Instytut Informatyki Teoretycznej i Stosowanej
Polskiej Akademii Nauk

Probabilistyczne kwantowe kody korekcyjne w zastosowaniu do ogólnych kanałów szumu

Rozprawa doktorska

mgr Ryszard Kukulski
Promotor:
dr hab. Zbigniew Puchała, prof. IITiS PAN
Promotor pomocniczy:
dr hab. inż. Łukasz Pawela, prof. IITiS PAN

Gliwice, 2023

Institute of Theoretical and Applied Informatics, Polish Academy of Sciences

Probabilistic quantum error correction codes for general noise channels

Doctoral dissertation

Ryszard Kukulski, MSc
Supervisor:
Zbigniew Puchała, PhD
Co-supervisor:
Łukasz Pawela, PhD

Gliwice, 2023

"The essence of mathematics is in its freedom."

– Georg Cantor

Contents

List of publications

Published work

1. Glos, A., Krawiec, A., Kukulski, R., & Puchała, Z. (2018). *Vertices cannot be hidden from quantum spatial search for almost all random graphs*. Quantum Information Processing, 17(4), 1-15., arXiv:1709.06829, DOI:10.1007/s11128-018-1844-7,

2. Puchała, Z., Pawela, Ł., Krawiec, A., & Kukulski, R. (2018). *Strategies for optimal single-shot discrimination of quantum measurements*. Physical Review A, 98(4), 042103., arXiv:1804.05856, DOI:10.1103/PhysRevA.98.042103,

3. Sadowski, P., Pawela, Ł., Lewandowska, P., & Kukulski, R. (2019). *Quantum walks on hypergraphs*. International Journal of Theoretical Physics, 58(10), 3382-3393., arXiv:1809.04521, DOI:10.1007/s10773-019-04212-6,

4. Lewandowska, P., Kukulski, R., & Pawela, Ł. (2020). *Optimal representation of quantum channels*. In International Conference on Computational Science (pp. 616-626). Springer, Cham., arXiv:2002.05507, DOI:10.1007/978-3-030-50433-5_47,

5. Kukulski, R., Lewandowska, P., & Pawela, Ł. (2020). *Perturbation of the numerical range of unitary matrices*. In International Conference on Computational Science (pp. 627-637). Springer, Cham., arXiv:2002.05553, DOI:10.1007/978-3-030-50433-5_48,

6. Kukulski, R., & Wojewódka-Ściążko, H. (2021). *The e-property of asymptotically stable Markov–Feller operators*. Colloquium Mathematicum, 165, 269-283, arXiv:2001.05785, DOI:10.4064/cm8165-6-2020,

7. Puchała, Z., Pawela, Ł., Krawiec, A., Kukulski, R., & Oszmaniec, M. (2021). *Multiple-shot and unambiguous discrimination of von Neumann measurements*. Quantum, 5, 425., arXiv:1810.05122, DOI:10.22331/q-2021-04-06-425,

8. **Kukulski, R., Nechita, I., Pawela, Ł., Puchała, Z., & Życzkowski, K. (2021).** ***Generating random quantum channels*****. Journal of Mathematical Physics, 62(6), 062201.**, arXiv:2011.02994, DOI:10.1063/5.0038838,

9. Lewandowska, P., Krawiec, A., Kukulski, R., Pawela, Ł., & Puchała, Z. (2021). *On the optimal certification of von Neumann measurements.* Scientific reports, 11(1), 1-16., arXiv:2009.06776, DOI:10.1038/s41598-021-81325-1,

10. Lewandowska, P., Kukulski, R., Pawela, Ł., & Puchała, Z. (2022). *Storage and retrieval of von Neumann measurements.* Physical Review A, 106(5), 052423., arXiv:2204.03029, DOI:10.1103/PhysRevA.106.052423,

11. **Kukulski, R., Pawela, Ł., & Puchała, Z. (2023).** ***On the probabilistic quantum error correction.*** **IEEE Transactions on Information Theory.** , arXiv:2206.05232, DOI:10.1109/TIT.2023.3254054,

Preprints

1. Kukulski, R., & Glos, A. (2020). *Comment to Spatial Search by Quantum Walk is Optimal for Almost all Graphs.* arXiv preprint arXiv:2009.13309.

2. Kukulski, R., & Wojewódka-Ściążko, H. (2022). *The e-property of asymptotically stable Markov semigroups.* arXiv preprint arXiv:2211.16424.

3. Shahbeigi, F., Chubb, C., Kukulski, R., Pawela, Ł., & Korzekwa, K. *Quantum-embeddable stochastic matrices* (available soon)

Publications related to the dissertation are highlighted in **bold font**.

Streszczenie w języku polskim

Jeszcze czterdzieści lat temu naukowcy na całym świecie nie podejrzewali, że możliwe będzie wykorzystanie praw fizyki kwantowej w celu przetwarzania informacji. Rozważano w tym czasie różne metody uogólnienia klasycznej teorii informacji Shannona. Przełom nastąpił w roku 1976 za sprawą polskiego fizyka, Romana Ingardena, który zdefiniował podstawy kwantowej teorii informacji. Opublikował on przełomowy artykuł "Kwantowa teoria informacji", wyprzedzający rewolucje wokół komputerów kwantowych o prawie ćwierć wieku. Kilka lat później, w 1981 roku, Richard Feynman stwierdził, że takowy komputer kwantowy mógłby być wydajnym narzędziem do symulacji fizyki kwantowej wielu ciał. Jego pomysł zapoczątkował owocne badania w dziedzinie obliczeń kwantowych, które utorowały drogę do budowy pierwszych urządzeń kwantowych typu NISQ. Szybki rozwój technologii i wciąż rosnące zainteresowanie komputerami kwantowymi sprawiły, że w dzisiejszych czasach każdy może uzyskać dostęp i wykonywać obliczenia na urządzeniach NISQ.

Powszechnie wiadomo, że urządzenia typu NISQ mają ograniczenia. Komputery kwantowe wciąż wykonują niedoskonałe obliczenia, a występujące na nim błędy są widoczne nawet w najprostszych algorytmach. Poprawność obliczeń kwantowych zakłócają trzy rodzaje błędów: dekoherencja, niepoprawna implementacja bramek kwantowych oraz błędy odczytu. Z tych powodów nasuwa się pytanie, jak pozbyć się tych błędów? Albo przynajmniej w jaki sposób je zminimalizować? Odpowiedzią na to pytanie jest kwantowa korekcja błędów.

Kwantowa korekcja błędów (QEC) to procedura kodowania-dekodowania, która chroni informację kwantową przed wcześniej wymienionymi błędami. Podobnie jak w przypadku obliczeń klasycznych, procedura ta jest niezbędna do zbudowania w pełni działających komputerów kwantowych. Chociaż wybór podejścia zależy od kilku czynników, główny kierunek badań kodów korekcyjnych koncentruje się na perfekcyjnych kodach QEC. Procedury te zakładają, że stan zdekodowany jest dokładnie taki sam jak stan zakodowany. Dynamiczny rozwój tego typu kodów nastąpił wraz z pojawieniem się kodu Shora pozwalającego zakodować jeden logiczny kubit informacji na dziewięciu kubitach danych.

Kiedy jednak możemy uzyskać perfekcyjny kod korekcyjny? Znanych jest kilka warunków dotyczących rodzaju szumów, które można doskonale skorygować. Większość z nich opiera się na twierdzeniu Knilla-Laflamme'a. Jeśli spełnione są warunki tego twierdzenia, możemy skonstruować schemat dekodowania i odtworzyć zakodowaną informację perfekcyjnie. Spośród szerokiego wachlarza metod kodowania warto wymienić kody inspirowane klasycznymi kodami korekcji błędów Calderbank, Shora, Steane (CSS) czy kody stabilizacyjne. W literaturze też zbadano wiele innych technik kodowania z różnymi konfiguracjami, na przykład kody wzmocnione splątaniem, czy kody zdefiniowane na podsystemach kwantowych.

W rzeczywistym scenariuszu niemożliwe jest jednak całkowite zniwelowanie efektów działania danego kanału szumu. Kluczowe jest więc wdrożenie przybliżonych kodów korekcji błędów. W tym podejściu, stan zdekodowany różni się w jakiś sposób od stanu zakodowanego. Dokładność takich procedur możemy mierzyć za pomocą różnych narzędzi, takich jak funkcja wierności. Możliwe jest również użycie procedur zwracających klasyczną informację, sugerującą, czy powinniśmy odrzucić zdekodowany stan. Podejście to rozważano na przykład w wykrywaniu błędów kwantowych czy probabilistycznej korekcji błędów (pQEC).

Probabilistyczna korekcja błędów pQEC to procedura, która wykorzystuje postselekcję w celu określenia, czy zakodowane informacje zostały pomyślnie zdekodowane. Dane wyjściowe procedury pQEC składają się z dwóch komponentów: stanu kwantowego i klasycznej etykiety binarnej. Etykieta ta informuje, czy korekcja błędu zakończyła się powodzeniem, a stan zdekodowany powinien zostać zaakceptowany. Chociaż metody pQEC wzbudziły pewne zainteresowanie w literaturze, nie zostały one uogólnione dla dowolnych modeli szumu, tak jak w przypadku kodów perfekcyjnych i twiedzenia Knilla-Laflamme'a. Głównym celem tej rozprawy jest wypełnienie tej luki badawczej i opracowanie teorii kodów pQEC dla ogólnych modeli szumu. Pokażemy, że kody probabilistyczne są odpowiednie dla zaszumionych układów kwantowych. W związku z tym formułujemy następującą **Hipotezę**:

Zastosowanie probabilistycznych kwantowych kodów korekcyjnych może poprawić jakość zaszumionych układów kwantowych.

Aby potwierdzić tę hipotezę, uogólnimy twierdzenie Killa-Laflamme'a i sformułujemy warunki konieczne i wystarczające do sprawdzenia, czy dany kanał szumu jest probabilistycznie korygowalny. Dzięki temu pokażemy, że kody pQEC umożliwiają korekcję szerszej klasy szumów, niż jest to możliwe w przypadku użycia kodów QEC. Dla kodów pQEC wskażemy relację pomiędzy prawdopodobieństwem, że procedura zakończy się sukcesem, a jakością otrzymanego kodu.

Jako inżynierski aspekt tej pracy potwierdzający hipotezę, stworzymy efektywny numerycznie algorytm do konstruowania przybliżonego kodu pQEC oraz zbadamy jego skuteczność. Zauważymy, że każdy kanał szumu o "niskim" rzędzie macierzy Choi prawie na pewno może być perfekcyjnie skorygowany przy użyciu zaproponowanej procedury. Dla pozostałych szumów, nasza konstrukcja zapewni nam schemat o stosunkowo dużej wartości funkcji wierności. Głównym narzędziem wykorzystywanym w symulacjach numerycznych są losowe kanały. Użyjemy ich do sprawdzenia skuteczności zaproponowanej procedury pQEC. Ulepszymy również techniki generowania kanałów kwantowych i pokażemy, jak skutecznie generować losowe podkanały, instrumenty i superkanały kwantowe.

Praca składa się z sześciu rozdziałów. Pierwszy rozdział zawiera wstęp do kwantowej korekcji błędów. W Rozdziale 2 zamieszczono wprowadzenie do matematycznego języka informatyki kwantowej. Pozostała część dysertacji została napisana na podstawie dwóch opublikowanych artykułów naukowych oraz własnych nieopublikowanych wyników. Rozdział 3 skupia się na losowych kanałach kwantowych. Rozdział ten został częściowo napisany w oparciu o prace [1]. Rozważania na temat metod losowania i własności superkanałów, podkanałów czy kwantowych instrumentów stanowi mój autorski wkład w dysertację. Rozdział 4, stanowiący główną część mojej rozprawy, opisuje probabilistyczne kody korekcyjne. W tym rozdziale pokazujemy zalety użycia probabilistycznych kodów korekcyjnych. Rozdział ten został napisany w oparciu o pracę [2]. W rozdziale 5 implementujemy zaproponowaną procedurę pQEC, która generuje efektywne aproksymacyjne kody korekcyjne. Aby pokazać potencjał pQEC testujemy ów procedurę na losowo wygenerowanych kanałach kwantowych. Ten rozdział stanowi mój autorski wkład w tę pracę. Rozdział 6 zawiera wnioski z rozprawy i je podsumowuje.

Abstract in English

Forty years ago, scientists did not suspect that in the near future, it would be possible to utilize the laws of quantum mechanics to process information. They were considering different approaches to generalize Shannon's classical information theory at that time. The breakthrough came in 1976 with the work of Polish researcher Roman Ingarden, who defined the concept of quantum information theory. He published the paper "Quantum Information Theory", which predated the explosion of interest in quantum information and quantum computing by almost a quarter of a century. A few years later, in 1981, Richard Feynman expressed that a quantum computer would be an efficient tool for simulating many-body quantum physics. His idea sparked fruitful research in quantum computation, which paved the way for the recent construction of the first quantum devices. These devices are called Noisy Intermediate-Scale Quantum (NISQ). The rapid development of technology and still growing interest in quantum computing made that nowadays everyone can make computations on NISQ devices.

It is well known that gate model-inspired NISQ devices have limitations. Quantum computers still make computations imperfectly, and it is beyond the capabilities of current technology to correct existing errors. Quantum computations are hindered by three kinds of errors: decoherence, poor implementation of quantum gates, and finally, readout errors. For these reasons, one may ask how to dispose of such errors or at least minimize them. The answer is quantum error correction.

Quantum error correction (QEC) is an encoding-decoding procedure that protects quantum information from errors caused by quantum noise. Similarly to classical computations, this procedure is essential to develop fully operational quantum computers. Until now, the main direction of research on quantum codes that correct errors focuses on perfect QEC codes. It covers the situation when the procedure always succeeds in protecting quantum information, which means the decoded state is precisely the same as the encoded one. The dynamic development of this type of codes came with the advent of Shor's code. He created the scheme that allows to encode one logical qubit of information into nine data qubits.

But under what assumptions, in general, can we achieve perfect error-correcting codes? For perfect QEC codes, several conditions were given for the type of noises which can be corrected perfectly. Most of them rely on the Knill-Laflamme theorem. If conditions of this theorem are satisfied, the receiver can construct a decoding scheme and perfectly restore the initial information. Covering a wide range of coding techniques, it is worth to mention the codes inspired by the classical error correction, Calderbank, Shor, Steane (CSS) codes or Quantum Reed–Solomon Codes and truly revolutionizing stabilizer QEC codes. Many more coding techniques were explored in the literature with different set-ups, for example, codes enhanced by shared entanglement, called entanglement-assisted error-correcting codes.

In the real case scenario, it is impossible to correct the given noise channel perfectly. Therefore, it is crucial to consider approximate quantum error-correcting codes. In this type of code, the decoded state is somehow different from the encoded one. We can measure the accuracy of such procedures by various figures of merit, for example the fidelity function. Another type of imperfection in quantum error correction codes comes with the classical post-processing procedure. Such a procedure returns classical information, suggesting when we should reject the decoded state. This approach was considered, for example, in quantum error detection and probabilistic quantum error correction (pQEC).

The pQEC procedure is an error-correcting procedure which uses postselection to determine if the encoded information was successfully restored. The output of the pQEC procedure consists of two objects: a quantum state and a classical binary label. The label informs if the error correction succeeded, and the output state should be accepted. Although pQEC drew some attention in the literature, there was no study of this procedure for general noise channels, like it was done by Knill and Laflamme for deterministic codes.

The main aim of this dissertation is to fill this research gap and develop a theory of pQEC codes for general noise channels. We will also show that the probabilistic codes are suitable for noisy quantum systems. Due to that we formulate the following **Hypothesis**:

The usage of probabilistic quantum error correction codes can improve the quality of quantum systems disturbed by general noise channels.

To confirm this hypothesis, we will generalize the Kill-Laflamme theorem, and formulate the necessary and sufficient conditions to check if a given noise channel is probabilistically correctable. We will use these conditions to show that the pQEC codes, in comparison to the QEC codes, can correct noise channels from a broader class of quantum channels. In particular, we will provide different families of noise channels for which it is possible to find pQEC code, but not the deterministic one. It will indicate a trade-off between the probability that our error-correcting procedure will successfully terminate and the quality of the code, which can be measured by *e.g.* the fidelity function.

As an engineering aspect confirming the hypothesis, we create a numerically efficient algorithm to construct an approximate pQEC code and investigate its effectiveness. This construction almost surely returns a perfect probabilistic error-correcting schemes for quantum channels with “low” Choi rank. For any other quantum channel, this construction provides a scheme with a relatively high value of the fidelity function.

The main tool used in numerical simulations is random channels. Due to the diversity of quantum channels these ensembles provide, they are suitable for numerical investigation of various quantum properties and testing the effectiveness of procedures. In this dissertation, we will use random quantum channels to check the effectiveness of the introduced construction of pQEC procedure. We will also generalize techniques of generating quantum channels and show how to effectively generate random quantum subchannels, instruments and super-channels.

The work consists of six chapters. The first Chapter introduces the theory of quantum error correction. Chapter 2 presents the necessary mathematical framework. The rest of the dissertation is based on two published articles and self-directed unpublished results. Chapter 3 concerns the overview of random quantum operations. This part of the chapter is written based on [1]. We extend the research to generating methods of subchannels, super-channels and instruments. This part, however, is my contribution to the dissertation. Next, the work [2] presented in Chapter 4 focuses on the pQEC codes. Here, we show the advantages of using the pQEC codes. In Chapter 5, basing on the introduced pQEC codes, we create an efficient method of constructing approximate quantum error correction codes. To show the potential of our approach, we test the proposed procedure on randomly generated quantum channels. This chapter is my contribution to the dissertation. Chapter 6 contains the conclusions and summarizes the results of the presented research.

Acknowledgment

A special thanks go to Zbigniew Puchała and Łukasz Pawela for continuous support of my research.

For various reasons, I would like to show gratitude to Paulina Lewnadowska and Karol Życzkowski.

I acknowledge support by the project „Near-term Quantum Computers: challenges, optimal implementations and applications" under Grant Number POIR.04.04.00-00-17C1/18-00, which is carried out within the Team-Net programme of the Foundation for Polish Science co-financed by the European Union under the European Regional Development Fund and European Unionscholarship through the European Social Fund, grant InterPOWER (POWR.03.05.00-00-Z305).

Chapter 1

Introduction

Forty years ago, scientists did not suspect that in the near future, it would be possible to utilize the laws of quantum mechanics to process information. They were considering different approaches to generalize Shannon's classical information theory [3] at that time. The breakthrough came in 1976 with the work of Polish researcher Roman Ingarden, who defined the concept of quantum information theory. He published the paper "Quantum Information Theory" [4], which predated the explosion of interest in quantum information and quantum computing by almost a quarter of a century. A few years later, in 1981, Richard Feynman expressed that a quantum computer would be an efficient tool for simulating many-body quantum physics. His idea sparked fruitful research in quantum computation, which paved the way for the recent construction of the first quantum devices. These devices are called Noisy Intermediate-Scale Quantum (NISQ) [5]. The rapid development of technology and still growing interest of quantum computing made that nowadays everyone can make computations on NISQ devices. For gate model-inspired NISQ devices, one could mention IBMQ [6], Rigetti [7], Xanadu [8] or IonQ [9].

It is well known that gate model-inspired NISQ devices have limitations [10]. Quantum computers still make computations imperfectly, and it is beyond the capabilities of current technology to correct existing errors. Quantum computations are hindered by three kinds of errors: decoherence, responsible for the loss of information; poor implementation of quantum gates, which diverts the computation process; and finally, readout errors. For these reasons, one may ask how to dispose of such errors or at least minimize them. The answer is quantum error correction, an indispensable ingredient for noise-tolerant and scalable quantum computing.

Quantum error correction (QEC) is an encoding-decoding procedure that protects quantum information from errors caused by quantum noise. Similarly to classical computations, this procedure is essential to develop fully operational quantum computers [5]. A general review of QEC methods we can see in [11]. The choice of a method depends on a few factors: the number of logical qubits, which we want to encode, quantum resources, that is, for example, the number of data qubits, or type of noise. Until now, the main direction of research on codes that correct errors focused on perfect QEC codes [12, 13]. It covers the situation when the procedure always succeeds in protecting quantum information, which means the decoded state is precisely the same as the encoded one. The dynamic development of this type of codes came with the advent of Shor's code [14]. He created the scheme that allows to encode one logical qubit of information into nine data qubits. His code can correct one-qubit errors perfectly.

