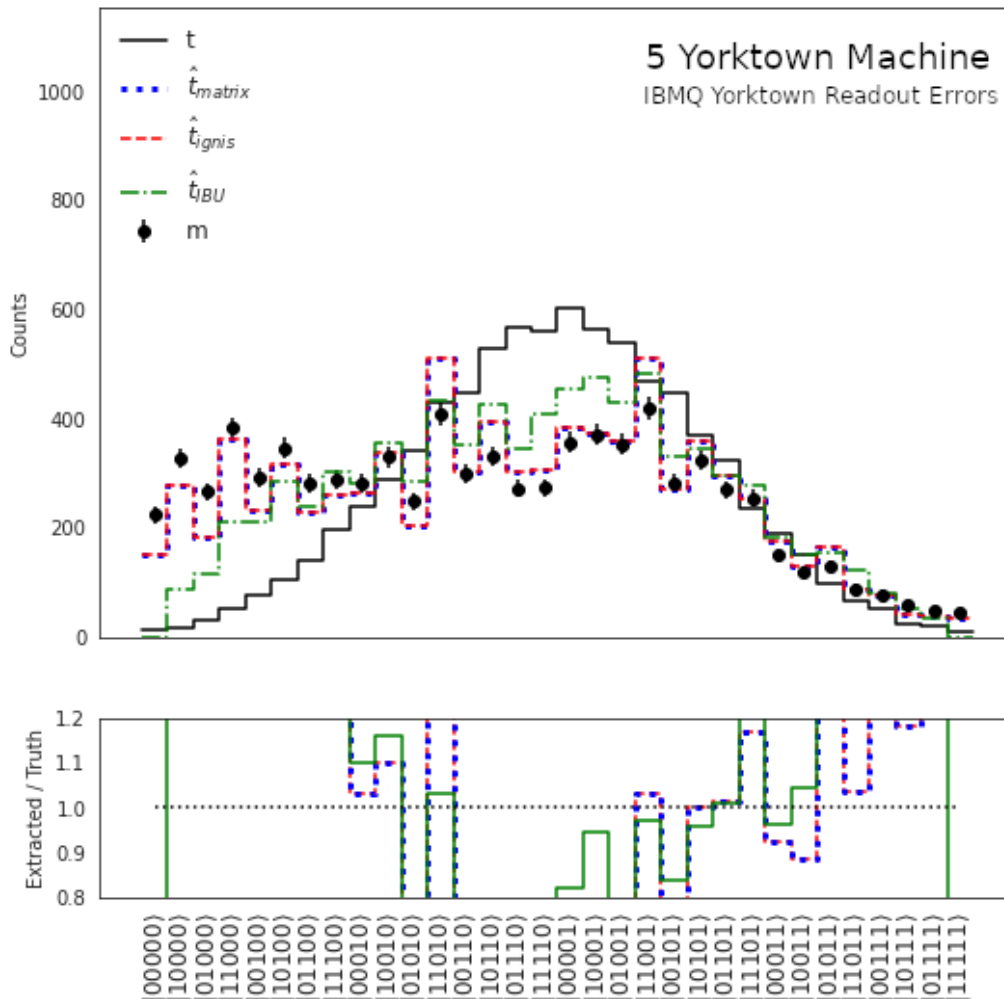


Figure 3.13: Unfolding results of a Gaussian distribution using three methods: Matrix inversion, ignis, and IBU with one iteration and a prior truth spectrum $t_0=t_s$.

Note that the values of `MatrixInversion` and `ignis` are identical. Moreover, the unfolding amount is very low. In the case of `IBU`, its values are closer to the true distribution and give a more important unfolding compared the other two methods, and for only 1 iteration.

In the second graph (`Extracted/truth`) we are dividing the components of each resulting vector by the component of the vector `t_g`, so if they are equal we get 1. Hence, we are representing a Gaussian distribution as a uniform one. If we look at this graph we see that the values closest to 1 in the majority of the time are those of `IBU`.

The results obtained above, concern the choice $t_0=t_s$. But what if we tried another vector t_0 , knowing that taking the vector $\lambda \cdot t_0$, where $\lambda \in \mathbb{R}^*$, will not change the results, this fact can be directly verified from the formula of `IBU` for one iteration

$$t_i^1 = \sum_j \left(\frac{\mathcal{R}_{ji} t_i^0}{\sum_k \mathcal{R}_{jk} t_k^0} \right) m_j \longrightarrow \sum_j \left(\frac{\mathcal{R}_{ji} (\lambda \cdot t_i^0)}{\sum_k \mathcal{R}_{jk} (\lambda \cdot t_k^0)} \right) m_j \equiv t_i^1. \quad (3.9)$$

So we have to try a different vector, for example, a vector which takes *triangular numbers*¹¹

¹¹Triangular numbers are generated by the formula: $\mathfrak{T}_n = \sum_{p=0}^n p = 0 + 1 + \dots + n = \frac{n(n+1)}{2}$.

as elements, where the values are increasing until the 16th element, then decreasing, i.e.

$$t_trig = [0, 1, 3, 6, \dots, 6, 3, 1, 0]$$

This can be obtained by the following algorithm:

```

1 b =0; p =[]
2 for n in range(16):
3     b = b + n
4     p.append(b)
5 t_trig = [*p,*p[::-1]]
6 print(t_trig)

```

Output

```

1 [0, 1, 3, 6, 10, 15, 21, 28, 36, 45, 55, 66, 78, 91, 105, 120, 120, 105,
2  91, 78, 66, 55, 45, 36, 28, 21, 15, 10, 6, 3, 1, 0]

```

Let's try this new *prior truth distribution* in the previous algorithm, where it suffices to change t_s to t_trig :

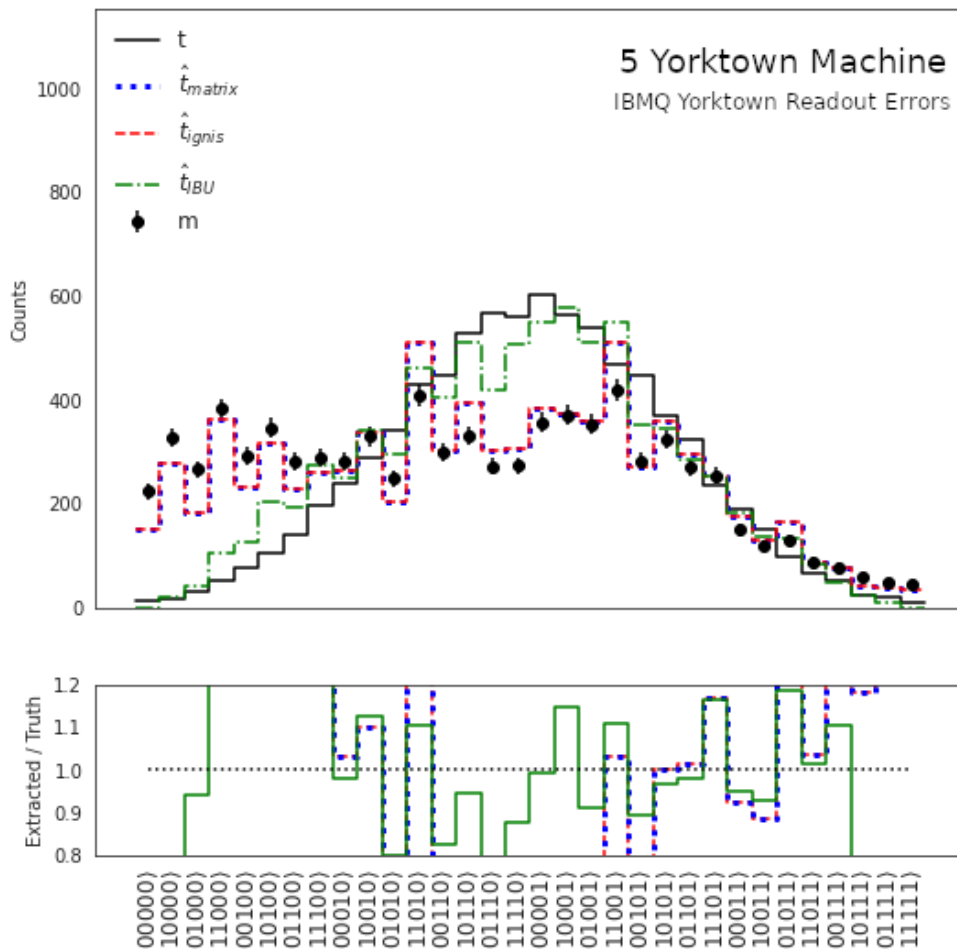


Figure 3.14: Unfolding results of a Gaussian distribution using three methods: Matrix inversion, ignis, and IBU with one iteration and a prior truth spectrum $t_0=t_trig$.