But under what assumptions, in general, can we achieve perfect error-correcting codes? For perfect QEC codes, several conditions were given for the type of noise channels which can be corrected perfectly. Most of them rely on the Knill-Laflamme theorem [12]. If conditions of this theorem are satisfied, the receiver can construct a decoding scheme and perfectly restore the initial information. Covering a wide range of coding techniques, it is worth mentioning the codes inspired by the classical theory of error correction, Calderbank, Shor, Steane (CSS) codes [15–18] or Quantum Reed–Solomon Codes [19] and truly revolutionizing stabilizer QEC codes [20, 21], which make use of the results from group theory. A stabilizer QEC code is defined by introducing a commutative subgroup of the Pauli group on n qubit system. Due to the compact description of stabilizer codes, the theory is an active field of research [22, 23]. An interesting construction of such a subgroup came from utilizing algebraic topology methods. They gave a rise to topological codes [24–26]. Moreover, thanks to the stabilizer formalism, many quantum maximum distance separable codes [27–29] were found. Another remarkable codes, raised from CSS codes, are quantum low-density parity-check (LDPC) codes [30–32]. Much of the interest in quantum LPDC codes started after Gottesman's result [31] showing that these codes can reduce the overhead of fault-tolerant quantum computation to be constant, in contrast to other quantum fault tolerance schemes [33]. Many more coding techniques were explored in the literature, in particular schemes with different set-ups, for example, codes enhanced by shared entanglement, called entanglement-assisted quantum error-correcting codes [34–36].

In the real case scenario, it is impossible to correct the given noise channel perfectly. Therefore, it is crucial to consider approximate quantum error-correcting codes [37–39]. In this type of code, the decoded state is somehow different from the encoded one. We can measure the accuracy of such procedures by various figures of merit. In the literature, the following measures were considered: the average channel (entanglement) fidelity [40, 41], the diamond norm [42, 43], the average output fidelity [44], the worst-case output fidelity [45, 46]. Another type of imperfection in quantum error correction codes comes with the classical post-processing procedure. Such a procedure returns classical information, suggesting when we should reject the decoded state. This approach was considered, for example, in quantum error detection [47–49] and *probabilistic* quantum error correction (pQEC) [50–53].

Probabilistic quantum error correction is an error-correcting procedure which uses postselection to determine if the encoded information was successfully restored. The working of the pQEC procedure relies on non-deterministic decoding operations [54, 55]. The output of the pQEC procedure consists of two objects: a quantum state and a classical binary label. The label informs if the error correction succeeded, and the output state should be accepted. In the context of QEC, probabilistic decoding operations have found application in stabilizer codes [51, 56], iterative probabilistic decoding in LDPC codes [57] or environment-assisted error correction [52]. It was noted that pQEC has the potential to increase the spectrum of correctable errors [51] and is useful when the number of qubits is limited [50]. It is also worth mentioning that probabilistic strategies were used with success in other fields of quantum information theory, *e.g.* for probabilistic cloning [58], learning of unknown quantum operations [59] or measurement discrimination [60].

Although pQEC drew some attention in the literature, there was no study of this procedure for general noise channels, like it was done by Knill and Laflamme in [12] or [61] for deterministic codes. It would allow us to establish an upper-bound on the amount of quantum information that potentially can be transferred through a particular noise channel. Additionally, it would provide new error-correcting schemes.

The main aim of this dissertation is to fill this research gap and develop a theory of pQEC codes for general noise channels. By general noise channel, we define quantum operations with no particular structure that can be described by open quantum systems and the Stinespring dilation theorem [62]. In this thesis, we will also show that the probabilistic codes are suitable for noisy quantum systems, which we formulate as the following **Hypothesis**:

The usage of probabilistic quantum error correction codes can improve the quality of quantum systems disturbed by general noise channels.

To confirm this hypothesis, we will generalize the Kill-Laflamme theorem and formulate the necessary and sufficient conditions to check if a given noise channel is probabilistically correctable. We will use these conditions to show that the pQEC codes, in comparison to the QEC codes, can correct noise channels from a broader class of quantum channels. In particular, we will provide different families of noise channels for which it is possible to find pQEC codes, but not the deterministic ones. It will indicate a trade-off between the probability that our error-correcting procedure will successfully terminate and the quality of the code, which can be measured by, for example, the fidelity function.

As an engineering aspect confirming the dissertation hypothesis, we create a numerically efficient algorithm to construct an approximate pQEC code and investigate its effectiveness. This construction has two essential properties. First, it almost surely returns a perfect probabilistic error-correcting scheme for any quantum channel with “low” Choi rank. Second, for any other quantum channel, this construction provides a scheme with a relatively high value of the fidelity function. We will prove the first property while we will provide convincing numerical evidence for the second one. The numerical investigation of the pQEC codes were possible due to advancements in generating random quantum channels.

The main tool used in this dissertation for numerical simulations is random channels. Such an ensemble of random quantum operations was introduced in [63]. Due to the diversity of quantum channels these ensembles provide, they are suitable for numerical investigation of various quantum properties and testing the effectiveness of procedures. We can find modern expositions of these results and many developments in monographs such as [64, Chapter 8], [65, Chapters 10 and 11], or [62, Chapter 2.2]. In this dissertation, we will use random quantum channels to check the effectiveness of the introduced construction of pQEC procedure. Moreover, we will show that the random ensembles proposed in [63] are equivalent, and we can choose the one which is the most suitable for the numerical simulation. We will generalize techniques of generating quantum channels and show how to effectively generate random quantum subchannels, quantum instruments and super-channels, and how to obtain the uniform measure in all the cases.

The work consists of six chapters. The first Chapter introduces the theory of quantum error correction and the motivation for my research. Chapter 2 presents the necessary mathematical framework. The rest of the dissertation is based on two published articles and my self-directed unpublished results. Chapter 3 concerns the overview of random quantum operations. First, we explore the methods of generating random channels and then analyze their properties. This part of the chapter is written based on [1]. We extend the research to generating methods of subchannels, super-channels and instruments. This part, however, is my contribution to the dissertation. Next, the work [2] presented in Chapter 4 focuses on the pQEC codes. Here, we present a general problem formulation, motivation and necessary theoretical framework. Most importantly, we show the advantages of using the pQEC codes. In Chapter 5, basing on the pQEC procedure, we create an efficient method of constructing approximate probabilistic quantum error correction codes. To show the potential of this approach, we test the proposed procedure on randomly generated quantum channels. This chapter is my contribution to the dissertation. Chapter 6 contains the conclusions of this dissertation and summarizes the results of the presented research.

Chapter 2

Mathematical preliminaries

This chapter is intended to serve as a review of mathematical concepts used throughout this dissertation. First, we will introduce a notation and mathematical concepts used in this thesis. Secondly, we will recall necessary basic facts of quantum information theory. This chapter is based on the book [62] by John Watrous.

2.1 Vector space and Dirac notation

This dissertation relies heavily on linear algebra in finite-dimensional spaces. We consider complex Euclidean space denoted by scripted letters $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \ldots$. The dimension of a space $\mathcal{X}$ will be denoted by $\dim(\mathcal{X})$. In this section we fix the notation $d = \dim(\mathcal{X})$. Henceforth, we may write $\mathcal{X} = \mathbb{C}^d$, where $\mathbb{C}$ denotes the set of complex numbers. In Dirac notation, a column vector $|\psi\rangle \in \mathcal{X}$ is called ket and denoted by

$$|\psi\rangle \coloneqq (\psi_0, \ldots, \psi_{d-1})^\top, \;\; \psi_i \in \mathbb{C}, \, i \in \{0, \ldots, d-1\}, \tag{2.1}$$

whereas a row vector $\langle\psi| \in \mathcal{X}^*$ is called bra and denoted by

$$\langle\psi| \coloneqq (\overline{\psi_0}, \ldots, \overline{\psi_{d-1}}), \tag{2.2}$$

where $\overline{\psi_i}$ denotes complex conjugate of ψ_i. From Riesz's representation theorem, each complex Euclidean space $\mathcal{X}$ and its dual space are isometrically isomorphic. Then, an isomorphism defining a one-to-one mapping between kets and bras is indicated as $\dagger$, that is $|\psi\rangle^\dagger \coloneqq \langle\psi|$.

In the space $\mathcal{X}$, we distinguish the standard basis represented by a collection of vectors $\{|i\rangle\}_{i=0}^{d-1}$, where the entry 1 appears in $(i+1)^{\text{th}}$ position and $i \in \{0, \ldots, d-1\}$. Each element $|\psi\rangle \in \mathcal{X}$ has a unique representation $|\psi\rangle = \sum_{i=0}^{d-1} \psi_i |i\rangle$ for $\psi_i \in \mathbb{C}$,

$i \in \{0, \ldots, d-1\}$. We introduce the inner product $\langle\psi|\phi\rangle$ of two vectors $|\psi\rangle, |\phi\rangle \in \mathcal{X}$ given by

$$\langle\psi|\phi\rangle = \sum_{i=0}^{d-1} \overline{\psi_i}\phi_i. \tag{2.3}$$

Hence, the Euclidean norm of a vector $|\psi\rangle \in \mathcal{X}$ is defined as

$$|||\psi\rangle||_2 = \sqrt{\langle\psi|\psi\rangle}. \tag{2.4}$$

The Euclidean norm represents the case $p = 2$ of the class of p-norms defined for each $|\psi\rangle \in \mathcal{X}$ as

$$|||\psi\rangle||_p = \left(\sum_{i=0}^{d-1} |\psi_i|^p\right)^{\frac{1}{p}}, \tag{2.5}$$

for $p < \infty$, and

$$|||\psi\rangle||_\infty = \max\{|\psi_i| : i \in \{0, \ldots, d-1\}\}. \tag{2.6}$$

Two vectors $|\psi\rangle, |\phi\rangle \in \mathcal{X}$ are said to be orthogonal if $\langle\psi|\phi\rangle = 0$. A collection vectors $\{|\psi_i\rangle\}_{i=0}^{n-1}$ is linearly independent if $\sum_{i=0}^{n-1} c_i|\psi_i\rangle = 0$ implies that $c_i = 0$ for all $i \in \{0, \ldots, n-1\}$. We say that a collection of vectors is orthonormal if they are pairwise orthogonal and have unit Euclidean norm.

Direct sum and tensor product of complex Euclidean spaces

Let $|x\rangle = (x_0, \ldots, x_k)^\top \in \mathcal{X}$ and $|y\rangle = (y_0, \ldots, y_l)^\top \in \mathcal{Y}$. We define the direct sum of $|x\rangle$ and $|y\rangle$ as

$$|x\rangle \oplus |y\rangle = (x_0, \ldots, x_k, y_0, \ldots, y_l)^\top. \tag{2.7}$$

We write

$$\mathcal{W} = \mathcal{X} \oplus \mathcal{Y} \tag{2.8}$$

if and only if or every $|w\rangle \in \mathcal{W}$ there exist unique vectors $|x\rangle \in \mathcal{X}$ and $|y\rangle \in \mathcal{Y}$ such that $|w\rangle = |x\rangle \oplus |y\rangle$.

We define the tensor product of $|x\rangle$ and $|y\rangle$ as

$$|x\rangle \otimes |y\rangle = (x_0 \cdot y_0, \ldots, x_0 \cdot y_l, \ldots, x_k \cdot y_0, \ldots, x_k \cdot y_l)^\top, \tag{2.9}$$

and denote as $|x, y\rangle$. We write

$$\mathcal{W} = \mathcal{X} \otimes \mathcal{Y}, \tag{2.10}$$

if and only if every $|w\rangle \in \mathcal{W}$ can be represented as $|w\rangle = \sum_i |x_i\rangle \otimes |y_i\rangle$, where $|x_i\rangle \in \mathcal{X}$ and $|y_i\rangle \in \mathcal{Y}$.

We can define the direct sum and tensor product for more then two complex Euclidean spaces in a similar same way. For example, if $\mathcal{X}$ is is a complex Euclidean space, $|\psi\rangle \in \mathcal{X}$ is a vector and $n \in \mathbb{N}$, then the notations $\mathcal{X}^{\otimes n}$ and $|\psi\rangle^{\otimes n}$ refer to the n-fold tensor product of either $\mathcal{X}$ or $|\psi\rangle$ with itself. Analogously for direct sum.

2.2 Linear operators

The set of linear operators $M : \mathcal{X} \to \mathcal{Y}$ will be written as $\mathcal{M}(\mathcal{X}, \mathcal{Y})$. For simplify notation, let $\mathcal{M}(\mathcal{X}) \coloneqq \mathcal{M}(\mathcal{X}, \mathcal{X})$. The identity operators will be denoted by $\mathbb{1}_{\mathcal{X}} \in \mathcal{M}(\mathcal{X})$. In analogous way, we can also define the direct sum and tensor product of two and more linear operators. For any choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, there is a bijective linear correspondence between the set of operators $\mathcal{M}(\mathcal{X}, \mathcal{Y})$ and the collection of all matrices of the form $(M_{j,i})_{\substack{i=0,\dots,\dim(\mathcal{X})-1 \\ j=0,\dots,\dim(\mathcal{Y})-1}}$. Hereafter in this dissertation, linear operators will be associated with matrices implicitly, and hence we will be using the words operator and matrix interchangeably.

Direct sum and tensor product of linear operators

Let $\mathcal{X}_0, \dots, \mathcal{X}_n, \mathcal{Y}_0, \dots, \mathcal{Y}_n$ be complex Euclidean spaces and let $A_0 \in \mathcal{M}(\mathcal{X}_0, \mathcal{Y}_0), \dots, A_n \in \mathcal{M}(\mathcal{X}_n, \mathcal{Y}_n)$ be linear operators.

The direct product

$$A_0 \oplus \ldots \oplus A_n \in \mathcal{M}\left(\bigoplus_{i=0}^{n} \mathcal{X}_i, \bigoplus_{i=0}^{n} \mathcal{Y}_i\right), \tag{2.11}$$

of these operators is the unique operator that satisfies the equation

$$(A_0 \oplus \ldots \oplus A_n)(|x_0\rangle \oplus \ldots \oplus |x_n\rangle) = (A_0|x_0\rangle) \oplus \ldots \oplus (A_n|x_n\rangle), \tag{2.12}$$

for all choices $|x_0\rangle \in \mathcal{X}_0, \dots, |x_n\rangle \in \mathcal{X}_n$.

The tensor product

$$A_0 \otimes \ldots \otimes A_n \in \mathcal{M}\left(\bigotimes_{i=0}^{n} \mathcal{X}_i, \bigotimes_{i=0}^{n} \mathcal{Y}_i\right), \tag{2.13}$$

of these operators is the unique operator that satisfies the equation

$$(A_0 \otimes \ldots \otimes A_n)(|x_0\rangle \otimes \ldots \otimes |x_n\rangle) = (A_0|x_0\rangle) \otimes \ldots \otimes (A_n|x_n\rangle), \tag{2.14}$$

for all choices $|x_0\rangle \in \mathcal{X}_0, \dots, |x_n\rangle \in \mathcal{X}_n$.

Rank

An image of $A \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ is the subspace of $\mathcal{Y}$ defined as $\mathrm{im}(A) = \{A|x\rangle : |x\rangle \in \mathcal{X}\}$. The rank of an operator A is the dimension of the image of A, that is $\mathrm{rank}(A) = \dim(\mathrm{im}(A))$.

Trace

The trace of an operator $X \in \mathcal{M}(\mathcal{X})$ is defined as the sum of its diagonal entries, that is $\operatorname{tr}(X) = \sum_{i=0}^{\dim(\mathcal{X})-1} x_{i,i}$. Equivalently, the trace is the unique linear function $\operatorname{tr} : \mathcal{M}(\mathcal{X}) \to \mathbb{C}$ such that

$$\operatorname{tr}(|x\rangle\langle y|) = \langle y|x\rangle, \tag{2.15}$$

for all vectors $|x\rangle, |y\rangle \in \mathcal{X}$.

Vectorization

For any operator $M \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ we will consider its vectorization $|M\rangle\rangle \in \mathcal{Y} \otimes \mathcal{X}$, which is defined as

$$|M\rangle\rangle := \left(\mathbb{1}_{\mathcal{Y}} \otimes M^\top\right) \sum_{i=0}^{\dim(\mathcal{Y})-1} |i\rangle \otimes |i\rangle. \tag{2.16}$$

The crucial property of the vectorization mapping which will be useful throughout the dissertation is

$$(A \otimes B)|C\rangle\rangle = |ACB^\top\rangle\rangle, \tag{2.17}$$

for $A \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$, $B \in \mathcal{M}(\mathcal{W}, \mathcal{Z})$ and $C \in \mathcal{M}(\mathcal{W}, \mathcal{X})$. This property is sometimes called the telegraphic notation.

Eigenvectors and eigenvalues

Let $A \in \mathcal{M}(\mathcal{X})$ be an operator and $|x\rangle \in \mathcal{X}$ be a nonzero vector for which it holds that

$$A|x\rangle = \lambda|x\rangle, \tag{2.18}$$

for some $\lambda \in \mathbb{C}$, then $|x\rangle$ is called an eigenvector of A and λ is its corresponding eigenvalue. Let us define the characteristic polynomial of A as

$$p(\alpha) = \det(\alpha \mathbb{1}_{\mathcal{X}} - A). \tag{2.19}$$

The spectrum of A, denoted $\operatorname{spec}(A)$ is the multiset containing the roots of the polynomial $p(\alpha)$ being the eigenvalues of A.

Types of linear operators

The following classes of operators have particular importance in the theory of quantum information.

The commutator, or Lie bracket, of a given two operators $A, B \in \mathcal{M}(\mathcal{X})$, that is $AB - BA$, will be denoted by $[A, B]$. We call an operator $A \in \mathcal{M}(\mathcal{X})$ is normal if it commutes with its Hermitian conjugate ($A^\dagger = \overline{A^\top}$), that is $[A, A^\dagger] = 0$.

In the space $\mathcal{M}(\mathcal{X})$, we distinguish several classes of linear operators. A normal operator $A \in \mathcal{H}(\mathcal{X})$ is said to be Hermitian if satisfies the equation $A = A^\dagger$, while an Hermitian operator $A \in \mathcal{P}(\mathcal{X})$ is called positive semi-definite if $\langle\psi|A|\psi\rangle \geq 0$ for every $|\psi\rangle \in \mathcal{X}$. Another important operator is a projection operator that is a positive semi-definite operator $\Pi \in \mathcal{P}(\mathcal{X})$ satisfying the equation $\Pi^2 = \Pi$.

An operator $V \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$, is an isometry if it preserves the Euclidean norm of vectors that is $\|V|x\rangle\|_2 = \||x\rangle\|_2$ for all $|x\rangle \in \mathcal{X}$. We will consider also an unitary operator $U \in \mathcal{M}(\mathcal{X})$ that is satisfying the equation $UU^\dagger = U^\dagger U = \mathbb{1}_\mathcal{X}$. We will denote the set of all isometries as $\mathcal{U}(\mathcal{X}, \mathcal{Y})$ whereas the set of all unitary operators by $\mathcal{U}(\mathcal{X})$.

By $\mathcal{D}(\mathcal{X})$ we denote the set of quantum states, that is, the set of positive semi-definite operators with unit trace. We say that a quantum state ρ is a pure state if $\operatorname{rank}(\rho) = 1$, otherwise, if $\operatorname{rank}(\rho) > 1$, we say that ρ is a mixed state. The maximally mixed state will be denoted by $\rho^*_\mathcal{X} \coloneqq \frac{1}{\dim(\mathcal{X})}\mathbb{1}_\mathcal{X}$.

An operator $X = (x_{i,j})_{i,j=0}^{\dim(\mathcal{X})-1} \in \mathcal{M}(\mathcal{X})$ is diagonal if $x_{i,j} = 0$ for all $i \neq j$. We will also introduce the operator $\operatorname{diag} : \mathcal{X} \to \mathcal{M}(\mathcal{X})$ given by

$$\operatorname{diag}(|x\rangle) \coloneqq \begin{pmatrix} x_0 & 0 & \dots & 0 \\ 0 & x_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & x_{\dim(\mathcal{X})-1} \end{pmatrix}, \tag{2.20}$$

where $|x\rangle = (x_0, \dots, x_{\dim(\mathcal{X})-1})^\top$ and consider the adjoint operator $\operatorname{diag}^\dagger : \mathcal{M}(\mathcal{X}) \to \mathcal{X}$, that is $\operatorname{diag}^\dagger(M) = (M_{0,0}, \dots, M_{\dim(\mathcal{X})-1,\dim(\mathcal{X})-1})^\top$.

Operator decompositions

Below we cite the theorems regarding operator decompositions, which we use later in the work. Proofs of the following theorems can be found, for example, in [62].

Theorem 2.1 (Spectral decomposition). *Let $X \in \mathcal{M}(\mathcal{X})$ be a normal operator. There exists a positive integer m, distinct complex numbers $\lambda_0, \dots, \lambda_{m-1} \in \mathbb{C}$ and nonzero projector operators $\Pi_0, \dots, \Pi_{m-1} \in \mathcal{P}(\mathcal{X})$ satisfying $\Pi_0 + \dots + \Pi_{m-1} = \mathbb{1}_\mathcal{X}$, such that*

$$X = \sum_{i=0}^{m-1} \lambda_i \Pi_i. \tag{2.21}$$

The scalars $\lambda_0, \dots, \lambda_{m-1}$ and projection operators $\Pi_0, \dots, \Pi_{m-1}$ are uniquely determined, *i.e.* each scalar λ_k is an eigenvalue of X with multiplicity equal to the rank of Π_k, and Π_k is the projection operator onto the space spanned by the eigenvectors of X corresponding to the eigenvalue λ_k. The set of all λ_k of an operator X will be denoted as $\lambda(X)$.

Corollary 2.2. *Let* $\dim(\mathcal{X}) = d$. *Let* $X \in \mathcal{M}(\mathcal{X})$ *be a normal operator and assume that multiset* $\{\lambda_0, \ldots, \lambda_{d-1}\}$ *be the spectrum of* X. *There exists an orthonormal basis* $\{|x_0\rangle, \ldots, |x_{d-1}\rangle\} \in \mathcal{X}$ *such that*

$$X = \sum_{i=0}^{d-1} \lambda_i |x_i\rangle\langle x_i|. \tag{2.22}$$

Now, we recall the singular value theorem. The singular value theorem has a close relationship to the spectral theorem. Unlike the spectral theorem, however, the singular value theorem holds for arbitrary (nonzero) operators, as opposed to just normal operators.

Theorem 2.3 (Singular value decomposition)**.** *Let* $\mathcal{X}, \mathcal{Y}$ *be complex Euclidean spaces. Let* $A \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ *be a nonzero operator having rank equal to* $r \in \mathbb{N}$. *There exist orthonormal sets* $\{|x_0\rangle, \ldots, |x_{r-1}\rangle\} \subset \mathcal{X}$ *and* $\{|y_0\rangle, \ldots, |y_{r-1}\rangle\} \subset \mathcal{Y}$ *along with positive real numbers* $s_0, \ldots, s_{r-1}$ *such that*

$$A = \sum_{k=0}^{r-1} s_k |y_k\rangle\langle x_k|. \tag{2.23}$$

An expression of a given operator A *in the form of Eq.* (2.23) *is said to be a singular value decomposition of* A*. The numbers* $s_1, \ldots, s_r$ *are called singular values of* A*, whereas the collection of vectors* $|x_0\rangle, \ldots, |x_{r-1}\rangle$ *and* $|y_0\rangle, \ldots, |y_{r-1}\rangle$ *are called right and left singular vectors of* A*, respectively. The set of all singular values of* A *will denoted by* $\sigma(A)$.

Corollary 2.4. *Let* $A \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ *be a nonzero operator with rank* $r \in \mathbb{N}$. *Then, there exists a diagonal, positive definite operator* $D \in \mathcal{P}(\mathbb{C}^r)$ *of the form* $D = \mathrm{diag}(\sigma(A))$ *and isometries* $U \in \mathcal{U}(\mathbb{C}^r, \mathcal{Y})$, $V \in \mathcal{U}(\mathbb{C}^r, \mathcal{X})$*, such that*

$$A = UDV^\dagger. \tag{2.24}$$

Theorem 2.5 (Jordan–Hahn decompositions)**.** *Let* $H \in \mathcal{H}(\mathcal{X})$. *The operator* H *can be expressed as*

$$H = P - Q, \tag{2.25}$$

where $P, Q \in \mathcal{P}(\mathcal{X})$ *and it holds* $PQ = 0$. *The operators* P *and* Q *are uniquely defined for a given operator* H. *The expression given by Eq.* (2.25) *is called the Jordan–Hahn decomposition of* H.

Theorem 2.6. *An arbitrary square matrix* $X \in \mathcal{M}(\mathcal{X})$ *can be written as the sum of Hermitian matrices* $A, B \in \mathcal{H}(\mathcal{X})$ *as*

$$X = A + i\,B, \tag{2.26}$$

where the matrices A and B have the following forms

$$A = \frac{1}{2}\left(X + X^{\dagger}\right), \tag{2.27}$$

and

$$B = \frac{1}{2\,i}(X - X^{\dagger}). \tag{2.28}$$

Power of positive operator

For a positive semidefinite operator $P \in \mathcal{P}(\mathcal{X})$ having spectral decomposition

$$P = \sum_{i=0}^{m-1} \lambda_i \Pi_i, \tag{2.29}$$

where $\lambda_i \geq 0$ for all $i \in \{0, \ldots, m-1\}$, we can also define k-th power of the operator P as

$$P^k = \sum_{i=0}^{m-1} \lambda_i^k \, \Pi_i, \tag{2.30}$$

for $k \in \mathbb{R}$.

Moore–Penrose pseudo-inverse

Let $A \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ be a nonzero operator of the form

$$A = \sum_{k=0}^{r-1} s_k |y_k\rangle\langle x_k|, \tag{2.31}$$

where $s_k, |y_k\rangle, |x\rangle_k$ are defined as in Theorem 2.3. We define an operator $A^{-1} \in \mathcal{M}(\mathcal{Y}, \mathcal{X})$, known as the Moore–Penrose pseudo-inverse of A, as the unique operator given by

$$A^{-1} = \sum_{k=0}^{r-1} \frac{1}{s_k} |x_k\rangle\langle y_k|. \tag{2.32}$$

Operator norms

Many interesting and useful norms can be defined on spaces of operators, but in quantum information theory we mostly use a single family of norms called Schatten p-norms.

Definition 2.7. *For any operator $A \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ and real number $p \geq 1$, we define the p-Schatten of A as*

$$\|A\|_p = \left(\operatorname{tr}\left(\left(A^\dagger A\right)^{\frac{p}{2}}\right)\right)^{\frac{1}{p}}. \tag{2.33}$$

The Schatten $\infty-$norm is defined as

$$\|A\|_\infty = \max\{\|A|\psi\rangle\|_2 : |\psi\rangle \in \mathcal{X}, \langle\psi|\psi\rangle = 1\}, \tag{2.34}$$

The Schatten norm can be rewritten by using singular values and vector norms in the way

$$\|A\|_p = \|\sigma(A)\|_p. \tag{2.35}$$

2.3 Linear transformations

We also consider transformations, called superoperator, between linear operators. More precisely we will consider mapping of the form

$$\Phi : \mathcal{M}(\mathcal{X}) \to \mathcal{M}(\mathcal{Y}). \tag{2.36}$$

The set of all such maps is denoted $\mathcal{L}(\mathcal{X}, \mathcal{Y})$. According to the introduced convention, let $\mathcal{L}(\mathcal{X}) \coloneqq \mathcal{L}(\mathcal{X}, \mathcal{X})$. The identity mapping on the space $\mathcal{L}(\mathcal{X})$ will be denoted by $\mathcal{I}_\mathcal{X}$. In analogues way to linear superoperators, we can also define the tensor product and direct sum of linear superoperators.

Definition 2.8. *For a given $\Phi \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ the adjoint of Φ is defined to be the unique map $\Phi^\dagger \in \mathcal{L}(\mathcal{Y}, \mathcal{X})$ such that*

$$\operatorname{tr}\left((\Phi^\dagger(Y))^\dagger X\right) = \operatorname{tr}\left(Y^\dagger \Phi(X)\right), \tag{2.37}$$

for all $X \in \mathcal{M}(\mathcal{X})$ and $Y \in \mathcal{M}(\mathcal{Y})$.

Direct sum and tensor product of superoperators

The direct sum between the superoperators of the set $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ are defined in a similar way to direct sum of operators. For the given superoperators $\Phi \in \mathcal{L}(\mathcal{X}_0, \mathcal{Y}_0)$ and $\Psi \in \mathcal{L}(\mathcal{X}_1, \mathcal{Y}_1)$ we define the direct sum of superoperators as

$$(\Psi \oplus \Phi) \in \mathcal{L}(\mathcal{X}_0 \oplus \mathcal{X}_1, \mathcal{Y}_0 \oplus \mathcal{Y}_1), \tag{2.38}$$

to be the unique linear mapping that satisfies the following equation

$$(\Phi \oplus \Psi)(X \oplus Y) = \Phi(X) \oplus \Psi(Y), \tag{2.39}$$

for all $X \in \mathcal{M}(\mathcal{X}_0)$ and $Y \in \mathcal{M}(\mathcal{X}_1)$.

Similarly, let us define the tensor product between the superoperators $\Phi \in \mathcal{L}(\mathcal{X}_0, \mathcal{Y}_0)$ and $\Psi \in \mathcal{L}(\mathcal{X}_1, \mathcal{Y}_1)$ as the unique linear mapping

$$(\Phi \otimes \Psi) \in \mathcal{L}(\mathcal{X}_0 \otimes \mathcal{X}_1, \mathcal{Y}_0 \otimes \mathcal{Y}_1), \tag{2.40}$$

which satisfies the following equation

$$(\Phi \otimes \Psi)(X \otimes Y) = \Phi(X) \otimes \Psi(Y), \tag{2.41}$$

for all $X \in \mathcal{M}(\mathcal{X}_0)$ and $Y \in \mathcal{M}(\mathcal{X}_1)$. As for vectors and operators, the notation $\Phi^{\oplus n}$ and $\Phi^{\otimes n}$ denotes the n-fold direct sum and tensor product of a map Φ with itself, respectively.

Partial trace

Let us consider the map

$$\mathrm{tr} \otimes \mathcal{I}_{\mathcal{Y}} \in \mathcal{L}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{Y}), \tag{2.42}$$

defined as a unique linear map satisfying

$$(\mathrm{tr} \otimes \mathcal{I}_{\mathcal{Y}})(X \otimes Y) = \mathrm{tr}(X)Y. \tag{2.43}$$

This map is called the partial trace, and is more commonly denoted $\mathrm{tr}_{\mathcal{X}}$. Along similar ways, the map $\mathrm{tr}_{\mathcal{Y}} \in \mathcal{L}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{X})$ is defined as

$$\mathrm{tr}_{\mathcal{Y}} = \mathcal{I}_{\mathcal{X}} \otimes \mathrm{tr}. \tag{2.44}$$

Now let us consider an operator $A \in \mathcal{M}(\mathcal{X} \otimes \mathcal{Y})$ not being a tensor product of two operator. Then, from the definition we can express the partial trace of A over space $\mathcal{X}$ as

$$\mathrm{tr}_{\mathcal{X}}(A) = \sum_{i=0}^{\dim(\mathcal{X})-1} \left(\langle i| \otimes \mathbb{1}_{\mathcal{Y}}\right) A \left(|i\rangle \otimes \mathbb{1}_{\mathcal{Y}}\right). \tag{2.45}$$

Along similar lines, the partial trace over space $\mathcal{Y}$ can be written as

$$\mathrm{tr}_{\mathcal{Y}}(A) = \sum_{i=0}^{\dim(\mathcal{Y})-1} \left(\mathbb{1}_{\mathcal{X}} \otimes \langle i|\right) A \left(\mathbb{1}_{\mathcal{X}} \otimes |i\rangle\right). \tag{2.46}$$

We can generalize the concept of partial trace mappings for more then two linear operators and define them in an analogous way.

Partial transpose

Let us consider the map

$$\cdot^{T_{\mathcal{X}}} \in \mathcal{L}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{X} \otimes \mathcal{Y}), \tag{2.47}$$

defined as unique linear map satisfying

$$(X \otimes Y)^{T_{\mathcal{X}}} = X^{\top} \otimes Y. \tag{2.48}$$

Now let us consider an operator $A \in \mathcal{M}(\mathcal{X} \otimes \mathcal{Y})$ not being a tensor product of two operator. Let us define $A \in \mathcal{M}(\mathcal{X} \otimes \mathcal{Y})$ as

$$A = \sum_{i,j=0}^{\dim(\mathcal{X})-1} \sum_{k,l=0}^{\dim(\mathcal{Y})-1} \alpha_{ijkl} |i\rangle\langle j| \otimes |k\rangle\langle l|. \tag{2.49}$$

Then, $A^{T_{\mathcal{X}}}$ is said to be a transpose trace of A on the space $\mathcal{X}$ and has the following form

$$A^{T_{\mathcal{X}}} = \sum_{i,j=0}^{\dim(\mathcal{X})-1} \sum_{k,l=0}^{\dim(\mathcal{Y})-1} \alpha_{ijkl} (|i\rangle\langle j|)^{\top} \otimes |k\rangle\langle l| = \sum_{i,j=0}^{\dim(\mathcal{X})-1} \sum_{k,l=0}^{\dim(\mathcal{Y})-1} \alpha_{ijkl} |j\rangle\langle i| \otimes |k\rangle\langle l|. \tag{2.50}$$

Analogously, we can define $A^{T_{\mathcal{Y}}}$ by replacing $|k\rangle\langle l|$ to $|l\rangle\langle k|$.

Classes of linear maps

The following classes of linear maps will be used later in this dissertation.

A linear map $\Phi \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ is Hermiticity-preserving if it holds that $\Phi(A) \in \mathcal{H}(\mathcal{Y})$ for all $A \in \mathcal{H}(\mathcal{X})$. A linear map $\Phi \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ is trace-preserving if it holds that $\operatorname{tr}(\Phi(A)) = \operatorname{tr}(A)$ for all $A \in \mathcal{M}(\mathcal{X})$. A map $\Phi \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ is positive if it holds $\Phi(A) \in \mathcal{P}(\mathcal{Y})$ for every $A \in \mathcal{P}(\mathcal{X})$. A map $\Phi \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ is completely positive if it holds that $(\Phi \otimes \mathcal{I}_{\mathcal{X}})(A) \in \mathcal{P}(\mathcal{Y} \otimes \mathcal{X})$ for every $A \in \mathcal{P}(\mathcal{X} \otimes \mathcal{X})$. A map $\Phi \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ is unital if it holds that $\Phi(\mathbb{1}_{\mathcal{X}}) = \mathbb{1}_{\mathcal{Y}}$.

2.4 Subchannels and channels

For a positive map Φ we additionally define the term of trace non-increasing map. A positive map Φ is trace non-increasing if $\operatorname{tr}(\Phi(\rho)) \leq 1$ for any $\rho \in \mathcal{D}(\mathcal{X})$.

A linear map Φ, which is completely positive and trace non-increasing is said to be a quantum subchannel. The set of all quantum subchannels will be denoted by $s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ [66]. In particular, the subchannel Φ which saturates $\operatorname{tr}(\Phi(\rho)) = 1$ for all $\rho \in \mathcal{D}(\mathcal{X})$ is known as a quantum channel. The set of quantum channels will be denoted by $\mathcal{C}(\mathcal{X}, \mathcal{Y})$.

One of the most popular quantum channels used in this dissertation are unitary channels Φ_U given by

$$\Phi_U(X) = UXU^* \tag{2.51}$$

where $U \in \mathcal{U}(\mathcal{X})$ and completely depolarizing channel Φ_* described as

$$\Phi_*(X) = \mathrm{tr}(X)\frac{\mathbb{1}_{\mathcal{X}}}{\dim(\mathcal{X})}. \tag{2.52}$$

Representations of subchannels

In this work, we will consider the following representations of subchannels.

- Kraus representation: Each subchannel $\Phi \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ can be defined by a collection of Kraus operators $(K_i)_{i=0}^{r-1} \subset \mathcal{M}(\mathcal{X}, \mathcal{Y})$, such that $\Phi(X) = \sum_{i=0}^{r-1} K_i X K_i^\dagger$ for $X \in \mathcal{M}(\mathcal{X})$ and $r \in \mathbb{N}$. The operators K_i satisfy the condition $\sum_{i=0}^{r-1} K_i^\dagger K_i \leq \mathbb{1}_{\mathcal{X}}$. We say that the subchannel Φ is given in a canonical Kraus representation $(K_i)_{i=0}^{r-1}$, if it holds that $\mathrm{tr}(K_j^\dagger K_i) \propto \delta_{ij}$ and $K_i \neq 0$ for each $i \leq r$. To represent the subchannel Φ by its Kraus representation $(K_i)_{i=0}^{r-1}$, we introduce the notation $\mathcal{K} : \mathcal{M}(\mathcal{X}, \mathcal{Y})^{\times r} \to s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ given by $\Phi = \mathcal{K}\left((K_i)_{i=0}^{r-1}\right)$. Finally, the Kraus representation is not unique. It holds that $\mathcal{K}\left((K_i)_{i=0}^{r-1}\right) = \mathcal{K}\left((K_i')_{i=0}^{r-1}\right)$ if and only if there exists $U \in \mathcal{U}(\mathbb{C}^r)$ such that $K_i' = \sum_j U_{ij} K_j$ for any i. Moreover, $\Phi = \mathcal{K}\left((K_i)_{i=0}^{r-1}\right)$ is a quantum channel belonging to $\mathcal{C}(\mathcal{X}, \mathcal{Y})$ if and only if it holds $\sum_{i=0}^{r-1} K_i^\dagger K_i = \mathbb{1}_{\mathcal{X}}$.

- Choi-Jamiołkowski representation: Each subchannel $\Phi \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ can be uniquely described by its Choi-Jamiołkowski operator $J(\Phi) \in \mathcal{M}(\mathcal{Y} \otimes \mathcal{X})$, which is defied as $J(\Phi) := (\Phi \otimes \mathcal{I}_{\mathcal{X}})(|\mathbb{1}_{\mathcal{X}}\rangle\rangle\langle\langle\mathbb{1}_{\mathcal{X}}|)$. The rank of $J(\Phi)$ is called the Choi rank and it determines the minimal number r of Kraus operators K_i needed to describe Φ in the Kraus form $\Phi = \mathcal{K}\left((K_i)_{i=0}^{r-1}\right)$. Therefore, if the Kraus representation $(K_i)_{i=0}^{r-1}$ is canonical, then $r = \mathrm{rank}(J(\Phi))$. One can retrieve the action of subchannel Φ on a state ρ by using Choi-Jamiołkowski matrix $J(\Phi)$ in the following way $\Phi(\rho) = \mathrm{tr}_{\mathcal{X}}(J(\Phi)(\mathbb{1}_{\mathcal{Y}} \otimes \rho^\top))$. For $\Phi \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ it holds that $\mathrm{tr}_{\mathcal{Y}}(J(\Phi)) = \mathbb{1}_{\mathcal{X}}$.

- Stinespring representation: By the Stinespring Dilatation Theorem, any subchannel $\Phi \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ can be defined as $\Phi(X) = \mathrm{tr}_{\mathcal{Z}}\left(AXA^\dagger\right)$ for $X \in \mathcal{M}(\mathcal{X})$, where $A \in \mathcal{M}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$. The minimal dimension of the space $\mathcal{Z}$ of the auxiliary system is equal to the Choi rank. In particular, for $\Phi \in \mathcal{C}(\mathcal{X})$, the Stinespring representation of Φ can be written in the form $\Phi(X) = \mathrm{tr}_{\mathcal{Z}}\left(U(X \otimes |\psi\rangle\langle\psi|)U^\dagger\right)$, where $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{Z})$ and $U \in \mathcal{U}(\mathcal{X} \otimes \mathcal{Z})$.

Diamond norm

In this section, we will introduce the norm on the linear space $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ called diamond norm (or the completely bounded trace norm) [67, 68].

Definition 2.9. *Let $\mathcal{X}$, $\mathcal{Y}$ be complex Euclidean spaces. The diamond norm $||\cdot||_\diamond : \mathcal{L}(\mathcal{X}, \mathcal{Y}) \to \mathbb{R}$ of a mapping Φ is defined as*

$$||\Phi||_\diamond = \max\left\{||(\Phi \otimes \mathcal{I}_\mathcal{X})(X)||_1 : X \in \mathcal{M}(\mathcal{X} \otimes \mathcal{X}), ||X||_1 \leq 1\right\}. \tag{2.53}$$

One can easily show that for all Hermiticity-preserving maps $\Phi \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ the following equation holds

$$||\Phi||_\diamond = \left\{||(\mathbb{1}_\mathcal{Y} \otimes \sqrt{\rho})\, J(\Phi)\, (\mathbb{1}_\mathcal{Y} \otimes \sqrt{\rho})||_1 : \rho \in \mathcal{D}(\mathcal{X})\right\}. \tag{2.54}$$

where $J(\Phi)$ is the Choi operator of Φ.

2.5 Random matrices

Some basic facts from the theory of random matrices will be relevant in this dissertation. For an in-depth introduction, we refer the reader to the classical textbook [69] or to modern presentations [70, 71]. In this dissertation, the most of introduced probabilistic measures describing random quantum objects will be induced from the standard complex Gaussian distribution [72]. In detail, let X and Y be independent standard normal real variables (both having mean 0 and variance 1). We define the standard complex Gaussian variable $Z = \frac{X+iY}{\sqrt{2}}$. The probability density function of Z is equal $f(z) = \frac{1}{\pi}\exp(-|z|^2)$.