With this specific initial truth spectrum (t_0), IBU really gives the best unfolding results, the amount of unfolding is very important. From these two results, we conclude that IBU is very sensitive to the choice of t_0 . To observe this fact let's take 3 possible t_0 's:

$$t_{\text{unif}} = [1, 1, 1, \dots, 1, 1, 1]$$

$$t_{\text{s}} = [0, 1, 2, \dots, 2, 1, 0]$$

$$t_{\text{trig}} = [0, 1, 3, \dots, 3, 1, 0]$$

We will apply unfolding via IBU with these 3 t_0 's, and for $n=1$:

```

1 unfolded_IBU_a = IBU(m_g, t_unif, R, 1)
2 unfolded_IBU_b = IBU(m_g, t_s, R, 1)
3 unfolded_IBU_c = IBU(m_g, t_trig, R, 1)

```

After visualizing the results using the same algorithm, we get:

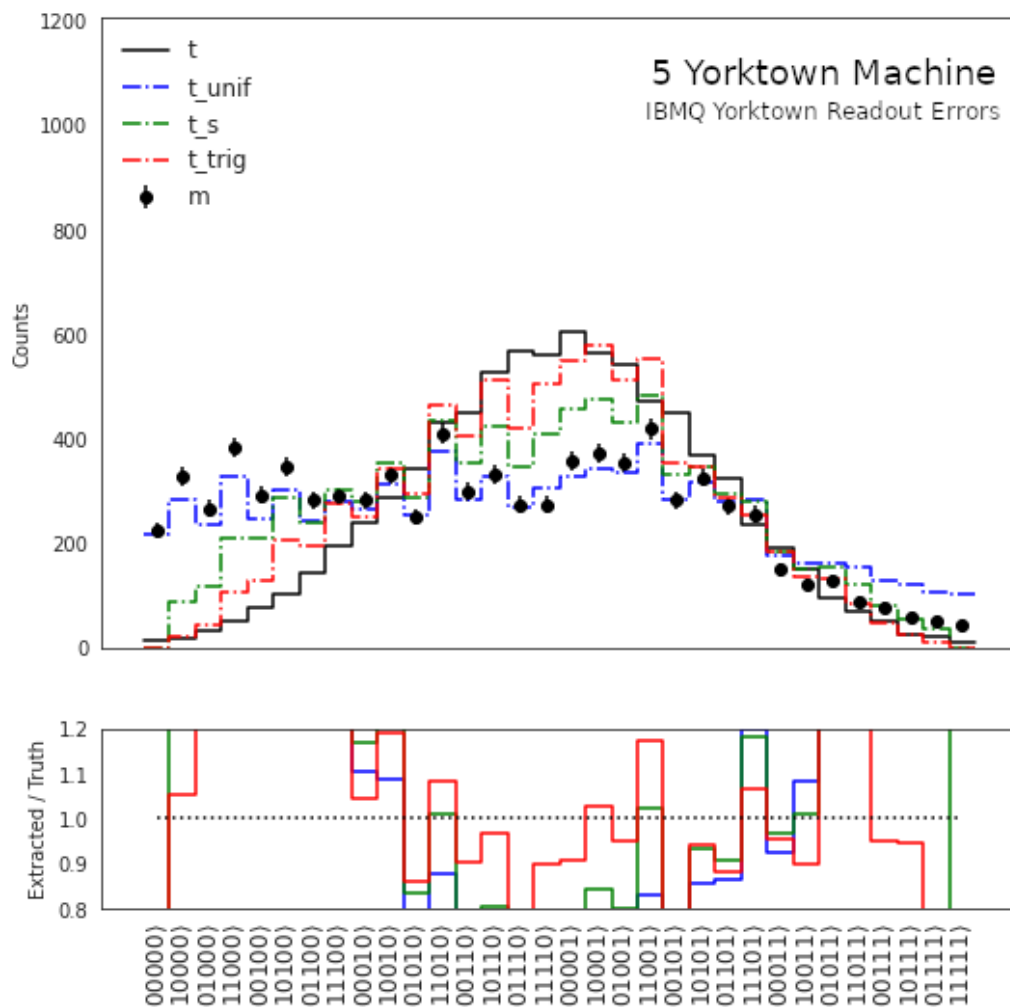


Figure 3.15: A comparison of unfolding results for a Gaussian distribution using IBU with three prior truth distributions: t_{unif} , t_{s} , t_{trig} and with 1 iteration.

The worst unfolding results are when we take t_0 as a *uniform distribution*. In our case, we know t_g , but in real experiences, we ignore it, so a good idea is to take the measured distribution as the prior truth distribution i.e., $t_0 = m_g$. But its effectiveness depends on the amount of noise.

Another property to note and which appears when the number of iterations increases, where the three prior truth distributions converge towards the same result, as seen below for 10 iterations:

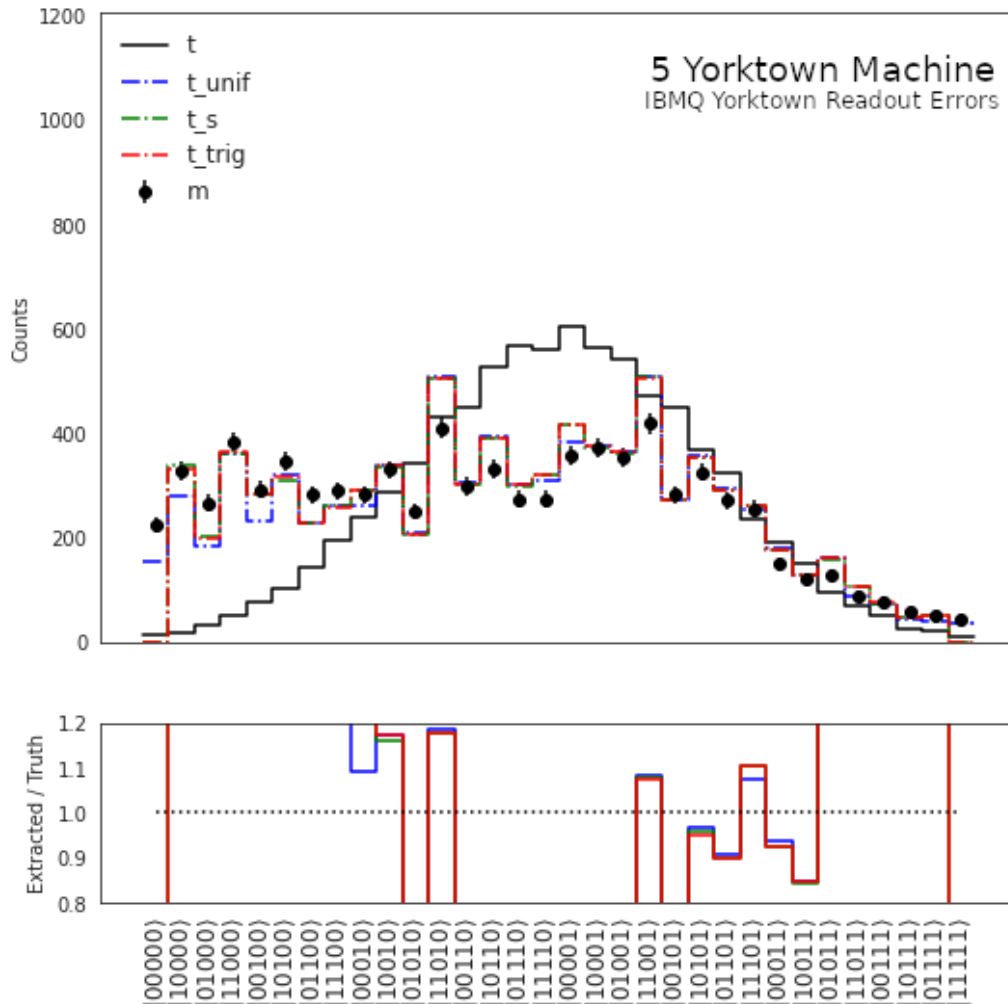


Figure 3.16: A comparison of unfolding results for a Gaussian distribution using IBU with three prior truth distributions: t_{unif} , t_s , t_{trig} and with 10 iterations.

Now let's investigate the behavior of the IBU, depending on the number of iterations. For this purpose we first do an unfolding on a uniform distribution, then we tackle the Gaussian one, keeping the same parameters (t_u , m_u , t_g , m_g), and we just vary the number of iterations, where we will take 1, 2, 3, and 10 iterations. Thus, for the first distribution, we have

```

1 t0=np.ones(len(m_u))
2
3 unfolded_IBU1 = IBU(m_u,t0,R, 1)
4 unfolded_IBU2 = IBU(m_u,t0,R, 2)

```

```

5  unfolded_IBU3  = IBU(m_u,t0,R, 3)
6  unfolded_IBU10 = IBU(m_u,t0,R,10)

```

The visualization algorithm:

```

1  bincenters_r = np.array(xvals_r)+0.5
2  sns.set_style("white")
3  f = plt.figure(figsize=(8.5, 8.5))
4  gs = GridSpec(2, 1, width_ratios=[1], height_ratios=[3.5, 1])
5  ax1 = plt.subplot(gs[0])
6  ax1.set_xticklabels( () )
7
8  mybincenters = np.linspace(-0.5,31.5,33)
9
10 plt.step(mybincenters,myre(t_u),color='black',label=r"t")
11 plt.errorbar(0.5*(mybincenters[1:]+mybincenters[:-1]),m_u,yerr=np.sqrt(
12     m_u),label='m',marker='o',linestyle='none',color='black')
13 plt.step(mybincenters,myre(unfolded_IBU1),color='red',linestyle='-.',la
14     bel=r"$\hat{t}_{IBU}\,,$(n=1)")
15 plt.step(mybincenters,myre(unfolded_IBU2),color='blue',linestyle='-.',l
16     abel=r"$\hat{t}_{IBU}\,,$(n=2)")
17 plt.step(mybincenters,myre(unfolded_IBU3),color='orange',linestyle='-. '
18     ,label=r"$\hat{t}_{IBU}\,,$(n=3)")
19 plt.step(mybincenters,myre(unfolded_IBU10),color='green',linestyle='-. '
20     ,label=r"$\hat{t}_{IBU}\,,$(n=10)")
21 _=plt.xticks(xvals_r,xlabs0_r,rotation='vertical',fontsize=10)
22
23 plt.ylabel('Counts')
24 plt.ylim([0,2.*max(t_u)])
25 plt.legend(loc='upper left',fontsize=8,frameon=False)
26 _=plt.text(20., max(t_u)*1.8, "5 Yorktown Machine", fontsize=18)
27 _=plt.text(20., max(t_u)*1.7, "IBMQ Yorktown Readout Errors",fontsize=12)
28
29 ax2 = plt.subplot(gs[1])
30 ratio_unfold1 = np.zeros(2**nqubits)
31 ratio_unfold2 = np.zeros(2**nqubits)
32 ratio_unfold3 = np.zeros(2**nqubits)
33 ratio_unfold10 = np.zeros(2**nqubits)
34
35 for i in range(len(ratio_unfold1)):
36     ratio_unfold1[i] =1.
37     ratio_unfold2[i] =1.
38     ratio_unfold3[i] =1.
39     ratio_unfold10[i]=1.
40     if (t_u[i] > 0):
41         ratio_unfold1[i] = unfolded_IBU1[i]/t_u[i]
42         ratio_unfold2[i] = unfolded_IBU2[i]/t_u[i]
43         ratio_unfold3[i] = unfolded_IBU3[i]/t_u[i]
44         ratio_unfold10[i] = unfolded_IBU10[i]/t_u[i]
45     pass
46 pass
47

```

```

48 plt.ylim([0.8,1.2])
49 plt.step(mybincenters,myre(ratio_unfold1 ),color='red' )
50 plt.step(mybincenters,myre(ratio_unfold2 ),color='blue' )
51 plt.step(mybincenters,myre(ratio_unfold3 ),color='orange' )
52 plt.step(mybincenters,myre(ratio_unfold10),color='green' )
53 plt.ylabel('Extracted / Truth')
54
55 xx2 = [1.,1.]
56 plt.plot([xvals_r[0],xvals_r[len(xvals_r)-1]],xx2,color='black'
57         ,linestyle=':')
58 _=plt.xticks(xvals_r,xlabs_r,rotation='vertical',fontsize=12)

```

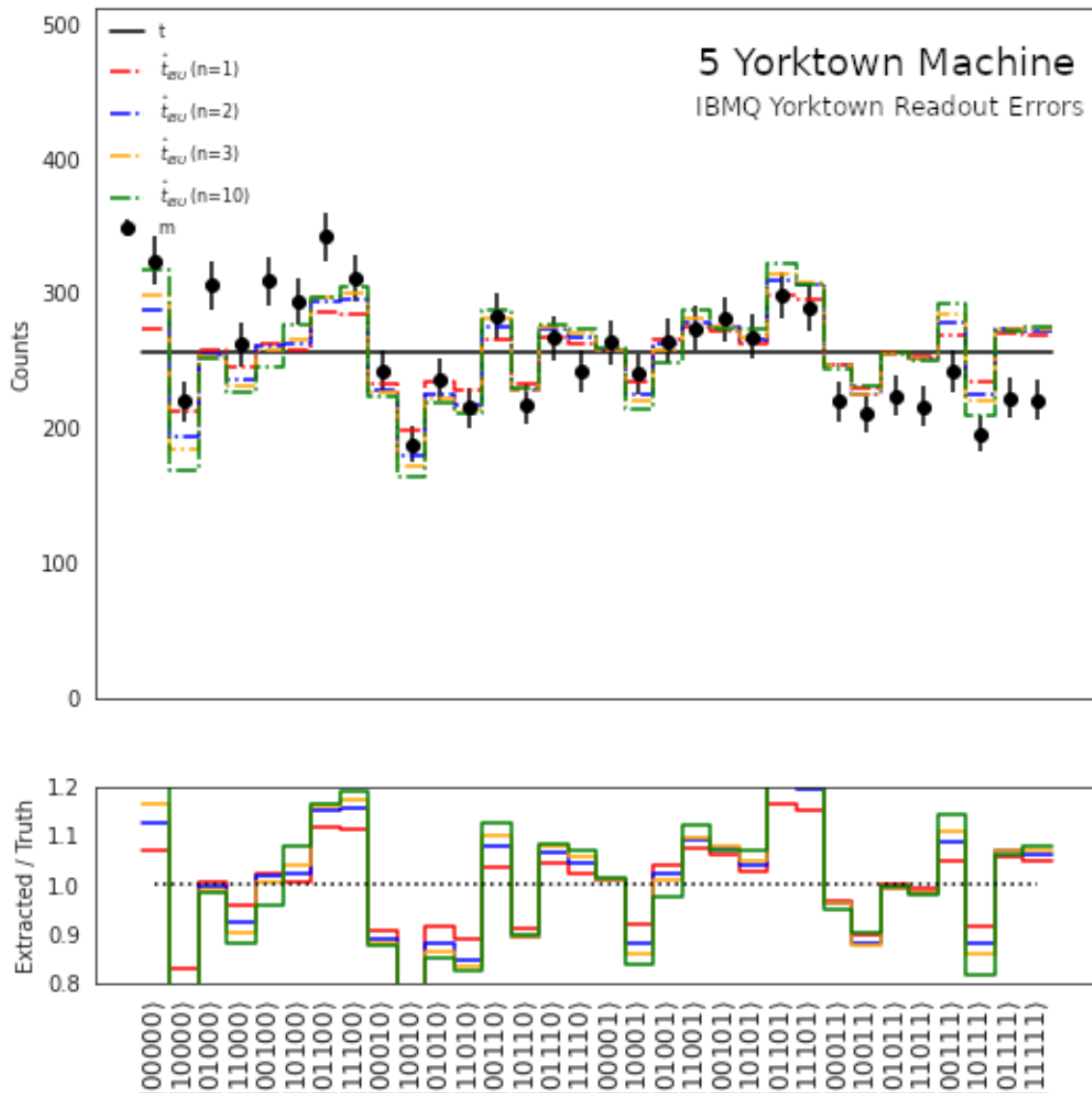


Figure 3.17: Unfolding results for a uniform distribution using IBU with 1,2,3,10 iterations and a uniform prior truth spectrum.

Now, for the Gaussian distribution:

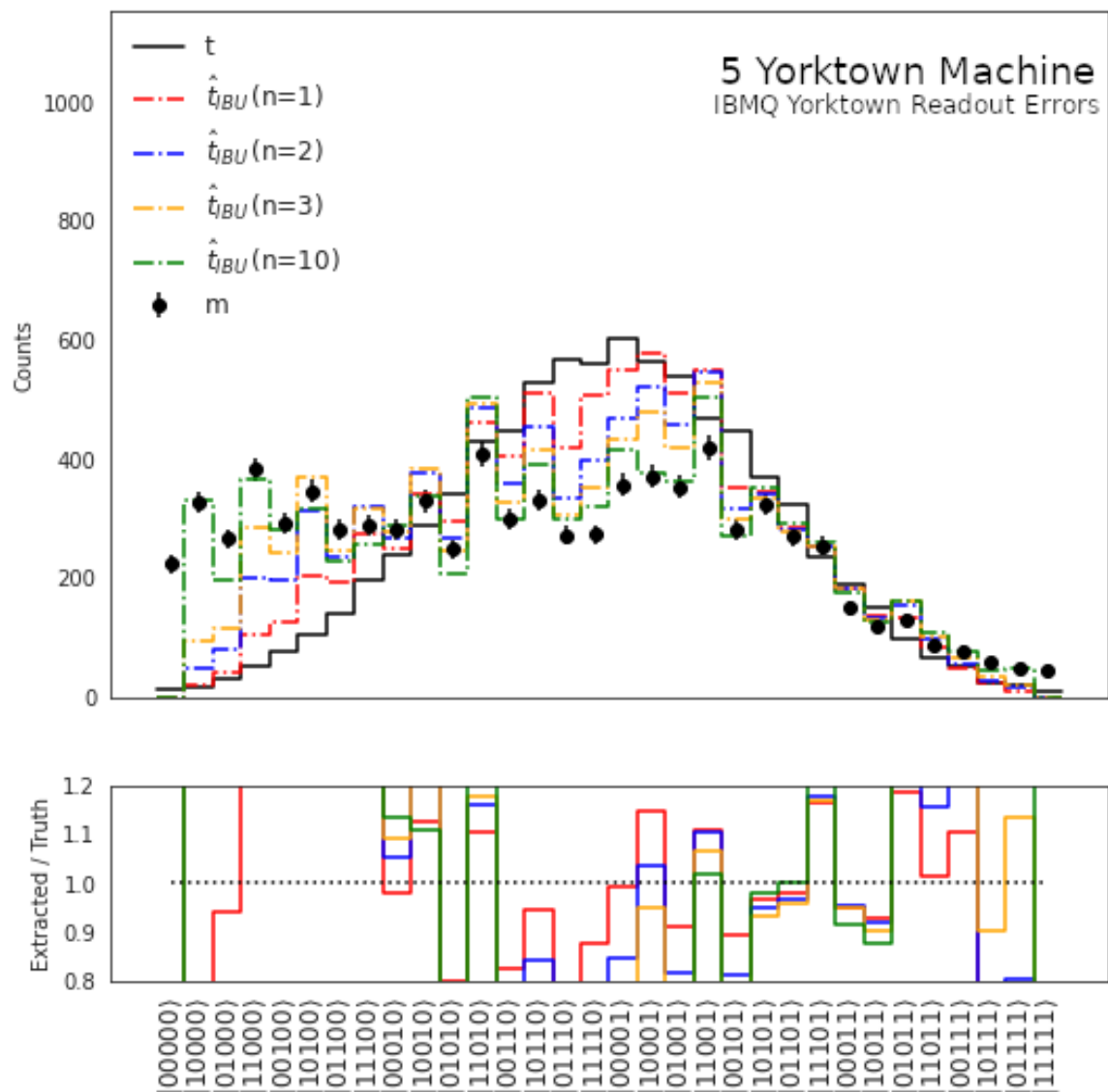


Figure 3.18: Unfolding results for a Gaussian distribution using IBU with 1,2,3,10 iterations and a prior truth spectrum $t_0=t_{\text{trig}}$.

Note that for the two distributions, the most efficient unfolding is for $n = 1$, where further iterations *diverge* more and more from the truth t until we finish with results identical to that obtained by MI and Ignis. Therefore, just 1 iteration is needed to get satisfactory results.