We will chronically use the notation of Dirac delta function [73] $\mathbb{R} \ni x \mapsto \delta(x)$, that is the limit probability distribution that is concentrated in 0. Formally speaking, δ can be seen as a distribution - not a function in itself but only about how it affects other functions when integrated against them, that is $\int \delta(x) f(x) dx = f(0)$. Informally, the delta function is often defined as

$$\delta(x) = \begin{cases} +\infty, & x = 0, \\ 0, & x \neq 0. \end{cases} \tag{2.55}$$

We extend the definition of the Dirac delta function to the case of complex matrices and distinguish two cases:

- If $H \in \mathcal{H}(\mathcal{X})$ is a Hermitian matrix, then $\delta(H)$ is defined as

$$\delta(H) = \prod_{i=0}^{\dim(\mathcal{X})-1} \delta(H_{i,i}) \prod_{i=0}^{\dim(\mathcal{X})-2} \prod_{j=i+1}^{\dim(\mathcal{X})-1} \delta(\mathrm{Re}(H_{i,j}))\delta(\mathrm{Im}(H_{i,j})) \tag{2.56}$$

- If $A \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ is a complex matrix, then $\delta(A)$ is defined as

$$\delta(A) = \prod_{i=0}^{\dim(\mathcal{Y})-1} \prod_{j=0}^{\dim(\mathcal{X})-1} \delta(\mathrm{Re}(A_{i,j}))\delta(\mathrm{Im}(A_{i,j})) \tag{2.57}$$

In this work we will use the following fact about Delta function of a matrix argument.

Lemma 2.10 ([73]). *Let $H \in \mathcal{H}(\mathcal{X})$ be a Hermitian matrix and $A \in \mathcal{M}(\mathcal{X})$ be some complex and invertible matrix. Then, it holds*

$$\delta(AHA^\dagger) = \det(AA^\dagger)^{-\dim(\mathcal{X})}\delta(H). \tag{2.58}$$

We will also use the following fact about Jacobian determinant of a matrix substitution.

Lemma 2.11 ([73]). *Let $A, A' \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ be some complex matrices and let $X \in \mathcal{M}(\mathcal{X})$ and $Y \in \mathcal{M}(\mathcal{Y})$ be some complex and invertible matrices. If $A' = YAX$, then $dA' = \det(XX^\dagger)^{\dim(\mathcal{Y})} \det(YY^\dagger)^{\dim(\mathcal{X})} dA$.*

In this work we are going to use the following ensembles of random matrices:

- **the complex Ginibre matrices** consisting of matrices G with independent complex entries distributed according to the standard complex Gaussian distribution. Note that the Ginibre matrices can be rectangular. The complex Ginibre matrices almost surely have the maximal possible rank, that is; for $G \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ almost surely it holds $\mathrm{rank}(G) = \min(\dim(\mathcal{X}), \dim(\mathcal{Y}))$. Moreover, the probability density function of G is equal to $\frac{1}{\pi^{\dim(\mathcal{X}\otimes\mathcal{Y})}} \exp(-\mathrm{tr}(GG^\dagger))$.
- **the complex Wishart matrices** consisting of matrices $W_r \in \mathcal{P}(\mathcal{X})$ of parameter $r \in \mathbb{N}$ that are defined as $W_r = GG^\dagger$, where $G \in \mathcal{M}(\mathbb{C}^r, \mathcal{X})$ is a complex Ginibre matrix [71]. By the construction almost surely we have $\mathrm{rank}(W_r) = \min(r, \dim(\mathcal{X}))$. Moreover, one can observe that the following two ensembles have the same distribution:
  - $\mathrm{tr}_{\mathcal{Y}}(W_r)$, where $W_r \in \mathcal{P}(\mathcal{Y} \otimes \mathcal{X})$ is the complex Wishart matrix of parameter r,
  - $W'_{r\dim(\mathcal{Y})}$, where $W'_{r\dim(\mathcal{Y})} \in \mathcal{P}(\mathcal{X})$ is the complex Wishart matrix of parameter $r\dim(\mathcal{Y})$.
- **the random isometry ensemble** consisting of Haar-distributed random operators $V \in \mathcal{U}(\mathcal{X}, \mathcal{Y})$ [74], where the Haar distribution is the unique

probability measure invariant with respect to left and right multiplication with fixed unitary matrices. The special case of random isometry ensamble constitute **the circular unitary ensemble** (CUE) consisting of random unitary matrices $U \in \mathcal{U}(\mathcal{X})$ distributed according to the Haar measure on the unitary group [69, 75]. In this dissertation we will use the following method of generating Haar-distributed random isometry matrices. Let $G \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ be the complex Ginibre matrix and define $V = G(G^\dagger G)^{-1/2}$. Almost surely, V is well-defined, random Haar isometry matrix, $V \in \mathcal{U}(\mathcal{X}, \mathcal{Y})$ [76].

2.6 Semidefinite programming

The paradigm of semidefinite programming finds numerous applications in the theory of quantum information, both analytical and computational. This section describes a formulation of semidefinite programming (SDP).

Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, and let $\Phi \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ be a Hermiticity-preserving map. Let $A \in \mathcal{H}(\mathcal{X})$ and $B \in \mathcal{H}(\mathcal{Y})$ be Hermitian operators. A semidefinite program is defined as a triple (Φ, A, B) with which the following pair of optimization problems is associated.

Primal problem		Dual problem	
maximize:	$\mathrm{tr}\left(A^\dagger X\right)$	minimize:	$\mathrm{tr}\left(B^\dagger Y\right)$
subject to:	$\Phi(X) = B$,	subject to:	$\Phi^\dagger(Y) \geq A$,
	$X \in \mathcal{P}(\mathcal{X})$.		$Y \in \mathcal{H}(\mathcal{Y})$.

Table 2.1: Formulation of primal and dual problem.

Let us define the primal feasible set $\mathcal{A}$ and the dual feasible set $\mathcal{B}$ of (Φ, A, B) as follows

$$\begin{aligned} \mathcal{A} &= \{X \in \mathcal{P}(\mathcal{X}) : \Phi(X) = B\}, \\ \mathcal{B} &= \{Y \in \mathcal{H}(\mathcal{Y}) : \Phi^\dagger(X) \geq A\}. \end{aligned} \tag{2.59}$$

Operators $X \in \mathcal{A}$ and $Y \in \mathcal{B}$ are also said to be primal feasible and dual feasible, respectively. The optimum values associated with the primal and dual problems are defined as

$$\alpha = \sup\{\mathrm{tr}\left(A^\dagger X\right) : A \in \mathcal{A}\}, \tag{2.60}$$

and

$$\beta = \inf\{\mathrm{tr}\left(B^\dagger Y\right) : B \in \mathcal{B}\}, \tag{2.61}$$

respectively.

Semidefinite programs have associated with them a notion of duality, which refers to the special relationship between the primal and dual problems. For many semidefinite programs, it happens that primary and dual problem values are equal. This situation is called strong duality. Slater's theorem provides one set of conditions under which strong duality is guaranteed.

Theorem 2.12. *(Slater's theorem for semidefinite programs) Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $\Phi \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ be a Hermitian-preserving map, and let $A \in \mathcal{H}(\mathcal{X})$ and $B \in \mathcal{H}(\mathcal{Y})$ be Hermitian operators. Let $\mathcal{A}, \mathcal{B}, \alpha, \beta$ be as defined above for the semidefinite program (Φ, A, B) one has the following two implications:*

1. *If α is finite and there exists a Hermitian operator $Y \in \mathcal{H}(\mathcal{Y})$ such that $\Phi^\dagger(Y) > A$, then $\alpha = \beta$, and moreover there exists a primal-feasible operator $X \in \mathcal{A}$ such that $\operatorname{tr}\left(A^\dagger X\right) = \alpha$.*

2. *If β is finite and there exists a positive definite operator $X > 0$ such that $\Phi(X) = B$, then $\alpha = \beta$, and moreover there exists a dual-feasible operator $Y \in \mathcal{B}$ such that $\operatorname{tr}\left(B^\dagger Y\right) = \beta$.*

Chapter 3

Random Quantum Operations

The goal of this chapter is to introduce ensembles of random quantum operations; mostly quantum channels, but also quantum subchannels, instruments, super-channels and quantum networks. Due to the diversity of quantum operations these ensembles provide, they are suitable for numerical investigation of various quantum properties and testing effectiveness of various procedures. In this dissertation, we will use random quantum operations to check the effectiveness of probabilistic quantum error correction codes. In the second part of this chapter we will also investigate some properties of random quantum channels.

This chapter is based mostly on [1]. Additionally, the Section 3.1.6 and Section 3.1.7 include unpublished, author results concerning sampling methods of random quantum instruments, subchannels and super-channels.

3.1 Distributions of random quantum operations

We introduce next three methods to generate random channels from the set $\mathcal{C}(\mathcal{X}, \mathcal{Y})$. We show for which particular choices of parameters the methods become equivalent and when they induce the flat measure on the set $\mathcal{C}(\mathcal{X}, \mathcal{Y})$.

3.1.1 Random Choi-Jamiołkowski matrix

Definition 3.1 ([63]). *Let $\mathcal{X}, \mathcal{Y}$ be given complex Euclidean spaces and let $r \in \mathbb{N}$ be a parameter satisfying*

$$r \geq \frac{\dim(\mathcal{X})}{\dim(\mathcal{Y})}. \tag{3.1}$$

We define $\mu^{Choi}_{\mathcal{X},\mathcal{Y};r}$ to be the probability measure of the random quantum channel $\Phi \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ given by a sampling procedure:

1. *Generate a random complex Wishart matrix $W_r \in \mathcal{P}(\mathcal{Y} \otimes \mathcal{X})$ of parameter r;*

2. *Calculate* $Q = \mathrm{tr}_{\mathcal{Y}}(W_r)$*;*

3. *Write the Choi-Jamiołkowski matrix of the random channel* $\Phi \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ *as*

$$J(\Phi) = \left(\mathbb{1}_{\mathcal{Y}} \otimes Q^{-1/2}\right) W_r \left(\mathbb{1}_{\mathcal{Y}} \otimes Q^{-1/2}\right). \tag{3.2}$$

Several remarks are in order here. First, the lower bound on the integer value of r, $\dim(\mathcal{Y})r \geq \dim(\mathcal{X})$, implies that the random matrix Q is, generically, invertible. Indeed, $Q \in \mathcal{P}(\mathcal{X})$ follows a Wishart distribution of the parameter $\dim(\mathcal{Y})r$. Hence, with probability one, the random matrix $J(\Phi)$ is constructed to be positive semi-definite and satisfies the normalization condition $\mathrm{tr}_{\mathcal{Y}}(J(\Phi)) = \mathbb{1}_{\mathcal{X}}$. The rank of the Choi matrix $J(\Phi)$ (and thus the Choi rank of Φ) is, almost surely equal to

$$\mathrm{rank}(J(\Phi)) = \min(\dim(\mathcal{X})\dim(\mathcal{Y}), r). \tag{3.3}$$

3.1.2 Random Kraus operators

Definition 3.2 ([63]). *Let* $\mathcal{X}, \mathcal{Y}$ *be given complex Euclidean spaces and let* $r \in \mathbb{N}$ *be a parameter satisfying*

$$r \geq \frac{\dim(\mathcal{X})}{\dim(\mathcal{Y})}. \tag{3.4}$$

We define $\mu^{Kraus}_{\mathcal{X},\mathcal{Y};r}$ *to be the probability measure of the random quantum channel* $\Phi \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ *given by a sampling procedure:*

1. *Generate* r *independent complex Ginibre matrices* $G_1, \ldots, G_r \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$*;*

2. *Compute* $Q = \sum_{i=1}^{r} G_i^\dagger G_i \geq 0$*;*

3. *The channel* Φ *is defined via its Kraus decomposition* $\Phi = \mathcal{K}\left((G_i Q^{-1/2})_{i=1}^{r}\right)$.

Let us first justify the validity of the construction. As in the random Choi matrix setting above, the matrix $Q \in \mathcal{P}(\mathcal{X})$ has a Wishart distribution of the parameter $\dim(\mathcal{Y})r$, hence, it is generically invertible. Therefore, it holds $\sum_{i=1}^{r} Q^{-1/2} G_i^\dagger G_i Q^{-1/2} = \mathbb{1}_{\mathcal{X}}$ and Φ is a legitimate quantum channel.

3.1.3 Random environmental form

Definition 3.3 ([63]). *Let* $\mathcal{X}, \mathcal{Y}$ *be given complex Euclidean spaces and let* $r \in \mathbb{N}$ *be a parameter satisfying*

$$r \geq \frac{\dim(\mathcal{X})}{\dim(\mathcal{Y})}. \tag{3.5}$$

We define $\mu^{Stinespring}_{\mathcal{X},\mathcal{Y};r}$ *to be the probability measure of the random quantum channel* $\Phi \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ *given by a sampling procedure:*

1. Generate a random Haar isometry $V \in \mathcal{U}(\mathcal{X}, \mathcal{Y} \otimes \mathbb{C}^r)$;

2. The channel $\Phi \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ is defined by its Stinespring decomposition

$$\Phi(\cdot) = \operatorname{tr}_{\mathbb{C}^r}\left(V \cdot V^\dagger\right). \tag{3.6}$$

3.1.4 The Lebesgue (flat) measure

Finally, a natural probability measure is the (normalized) Lebesgue (or flat) measure on the set of quantum channels. Since the set $\mathcal{C}(\mathcal{X}, \mathcal{Y})$ is a convex compact set, one can endow it with the probability measure obtained by normalizing the volume to have total mass 1. We denote the flat measure on the set of quantum channels $\mathcal{C}(\mathcal{X}, \mathcal{Y})$ by $\mu^{Lebesgue}_{\mathcal{X},\mathcal{Y}}$.

3.1.5 Equivalence of sampling methods and their comparison

Proposition 3.4 ([1]). *For all integers r such that $\dim(\mathcal{Y})r \geq \dim(\mathcal{X})$, we have the equality of probability measures*

$$\mu^{Choi}_{\mathcal{X},\mathcal{Y};r} = \mu^{Kraus}_{\mathcal{X},\mathcal{Y};r} = \mu^{Stinespring}_{\mathcal{X},\mathcal{Y};r}. \tag{3.7}$$

Proof. First, we will show that $\mu^{Choi}_{\mathcal{X},\mathcal{Y};r} = \mu^{Kraus}_{\mathcal{X},\mathcal{Y};r}$. Let $\Phi = \mathcal{K}\left((G_i Q^{-1/2})_{i=1}^r\right)$ be a random quantum channel defined as in Definition 3.2. The Choi-Jamiołkowski matrix $J(\Phi)$ can be expressed in the terms of given Kraus operators as

$$J(\Phi) = \sum_{i=1}^r |G_i Q^{-1/2}\rangle\rangle\langle\langle G_i Q^{-1/2}| = (\mathbb{1}_\mathcal{Y} \otimes (Q^\top)^{-1/2}) \sum_{i=1}^r |G_i\rangle\rangle\langle\langle G_i| (\mathbb{1}_\mathcal{Y} \otimes (Q^\top)^{-1/2}). \tag{3.8}$$

We can observe that the matrix $\sum_{i=1}^r |G_i\rangle\rangle\langle\langle G_i| \in \mathcal{P}(\mathcal{Y} \otimes \mathcal{X})$ follows the complex Wishart distribution with the parameter r. Moreover, we have

$$\operatorname{tr}_\mathcal{Y}\left(\sum_{i=1}^r |G_i\rangle\rangle\langle\langle G_i|\right) = \sum_{i=1}^r G_i^\top \overline{G_i} = Q^\top, \tag{3.9}$$

which proves the Φ is sampled according to Definition 3.1.

Now, we will show $\mu^{Kraus}_{\mathcal{X},\mathcal{Y};r} = \mu^{Stinespring}_{\mathcal{X},\mathcal{Y};r}$. Let $G \in \mathcal{M}(\mathcal{X}, \mathcal{Y} \otimes \mathbb{C}^r)$ be a random complex Ginibre matrix. Define a matrix V in the following way

$$V = G(G^\dagger G)^{-1/2}. \tag{3.10}$$

Almost surely, V is well-defined, random Haar isometry matrix, $V \in \mathcal{U}(\mathcal{X}, \mathcal{Y} \otimes \mathbb{C}^r)$. Let Φ be a random channel generated according to Definition 3.3. If $V =$

$G(G^\dagger G)^{-1/2}$ is the isometry matrix defining Φ, then the Kraus operators of $\Phi = \mathcal{K}\left((A_i)_{i=1}^r\right)$ are of the form $A_i = (\mathbb{1}_\mathcal{Y} \otimes \langle i|)V$ for $i = 1, \ldots, r$. The matrix G follows the complex Ginibre distribution, therefore it can be written as $G = \sum_{i=1}^r G_i \otimes |i\rangle$, where $G_i \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ are independent complex Ginibre matrices. Hence, we write $G^\dagger G = \sum_{i=1}^r G_i^\dagger G_i$ and eventually, $A_i = G_i \left(\sum_{i=1}^r G_i^\dagger G_i\right)^{-1/2}$. That means, Φ is generated according to Definition 3.2. □

Proposition 3.5 ([1]). *For all $\mathcal{X}, \mathcal{Y}$ it holds that the flat measure $\mu_{\mathcal{X},\mathcal{Y}}^{Lebesgue}$ on the set of quantum channels $\mathcal{C}(\mathcal{X}, \mathcal{Y})$ is a particular case of the constructions in Definitions 3.1, 3.2, 3.3, obtained for the parameter $r = \dim(\mathcal{X})\dim(\mathcal{Y})$, that is*

$$\mu_{\mathcal{X},\mathcal{Y}}^{Lebesgue} = \mu_{\mathcal{X},\mathcal{Y};\dim(\mathcal{X})\dim(\mathcal{Y})}^{Choi} = \mu_{\mathcal{X},\mathcal{Y};\dim(\mathcal{X})\dim(\mathcal{Y})}^{Kraus} = \mu_{\mathcal{X},\mathcal{Y};\dim(\mathcal{X})\dim(\mathcal{Y})}^{Stinespring}. \tag{3.11}$$

Proof. We will show that $\mu_{\mathcal{X},\mathcal{Y}}^{Lebesgue} = \mu_{\mathcal{X},\mathcal{Y};\dim(\mathcal{X})\dim(\mathcal{Y})}^{Choi}$. We use standard calculus methods to obtain the distribution of $J(\Phi)$ for Φ sampled according to Definition 3.1. Let $f_{J(\Phi)}(D)$ be the probability density function of the random Choi matrix $J(\Phi)$ at the point $D \in \mathcal{P}(\mathcal{Y} \otimes \mathcal{X})$. The distribution of matrix $J(\Phi)$ is induced by the distribution of the complex Wishart matrix $W = W_{\dim(\mathcal{X})\dim(\mathcal{Y})} \in \mathcal{P}(\mathcal{Y} \otimes \mathcal{X})$, that is

$$J(\Phi) = \left(\mathbb{1}_\mathcal{Y} \otimes (\mathrm{tr}_\mathcal{Y}(W))^{-1/2}\right) W \left(\mathbb{1}_\mathcal{Y} \otimes (\mathrm{tr}_\mathcal{Y}(W))^{-1/2}\right). \tag{3.12}$$

What is more, the distribution of W is induced by the distribution of the complex Ginibre matrix $G \in \mathcal{M}(\mathcal{Y} \otimes \mathcal{X})$, that is $W = GG^\dagger$. Therefore, by the properties of Dirac delta function we have

$$\begin{aligned} f_{J(\Phi)}(D) &\propto \int \delta(J(\Phi) - D)\exp(-\mathrm{tr}GG^\dagger)dG \\ &= \int\int \delta((\mathbb{1}_\mathcal{Y} \otimes Q^{-1/2})GG^\dagger(\mathbb{1}_\mathcal{Y} \otimes Q^{-1/2}) - D)\delta(Q - \mathrm{tr}_\mathcal{Y}(GG^\dagger))\exp(-\mathrm{tr}Q)dQdG. \end{aligned} \tag{3.13}$$

As Q is almost surely strictly positive we may use the substitution $G = (\mathbb{1}_\mathcal{Y} \otimes \sqrt{Q})\sqrt{D}G'$, which leads to $dG = \det(Q)^{\dim(\mathcal{X})\dim(\mathcal{Y})^2}\dim(D)^{\dim(\mathcal{X})\dim(\mathcal{Y})}dG'$ and we get

$$
\begin{aligned}
&f_{J(\Phi)}(D)\\
&\propto \int\int \delta(\sqrt{D}GG^\dagger\sqrt{D}-D)\delta(Q-\sqrt{Q}\mathrm{tr}_{\mathcal{Y}}(\sqrt{D}GG^\dagger\sqrt{D})\sqrt{Q})\\
&\qquad \exp(-\mathrm{tr}Q)\det(Q)^{\dim(\mathcal{X})\dim(\mathcal{Y})^2}\dim(D)^{\dim(\mathcal{X})\dim(\mathcal{Y})}dQdG\\
&=\int\int \delta(GG^\dagger-\mathbb{1}_{\mathcal{Y}\otimes\mathcal{X}})\delta(Q-\sqrt{Q}\mathrm{tr}_{\mathcal{Y}}(\sqrt{D}GG^\dagger\sqrt{D})\sqrt{Q})\\
&\qquad \exp(-\mathrm{tr}Q)\det(Q)^{\dim(\mathcal{X})\dim(\mathcal{Y})^2}dQdG\\
&=\int\int \delta(GG^\dagger-\mathbb{1}_{\mathcal{Y}\otimes\mathcal{X}})\delta(\mathbb{1}_{\mathcal{X}}-\mathrm{tr}_{\mathcal{Y}}(D))\exp(-\mathrm{tr}Q)\det(Q)^{\dim(\mathcal{X})(\dim(\mathcal{Y})^2-1)}dQdG\\
&\propto\delta(\mathbb{1}_{\mathcal{X}}-\mathrm{tr}_{\mathcal{Y}}(D)).
\end{aligned}
\tag{3.14}
$$

Hence, the density function is constant for each $D \geq 0$, such that $\mathrm{tr}_{\mathcal{Y}}(D) = \mathbb{1}_{\mathcal{X}}$, which proves the claim. □

Finally, let us discuss computational complexity of sampling methods (Definitions 3.1, 3.2, 3.3) in numerical analyzes. For the sake of simplicity, consider the case $\dim(\mathcal{X}) = \dim(\mathcal{Y})$ and $r \leq \dim(\mathcal{X})^2$. We assume that the complexity of a matrix multiplication, calculating matrix inverse, the spectral decomposition and matrix power in typical implementations is $\mathcal{O}(n^3)$, where $M \in \mathcal{M}(\mathbb{C}^n)$, although the author is aware that there exist faster algorithms, for example for matrix multiplication [77] with complexity $\Theta(n^{2.37188})$. However, most of them are not practical for reasonable range of values n due to large constant overhead.