3.6 Unfolding Readout Errors from The Melbourne IBM Q Machine

In this section, we want to explore more about unfolding methods for the IBM Melbourne quantum computer. This quantum computer has 15 qubits, but we will construct a response matrix for 7 qubits (we can not visualize clearly the result for more than 7 qubits). Then, we will follow the same steps to unfold the readout errors from the Melbourne quantum machine and represent the result graphically.

The Response Matrix

To construct the response matrix we follow the same steps as the Yorktown QC with the only difference is that we should split our measurement calibration circuits generated by `complete_meas_cal`. We do this step because the number of experiments supported by the device is 75 which means for a number of qubits greater than 6 we should split the experiment. (75 is between $2^6 = 64$ and $2^7 = 128$).

We split our measurement calibration circuits generated by `complete_meas_cal` into two batches:

```

1 device=provider.get_backend('ibmq_16_melbourne')
2 job1_res = qiskit.execute(meas_calibs[0:64], backend=device,
  ↪ shots=8192,optimization_level=0).result()
3 job2_res = qiskit.execute(meas_calibs[64:128], backend=device,
  ↪ shots=8192,optimization_level=0).result()

```

Now, we use the results to initialize a measurement correction fitter with `CompleteMeasFitter` with the first 64 circuits:

```

1 meas_fitter = CompleteMeasFitter(job1_res, state_labels,circlabel='mcal')

```

In this step we use the `CompleteMeasFitter.add_data` function to update the measurement correction fitter with the rest of calibration circuit batches.

```

1 meas_fitter.add_data(new_results=job2_res)

```

We need to update the fitter with the last batch of calibration circuits to see the full calibration matrix. If we plot the calibration matrix after initializing the measurement correction fitter with the first batch of calibration circuits, we will see a half-calibration matrix.

Finally, we use the same code as Yorktown QC to plot the heatmap of the Melbourne response matrix, the only change is the number of the qubits. The matrix will have a dimension of $2^7 \times 2^7$.

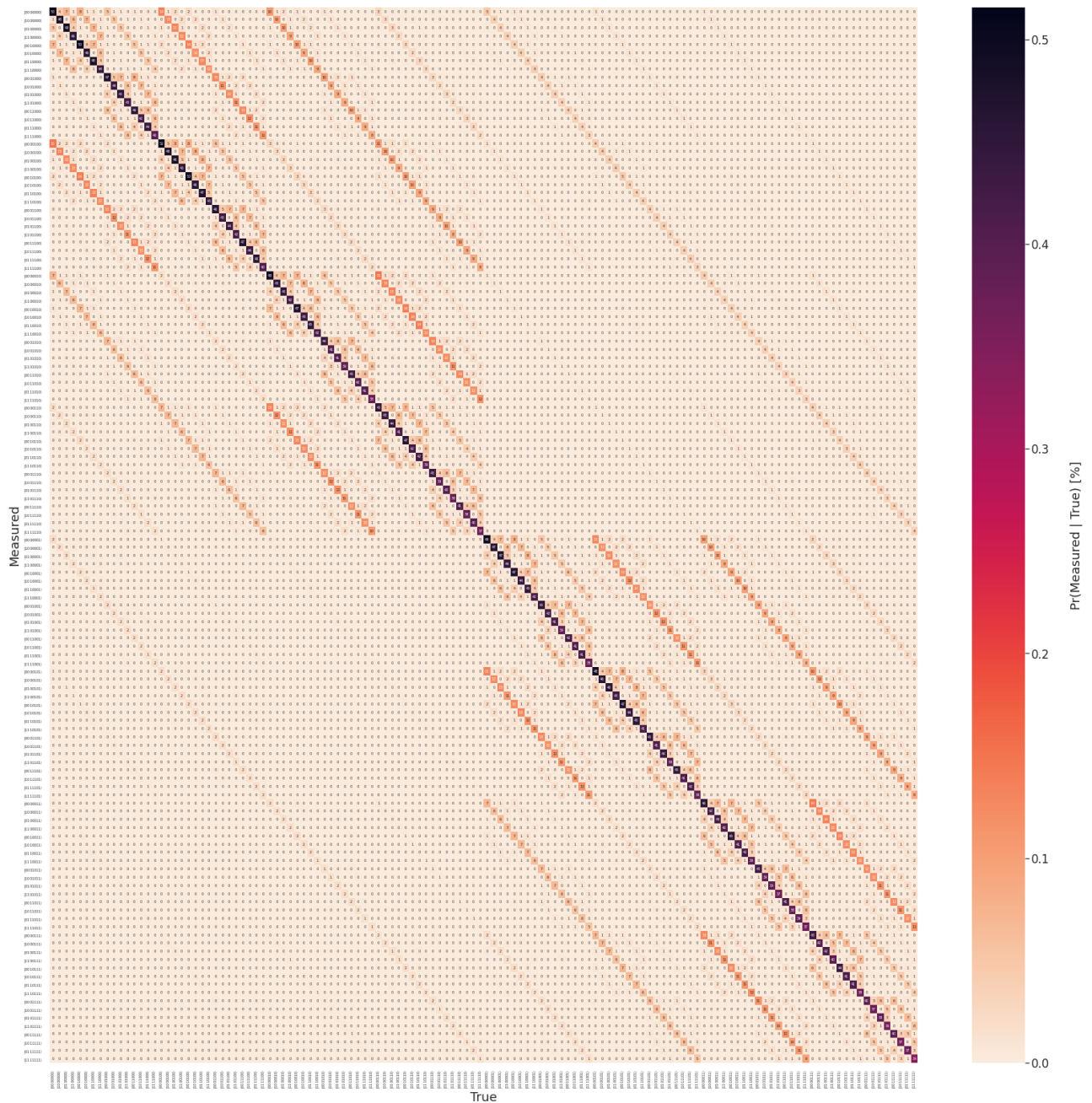


Figure 3.19: Response matrix for 7 qubits from the IBM Q Melbourne machine.

We notice that as the preview calibration matrix, there is a significant diagonal component with high probability and also many off-diagonal components with a non-zero probability of transitions.

Unfolding a Uniform Distribution

First, we create a quantum circuit that generates a uniform distribution. The output will be a uniform distribution only if quantum computers was perfect, but this is not the case. So, the distribution will be distorted by the noisy Melbourne quantum machine.

The circuit for seven qubits which generate uniform distribution is represented as follow:

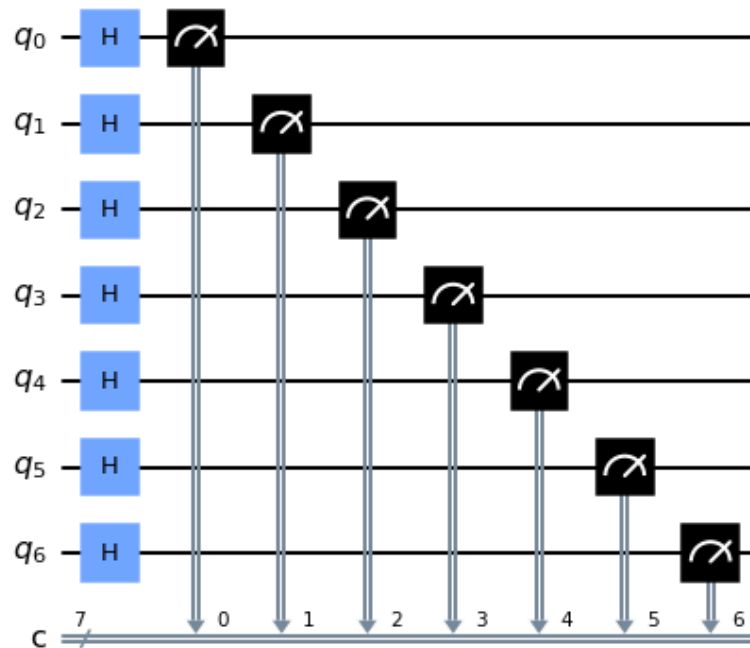


Figure 3.20: A circuit that generates a uniform distribution based on 7 qubits.

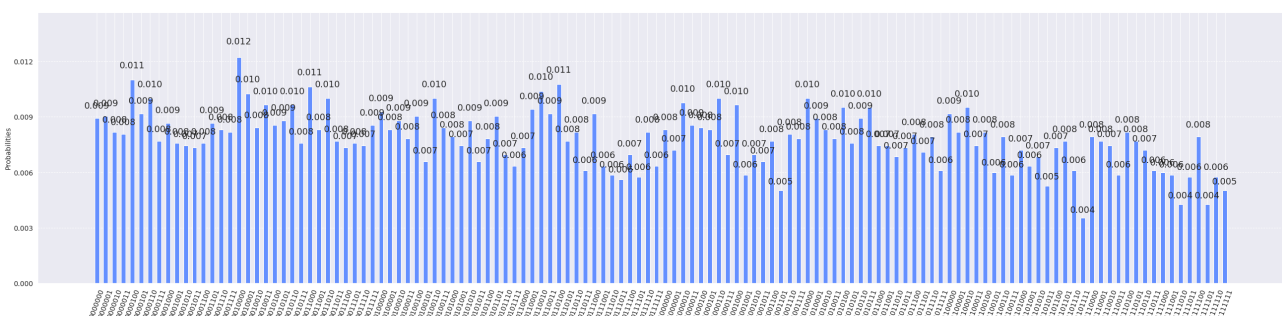
After the measurement, we organize the data and, extract the count. We get the following data from IBM Melbourne quantum computer:

```
1 [73, 74, 100, 84, 69, 79, 70, 72, 79, 62, 87, 68, 82, 63, 60, 62, 61, 70, 67, 76,
  ↪ 68, 72, 64, 74, 54, 82, 69, 65, 61, 72, 54, 64, 74, 57, 52, 66, 60, 77, 85, 75,
  ↪ 88, 63, 67, 50, 75, 52, 48, 46, 57, 47, 67, 52, 90, 68, 59, 80, 70, 69, 68, 82,
  ↪ 57, 79, 48, 57, 54, 63, 41, 66, 64, 75, 82, 73, 68, 64, 78, 62, 73, 78, 61, 61,
  ↪ 56, 60, 66, 58, 65, 50, 82, 75, 67, 78, 61, 67, 49, 65, 48, 59, 52, 56, 43, 60,
  ↪ 63, 50, 29, 63, 65, 63, 61, 48, 67, 63, 59, 50, 49, 48, 35, 47, 65, 35, 47, 41,
  ↪ 71, 62, 61, 60, 62, 71, 68, 67]
```