- Random Choi-Jamiołkowski matrix (Definition 3.1): First we define random complex Ginibre matrix $G \in \mathcal{M}(\mathbb{C}^r, \mathcal{Y}\otimes\mathcal{X})$ - $\dim(\mathcal{X})^2 r$ random complex variables in total. Then, we calculate the Wishart matrix $W_r = GG^\dagger \in \mathcal{P}(\mathcal{Y}\otimes\mathcal{X})$ - matrix multiplication complexity $\mathcal{O}(\dim(\mathcal{X})^4 r)$ and define $Q = \mathrm{tr}_{\mathcal{Y}}(W_r)$ - matrix addition complexity $\mathcal{O}(\dim(\mathcal{X})^3)$. Finally, to define $J(\Phi)$ it is necessary to define $Q^{-1/2}$ (complexity $O(\dim(\mathcal{X})^3)$) and multiple $(\mathbb{1}_{\mathcal{Y}}\otimes Q^{-1/2})$ with W_r (without using the fact that the first matrix is sparse, the complexity is $\mathcal{O}(\dim(\mathcal{X})^6)$). Finally, to compute $\Phi(\rho)$ for some $\rho \in \mathcal{D}(\mathcal{X})$ one can use the equation $\Phi(\rho) = \mathrm{tr}_{\mathcal{X}}(J(\Phi)(\mathbb{1}_{\mathcal{Y}}\otimes\rho^\top))$ - the complexity of the operation is $\mathcal{O}(\dim(\mathcal{X})^6)$.
- Random Kraus operators (Definition 3.2): We generate r complex Ginibre matrices $G_i \in \mathcal{M}(\mathcal{X})$ - $\dim(\mathcal{X})^2 r$ random complex variables in total. To compute $Q = \sum_{i=1}^r G_i^\dagger G_i$ we need $\mathcal{O}(\dim(\mathcal{X})^3 r)$ operations. To define $(G_i Q^{-1/2})_{i=1}^r$ we use $\mathcal{O}(\dim(\mathcal{X})^3)$ operations to define $Q^{-1/2}$ and then $\mathcal{O}(\dim(\mathcal{X})^3 r)$ operations

to get Φ. Calculating the action of Φ on $\rho \in \mathcal{D}(\mathcal{X})$, that is $\Phi(\rho) = \sum_{i=1}^{r} G_i \rho G_i^\dagger$ has once more the complexity $\mathcal{O}(\dim(\mathcal{X})^3 r)$.

- Random environmental form (Definition 3.3): We can define $V \in \mathcal{U}(\mathcal{X}, \mathcal{Y} \otimes \mathbb{C}^r)$ by first sampling random complex Ginibre matrix $G \in \mathcal{M}(\mathcal{X}, \mathcal{Y} \otimes \mathbb{C}^r)$ - $\dim(\mathcal{X})^2 r$ random complex variables in total, and then, by using the SVD procedure - in that case having the complexity $\mathcal{O}(\dim(\mathcal{X})^3 r)$. The action of the channel Φ on some state $\rho \in \mathcal{D}(\mathcal{X})$ is given as $\Phi(\rho) = \mathrm{tr}_{\mathbb{C}^r}(V \rho V^\dagger)$, hence, we need $\mathcal{O}(\dim(\mathcal{X})^3 r^2)$ operations to perform.

We summarize the discussion about computational complexity in the table below, where for each method we present: the number of random complex variables required to define quantum channel; the computational complexity of sampling quantum channel in considered representation; the computational complexity of performing the given quantum channel on arbitrary input:

Method	Choi	Kraus	Stinespring
Number of variables	$\dim(\mathcal{X})^2 r$	$\dim(\mathcal{X})^2 r$	$\dim(\mathcal{X})^2 r$
Complexity of sampling	$\dim(\mathcal{X})^6$	$\dim(\mathcal{X})^3 r$	$\dim(\mathcal{X})^3 r$
Complexity of using	$\dim(\mathcal{X})^6$	$\dim(\mathcal{X})^3 r$	$\dim(\mathcal{X})^3 r^2$

As we can see, the random Kraus operators form is the most suitable for numerical investigation.

3.1.6 Random instruments and subchannels

We may use the techniques proposed in Section 3.1 to provide a method for sampling random quantum subchannels and instruments. We show how to obtain a flat measure in both cases and indicate the differences in sampling.

Random instruments

A quantum instrument is a tuple of quantum subchannels that sum is equal to some channel [62]. More formally, let $N \in \mathbb{N}$ and $\mathcal{X}, \mathcal{Y}$ be some Euclidean spaces. We say that a tuple $(\Phi_n)_{n=1}^N \subset s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ is a quantum instrument if $\sum_{n=1}^N \Phi_n \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$.

Definition 3.6. *Let $\mathcal{X}, \mathcal{Y}$ be given complex Euclidean spaces and let $n \in \mathbb{N}$ and $r_n \in \mathbb{N}$ for $n = 1, \ldots, N$ be some parameters, such that*

$$\sum_{n=1}^{N} r_n \geq \frac{\dim(\mathcal{X})}{\dim(\mathcal{Y})}. \tag{3.15}$$

We define a procedure of generating random quantum instruments $(\Phi_n)_{n=1}^N \subset s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ as follows:

1. *Generate a sequence of random and independent complex Wishart matrices* $W_{r_n}^{(n)} \in \mathcal{P}(\mathcal{Y} \otimes \mathcal{X})$ *of parameter* r_n *for* $n = 1, \ldots, N$;
2. *Calculate* $Q = \mathrm{tr}_{\mathcal{Y}} \left(\sum_{n=1}^{N} W_{r_n}^{(n)} \right)$;
3. *Define* $(\Phi_n)_{n=1}^{N} \subset s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ *by the Choi-Jamiołkowski matrices in the following way:*
$$J(\Phi_n) = \left(\mathbb{1}_{\mathcal{Y}} \otimes Q^{-1/2}\right) W_{r_n}^{(n)} \left(\mathbb{1}_{\mathcal{Y}} \otimes Q^{-1/2}\right). \tag{3.16}$$

Due to the condition $\sum_{n=1}^{N} r_n \dim(\mathcal{Y}) \geq \dim(\mathcal{X})$ almost surely the matrix Q in the above definition is invertible and therefore the definition of quantum instrument is correct.

Proposition 3.7. *For all Euclidean spaces* $\mathcal{X}, \mathcal{Y}$ *and* $N \in \mathbb{N}$ *random quantum instruments* $(\Phi_n)_{n=1}^{N} \subset s\mathcal{C}(\mathcal{X}, \mathcal{Y})$, *generated according to Definition 3.6 for* $r_n = \dim(\mathcal{X}) \dim(\mathcal{Y})$, *where* $n = 1, \ldots, N$, *are uniformly distributed.*

Proof. We use the similar proof technique as in Proposition 3.5. Let $f(D_1, \ldots, D_N)$ be the probability density function of the random instrument $(J(\Phi_n))_{n=1}^{N}$ at the point $(D_n)_{n=1}^{N} \subset \mathcal{P}(\mathcal{Y} \otimes \mathcal{X})$. This distribution is induced by the product of distributions of complex Ginibre matrices $G_n \in \mathcal{M}(\mathcal{Y} \otimes \mathcal{X})$, $n = 1, \ldots, N$. We have

$$\begin{aligned}
&f(D_1, \ldots, D_N) \propto \int \prod_{n=1}^{N} \delta(J(\Phi_n) - D_n) \exp(-\mathrm{tr} G_n G_n^\dagger) dG_n \\
&= \int \prod_{n=1}^{N} \delta(\left(\mathbb{1}_{\mathcal{Y}} \otimes Q^{-1/2}\right) G_n G_n^\dagger \left(\mathbb{1}_{\mathcal{Y}} \otimes Q^{-1/2}\right) - D_n) \delta\left(Q - \mathrm{tr}_{\mathcal{Y}}\left(\sum_{n=1}^{N} G_n G_n^\dagger\right)\right) \\
&\qquad \exp(-\mathrm{tr} Q) dG_n dQ \\
&= \int \prod_{n=1}^{N} \delta(G_n G_n^\dagger - D_n) \delta\left(Q - \sqrt{Q} \mathrm{tr}_{\mathcal{Y}}\left(\sum_{n=1}^{N} G_n G_n^\dagger\right) \sqrt{Q}\right) \\
&\qquad \det(Q)^{N \dim(\mathcal{Y})^2 \dim(\mathcal{X})} \exp(-\mathrm{tr} Q) dG_n dQ \\
&= \int \prod_{n=1}^{N} \delta(G_n G_n^\dagger - D_n) \delta\left(\mathbb{1}_{\mathcal{X}} - \mathrm{tr}_{\mathcal{Y}}\left(\sum_{n=1}^{N} D_n\right)\right) \\
&\qquad \det(Q)^{(N \dim(\mathcal{Y})^2 - 1) \dim(\mathcal{X})} \exp(-\mathrm{tr} Q) dG_n dQ \\
&= \int \prod_{n=1}^{N} \delta(G_n G_n^\dagger - \mathbb{1}_{\mathcal{Y} \otimes \mathcal{X}}) \delta\left(\mathbb{1}_{\mathcal{X}} - \mathrm{tr}_{\mathcal{Y}}\left(\sum_{n=1}^{N} D_n\right)\right) \\
&\qquad \det(Q)^{(N \dim(\mathcal{Y})^2 - 1) \dim(\mathcal{X})} \exp(-\mathrm{tr} Q) dG_n dQ \propto \delta\left(\mathbb{1}_{\mathcal{X}} - \mathrm{tr}_{\mathcal{Y}}\left(\sum_{n=1}^{N} D_n\right)\right).
\end{aligned} \tag{3.17}$$

□

Random subchannels

Although it is possible to generate random subchannel by at first sampling a random instrument and then taking the first element of the collection, this method has two drawbacks: it is inefficient; it is not obvious how to obtain the flat measure. Therefore, we present an alternative method of generating random quantum subchannels.

Definition 3.8. *Let $\mathcal{X}, \mathcal{Y}$ be given complex Euclidean spaces and let $r, r' \in \mathbb{N}$ be some parameters. We define a procedure of generating random quantum subchannels $\Phi \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ as follows:*

1. *Generate two random and independent complex Wishart matrices: $W_r \in \mathcal{P}(\mathcal{Y} \otimes \mathcal{X})$ of parameter r and $W'_{r'} \in \mathcal{P}(\mathcal{X})$ of parameter r';*

2. *Calculate $Q = \mathrm{tr}_{\mathcal{Y}}(W_r) + W'_{r'}$;*

3. *Write the Choi-Jamiołkowski matrix of random subchannel $\Phi \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ as*

$$J(\Phi) = \left(\mathbb{1}_{\mathcal{Y}} \otimes Q^{-1/2}\right) W_r \left(\mathbb{1}_{\mathcal{Y}} \otimes Q^{-1/2}\right). \tag{3.18}$$

In that definition there is no particular restrictions on the parameters r, r', henceforth the matrix $Q + W'_{r'}$ may not be invertible (as a reminder in that case $\cdot^{-1}$ means Moore-Penrose pseudo-inverse). However, the construction returns a valid Choi-Jamiołkowski matrix of a subchannel Φ, that is

$$\mathrm{tr}_{\mathcal{Y}}(J(\Phi)) = Q^{-1/2}\mathrm{tr}_{\mathcal{Y}}(W_r)Q^{-1/2} \leq \mathbb{1}_{\mathcal{X}}. \tag{3.19}$$

Proposition 3.9. *For all Euclidean spaces $\mathcal{X}, \mathcal{Y}$ random quantum subchannels $\Phi \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$, generated according to Definition 3.8 for $r = \dim(\mathcal{X})\dim(\mathcal{Y})$ and $r' = \dim(\mathcal{X})$, are uniformly distributed.*

Proof. We use the similar proof technique as in Proposition 3.5. Let $f(D_1, D_2)$ be the probability density function of the tuple $(J(\Phi), Q^{-1/2}W'_{\dim(\mathcal{X})}Q^{-1/2})$ at the point $(D_1, D_2) \in \mathcal{P}(\mathcal{Y} \otimes \mathcal{X}) \times \mathcal{P}(\mathcal{X})$, where $J(\Phi), Q$ and $W'_{\dim(\mathcal{X})}$ are defined as in Definition 3.8. The distribution of this tuple is induced by the distribution of two independent complex Ginibre matrices $G \in \mathcal{M}(\mathcal{Y} \otimes \mathcal{X})$ and $G' \in \mathcal{M}(\mathcal{X})$. We have

$$
\begin{aligned}
& f(D_1, D_2) \propto \\
\propto & \int \delta(J(\Phi) - D_1)\delta(Q^{-1/2} W'_{\dim(\mathcal{X})} Q^{-1/2} - D_2) \exp(-\mathrm{tr} G G^\dagger) \exp(-\mathrm{tr} G' G'^\dagger) dG dG' \\
= & \int \delta(\left(\mathbb{1}_\mathcal{Y} \otimes Q^{-1/2}\right) G G^\dagger \left(\mathbb{1}_\mathcal{Y} \otimes Q^{-1/2}\right) - D_1)\delta(Q^{-1/2} G' G'^\dagger Q^{-1/2} - D_2) \\
& \quad \delta(Q - (\mathrm{tr}_\mathcal{Y}(G G^\dagger) + G' G'^\dagger)) \exp(-\mathrm{tr} Q) dG dG' dQ \\
= & \int \delta(G G^\dagger - D_1)\delta(G' G'^\dagger - D_2)\delta(Q - \sqrt{Q}(\mathrm{tr}_\mathcal{Y}(G G^\dagger) + G' G'^\dagger)\sqrt{Q}) \\
& \quad \det(Q)^{(\dim(\mathcal{Y})^2 + 1)\dim(\mathcal{X})} \exp(-\mathrm{tr} Q) dG dG' dQ \\
= & \int \delta(G G^\dagger - D_1)\delta(G' G'^\dagger - D_2)\delta(\mathbb{1}_\mathcal{X} - (\mathrm{tr}_\mathcal{Y}(D_1) + D_2)) \\
& \quad \det(Q)^{\dim(\mathcal{Y})^2 \dim(\mathcal{X})} \exp(-\mathrm{tr} Q) dG dG' dQ \propto \delta(\mathbb{1}_\mathcal{X} - (\mathrm{tr}_\mathcal{Y}(D_1) + D_2)).
\end{aligned}
\tag{3.20}
$$

Therefore, we obtained the uniform measure on the compact set $S = \{(D_1, D_2) \in \mathcal{P}(\mathcal{Y} \otimes \mathcal{X}) \times \mathcal{P}(\mathcal{X}) : \mathrm{tr}_\mathcal{Y}(D_1) + D_2 = \mathbb{1}_\mathcal{X}\}$. Now, let us define an affine transformation $L : \mathcal{H}(\mathcal{Y} \otimes \mathcal{X}) \times \mathcal{H}(\mathcal{X}) \to \mathcal{H}(\mathcal{Y} \otimes \mathcal{X}) \times \mathcal{H}(\mathcal{X})$ given by the equation

$$
L(D_1, D_2) = (D_1, \mathbb{1}_\mathcal{X} - \mathrm{tr}_\mathcal{Y}(D_1) - D_2). \tag{3.21}
$$

Observe, that L is bijection and $L[S] = \{(D_1, 0) : D_1 \in \mathcal{P}(\mathcal{Y} \otimes \mathcal{X}), \mathrm{tr}_\mathcal{Y}(D_1) \leq \mathbb{1}_\mathcal{X}\}$. As affine transformations preserve flat measures we obtain the flat measure on the set $\{J(\Phi) : \Phi \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})\}$, which proves the claim. □

3.1.7 Random super-channels and beyond

In this section we will show a method of generating random super-channels [78] and present how to obtain the flat measure in this case. Loosely speaking, a super-channel Υ is a higher-order linear map that for any quantum channel $\Phi \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ returns quantum channel $\Upsilon(\Phi) \in \mathcal{C}(\mathcal{Z}, \mathcal{T})$. Additionally, this operation preserves quantum channels that are only partially under the action of Υ, that is $(\Upsilon \otimes \mathcal{I}_{\mathcal{X}'})(\Phi) \in \mathcal{C}(\mathcal{Z} \otimes \mathcal{X}', \mathcal{T} \otimes \mathcal{X}')$ for $\Phi \in \mathcal{C}(\mathcal{X} \otimes \mathcal{X}', \mathcal{Y} \otimes \mathcal{X}')$. Set of super-channels is compact and convex. Let $J(\Upsilon) \in \mathcal{M}(\mathcal{T} \otimes \mathcal{Y} \otimes \mathcal{X} \otimes \mathcal{Z})$ be Choi-Jamiołkowski matrix that characterizes Υ. Then, Υ is a super-channel if and only if [78]

- $J(\Upsilon) \in \mathcal{P}(\mathcal{T} \otimes \mathcal{Y} \otimes \mathcal{X} \otimes \mathcal{Z})$,

- $\mathrm{tr}_{\mathcal{T}}(J(\Upsilon)) = \mathbb{1}_{\mathcal{Y}} \otimes \frac{1}{\dim(\mathcal{Y})}\mathrm{tr}_{\mathcal{T}\otimes\mathcal{Y}}(J(\Upsilon))$,

- $\frac{1}{\dim(\mathcal{Y})}\mathrm{tr}_{\mathcal{T}\otimes\mathcal{Y}\otimes\mathcal{X}}(J(\Upsilon)) = \mathbb{1}_{\mathcal{Z}}$.

Definition 3.10. *Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{T}$ be given complex Euclidean spaces and let $r, r' \in \mathbb{N}$ be some parameters satisfying*

$$r \geq \frac{\dim(\mathcal{Y} \otimes \mathcal{X} \otimes \mathcal{Z})}{\dim(\mathcal{T})} \quad \texttt{and} \quad r' \geq \frac{\dim(\mathcal{Z})}{\dim(\mathcal{X})}. \tag{3.22}$$

We define a procedure of generating random quantum super-channels $\Upsilon : \mathcal{C}(\mathcal{X}, \mathcal{Y}) \ni \Phi \mapsto \Upsilon(\Phi) \in \mathcal{C}(\mathcal{Z}, \mathcal{T})$ as follows:

1. *Generate two random and independent matrices: complex Wishart matrix $W_r \in \mathcal{P}(\mathcal{T} \otimes \mathcal{Y} \otimes \mathcal{X} \otimes \mathcal{Z})$ of parameter r and Choi-Jamiołkowski matrix of quantum channel $J(\Phi) \in \mu^{Choi}_{\mathcal{Z},\mathcal{X};r'}$;*

2. *Calculate $Q = \mathrm{tr}_{\mathcal{T}}(W_r)$;*

3. *Write the Choi-Jamiołkowski matrix of random super-channel Υ as*

$$J(\Upsilon) = \left(\mathbb{1}_{\mathcal{T}\otimes\mathcal{Y}} \otimes \sqrt{J(\Phi)}\right)\left(\mathbb{1}_{\mathcal{T}} \otimes Q^{-1/2}\right) W_r \left(\mathbb{1}_{\mathcal{T}} \otimes Q^{-1/2}\right)\left(\mathbb{1}_{\mathcal{T}\otimes\mathcal{Y}} \otimes \sqrt{J(\Phi)}\right). \tag{3.23}$$

We may check that Definition 3.10 is correct. Indeed, $J(\Upsilon)$ is positive semi-definite and almost surely it holds

$$\mathrm{tr}_{\mathcal{T}}(J(\Upsilon)) = \left(\mathbb{1}_{\mathcal{Y}} \otimes \sqrt{J(\Phi)}\right) Q^0 \left(\mathbb{1}_{\mathcal{Y}} \otimes \sqrt{J(\Phi)}\right) = \mathbb{1}_{\mathcal{Y}} \otimes J(\Phi) \tag{3.24}$$

as well as

$$\mathrm{tr}_{\mathcal{X}}(J(\Phi)) = \mathbb{1}_{\mathcal{Z}}. \tag{3.25}$$

Proposition 3.11. *For all Euclidean spaces $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{T}$ random quantum super-channels $\Upsilon : \mathcal{C}(\mathcal{X}, \mathcal{Y}) \ni \Phi \mapsto \Upsilon(\Phi) \in \mathcal{C}(\mathcal{Z}, \mathcal{T})$, generated according to Definition 3.10 for $r = \dim(\mathcal{T}\otimes\mathcal{Y}\otimes\mathcal{X}\otimes\mathcal{Z})$ and $r' = \dim(\mathcal{X}\otimes\mathcal{Z})(1+\dim(\mathcal{Y})^2(\dim(\mathcal{T})^2 - 1))$, are uniformly distributed.*

Proof. We use the similar proof technique as in Proposition 3.9. Let $f(D_1, D_2)$ be the probability density function of the tuple $(J(\Upsilon), J(\Phi))$ at the point $(D_1, D_2) \in \mathcal{P}(\mathcal{T}\otimes\mathcal{Y}\otimes\mathcal{X}\otimes\mathcal{Z}) \times \mathcal{P}(\mathcal{X}\otimes\mathcal{Z})$, where $J(\Upsilon), J(\Phi)$ are defined as in Definition 3.10. The distribution of this tuple is induced by the distribution of two independent complex Ginibre matrices $G \in \mathcal{M}(\mathcal{T} \otimes \mathcal{Y} \otimes \mathcal{X} \otimes \mathcal{Z})$ and $G' \in \mathcal{M}(\mathbb{C}^{r'}, \mathcal{X} \otimes \mathcal{Z})$, where $r' = \dim(\mathcal{X} \otimes \mathcal{Z})(1 + \dim(\mathcal{Y})^2(\dim(\mathcal{T})^2 - 1))$. We have

$$
\begin{aligned}
& f(D_1, D_2) \propto \\
\propto & \int \delta(J(\Upsilon) - D_1)\delta(J(\Phi) - D_2) \exp(-\mathrm{tr} GG^\dagger) \exp(-\mathrm{tr} G'G'^\dagger) dG dG' \\
= & \int \delta\left(\left(\mathbb{1}_{\mathcal{T}\otimes\mathcal{Y}} \otimes \sqrt{J(\Phi)}\right)\left(\mathbb{1}_{\mathcal{T}} \otimes Q^{-1/2}\right) GG^\dagger \left(\mathbb{1}_{\mathcal{T}} \otimes Q^{-1/2}\right)\left(\mathbb{1}_{\mathcal{T}\otimes\mathcal{Y}} \otimes \sqrt{J(\Phi)}\right) - D_1\right) \\
& \delta\left(Q - \mathrm{tr}_{\mathcal{T}}(GG^\dagger)\right) \delta(J(\Phi) - D_2) \exp(-\mathrm{tr} Q) \exp(-\mathrm{tr} G'G'^\dagger) dG dG' dQ \\
= & \int \delta\left(GG^\dagger - D_1\right) \delta\left(Q - \sqrt{Q}\left(\mathbb{1}_{\mathcal{Y}} \otimes J(\Phi)^{-1/2}\right) \mathrm{tr}_{\mathcal{T}}(GG^\dagger)\left(\mathbb{1}_{\mathcal{Y}} \otimes J(\Phi)^{-1/2}\right)\sqrt{Q}\right) \\
& \delta(J(\Phi) - D_2) \exp(-\mathrm{tr} Q) \exp(-\mathrm{tr} G'G'^\dagger) \det(J(\Phi))^{-\dim(\mathcal{T}^{\otimes 2}\otimes\mathcal{Y}^{\otimes 2}\otimes\mathcal{X}\otimes\mathcal{Z})} \\
& \det(Q)^{\dim(\mathcal{T}^{\otimes 2}\otimes\mathcal{Y}\otimes\mathcal{X}\otimes\mathcal{Z})} dG dG' dQ \\
\propto & \int \delta\left(Q - \sqrt{Q}\left(\mathbb{1}_{\mathcal{Y}} \otimes J(\Phi)^{-1/2}\right) \mathrm{tr}_{\mathcal{T}}(D_1)\left(\mathbb{1}_{\mathcal{Y}} \otimes J(\Phi)^{-1/2}\right)\sqrt{Q}\right) \delta(J(\Phi) - D_2) \\
& \exp(-\mathrm{tr} Q) \exp(-\mathrm{tr} G'G'^\dagger) \det(J(\Phi))^{-\dim(\mathcal{T}^{\otimes 2}\otimes\mathcal{Y}^{\otimes 2}\otimes\mathcal{X}\otimes\mathcal{Z})} \\
& \det(Q)^{\dim(\mathcal{T}^{\otimes 2}\otimes\mathcal{Y}\otimes\mathcal{X}\otimes\mathcal{Z})} dG' dQ \\
\propto & \int \delta\left(\mathbb{1}_{\mathcal{Y}\otimes\mathcal{X}\otimes\mathcal{Z}} - \left(\mathbb{1}_{\mathcal{Y}} \otimes D_2^{-1/2}\right) \mathrm{tr}_{\mathcal{T}}(D_1)\left(\mathbb{1}_{\mathcal{Y}} \otimes D_2^{-1/2}\right)\right) \delta(J(\Phi) - D_2) \\
& \exp(-\mathrm{tr} G'G'^\dagger) \det(D_2)^{-\dim(\mathcal{T}^{\otimes 2}\otimes\mathcal{Y}^{\otimes 2}\otimes\mathcal{X}\otimes\mathcal{Z})} dG' \\
= & \int \delta\left(\mathbb{1}_{\mathcal{Y}} \otimes D_2 - \mathrm{tr}_{\mathcal{T}}(D_1)\right) \det(D_2)^{\dim(\mathcal{Y}^{\otimes 2}\otimes\mathcal{X}\otimes\mathcal{Z}) - \dim(\mathcal{T}^{\otimes 2}\otimes\mathcal{Y}^{\otimes 2}\otimes\mathcal{X}\otimes\mathcal{Z})} \exp(-\mathrm{tr} Q') \\
& \delta\left(\left(\mathbb{1}_{\mathcal{X}} \otimes Q'^{-1/2}\right) G'G'^\dagger \left(\mathbb{1}_{\mathcal{X}} \otimes Q'^{-1/2}\right) - D_2\right) \delta\left(Q' - \mathrm{tr}_{\mathcal{X}}(G'G'^\dagger)\right) dG' dQ' \\
= & \int \delta\left(\mathbb{1}_{\mathcal{Y}} \otimes D_2 - \mathrm{tr}_{\mathcal{T}}(D_1)\right) \det(D_2)^{\dim(\mathcal{Y}^{\otimes 2}\otimes\mathcal{X}\otimes\mathcal{Z}) - \dim(\mathcal{T}^{\otimes 2}\otimes\mathcal{Y}^{\otimes 2}\otimes\mathcal{X}\otimes\mathcal{Z})} \\
& \delta\left(G'G'^\dagger - D_2\right) \delta\left(\mathbb{1}_{\mathcal{Z}} - \mathrm{tr}_{\mathcal{X}}(G'G'^\dagger)\right) \exp(-\mathrm{tr} Q') \det(Q')^{\dim(\mathcal{X})r' - \dim(\mathcal{Z})} dG' dQ' \\
\propto & \int \delta\left(\mathbb{1}_{\mathcal{Y}} \otimes D_2 - \mathrm{tr}_{\mathcal{T}}(D_1)\right) \det(D_2)^{\dim(\mathcal{Y}^{\otimes 2}\otimes\mathcal{X}\otimes\mathcal{Z}) - \dim(\mathcal{T}^{\otimes 2}\otimes\mathcal{Y}^{\otimes 2}\otimes\mathcal{X}\otimes\mathcal{Z})} \\
& \delta\left(\mathbb{1}_{\mathcal{Z}} - \mathrm{tr}_{\mathcal{X}}(D_2)\right) \delta\left(G'G'^\dagger - D_2\right) dG' \\
\propto & \, \delta\left(\mathbb{1}_{\mathcal{Y}} \otimes D_2 - \mathrm{tr}_{\mathcal{T}}(D_1)\right) \delta\left(\mathbb{1}_{\mathcal{Z}} - \mathrm{tr}_{\mathcal{X}}(D_2)\right).
\end{aligned}
\tag{3.26}
$$

Therefore, we obtained the uniform measure on the compact set $S = \{(D_1, D_2) \in \mathcal{P}(\mathcal{T} \otimes \mathcal{Y} \otimes \mathcal{X} \otimes \mathcal{Z}) \times \mathcal{P}(\mathcal{X} \otimes \mathcal{Z}) : \mathrm{tr}_{\mathcal{T}}(D_1) = \mathbb{1}_{\mathcal{Y}} \otimes D_2, \mathrm{tr}_{\mathcal{X}}(D_2) = \mathbb{1}_{\mathcal{Z}}\}$. Similarly as in the proof of Proposition 3.9 this proves the claim. □

To sum up, in the Section 3.1 we showed how to effectively generate random quantum operations (quantum channels, quantum subchannels, quantum instruments, quantum super-channels) according to the uniform measure. One may combine the presented ideas (Propositions 3.5, 3.7, 3.9, 3.11) and use them to generate other higher-order operations, like quantum deterministic networks, quantum probabilistic networks, quantum testers [79] with the flat measure.

3.2 Some properties of random quantum channels

In this part of the chapter we investigate various properties of random quantum channels. We look into the behavior of the image of generic channels. In particular, we will be interested in the behavior of the invariant states. Also, some operational properties will be checked; purity, unitarity, Lipschitz constant.

3.2.1 Random extremal channels

A quantum channel $\Phi = \mathcal{K}\left((K_i)_{i=1}^r\right)$ is extremal, precisely when r is the Choi rank of Φ and the collection of matrices $\{K_j^\dagger K_i : i,j = 1,\ldots,r\}$ is linearly independent [62]. If $\Phi \in \mu^{Kraus}_{\mathcal{X},\mathcal{Y};r}$ is a random channel, then almost surely its Choi rank is $\min(r, \dim(\mathcal{Y})\dim(\mathcal{X}))$. In the next proposition we will show that if $r \leq \dim(\mathcal{X})$, then almost surely $\Phi \in \mu^{Kraus}_{\mathcal{X},\mathcal{Y};r}$ is extremal.

Proposition 3.12 ([2]). *Let $\Phi \in \mu^{Kraus}_{\mathcal{X},\mathcal{Y};r}$ be a random quantum channel defined according to Definition 3.2. Then, almost surely it holds*

$$\operatorname{rank}(J(\Phi^\dagger\Phi)) = \min(r^2, \dim(\mathcal{X})^2). \tag{3.27}$$

Proof. The channel $\Phi \in \mu^{Kraus}_{\mathcal{X},\mathcal{Y};r}$ is given as $\Phi = \mathcal{K}\left((G_i Q^{-1/2})_{i=1}^r\right)$, where $Q = \sum_{i=1}^r G_i^\dagger G_i$ and $(G_i)_{i=1}^r \subset \mathcal{M}(\mathcal{X},\mathcal{Y})$ is a tuple of random and independent Ginibre matrices. As $\operatorname{rank}(Q) = \dim(\mathcal{X})$ almost surely, then

$$\begin{split}\operatorname{rank}(J(\Phi^\dagger\Phi)) &= \operatorname{rank}\left(\sum_{i,j=1}^r |Q^{-1/2}G_j^\dagger G_i Q^{-1/2}\rangle\!\rangle\langle j,i|\right) = \operatorname{rank}\left(\sum_{i,j=1}^r |G_j^\dagger G_i\rangle\!\rangle\langle j,i|\right)\\ =&\operatorname{rank}\left(\sum_{a=0}^{\dim(\mathcal{Y})-1}\sum_{i,j=1}^r (G_j^\dagger|a\rangle \otimes \overline{G_i^\dagger|a\rangle})\langle j,i|\right) = \operatorname{rank}\left(\sum_{a=0}^{\dim(\mathcal{Y})-1} K_a \otimes \overline{K_a}\right),\end{split} \tag{3.28}$$

where $(K_a)_{a=0}^{\dim(\mathcal{Y})-1} \subset \mathcal{M}(\mathbb{C}^r, \mathcal{X})$ is a tuple of random and independent Ginibre matrices. As $\operatorname{rank}(K_a) = \min(r, \dim(\mathcal{X}))$ it follows that $\operatorname{rank}(J(\Phi^\dagger\Phi)) = \min(r^2, \dim(\mathcal{X})^2)$. □

3.2.2 Distribution of output states of random quantum channels

We consider in this section the output state of a random quantum channel, for a given input. We start by recalling the induced measures on the set of density matrices. This one-parameter family of probability measures $\nu_{\mathcal{Y};r}$ has been introduced in [80] and can be described in the following way. Let $W_r \in \mathcal{P}(\mathcal{Y})$ be a complex Wishart matrix of parameter $r \in \mathbb{N}$. We define random quantum state $\rho \in \nu_{\mathcal{Y};r}$ as

$$\rho = \frac{W_r}{\mathrm{tr}(W_r)}. \tag{3.29}$$

Remarkably, the uniform measure on the set $\mathcal{D}(\mathcal{Y})$ corresponds to the particular value $r = \dim(\mathcal{Y})$ [81].

Proposition 3.13 ([1]). *Let $\Phi \in \mu^{Stinespring}_{\mathcal{X},\mathcal{Y};r}$ be a random quantum channel. Then, for any given fixed pure input state $|\psi\rangle\langle\psi|$, the output state $\Phi(|\psi\rangle\langle\psi|)$ has the distribution $\nu_{\mathcal{Y};r}$.*

Proof. We have $\Phi(|\psi\rangle\langle\psi|) = \mathrm{tr}_{\mathbb{C}^r}(V|\psi\rangle\langle\psi|V^\dagger)$. As $V \in \mathcal{U}(\mathcal{X}, \mathcal{Y} \otimes \mathbb{C}^r)$ is Haar-distributed, the state $V|\psi\rangle$ is uniformly distributed on the unit complex sphere, hence it can be parametrized by using Gaussian distribution as $V|\psi\rangle = \frac{G}{\|G\|_2}$, where $G \in \mathcal{M}(\mathbb{C}, \mathcal{Y} \otimes \mathbb{C}^r)$. Therefore, $\Phi(|\psi\rangle\langle\psi|) = \mathrm{tr}_{\mathbb{C}^r}(W_1)/\mathrm{tr}(W_1) = W_r/\mathrm{tr}(W_r)$, which proves the claim. □

Remark 3.14. *Let* $\dim(\mathcal{X}) = \dim(\mathcal{Y})$. *For uniformly distributed random quantum channel, $r = \dim(\mathcal{X})^2$, the distribution of $\Phi(|\psi\rangle\langle\psi|)$ is more concentrated towards the maximally mixed state $\rho^*_{\mathcal{X}}$.*

From the proposition above, we can infer that the average of the output state (with respect to the randomness in the channel) for a fixed input is the maximally mixed state

$$\mathbb{E}\Phi(|\psi\rangle\langle\psi|) = \rho^*_{\mathcal{Y}}. \tag{3.30}$$

This fact is equally a consequence of the following result.

Proposition 3.15 ([1]). *The average of a random quantum channel $\Phi \in \mu^{Stinespring}_{\mathcal{X},\mathcal{Y};r}$ is the maximally depolarizing channel, $\mathbb{E}(\Phi)(\cdot) = \mathrm{tr}(\cdot)\rho^*_{\mathcal{Y}}$.*

Proof. For any $X \in \mathcal{M}(\mathcal{X})$ and any $U \in \mathcal{M}(\mathcal{Y})$, $\Phi(X)$ has the same distribution as $U\Phi(X)U^\dagger = \mathrm{tr}_{\mathbb{C}^r}\left((U \otimes \mathbb{1}_{\mathbb{C}^r})VXV^\dagger(U^\dagger \otimes \mathbb{1}_{\mathbb{C}^r})\right)$ as $V \in \mathcal{U}(\mathcal{X}, \mathcal{Y} \otimes \mathbb{C}^r)$ is Haar random. Therefore, $U\mathbb{E}\Phi(X)U^\dagger = \mathbb{E}\Phi(X)$, so $\mathbb{E}\Phi(X) \propto \mathbb{1}_{\mathcal{Y}}$ [62]. Moreover, $\mathrm{tr}\mathbb{E}\Phi(X) = \mathrm{tr}(X)$. It means $\mathbb{E}J(\Phi) = \rho^*_{\mathcal{Y}} \otimes \mathbb{1}_{\mathcal{X}}$. □

3.2.3 Average output purity and unitarity of random channels

Let us now consider two statistical quantities associated to an arbitrary quantum channel $\Phi \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$:

- the average output purity:

$$p(\Phi) = \mathbb{E}\mathrm{tr}\left(\Phi(|\psi\rangle\langle\psi|)^2\right), \tag{3.31}$$

- the unitarity [82]:

$$u(\Phi) = \frac{\dim(\mathcal{X})}{\dim(\mathcal{X}) - 1}\mathbb{E}\mathrm{tr}\left((\Phi(|\psi\rangle\langle\psi|) - \Phi(\rho^*_{\mathcal{X}}))^2\right), \tag{3.32}$$

The expectation values correspond to the choice of a uniform unit vector $|\psi\rangle$ on the sphere. Before we compute the averages of these two quantities, we provide the necessary lemma, which might be of an independent interest.

Lemma 3.16 ([1]). *Let $\Phi \in \mu^{Stinespring}_{\mathcal{X},\mathcal{Y};r}$ be a random quantum channel. Then, for any matrices $A, B \in M(\mathcal{X})$, we have*

$$\mathbb{E}\mathrm{tr}\left(\Phi(A)\Phi(B)\right) = \frac{\mathrm{tr}(A)\mathrm{tr}(B)\dim(\mathcal{Y})(r^2-1) + \mathrm{tr}(AB)r(\dim(\mathcal{Y})^2-1)}{(\dim(\mathcal{Y})r)^2 - 1}. \tag{3.33}$$

Proof. The proof is a standard application of the Weingarten calculus [83] together with its graphical representation introduced in [84]. The value $\mathbb{E}\mathrm{tr}\left(\Phi(A)\Phi(B)\right)$ is represented in diagrammatic notation in Figure 3.1. According to [84, Theorem

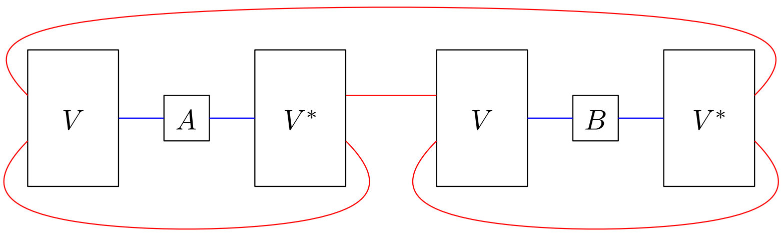


Figure 3.1: The diagram corresponding to $\mathrm{tr}(\Phi(A)\Phi(B))$. When computing the expectation of this diagram with respect to the Haar-distributed random isometry $V \in \mathcal{U}(\mathcal{X}, \mathcal{Y} \otimes \mathbb{C}^r)$, (visualized as a block with two inputs and one output), we use the permutation α to pair the outer (red) wires, and the permutation β to pair the inner (blue) wires. The space dimensions are as follows: outer, upper wires: $\dim(\mathcal{Y})$; outer, lower wires: r; inner wires: $\dim(\mathcal{X})$.