We can represent the result of the measurement in a histogram:

```
1 plot_histogram(m,figsize=(45,10))
```

The output of the plot:



We recall that our true distribution can be generated using the following code:

```
1 t_u = [8192/128 for i in range(128)]
```

The output after using the print function:

```
1 [64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0,
  ↪ 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0,
  ↪ 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0,
  ↪ 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0,
  ↪ 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0,
  ↪ 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0,
  ↪ 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0,
  ↪ 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0,
  ↪ 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0, 64.0]
```

Finally, after using the three unfolding methods, the result is represented in the following graphics:

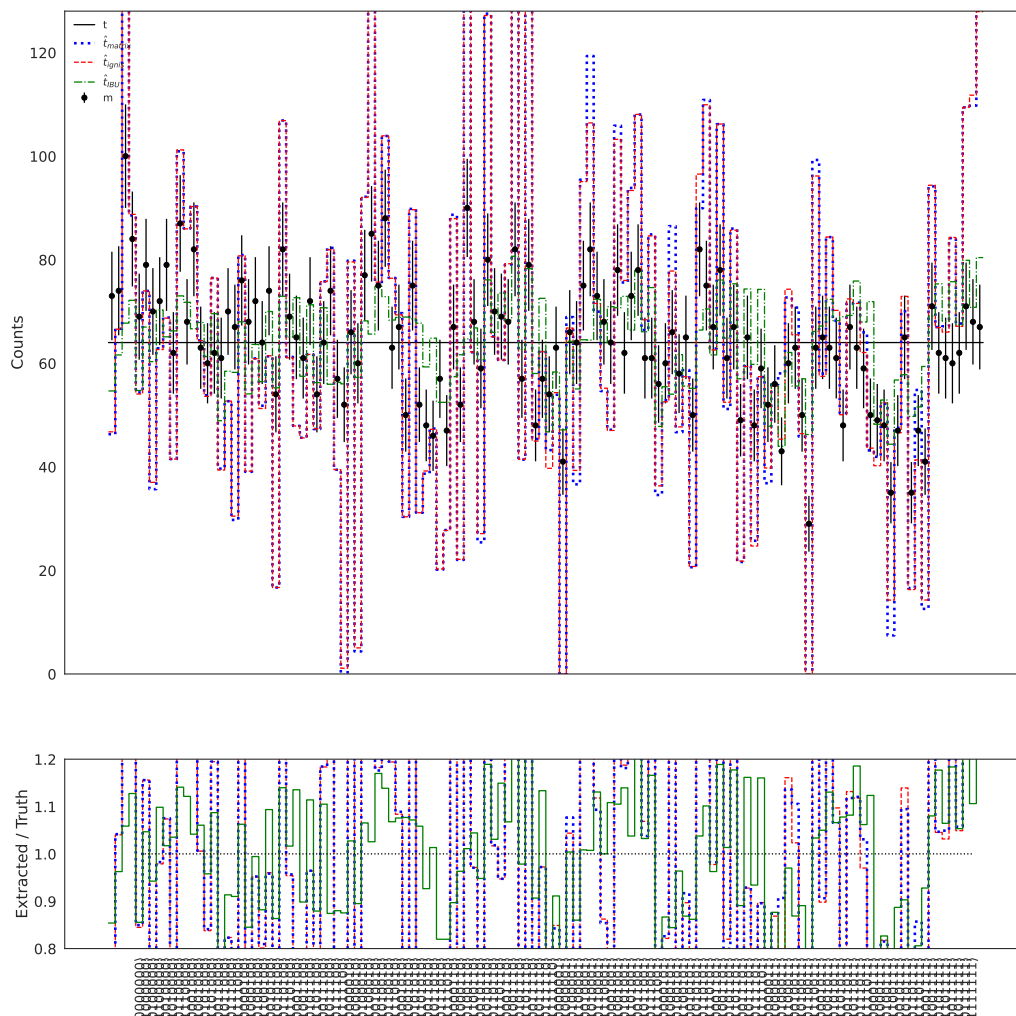


Figure 3.21: Unfolding results of a uniform distribution based on seven-qubits using 3 methods: Matrix inversion, ignis, and IBU with one iteration and a uniform prior truth spectrum.

This is clearly different from the previous situation of 5 Qubits. First, we notice that the uncertainties on the *measured distribution* values are very high. Secondly the two unfolding methods MI and **ignis** show *large oscillations*. Unlike the case of 5 Qubits, their values, in this case, are not completely *identical*. Neither technique was able to unfold this measured distribution. The IBU does not exhibit large oscillations, and it was able to unfold a certain extent although the great uncertainties on **m_u**.

So, after increasing the number of qubits we notice that the first two methods have difficulties to unfold; for the IBU, its ability to unfold has decreased but compared to the two other methods, its unfolding is important.

After these several experiences which generally aim to deal with readout noise of quantum computers, and especially to study the ITERATIVE BAYESIAN UNFOLDING method. We can recap what we have learned about this unfolding technique:

- IBU is very sensitive to the choice of the *prior truth distribution* \mathbf{t}^0 , as well as to the number of iterations n .
- Unfolding with the vector \mathbf{t}^0 or with $\lambda \cdot \mathbf{t}^0$ gives the same values, $\forall \lambda \in \mathbb{R}^*$.
- If the i^{th} component of the vector $\mathbf{t}^0 \equiv \mathbf{t}^0$ equal to 0 then the i^{th} component of the vector resulting from the unfolding process is equal to 0, i.e.

$$\mathbf{t}^0 = [\alpha, 0, \beta, \gamma, 0, 0] \xrightarrow{\text{IBU}} \hat{\mathbf{t}}_{\text{IBU}}^n = [\alpha', 0, \beta', \gamma', 0, 0]$$

- If the chosen \mathbf{t}^0 is of the same **nature** of the truth \mathbf{t} , then one have to take the **minimum** number of iterations, and vice versa.
- For a higher number of iterations ($n \gg 1$) all chosen *prior truth distributions* converge towards the same result. Moreover, this common result is identical to that obtained by MI and **ignis**.

Conclusion

The invention of the Quantum Computer represents a leap for human civilization, which will open new horizons and allow the emergence of a new era, not only in physics, but will contribute to the development of all sciences. However, current challenges limit its effectiveness, and one of the important limitations is the rate of errors.

In this master thesis, our objective was to mitigate Quantum Computer Readout Errors, using of Unfolding Methods from the field of High Energy Physics. To achieve this goal, we first had to understand the basic concepts of Classical and Quantum Information Science QIS at the same time. The final subject in this overview was Classical and Quantum Error Detection and Correction, where we explored the differences between the classical and quantum situation.

After this we visited the field of High Energy Physics, to understand the concept of unfolding, where we saw the simplest method called “Matrix Inversion” and a more advanced one called “Iterative Bayesian Unfolding” (IBU), which represented a promising technique for mitigating readout errors in QIS.

Then we investigated readout errors in the field of QIS, where we tried to understand their nature and their sources and how to construct the response matrix in order to mitigate these errors.

Finally, we studied three unfolding methods (MI, ignis, IBU) through a real Quantum Computer Readout Noise, for a uniform and Gaussian distribution, and we found that it gives an excellent unfolding compared to MI and ignis, which contribute to confirm the results of the researchers BENJAMIN NACHMAN et al. in 2020. After achieving this, we took a small step beyond these results to explore in particular the behavior of IBU in a larger system of qubits, where we found that IBU maintained its effectiveness.

Our work is a small step to move beyond the era of Noisy Intermediate-Scale Quantum computers. Further studies are needed to overcome the noise of quantum computers.

“We can only see a short distance ahead, but we can see plenty there that needs to be done.”

ALAN TURING

Appendices

A

Boolean Algebra

A boolean algebra represents the mathematical rules between two-state quantities to solve logical problems. These quantities are known as **Boolean logic variables** and can be represented by 0 and 1 or TRUE and FALSE. Boolean algebra use only three operators: NOT, AND, OR. The AND and OR operators work on two or more variable, the NOT operator operate on a single variable. These operators are represented in table A.1 as follow [30]:

Operator	Symbol	Usage	Spoken as
NOT	-	\bar{A}	A bar , or not A
OR	+	$A + B$	A or B
AND	.	$A.B$	A and B

Table A.1: Boolean variables and operators.

- **The NOT Operator:** A boolean variable can be 0 or 1. If we want to switch the state from 0 to 1, we need an operator that produces the complement of the variable 0. In the same way, we apply NOT to the state 1 to get 0. In the general form, every variable A has an opposite value denoted as A -not or A -bar, or A -prime. If $A = 1$ then $\bar{A} = 0$ or if $A = 0$ then $\bar{A} = 1$.

- **The OR Operator:** The OR operator acts on at least two inputs variables and produces only one output. The result depends on the variables; if there is at least one variable with the value 1, the result will be 1. If all the variables are 0, the result will be 0. The equation for OR operation (for two variables) is $A + B = R$ with A and B the inputs, and they take the value 0 or 1. The «+» sign stands for the OR operation and R for the boolean result. We can see that this logical function acts according to these two boolean conditions: $X + 1 = 1$ and $X + 0 = X$. The input X takes the value 0 or 1 [31]. We will see this condition in the *Boolean laws*.