4.1], the average trace in the statement can be decomposed as

$$\mathbb{E}\mathrm{tr}\left(\Phi(A)\Phi(B)\right) = \sum_{\alpha,\beta\in\mathcal{S}_2} \mathcal{D}_{\alpha,\beta}\mathrm{Wg}(\alpha^{-1}\beta, \dim(\mathcal{Y})r),$$

where Wg is the Weingarten function [83, 85], S_2 is a symmetric group of $\{1, 2\}$, namely $S_2 = \{(1, 2), (2, 1)\}$ and $D_{\alpha,\beta}$ is the number associated with a diagram removal (formal definition and derivation od $\mathcal{D}$ is presented in [84]). In our case we have

- $\mathrm{Wg}((1, 2), \dim(\mathcal{Y})r) = \frac{1}{\dim(\mathcal{Y})^2r^2-1}$,
- $\mathrm{Wg}((2, 1), \dim(\mathcal{Y})r) = \frac{-1}{\dim(\mathcal{Y})r(\dim(\mathcal{Y})^2r^2-1)}$,
- $\mathcal{D}_{(1,2),(1,2)} = \dim(\mathcal{Y})r^2\mathrm{tr}(A)\mathrm{tr}(B)$,
- $\mathcal{D}_{(1,2),(2,1)} = \dim(\mathcal{Y})r^2\mathrm{tr}(AB)$,
- $\mathcal{D}_{(2,1),(1,2)} = \dim(\mathcal{Y})^2r\mathrm{tr}(A)\mathrm{tr}(B)$,
- $\mathcal{D}_{(2,1),(2,1)} = \dim(\mathcal{Y})^2r\mathrm{tr}(AB)$,

which proves the claim. □

Proposition 3.17 ([1]). *Let $\Phi \in \mu^{Stinespring}_{\mathcal{X},\mathcal{Y};r}$ be a random quantum channel. Then, the expectation values of the average output purity and unitarity read,*

$$\mathbb{E}p(\Phi) = \frac{\dim(\mathcal{Y}) + r}{\dim(\mathcal{Y})r + 1}, \tag{3.34}$$

$$\mathbb{E}u(\Phi) = \frac{r(\dim(\mathcal{Y})^2 - 1)}{(\dim(\mathcal{Y})r)^2 - 1}. \tag{3.35}$$

Proof. First of all, note that the expectation over the random pure state $|\psi\rangle$ in the definition of $p(\Phi)$ and $u(\Phi)$ is absorbed in the expectation over the random channel Φ. Hence, we can fix $|\psi\rangle\langle\psi|$ and utilize Lemma 3.16. For $\mathbb{E}p(\Phi)$ we take $A = B = |\psi\rangle\langle\psi|$ and for $\mathbb{E}u(\Phi)$ we take $A = B = |\psi\rangle\langle\psi| - \rho^*_{\mathcal{X}}$. □

Note that in the regime where $\dim(\mathcal{Y}) \to \infty$, the average purity and the average unitarity of a random quantum channels scales as $1/r$.

Corollary 3.18 ([1]). *The average output purity and the average unitarity of a uniformly distributed random quantum channel $\Phi \in \mu^{Lebesgue}_{\mathcal{X},\mathcal{X}}$ are*

$$\mathbb{E}p(\Phi) = \frac{\dim(\mathcal{X})}{\dim(\mathcal{X})^2 - \dim(\mathcal{X}) + 1} \quad \text{and} \quad \mathbb{E}u(\Phi) = \frac{\dim(\mathcal{X})^2}{\dim(\mathcal{X})^4 + \dim(\mathcal{X})^2 + 1}. \tag{3.36}$$

3.2.4 Random probability vectors and stochastic matrices induced by random quantum channels

In this section we describe how classical objects like probability vectors and stochastic maps arise from random quantum channels. We will consider probability vectors of the form $p = (p_i)_{i=0}^{\dim(\mathcal{Y})-1}$ induced by a state $\rho \in \mathcal{D}(\mathcal{Y})$ as $p = \mathrm{diag}^\dagger(\rho)$. In the similar way, quantum channels $\Phi \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ will define stochastic matrices $T \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ as $T_{j,i} = \langle j|\Phi(|i\rangle\langle i|)|j\rangle$ for $i = 0, \ldots, \dim(\mathcal{X})-1$ and $j = 0, \ldots, \dim(\mathcal{Y})-1$.

Starting, with induced probability vectors, let us recall a particular distribution defined on the probability simplex – Dirichlet distribution with parameter r [86, Chapter XI.4]. We say that the probability vector $p = (p_0, \ldots, p_{\dim(\mathcal{Y})-1})$ is distributed according to the Dirichlet distribution with parameter r, if it has density $f(p) \propto \delta(\sum_i p_i - 1) p_0^{r-1} \cdots p_{\dim(\mathcal{Y})-1}^{r-1}$.

Corollary 3.19 ([1]). *Let $\Phi \in \mu_{\mathcal{X},\mathcal{Y};r}^{Stinespring}$ be a random quantum channel and $|\psi\rangle\langle\psi|$ be any given fixed pure input state. Then, the probability vector $p = \mathrm{diag}^\dagger(\Phi(|\psi\rangle\langle\psi|))$ has a Dirichlet distribution with parameter r.*

Proof. From Proposition 3.13 it follows that $\Phi(|\psi\rangle\langle\psi|)$ has the distribution $\nu_{\mathcal{Y};r}$, that is the distribution of $W_r/\mathrm{tr}(W_r) \in \mathcal{P}(\mathcal{Y})$. If we write $W_r = GG^\dagger$ for a complex Ginibre matrix $G \in \mathcal{M}(\mathbb{C}^r, \mathcal{Y})$ we obtain

$$p_i = \frac{(W_r)_{i,i}}{\mathrm{tr}(W_r)} = \frac{\sum_{j=0}^{r-1} |G_{i,j}|^2}{\sum_{i=0}^{\dim(\mathcal{Y})-1} \sum_{j=0}^{r-1} |G_{i,j}|^2}. \tag{3.37}$$

The random variables $2\sum_{j=0}^{r-1} |G_{i,j}|^2$ for $i = 0, \ldots, \dim(\mathcal{Y}) - 1$ are independent and each has a chi-squared distribution of parameter $2r$, or equivalently a Gamma distribution with parameters $(r, 2)$. It follows from [86, Theorem XI.4.1] that normalizing independent Gamma random variables with parameters $(r, 2)$ yields the Dirichlet distribution with parameter r. □

Remark 3.20. *Let $\dim(\mathcal{X}) = \dim(\mathcal{Y})$. The uniform distribution on the probability simplex (Dirichlet distribution with parameter $r = 1$) is obtained by considering the diagonal of random pure states, $\Phi(|\psi\rangle\langle\psi|) = V|\psi\rangle\langle\psi|V^\dagger$, where $V \in \mathcal{U}(\mathcal{X})$ is Haar-distributed. The diagonal of uniformly distributed random density matrices ($r = \dim(\mathcal{X})$) or uniformly distributed random quantum channel $r = \dim(\mathcal{X})^2$ does not yield uniform probability vectors, but the distribution that is more concentrated towards the central point $(1/\dim(\mathcal{X}), \ldots, 1/\dim(\mathcal{X}))$ of the given simplex.*

Proposition 3.21 ([1]). *Let* $\Phi \in \mu^{Stinespring}_{\mathcal{X},\mathcal{Y};r}$ *be a random quantum channel and let* $T \in \mathcal{M}(\mathcal{X},\mathcal{Y})$ *given as* $T_{j,i} = \langle j|\Phi(|i\rangle\langle i|)|j\rangle$ *be the induced stochastic matrix. Then, every column of* T *has a Dirichlet distribution with parameter* r*. However, the columns of* T *are not independent, the entries having covariances*

$$\mathbb{E}(T_{j,i_1}T_{j,i_2}) = \frac{\dim(\mathcal{Y})r^2 - 1}{\dim(\mathcal{Y})^3 r^2 - \dim(\mathcal{Y})} < \frac{1}{\dim(\mathcal{Y})^2}, \qquad i_1 \neq i_2, \tag{3.38}$$

$$\mathbb{E}(T_{j_1,i_1}T_{j_2,i_2}) = \frac{r^2}{(\dim(\mathcal{Y})r)^2 - 1} > \frac{1}{\dim(\mathcal{Y})^2}, \qquad j_1 \neq j_2, i_1 \neq i_2. \tag{3.39}$$

In particular, for $\dim(\mathcal{X}) > 1$*, the distribution of* T *induced by* $\mu^{Stinespring}_{\mathcal{X},\mathcal{Y};r}$ *is not uniform.*

Proof. First, the i-th column of T is given as $\mathrm{diag}^\dagger(\Phi(|i\rangle\langle i|))$. Therefore, from Corollary 3.19 it has a Dirichlet distribution with parameter r. Second, the covariance expressions can be readily obtained from spherical integration formulas or using the Weingarten calculus, see [87] or [85]:

$$\mathbb{E}|V_{a,i_1}|^2|V_{a,i_2}|^2 = \frac{1}{\dim(\mathcal{Y})r(\dim(\mathcal{Y})r + 1)}, \qquad i_1 \neq i_2, \tag{3.40}$$

$$\mathbb{E}|V_{a_1,i_1}|^2|V_{a_2,i_2}|^2 = \frac{1}{(\dim(\mathcal{Y})r)^2 - 1}, \qquad a_1 \neq a_2, i_1 \neq i_2, \tag{3.41}$$

where $V \in \mathcal{U}(\mathcal{X}, \mathcal{Y} \otimes \mathbb{C}^r)$ is Haar-distributed, such that $\Phi(X) = \mathrm{tr}_{\mathbb{C}^r}(VXV^\dagger)$. Hence, as $T_{j,i} = \sum_{a=0}^{r-1} |\langle j|\langle a|V|i\rangle|^2$ we obtain the desired covariances. Finally, we have $\mathbb{E}(T_{j,i}) = \frac{1}{\dim(\mathcal{Y})}$, hence the columns of T are not independent and in particular T cannot be uniformly distributed. □

3.2.5 Coherence of random quantum channels

In this section we briefly discuss "quantumness" of random channels derived from $\mu^{Lebesgue}_{\mathcal{X},\mathcal{Y}}$, where $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = d$, be means of coherence measures [88] defined for Choi-Jamiołkowski matrices. One can note that quantum channels that are classical are not distinguishable from stochastic operations T they induce (see Section 3.2.4 for a definition of T). More formally, such channels satisfy $\Phi(\rho) = \Delta(\Phi(\Delta(\rho)))$ for any ρ, where $\Delta(X) = \sum_{i=0}^{d-1} |i\rangle\langle i|X|i\rangle\langle i|$ and henceforth, the Choi-Jamiołkowski matrix of such Φ is diagonal and its coherence is zero.

Proposition 3.22 ([1]). *Let* $\Phi \in \mu^{Lebesgue}_{\mathcal{X},\mathcal{Y}}$ *be a random quantum channel and let*

$d = \dim(\mathcal{X}) = \dim(\mathcal{Y})$. As $d \to \infty$, it holds almost surely that

$$
\begin{aligned}
c_2(\Phi) &= \sum_{i \neq j} |(J(\Phi))_{i,j}|^2 \simeq 1, \\
c_1(\Phi) &= \sum_{i \neq j} |(J(\Phi))_{i,j}| \simeq d^2 \frac{\sqrt{\pi}}{2}, \\
c_e(\Phi) &= S(\mathrm{diag}^\dagger(J(\Phi))) - S(J(\Phi)) \simeq \frac{d}{2},
\end{aligned}
\tag{3.42}
$$

where $S(\cdot)$ is the von Neumann entropy.

Proof. Let $\Phi \in \mu^{\textit{Lebesgue}}_{\mathcal{X},\mathcal{Y}} = \mu^{\textit{Choi}}_{\mathcal{X},\mathcal{Y};d^2}$ and $W_{d^2} \in \mathcal{P}(\mathcal{Y} \otimes \mathcal{X})$ be the Wishart matrix used in construction of Φ according to Definition 3.1. According to Proposition 11 from [89] almost surely as $d \to \infty$ it holds

$$
\left\| J(\Phi) - W_{d^2}/d^3 \right\|_\infty = \mathcal{O}(d^{-2}). \tag{3.43}
$$

Therefore, the asymptotic behavior of coherence measures for random channels can be derived form the results for the random quantum states [90]. The correctness of this approximation can be easily proven for $c_1(\Phi)$ and $c_2(\Phi)$, while for $c_e(\Phi)$ one may use Fannes–Audenaert inequality [62]. □

3.2.6 Invariant states of random quantum channels

In this section we will focus on the properties of the invariant state of random quantum channels sampled according to $\mu^{\textit{Lebesgue}}_{\mathcal{X},\mathcal{Y}}$. One of the properties characterizing random quantum maps is their ability of mixing quantum states. The image of the maximally mixed state $\rho^*_{\mathcal{X}}$ under random quantum maps will be concentrated around the maximally mixed state $\rho^*_{\mathcal{Y}}$. Thus, one can expect that the invariant state ρ_{inv} of a random map Φ can be found in the neighborhood of the maximally mixed state. Before we formalize this statement let us first establish a few lemmas.

We first prove that random quantum operations are contraction maps. We examine behavior of the Lipschitz constant (with respect to the Schatten 1-norm) which is defined as minimum over constants L satisfying $\|\Phi(\rho - \sigma)\|_1 \leq L\|\rho - \sigma\|_1$ for all states $\rho, \sigma \in \mathcal{D}(\mathcal{X})$.

Lemma 3.23 ([1])**.** *Let L_Φ denote the Lipschitz constant of random quantum operation $\Phi \in \mu^{\textit{Lebesgue}}_{\mathcal{X},\mathcal{Y}}$, where $\dim(\mathcal{X}) = \dim(\mathcal{Y})$. Then, almost surely, as $\dim(\mathcal{X}) \to \infty$*

$$
L_\Phi \leq \frac{3\sqrt{3}}{2\pi} < 1. \tag{3.44}
$$

In particular, Φ almost surely has the unique invariant state ρ_{inv}.

Proof. One can obtain

$$
\begin{aligned}
L_\Phi &= \max\left\{\left\|\Phi\left(\frac{\rho-\sigma}{\|\rho-\sigma\|_1}\right)\right\|_1 : \rho\neq\sigma\right\} \\
&= \max\left\{\|\Phi\left(H\right)\|_1 : H\in\mathcal{H}(\mathcal{X}), \operatorname{tr}(H)=0, \|H\|_1=1\right\} \\
&= \frac{1}{2}\max\left\{\|\Phi(|x\rangle\langle x|-|y\rangle\langle y|)\|_1 : |x\rangle\langle x|, |y\rangle\langle y|\in\mathcal{D}(\mathcal{X}), \langle x|y\rangle=0\right\}.
\end{aligned}
\tag{3.45}
$$

In the next step we will use the diamond norm to bound this value:

$$
\begin{aligned}
L_\Phi &\leq \max\left\{\|\Phi(|x\rangle\langle x|)-\rho^*_{\mathcal{Y}}\|_1 : |x\rangle\langle x|\in\mathcal{D}(\mathcal{X})\right\} \\
&= \max\left\{\|(\Phi-\Phi_*)(|x\rangle\langle x|)\|_1 : |x\rangle\langle x|\in\mathcal{D}(\mathcal{X})\right\} \leq \|\Phi-\Phi_*\|_\diamond,
\end{aligned}
\tag{3.46}
$$

where Φ_* denotes the maximally depolarizing channel. By [89, Theorem 16] we have $\|\Phi-\Phi_*\|_\diamond \to \frac{3\sqrt{3}}{2\pi} < 1$, proving the claim. Finally, each quantum channel $\Phi\in\mathcal{C}(\mathcal{X},\mathcal{Y})$, where $\dim(\mathcal{X})=\dim(\mathcal{Y})$ has at least one invariant state [62]. As $L_\Phi<1$ it has to be unique. □

The next lemma gives an upper bound on the distance between the maximally mixed state and its image through a random quantum channel.

Lemma 3.24 ([1])**.** *Let* $\Phi\in\mu^{\text{Lebesgue}}_{\mathcal{X},\mathcal{Y}}$ *be a random quantum channel, where* $\dim(\mathcal{X})=\dim(\mathcal{Y})$. *Then, we have*

$$
\begin{aligned}
\mathbb{E}\operatorname{tr}\left(\left(\Phi(\rho^*_{\mathcal{X}})-\rho^*_{\mathcal{Y}}\right)^2\right) &= \frac{(\dim(\mathcal{X})^2-1)^2}{\dim(\mathcal{X})(\dim(\mathcal{X})^6-1)}, \\
\operatorname{Var}\left(\dim(\mathcal{X})^3\operatorname{tr}\left(\left(\Phi(\rho^*_{\mathcal{X}})-\rho^*_{\mathcal{Y}}\right)^2\right)\right) &= \frac{2}{\dim(\mathcal{X})^2}+\mathcal{O}(\dim(\mathcal{X})^{-4}),
\end{aligned}
$$

where Var *is a variance. In particular, almost surely,*

$$
\lim_{\dim(\mathcal{X})\to\infty} \dim(\mathcal{X})^3\|\Phi(\rho^*_{\mathcal{X}})-\rho^*_{\mathcal{Y}}\|_2^2 = 1.
$$

Proof. The first claim follows from Proposition 3.15 and Lemma 3.16 and with the choice $A=B=\rho^*_{\mathcal{X}}$ and for $r=\dim(\mathcal{X})^2$ (see Proposition 3.5).

The second claim is proven using the Weingarten calculus [83] to compute fourth moments of $\Phi(\rho^*_{\mathcal{X}})$. Due to tedious computation we provided in the Supplementary Material in [1] a `Mathematica` notebook which performs this computation using the RTNI package provided in [91].

Let $z_{\mathcal{X}}(\Phi)=\dim(\mathcal{X})^3\|\Phi(\rho^*_{\mathcal{X}})-\rho^*_{\mathcal{Y}}\|_2^2$ be a random variable. From the previous claims we have

$$
\begin{aligned}
\mathbb{E}z_{\mathcal{X}}(\Phi) &= 1+\mathcal{O}\left(\dim(\mathcal{X})^{-2}\right), \\
\operatorname{Var} z_{\mathcal{X}}(\Phi) &= \mathcal{O}\left(\dim(\mathcal{X})^{-2}\right).
\end{aligned}
\tag{3.47}
$$

From the Chebyshev's inequality for any ϵ we have

$$P\left(|z_{\mathcal{X}}(\Phi) - \mathbb{E}z_{\mathcal{X}}(\Phi)| \geq \epsilon\right) \leq \epsilon^{-2}\mathcal{O}\left(\dim(\mathcal{X})^{-2}\right). \tag{3.48}$$

Hence, for any ϵ it holds

$$\sum_{\dim(\mathcal{X})=2}^{\infty} P\left(|z_{\mathcal{X}}(\Phi) - \mathbb{E}z_{\mathcal{X}}(\Phi)| \geq \epsilon\right) < \infty. \tag{3.49}$$

From the Borel-Cantelli lemma, almost surely $|z_{\mathcal{X}}(\Phi) - \mathbb{E}z_{\mathcal{X}}(\Phi)| \to 0$, which proves the last claim. □

We can also show the following stronger result, which we can only prove in the asymptotical regime $1 \ll \dim(\mathcal{X}) \ll \dim(\mathcal{Y})$.