- **The AND Operator:** The AND operator works as a product of two or more variables. If all variables are 1, the result will also be 1. However, if at least one of the variables is 0, the result will be 0. The equation for AND operation (for two variables) is $A.B = R$ with A , B , the input, and they take the value 0 or 1. The «.» sign stands for the AND operation, and R is the boolean result. The AND function acts according to the following relation: $X.1 = X$ and $X.0 = 0$. The variable X takes the value 0 or 1.

B

Boolean Laws

In this part, we will see the boolean laws that describe how operators act on boolean variables. There are two categories of boolean laws; the first describes operations on a single variable, and the second, for multivariable [30].

1. Single variable

This part will explain the rules that describe the logic operation for only one variable. There are three laws for a single variable: idempotent, Inverse, Involution. These laws are represented as follow:

- **Idempotent**

Idem means “*same*” in Latin. These laws describe the effect of an operator when the same variable goes to all inputs. Let’s take two-input OR for example: $A + A = S$ this will give 1 if $A = 1$, and 0 if $A = 0$; therefore $A + A = A = S$.

- **Inverse elements**

This law is about the effect of an operator on a variable A with its complement \bar{A} . Let us take AND as example, the expression is written as follow: $R = A.\bar{A}$. If $R = 0$ then either A or \bar{A} must have 0 because they have a complementary values.

- **Involution law**

When we apply the NOT for the first time on a variable A , we obtain \bar{A} , but by applying it for the second time, this will return the variable to its original state $\bar{\bar{A}} = A$.

2. Multivariable

Multi-variable theorems are about operations on more than one variable, and these laws are: *commutative law, associative, distributive, De Morgan’s laws, Absorption.*

- **Commutative law**

when applying the operators OR, AND between two variables; the order of these variables does not matter. So $R = A.B = B.A$ and $S = A + B = B + A$.

- **Associative law**

These laws show how the operations can be grouped, and it is the same as for conventional algebra. For example: $A + (B + C) = (A + B) + C$ and $A.(B.C) = (A.B).C$.

- **Distributive law**

Distribution is used to expand and simplify the boolean expression. Example: $A.(B + C) = A.B + A.C$.

- **De Morgan's law**

This theorem is about *complementation* and announces that, complementing $A.B$ is equivalent to $\bar{A} + \bar{B}$ and complementing $A + B$ is equivalent to $\bar{A}.\bar{B}$.

- **Absorption**

These laws are used for *circuit simplification* to increase the *reliability* of the logic circuit and decrease the cost of manufacture. To understand the laws of Absorption, let us develop some expressions.

$$A + A.B = A.(1 + B) = A.1 = A$$

$$A.(A + B) = A.A + A.B = A + A.B = A.(1 + B) = A$$

$$A.(\bar{A} + B) = A.\bar{A} + A.B = 0 + A.B = A.B$$

$$A.B + \bar{B} = \bar{B}.(B.A.B + 1) = \bar{B}.(B.A + 1) = A + \bar{B}$$

$$A + \bar{A}.B = (\bar{A} + A).(A + B) = 1.(A + B) = A + B$$

C

Qubit Design

Due to the complexity of quantum mechanics, it is not easy to build a practical quantum computer. There is great competition and enormous effort from large companies to accelerate the development of quantum programming. Today several types of Qubits are available. We site the most common types of qubits:

- **Superconducting Loops**

Superconducting quantum computing is an implementation of a quantum computer in superconducting electronic circuits. Superconducting quantum circuits use Josephson junction to create a *non-linear* inductance which allows a design of *anharmonic oscillators*. A quantum harmonic oscillator cannot be used as a qubit, as there is no way to address only two of its states. A microwave pulse is injected which excites the current into a superposition of states. This design is used by IBM.

- **Trapped Ions**

Ion trap uses a combination of electric or magnetic fields to capture ions or charged atomic particles in a system (isolated from the external environment). Qubits are stored in stable electronic states of each ion, and quantum information can be transferred through interacting by the Coulomb force. To induce coupling between the qubit states, we apply lasers for single-qubit operations or entanglement between qubits.

- **Silicon Quantum Dots**

Quantum dots are semiconductor devices at the nanoscale. Electrons are vertically confined to the ground state of quantum *gallium arsenide* (GaAs) and form a two-dimensional electron gas (2DEG). This gas is free to move in two dimensions but tightly confined in the third. This confinement leads to quantized energy levels for movement in the third direction.

- **Diamond Vacancies**

This multiqubit system is based on a *nitrogen-vacancy* (NV) center in diamond. The NV center's electron spin constitutes the central qubit, while the nuclear spins of the nitrogen and surrounding carbon atoms form the other qubits. Electron spin responds optically, so the quantum state can be programmed and read out quickly using a laser, the electron spin is used as a bus, that can be coupled through magnetic interaction. For quantum memory and processing,

we use the nuclear spin, which does not respond optically and has longer coherence times.

Superconducting Qubit

Qubit design significantly impacts creating devices with high performance, like long coherence time and high controllability. There are many quantum computing platforms, but the most famous one is based on **superconducting Qubits**. Because of its advantages, superconducting Qubit became the leading candidate for a scalable quantum processor. Like any quantum mechanical object, superconducting qubits can be controlled, placed into quantum superposition states, exhibit quantum interference effects, and become entangled [32].

- **What is superconducting quantum computing?**

A superconducting quantum computer is a connected artificial atom (rather than natural atoms). Each artificial atom is represented as a non-linear inductor-capacitor circuit. The inductor-capacitor circuits are harmonic and the energy levels are equally spaced if we add a non-linear element in the circuit; therefore, the energy separation between pairs of levels differ from each other, therefore the transitions between pairs of energy levels can be separately addressed [33].

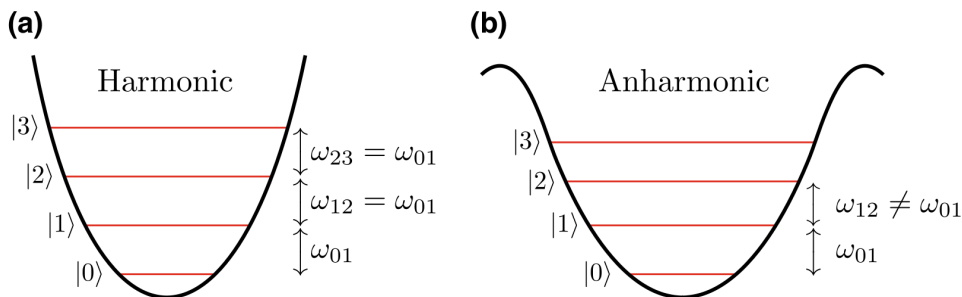


Figure C.1: Harmonic and anharmonic systems and their suitability as qubits [34].

(a) In the quadratic potential (black curve) of a harmonic system, the energy levels (red lines) are equally spaced, i.e. $\omega_{j,j+1} = \omega_{01}$, where ω_{jk} is the transition frequency between energy levels j and k . A signal at frequency ω_{01} will thus not only transfer population from $|0\rangle$ to $|1\rangle$, but also from $|1\rangle$ to $|2\rangle$, etc.

(b) In the potential of an anharmonic system, e.g., the cosine potential characteristic of a Josephson junction, $\omega_{01} \neq \omega_{12}$. A signal at frequency ω_{01} will thus only drive transitions between $|0\rangle$ to $|1\rangle$ and not affect any other levels in the system (provided that the signal is not too strong). This limits the dynamics to the two-level system formed by $|0\rangle$ and $|1\rangle$, which can be interpreted as a qubit [33].

- **Why we use superconductors?**

A superconductor enables an electric current in a circuit loop to circulate without resistance. Depends on the inductance and the capacitance of the electric circuit, the electric field oscillates

with a specific frequency and the electric current experiences no resistance during its oscillation and is called a *supercurrent*. If the supercurrent is counter-clockwise then it generates an upward magnetic flux, and a clockwise supercurrent generates a downward magnetic flux. Superposing clockwise and counter-clockwise supercurrents creates a superposition of up and down magnetic fields. For a small circuit, the magnetic flux is quantized and equally spaced which does not help make a two-level qubit. To make the spacing unequal, a Josephson junction is placed into the circuit.

- **What is Josephson junction?**

A Josephson junction contains two superconducting materials weakly linked via another medium, such as an insulator. The supercurrent in a loop faces an energy barrier at this weak link, which modifies the *harmonic supercurrent oscillation* to an *anharmonic dynamic*. This results in unequally spaced magnetic flux levels. For sufficiently high anharmonicity, just the two lowest levels play a role in the dynamics so we have an effective two-level atom that functions as a flux qubit. The downward magnetic flux serves as $|0\rangle$ and the upward flux as $|1\rangle$ [35].

The Basic Josephson-Junction Qubits

There are three basic designs for Josephson-junction qubits. The three are known as a charge qubit, a flux qubit, and a phase qubit. The charge qubit is a box for charge, controlled by an external voltage V_g ; the flux qubit is a loop controlled by an external magnetic flux ϕ_{ext} ; and the phase qubit is a Josephson junction biased by a current I_b [34].

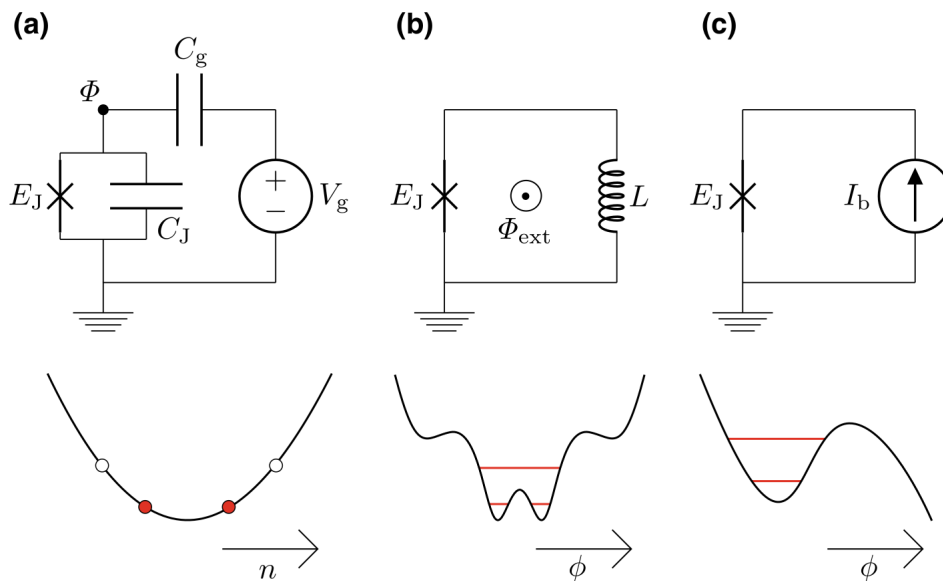


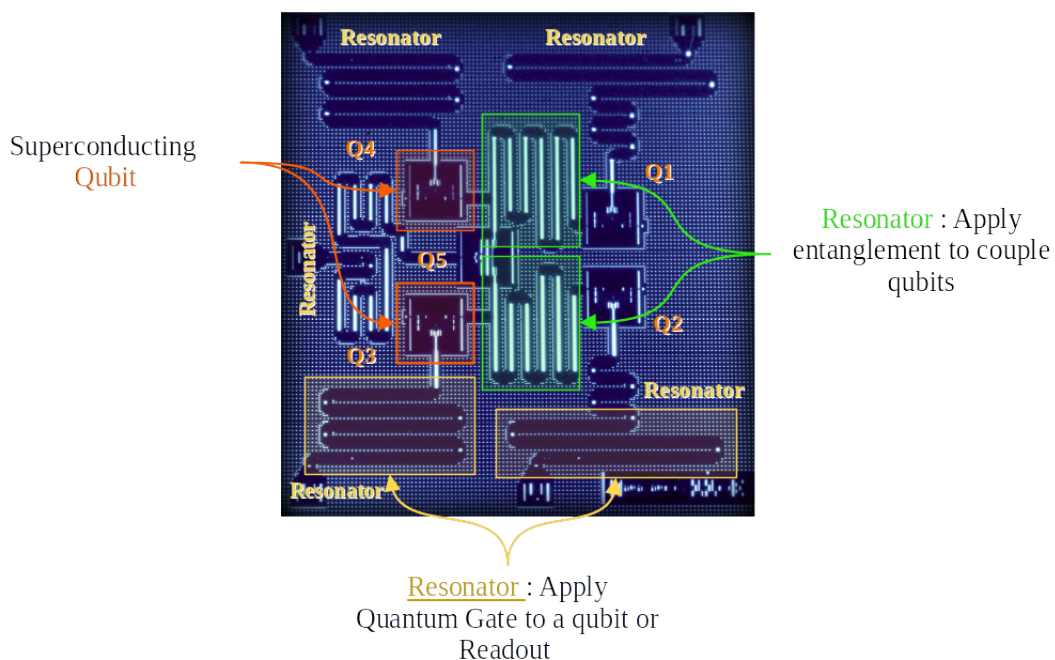
Figure C.2: The three basic **Josephson-junction** qubit circuits and their potential-energy landscapes, with the two lowest energy levels marked in red. **(a) Charge qubit.** **(b) Flux qubit.** **(c) Phase qubit.** For simplicity, the capacitance C_J is only shown in panel **(a)**, although it is also present in the circuits in panels **(b)** and **(c)** [34].

Quantum Processor

IBM's quantum chips consist of a **silicon wafer** that contains qubits and resonators composed of **aluminum** and **niobium**. A resonator is connected to each qubit to deliver the control pulses that govern the qubit's state.

An entangled qubit's state cannot be described independently of its paired qubit. If the state of the first qubit is measured, measuring the state of the second qubit is correlated to that of the first. Practical quantum computations often involve executing conditional logic operations in which one qubit's state depends on that of another qubit.

To determine the result of a qubit's operation, a measurement signal is sent as a **microwave pulse** to the qubit's resonator. The measurement destroys the qubit's superposition and collapses the qubit's state to 0 or 1. The reflected signal travels back up the **dilution refrigerator**, passing through *isolators* that protect the qubits from noise as well as *amplifiers* that boost the signal's strength enough to read the qubit's state accurately. The signal then passes back through the stack of **microwave electronics**, this time to be converted from **microwave frequency** to a digital signal that is sent over the internet to the local computer to display the result of the quantum calculation¹.



¹<https://electronics360.globalspec.com/article/13553/how-quantum-computers-work>

D

Python and Qiskit toolkit

In this part, we will see the most used python and qiskit functions.

Qiskit packages

- [Aer](#): is an element that provides quantum computing simulators with realistic noise models.
- [ClassicalRegister](#): Implement a classical register.
- [QuantumCircuit](#): Holds quantum operations and instructions for the quantum system.
- [QuantumRegister](#): Implement a quantum register.
- [qiskit.ignis.mitigation.measurement](#): The Qiskit Ignis Measurement Calibration module and mitigation is used to calibrate measurement errors via a series of basis state measurements and then apply that calibration to correct the average results of another experiment of interest.
- [qiskit.providers.aer](#): This module contains classes and functions to build a noise model for simulating a Qiskit quantum circuit in the presence of errors.
- [readout_error](#): Readout error class for Qiskit Aer noise model.

Python tools

- [NumPy](#): which stands for Numerical Python, is a library consisting of multidimensional array objects and a collection of routines for processing those arrays. Using NumPy, mathematical and logical operations on arrays can be performed.
- [SLSQP](#): Sequential Least-Squares Quadratic Programming algorithm for nonlinearly constrained, gradient-based optimization, supporting both equality and inequality constraints.
- [SciPy optimize](#): provides functions for minimizing (or maximizing) objective functions, possibly subject to constraints. It includes solvers for nonlinear problems (with support

for both local and global optimization algorithms), linear programming, constrained and nonlinear least-squares, root finding, and curve fitting.

- [len](#): Return the number of items in a list.
- [make_axes_locatable](#): takes an existing axes instance and creates a divider for it. It provides `append_axes()` method that creates a new axes on the given side (“top”, “right”, “bottom” and “left”) of the original axes.
- [matplotlib.gridspec](#): is used to specify the grid’s geometry to place a subplot. We need to set the number of rows and columns of the grid.
- [matplotlib.lines](#): This module contains all the 2D line class which can draw with a variety of line styles, markers and colors.
- [matplotlib.pyplot](#): is a collection of command style functions that make Matplotlib work like MATLAB. Each Pyplot function makes some change to a figure. For example, a function creates a figure, a plotting area in a figure, plots some lines in a plotting area, decorates the plot with labels, etc.
- [matplotlib](#): is a library for creating static, animated, and interactive visualizations in Python.
- [mpl_toolkits.axes_grid1](#): is a collection of helper classes to ease displaying (multiple) images with matplotlib. In matplotlib, the axes location (and size) is specified in the normalized figure coordinates, which may not be ideal for displaying images that needs to have a given aspect ratio.
- [np.clip](#): function is used to Clip (limit) the values in an array. Given an interval, values outside the interval are clipped to the interval edges. For example, if an interval of $[0, 1]$ is specified, values smaller than 0 become 0, and values larger than 1 become 1.
- [np.linspace](#): Return evenly spaced numbers over a specified interval.
- [np.matmul](#): is Matrix product of two arrays.
- [np.random.rand](#): Create an array of a given shape and populate it with random samples from a uniform distribution over $[0, 1)$.
- [np.random.randn](#): Return a sample (or samples) from the “standard normal” distribution.
- [np.random.seed\(\)](#): generate random numbers based on “pseudo-random number generators” algorithms. The `seed()` is used to initialize the random number generator to be able to generate a random number.

- `np.ravel`: return 1D array with all the input-array elements and with the same type as it.
- `plt.hist`: function in pyplot module of matplotlib library is used to plot a histogram.
- `random.uniform(low, high, size)`: Draw samples from a uniform distribution. Samples are uniformly distributed over the half-open interval [low, high) (includes low, but excludes high). In other words, any value within the given interval is equally likely to be drawn by uniform.
- `round`: returns a floating point number that is a rounded version of the specified number.
- `seaborn`: Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.
- `sum()`: to calculate the sum of an iterable like range and list.
- `tol`: is the tolerance for the stopping criteria. This tells scikit to stop searching for a minimum (or maximum) once some tolerance is achieved.
- `tuple`: Tuples are used to store multiple items in a single variable. Items are ordered, unchangeable, and allow duplicate values.

E

Glossary

|A|B|C|D|E|F|G|H|I|J|K|L|M|N|O|P|Q|R|S|T|U|V|W|X|Y|Z|

- **Aer:**

Aer is an element of **Qiskit**, which provides *high-performance quantum computing **simulators*** with realistic **noise models**.

- **Backend:**

The term backend can refer to either a **quantum computer** or a *high-performance classical **simulator*** of a quantum computer.

- **Classical Register:**

A classical register consists of bits that can be written to and read within the coherence time of the quantum circuit.

- **Frequency:**

In IBM Quantum, a *frequency* (or *counts*) of a **job** is defined as the number of times a given *basis state* is **observed** after measurement.

- **Job:**

A job ties together all of the relevant information about a computation on IBM Q: a **quantum circuit**, choice of **backend**, the choice of how many **shots** to execute on the backend, and the results upon executing the quantum circuit on the backend.

- **Jupyter Notebook:**

Jupyter Notebook is a *web-based interactive* computational environment for creating notebook documents.

- **Hub:**

A hub is the top level of a given organization and contains within it one or more groups. These groups are in turn populated with projects.

- **IBM Quantum Experience:**

IBM Q Experience is an *online platform* that allow public and premium access to **cloud-based quantum computing** services. Users interact with a quantum processor through the quantum circuit model of computation. Circuits can be created either using a *graphically* with the **Quantum Composer**, or *programmatically* within the **Jupyter notebooks** of the **Quantum Lab**.

- **Provider:**

Access to the various services offered by IBM Q is controlled by the providers to which you are assigned. A provider is defined by a *hierarchical organization* of **hub**, **group**, and **project**. Users can belong to more than one provider at any given time.

- **Python:**

Python is an interpreted, multi-paradigm, cross-platform **programming language**.

- **Qiskit:**

Qiskit is an *open-source software development kit* (SDK) for working with **quantum computers** at the level of **circuits**, **pulses**, and **algorithms**. It provides tools for creating and manipulating quantum programs and running them on prototype **quantum devices** on IBM QUANTUM EXPERIENCE or on **simulators** on a local computer.

- **Quantum Circuit:**

A quantum circuit is an ordered sequence of **quantum gates**, **measurements**, and **resets**.

- **Quantum Gate:**

A quantum gate is a **unitary** operation that can act on a single or multiple qubits.

- **Qubit:**

A qubit is the **basic** unit of *quantum information*. A qubit is defined as a *two-level* quantum system.

- **Quantum Register:**

A quantum register is a collection of qubits on which gates and other operations act.

- **Shot:**

The number of shots defines how many times a quantum circuit is **executed** in a quantum computer.

References

- [1] Richard P. Feynman. “Simulating physics with computers”. In: *Int J Theor Phys* 21.6 (June 1982). DOI: [10.1007/BF02650179](https://doi.org/10.1007/BF02650179).
- [2] Dan C. Marinescu. *Classical and Quantum Information*. Elsevier, Academic Press, Feb. 2011. DOI: [10.1016/C2009-0-64195-7](https://doi.org/10.1016/C2009-0-64195-7).
- [3] P. A. M. Dirac. *The Principles of Quantum Mechanics (International Series of Monographs on Physics)*. Oxford University Press, USA, Feb. 1967.
- [4] Ryszard Horodecki et al. “Quantum entanglement”. In: *Rev Mod Phys* 81.2 (June 2009). DOI: [10.1103/RevModPhys.81.865](https://doi.org/10.1103/RevModPhys.81.865).
- [5] Barton Zwiebach. *Lecture 19: Multiparticle States and Tensor Products (cont.)*, MIT OpenCourseWare. Dec. 2013. URL: https://ocw.mit.edu/courses/physics/8-05-quantum-physics-ii-fall-2013/lecture-notes/MIT8_05F13_Chap_08.pdf.
- [6] A. Einstein et al. “Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?” In: *Phys Rev* 47.10 (May 1935). DOI: [10.1103/PhysRev.47.777](https://doi.org/10.1103/PhysRev.47.777).
- [7] Gregg Jaeger. *Entanglement, Information, and the Interpretation of Quantum Mechanics (The Frontiers Collection)*. Springer, June 2009. DOI: [10.1007/978-3-540-92128-8](https://doi.org/10.1007/978-3-540-92128-8).
- [8] Robert S. Sutor. *Dancing with Qubits*. Packt Publishing, Nov. 2019.
- [9] P. W. Shor. *Algorithms for quantum computation: discrete logarithms and factoring*. IEEE, Nov. 1994. DOI: [10.1109/SFCS.1994.365700](https://doi.org/10.1109/SFCS.1994.365700).
- [10] Barry C. Sanders. *How to Build a Quantum Computer*. IOP Publishing, Nov. 2017. DOI: [10.1088/978-0-7503-1536-4](https://doi.org/10.1088/978-0-7503-1536-4).
- [11] Bilal Shaw et al. “Encoding One Logical Qubit Into Six Physical Qubits”. In: *arXiv* (). DOI: [10.1103/PhysRevA.78.012337](https://doi.org/10.1103/PhysRevA.78.012337).
- [12] David P. DiVincenzo et al. “The Physical Implementation of Quantum Computation”. In: *arXiv* (Feb. 2000). DOI: [10.1002/1521-3978\(200009\)48:9/11<771::AID-PROP771>3.0.CO;2-E](https://doi.org/10.1002/1521-3978(200009)48:9/11<771::AID-PROP771>3.0.CO;2-E).
- [13] Hiroyuki Sagawa et al. *Fundamentals of Quantum Information*. Nov. 2011. DOI: [10.1142/7915](https://doi.org/10.1142/7915).
- [14] Ashok Chatterjee. “Introduction to Quantum Computation”. In: *arXiv* (Dec. 2003). URL: <https://arxiv.org/abs/quant-ph/0312111v2>.
- [15] Christine Corbett Moran. *Mastering Quantum Computing with IBM QX*. ial xn-3ug : xn-2ug Packt Publishing, Jan. 2019.
- [16] Liu Ricky Ini. *Shor’s algorithm*. May 2018. URL: <https://riliu.math.ncsu.edu/437/notes3se4.html>.
- [17] Dieter Suter Joachim Stolze. *Quantum computing: a short course from theory to experiment*. Wiley-VCH, 2004.
- [18] Aziz Mouzali and Fatiha Merazka. “Simulation on maple of the nine qubit Shor code using Feynman program”. In: *ResearchGate* (Dec. 2009). URL: https://www.researchgate.net/publication/255661084_Simulation_on_maple_of_the_nine_qubit_Shor_code_using_Feynman_program.
- [19] Aziz Mouzali and Fatiha Merazka. “Quantum correction with three codes”. In: *arXiv* (Sept. 2014). URL: <https://arxiv.org/abs/1409.7853v1>.
- [20] Benjamin Nachman et al. “Unfolding quantum computer readout noise”. In: *npj Quantum Inf* 6.84 (Sept. 2020). DOI: [10.1038/s41534-020-00309-7](https://doi.org/10.1038/s41534-020-00309-7).

- [21] Frank Leymann et al. “The Bitter Truth About Quantum Algorithms in the NISQ Era”. In: *Quantum Sci Technol* (Sept. 2020). DOI: [10.1088/2058-9565/abae7d](https://doi.org/10.1088/2058-9565/abae7d).
- [22] Volker Blobel. “Unfolding methods in particle physics”. In: *CERN Document Server* (Jan. 2011). DOI: [10.5170/CERN-2011-006.240](https://doi.org/10.5170/CERN-2011-006.240).
- [23] Mikael Kuusela. *Introduction to Unfolding in High Energy Physics*. July 2014. URL: <https://indico.phys.ethz.ch/event/5/contributions/52/attachments/35/47/slides.pdf>.
- [24] Glen Cowan. “A survey of unfolding methods for particle physics”. In: *Psychology Journal* (Jan. 2002). URL: https://www.researchgate.net/publication/255672407_A_SURVEY_OF_UNFOLDING_METHODS_FOR_PARTICLE_PHYSICS.
- [25] Scott Oser. *Deconvolution and Unfolding*. Nov. 2010. URL: https://phas.ubc.ca/~oser/p509/Lec_21.pdf.
- [26] G. D’Agostini. “A multidimensional unfolding method based on Bayes’ theorem”. In: *Nucl Instrum Methods Phys Res., Sect A* 362.2 (Aug. 1995). DOI: [10.1016/0168-9002\(95\)00274-X](https://doi.org/10.1016/0168-9002(95)00274-X).
- [27] L. B. Lucy. “An iterative technique for the rectification of observed distributions”. In: *Astron. J* 79 (1974). DOI: [10.1086/111605](https://doi.org/10.1086/111605).
- [28] William Hadley Richardson. “Bayesian-Based Iterative Method of Image Restoration”. In: *Journal of the Optical Society of America* 62.1 (Jan. 1972). DOI: [10.1364/JOSA.62.000055](https://doi.org/10.1364/JOSA.62.000055).
- [29] Lydia Brenner et al. “Comparison of unfolding methods using RooFitUnfold”. In: *Int J Mod Phys A* 35.24 (Aug. 2020). DOI: [10.1142/S0217751X20501456](https://doi.org/10.1142/S0217751X20501456).
- [30] Barrie Hayes-Gill John Crowe. *Introduction to Digital Electronics*. Butterworth-Heinemann, 1998. DOI: [10.1016/B978-0-340-64570-3.X5000-X](https://doi.org/10.1016/B978-0-340-64570-3.X5000-X).
- [31] Mandal. *Digital Electronics: Principles And Applications*. MC GRAW HILL INDIA, Jan. 2010.
- [32] William D. Oliver et al. “Materials in superconducting quantum bits”. In: *MRS Bull* 38.10 (Oct. 2013). DOI: [10.1557/mrs.2013.229](https://doi.org/10.1557/mrs.2013.229).
- [33] P. Krantz et al. “A quantum engineer’s guide to superconducting qubits”. In: *Appl Phys Rev* 6.2 (June 2019). DOI: [10.1063/1.5089550](https://doi.org/10.1063/1.5089550).
- [34] Anton Frisk Kockum et al. “Quantum bits with Josephson junctions”. In: *SpringerLink* (Aug. 2019). DOI: [10.1007/978-3-030-20726-7_17](https://doi.org/10.1007/978-3-030-20726-7_17).
- [35] Barry C. Sanders. “How to Build a Quantum Computer”. In: (Nov. 2017). DOI: [10.1088/978-0-7503-1536-4ch1](https://doi.org/10.1088/978-0-7503-1536-4ch1).