Proposition 3.25 ([1]). *Let $\Phi \in \mu_{\mathcal{X},\mathcal{Y}}^{Lebesgue}$ be a random quantum channel. If $1 \ll \dim(\mathcal{X}) \ll \dim(\mathcal{Y})$, then $\dim(\mathcal{X})\dim(\mathcal{Y})(\Phi(\rho_{\mathcal{X}}^*) - \rho_{\mathcal{Y}}^*)$ converges in moments toward the standard semicircular distribution [92].*

Proof. We can write

$$\begin{aligned}\dim(\mathcal{X}\otimes\mathcal{Y})(\Phi(\rho_{\mathcal{X}}^*) - \rho_{\mathcal{Y}}^*) = \dim(\mathcal{Y})\mathrm{tr}_{\mathcal{X}}\left(J(\Phi) - \frac{W_{\dim(\mathcal{X}\otimes\mathcal{Y})}}{\dim(\mathcal{X})\dim(\mathcal{Y})^2}\right) \\ + \dim(\mathcal{X})\left(\frac{\mathrm{tr}_{\mathcal{X}}W_{\dim(\mathcal{X}\otimes\mathcal{Y})}}{\dim(\mathcal{X})^2\dim(\mathcal{Y})} - \mathbb{1}_{\mathcal{Y}}\right),\end{aligned} \tag{3.50}$$

where we introduced a Wishart matrix $W_{\dim(\mathcal{X}\otimes\mathcal{Y})} \in \mathcal{P}(\mathcal{Y}\otimes\mathcal{X})$, such that it is used to define the Choi matrix $J(\Phi)$ as in Definition 3.1. By the Corollary 2.5 in [92] we know that the matrix $\dim(\mathcal{X})\left(\frac{\mathrm{tr}_{\mathcal{X}}W_{\dim(\mathcal{X}\otimes\mathcal{Y})}}{\dim(\mathcal{X})^2\dim(\mathcal{Y})} - \mathbb{1}_{\mathcal{Y}}\right)$ converges to the standard semicircular distribution. To finish the proof we need to bound the first term of the sum:

$$\begin{aligned}&\left\|\dim(\mathcal{Y})\mathrm{tr}_{\mathcal{X}}\left(J(\Phi) - \frac{W_{\dim(\mathcal{X}\otimes\mathcal{Y})}}{\dim(\mathcal{X})\dim(\mathcal{Y})^2}\right)\right\|_{\infty} \\ &= \dim(\mathcal{Y})\left\|\mathrm{tr}_{\mathcal{X}}\left(W_{\dim(\mathcal{X}\otimes\mathcal{Y})}\left(\mathbb{1}_{\mathcal{Y}}\otimes\left((\mathrm{tr}_{\mathcal{Y}}(W_{\dim(\mathcal{X}\otimes\mathcal{Y})}))^{-1} - \frac{\mathbb{1}_{\mathcal{X}}}{\dim(\mathcal{X})\dim(\mathcal{Y})^2}\right)\right)\right)\right\|_{\infty} \\ &\leq \dim(\mathcal{Y})\|\mathrm{tr}_{\mathcal{X}}\left(W_{\dim(\mathcal{X}\otimes\mathcal{Y})}\right)\|_{\infty}\left\|(\mathrm{tr}_{\mathcal{Y}}(W_{\dim(\mathcal{X}\otimes\mathcal{Y})}))^{-1} - \frac{\mathbb{1}_{\mathcal{X}}}{\dim(\mathcal{X})\dim(\mathcal{Y})^2}\right\|_{\infty}.\end{aligned} \tag{3.51}$$

According to Theorem 2.7 in [92] we have that in the limit $\dim(\mathcal{X}), \dim(\mathcal{Y}) \to \infty$, it holds $\left\|\frac{\mathrm{tr}_{\mathcal{Y}}(W_{\dim(\mathcal{X}\otimes\mathcal{Y})})}{\dim(\mathcal{X}\otimes\mathcal{Y})} - \dim(\mathcal{Y})\mathbb{1}_{\mathcal{X}}\right\|_{\infty} \to 2$, therefore

$$\left\|(\mathrm{tr}_{\mathcal{Y}}(W_{\dim(\mathcal{X}\otimes\mathcal{Y})}))^{-1} - \frac{\mathbb{1}_{\mathcal{X}}}{\dim(\mathcal{X})\dim(\mathcal{Y})^2}\right\|_{\infty} = \mathcal{O}\left((\dim(\mathcal{X})\dim(\mathcal{Y})^3)^{-1}\right). \tag{3.52}$$

On the other hand $\|\mathrm{tr}_{\mathcal{X}}\left(W_{\dim(\mathcal{X}\otimes\mathcal{Y})}\right)\|_\infty = \mathcal{O}(\dim(\mathcal{X})^2\dim(\mathcal{Y}))$ (see Theorem 2.7 in [92]). Hence, the first term of the introduced sum is upper-bounded by $\mathcal{O}(\dim(\mathcal{X})/\dim(\mathcal{Y}))$. To finish the proof we use the Weyl's inequality. □

Theorem 3.26 ([1]). *Let $\Phi \in \mu_{\mathcal{X},\mathcal{Y}}^{Lebesgue}$ be a random quantum channel, where $\dim(\mathcal{X}) = \dim(\mathcal{Y})$ and let ρ_{inv} be the unique invariant state of Φ. As $\dim(\mathcal{X}) \to \infty$, the invariant state convergences almost surely in the trace norm to the maximally mixed state,*

$$\|\rho_{\mathrm{inv}} - \rho_{\mathcal{Y}}^*\|_1 = \mathcal{O}\left(\dim(\mathcal{X})^{-1}\right). \tag{3.53}$$

Proof. Using Lemmas 3.23 and 3.24, we have that, almost surely as $\dim(\mathcal{X}) \to \infty$,

$$\begin{aligned}
\|\rho_{\mathrm{inv}} - \rho_{\mathcal{Y}}^*\|_1 &= \|\Phi^k(\rho_{\mathrm{inv}}) - \Phi^k(\rho_{\mathcal{X}}^*) + \Phi^k(\rho_{\mathcal{X}}^*) - \Phi^{k-1}(\rho_{\mathcal{X}}^*) + \dots \\
&\quad - \Phi(\rho_{\mathcal{X}}^*) + \Phi(\rho_{\mathcal{X}}^*) - \rho_{\mathcal{Y}}^*\|_1 \\
&\leq 2L_\Phi^k + \frac{1-L_\Phi^k}{1-L_\Phi}\|\Phi(\rho_{\mathcal{X}}^*) - \rho_{\mathcal{Y}}^*\|_1 \xrightarrow{k\to\infty} \frac{\|\Phi(\rho_{\mathcal{X}}^*) - \rho_{\mathcal{Y}}^*\|_1}{1-L_\Phi} \\
&\leq \frac{\dim(\mathcal{X})^{1/2}\|\Phi(\rho_{\mathcal{X}}^*) - \rho_{\mathcal{Y}}^*\|_2}{1-L_\Phi} = \mathcal{O}(\dim(\mathcal{X})^{-1}).
\end{aligned} \tag{3.54}$$

□

Chapter 4

Probabilistic quantum error correction

Probabilistic quantum error correction is an error-correcting procedure which uses postselection to determine if the encoded information was successfully restored. It generalizes the notation of quantum error correction, where in the standard set-up it is always possible to recover perfectly the initial information. In the probabilistic approach, the output consists of not only the decoded initial state, but also a binary label that indicates if the error-correcting procedure succeed and the output state should be accepted.

The goal of this chapter is to introduce the the concept of probabilistic quantum error correction. We provide equivalent conditions for pQEC and show mixed-state encoding phenomenon related with this procedure. Last but not least we point out advantage of using probabilistic QEC. This chapter, containing theoretical results about pQEC, constitute a basis for considering applications of pQEC, which will be discussed in the next chapter.

This chapter is based mostly on the article [2]. The pQEC procedure presented in Algorithm 28 is the subject of an European patent application EP 22156614.

4.1 Problem formulation

We are given a noise channel $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ and an Euclidean space $\mathcal{X}$. The goal is to choose an appropriate encoding operation $\mathcal{S} \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ and decoding operation $\mathcal{R} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$, such that for any state $\rho \in \mathcal{D}(\mathcal{X})$ we have $\mathcal{RES}(\rho) \propto \rho$. The pair $(\mathcal{S}, \mathcal{R})$ represents the error-correcting scheme and the quantity $\operatorname{tr}(\mathcal{RES}(\rho))$ represents the probability of successful error correction. This protocol may fail with the probability $1 - \operatorname{tr}(\mathcal{RES}(\rho))$. In such a case, the output state is rejected. To exclude a trivial, null strategy, we add the constrain that a valid error-correcting scheme must satisfy $\operatorname{tr}(\mathcal{RES}(\rho)) > 0$ for any $\rho \in \mathcal{D}(\mathcal{X})$. We begin with the following lemma to standardize the definition of pQEC.

Lemma 4.1 ([93]). *Let $\Phi \in s\mathcal{C}(\mathcal{X})$. If for any pure state $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{X})$ it holds $\Phi(|\psi\rangle\langle\psi|) \propto |\psi\rangle\langle\psi|$, then there exists $p \in [0,1]$ such that $\Phi = p\mathcal{I}_{\mathcal{X}}$.*

Proof. For any unitary operator $U \in \mathcal{U}(\mathcal{X})$ and $i = 0, \ldots, \dim(\mathcal{X}) - 1$ let us define $p_{U,i} \in [0,1]$ by $\Phi(U|i\rangle\langle i|U^\dagger) = p_{U,i}U|i\rangle\langle i|U^\dagger$. We have $\Phi(\mathbb{1}_{\mathcal{X}}) = U\left(\sum_i p_{U,i}|i\rangle\langle i|\right)U^\dagger$ for any U and hence, there exists $p \in [0,1]$ such that $\Phi(\mathbb{1}_{\mathcal{X}}) = p\mathbb{1}_{\mathcal{X}}$. That means, $p_{U,i} = p$ for any U and i, so $\Phi(|\psi\rangle\langle\psi|) = p|\psi\rangle\langle\psi|$ for any $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{X})$. We obtain the thesis by noting that $\operatorname{span}_{\mathbb{C}}(|\psi\rangle\langle\psi|) = \mathcal{M}(\mathcal{X})$. □

As we can see, in our set-up, the probability of successful error correction does not depend on the input state ρ. We use this fact to provide the definition of pQEC.

Definition 4.2 ([2]). *We say that $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ is probabilistically correctable for $\mathcal{X}$, if there exists an error-correcting scheme $(\mathcal{S},\mathcal{R}) \in s\mathcal{C}(\mathcal{X},\mathcal{Y}) \times s\mathcal{C}(\mathcal{Y},\mathcal{X})$ such that*

$$0 \neq \mathcal{RES} \propto \mathcal{I}_{\mathcal{X}}. \tag{4.1}$$

We say that $\mathcal{E}$ is correctable perfectly if $\mathcal{RES} = \mathcal{I}_{\mathcal{X}}$.

For any Euclidean spaces $\mathcal{X}, \mathcal{Y}$ let us define two families of noise channels: these which are probabilistically correctable for $\mathcal{X}$, denoted as $\xi(\mathcal{X},\mathcal{Y})$, and these which are correctable perfectly for $\mathcal{X}$, denoted as $\xi_1(\mathcal{X},\mathcal{Y})$:

$$\begin{aligned} \xi(\mathcal{X},\mathcal{Y}) &:= \{\mathcal{E} \in \mathcal{C}(\mathcal{Y}) :\ \exists_{(\mathcal{S},\mathcal{R}) \in s\mathcal{C}(\mathcal{X},\mathcal{Y}) \times s\mathcal{C}(\mathcal{Y},\mathcal{X})}\ 0 \neq \mathcal{RES} \propto \mathcal{I}_{\mathcal{X}}\}, \\ \xi_1(\mathcal{X},\mathcal{Y}) &:= \{\mathcal{E} \in \mathcal{C}(\mathcal{Y}) :\ \exists_{(\mathcal{S},\mathcal{R}) \in s\mathcal{C}(\mathcal{X},\mathcal{Y}) \times s\mathcal{C}(\mathcal{Y},\mathcal{X})}\ \mathcal{RES} = \mathcal{I}_{\mathcal{X}}\}. \end{aligned} \tag{4.2}$$

We will denote by $p_{\mathcal{X}}(\mathcal{E})$ the maximal probability of successful error correction for a given $\mathcal{E}$ and $\mathcal{X}$, that is

$$p_{\mathcal{X}}(\mathcal{E}) := \max\left\{p :\ \mathcal{RES} = p\mathcal{I}_{\mathcal{X}},\ (\mathcal{S},\mathcal{R}) \in s\mathcal{C}(\mathcal{X},\mathcal{Y}) \times s\mathcal{C}(\mathcal{Y},\mathcal{X})\right\}. \tag{4.3}$$

Observe that $p_{\mathcal{X}}(\mathcal{E})$ is well-defined (we can wirte "max" rather than "sup") as the optimal value of this optimization is achievable.

Remark 4.3. *To simplify some statements of theoretical results, we will occasionally use Definition 4.2 in the context of subchannels, for example; $\mathcal{E} \in \xi(\mathcal{X},\mathcal{Y})$ for $\mathcal{E} \in s\mathcal{C}(\mathcal{Y})$. Moreover, we will use the notation $p_{\mathcal{X}}(\mathcal{E})$ for $\mathcal{E} \in s\mathcal{C}(\mathcal{Y}_0,\mathcal{Y}_1)$; in that example it would naturally mean*

$$\max\left\{p :\ \mathcal{RES} = p\mathcal{I}_{\mathcal{X}},\ (\mathcal{S},\mathcal{R}) \in s\mathcal{C}(\mathcal{X},\mathcal{Y}_0) \times s\mathcal{C}(\mathcal{Y}_1,\mathcal{X})\right\}.$$

4.2 Conditions on probabilistic quantum error correction

To inspect the pQEC procedure, first, we should state conditions which determine when a given noise channel is probabilistically correctable. For deterministic QEC, such conditions have been known for a long time as the Knill-Laflamme conditions [12]. Let $\mathcal{E} = \mathcal{K}\left((E_i)_i\right) \in \mathcal{C}(\mathcal{Y})$ be a given noise channel. Then, according to the Knill-Laflamme Theorem, $\mathcal{E}$ is perfectly correctable for $\mathcal{X}$ if and only if

$$S^\dagger E_j^\dagger E_i S \propto \mathbb{1}_\mathcal{X} \tag{4.4}$$

for all i, j and some isometry operator $S \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$. In the following theorem we generalize the above, to cover probabilistically correctable noise channels.

Theorem 4.4 ([2]). *Let $\mathcal{E} = \mathcal{K}\left((E_i)_i\right) \in \mathcal{C}(\mathcal{Y})$ and $\mathcal{S} = \mathcal{K}\left((S_k)_k\right) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$. Define $\mathcal{F} = \mathcal{E}\mathcal{S}$ given in the Kraus decomposition as $\mathcal{F} = \mathcal{K}\left((F_k)_k\right)$. Then, the following conditions are equivalent:*

(A) There exist $\mathcal{R} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ and $p > 0$ such that

$$\mathcal{R}\mathcal{F} = p\mathcal{I}_\mathcal{X}. \tag{4.5}$$

(B) There exists $R \in \mathcal{P}(\mathcal{Y})$ for which it holds

$$\mathcal{K}\left(\left(\sqrt{R}F_k\right)_k\right) = \mathcal{K}\left((A_i)_i\right): \quad A_i \neq 0, A_j^\dagger A_i \propto \delta_{ij}\mathbb{1}_\mathcal{X}. \tag{4.6}$$

(C) There exist $R \in \mathcal{P}(\mathcal{Y})$ and a matrix $M = (M_{l,k})_{l,k} \neq 0$, for which it holds

$$\forall_{k,l} \quad F_l^\dagger R F_k = M_{l,k}\mathbb{1}_\mathcal{X}. \tag{4.7}$$

$(*)$ *If point (A) holds for $\mathcal{R} = \mathcal{K}\left((R_l)_l\right)$, then $R \in \mathcal{P}(\mathcal{Y})$ from points (B) and (C) can be chosen to satisfy $R = \sum_l R_l^\dagger R_l$. It also holds that $R_l F_k \propto \mathbb{1}_\mathcal{X}$ for any k, l.*

$(**)$ *Moreover, $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$ if and only if there exist $S_* \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ and $R_* \in \mathcal{M}(\mathcal{Y}, \mathcal{X})$ such that*

$$\forall_i \quad R_* E_i S_* \propto \mathbb{1}_\mathcal{X} \tag{4.8}$$

and there exists i_0, for which it holds $R_* E_{i_0} S_* \neq 0$.

Proof. $(B) \implies (A)$

Let us define $\alpha_i > 0$ to satisfy $A_i^\dagger A_i = \alpha_i \mathbb{1}_\mathcal{X}$ and a map $\mathcal{R} : \mathcal{M}(\mathcal{Y}) \to \mathcal{M}(\mathcal{X})$ given by

$$\mathcal{R} = \mathcal{K}\left(\left(\alpha_i^{-1/2} A_i^\dagger \sqrt{R}\right)_i\right). \tag{4.9}$$

We will check that $\mathcal{R}$ is a subchannel. First, from the definition of $\mathcal{R}$, it follows that $\mathcal{R}$ is completely positive. Second, from the assumption (B), operators $\alpha_i^{-1} A_i A_i^\dagger \in \mathcal{P}(\mathcal{Y})$ are projectors satisfying $\left(\alpha_j^{-1} A_j A_j^\dagger\right)\left(\alpha_i^{-1} A_i A_i^\dagger\right) = 0$ for $i \neq j$. Hence, we have

$$\sum_i \alpha_i^{-1} \sqrt{R} A_i A_i^\dagger \sqrt{R} = \sqrt{R}\left(\sum_i \alpha_i^{-1} A_i A_i^\dagger\right)\sqrt{R} \leq R \leq \mathbb{1}_\mathcal{Y}. \tag{4.10}$$

It means that $\mathcal{R} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$. Finally, it holds

$$\mathcal{RF} = \mathcal{K}\left(\left(\alpha_j^{-1/2} A_j^\dagger \sqrt{R} F_i\right)_{i,j}\right) = \mathcal{K}\left(\left(\alpha_j^{-1/2} A_j^\dagger A_i\right)_{i,j}\right) = \mathcal{K}\left((\alpha_i^{1/2}\mathbb{1}_\mathcal{X})_i\right) = p\mathcal{I}_\mathcal{X}, \tag{4.11}$$

where we introduced $p = \sum_i \alpha_i > 0$.

$(A) \implies (B)$

Let $\mathcal{R} = \mathcal{K}((R_l)_l)$ and take $R = \sum_l R_l^\dagger R_l \in \mathcal{P}(\mathcal{Y})$. Define Π_R to be the projector onto the image of R. One can show that $R_l \Pi_R = R_l$ for each l. Indeed, if $\Pi_R|v\rangle = 0$ for some $|v\rangle \in \mathcal{Y}$, then $\langle v|\Pi_R|v\rangle = \langle v|R|v\rangle = 0$. That implies that for each l we have $\langle v|R_l^\dagger R_l|v\rangle$ which further implies $R_l|v\rangle = 0$.

Now, let us define $\widetilde{\mathcal{R}} = \mathcal{K}\left(\left(\widetilde{R_l}\right)_l\right)$, where $\widetilde{R_l} = R_l \sqrt{R}^{-1}$. From the definition of $\widetilde{\mathcal{R}}$ we have $\sum_l \widetilde{R_l}^\dagger \widetilde{R_l} = \sqrt{R}^{-1} R \sqrt{R}^{-1} = \Pi_R$. Using the assumption (A) we get

$$p\mathcal{I}_\mathcal{X} = \mathcal{RF} = \mathcal{K}((R_l F_k)_{l,k}) = \mathcal{K}\left((R_l \sqrt{R}^{-1} \sqrt{R} F_k)_{l,k}\right) = \widetilde{\mathcal{R}}\mathcal{K}\left((\sqrt{R} F_k)_k\right). \tag{4.12}$$

As it holds $p > 0$, we have $\mathcal{K}\left((\sqrt{R}F_k)_k\right) \neq 0$. Hence, there exists a canonical decomposition

$$\mathcal{K}\left((\sqrt{R}F_k)_k\right) = \mathcal{K}((A_i)_i): \quad A_i \neq 0,\ \operatorname{tr}(A_j^\dagger A_i) = 0 \text{ for } i \neq j. \tag{4.13}$$

From the relationship between Kraus representations, it follows that A_i satisfy $\Pi_R A_i = A_i$. Then, by the assumption (A) we get

$$p|\mathbb{1}_\mathcal{X}\rangle\rangle\langle\langle\mathbb{1}_\mathcal{X}| = (\mathcal{RF} \otimes \mathcal{I}_\mathcal{X})(|\mathbb{1}_\mathcal{X}\rangle\rangle\langle\langle\mathbb{1}_\mathcal{X}|) = \sum_i (\widetilde{\mathcal{R}} \otimes \mathcal{I}_\mathcal{X})(|A_i\rangle\rangle\langle\langle A_i|). \tag{4.14}$$

Therefore, from the extremality of the point $|\mathbb{1}_\mathcal{X}\rangle\rangle\langle\langle\mathbb{1}_\mathcal{X}|$ in $\mathcal{P}(\mathcal{X}\otimes\mathcal{X})$ we have $\widetilde{R_l}A_i \propto \mathbb{1}_\mathcal{X}$ for any i, l. We get in conclusion

$$A_j^\dagger A_i = A_j^\dagger \Pi_R A_i = \sum_l A_j^\dagger \widetilde{R_l}^\dagger \widetilde{R_l} A_i \propto \mathbb{1}_\mathcal{X}. \tag{4.15}$$

The above provides that $A_j^\dagger A_i = c_{j,i}\mathbb{1}_\mathcal{X}$, for some $c_{j,i}\in\mathbb{C}$. Then, for $i\neq j$ we have $0 = \operatorname{tr}(A_j^\dagger A_i) = c_{j,i}\dim(\mathcal{X})$ and finally $A_j^\dagger A_i = 0$.

$(B) \implies (C)$

Let us define $\alpha_i > 0$ to satisfy $A_i^\dagger A_i = \alpha_i \mathbb{1}_\mathcal{X}$. From the relationship between Kraus decompositions $\mathcal{K}\left((\sqrt{R}F_l)_l\right)$ and $\mathcal{K}\left((A_i)_i\right)$, there exists an isometry operator U, such that

$$\sqrt{R}F_k = \sum_i U_{k,i}A_i. \tag{4.16}$$

Therefore, it holds

$$F_l^\dagger R F_k = \sum_{i,j} U_{k,i}\overline{U_{l,j}}A_j^\dagger A_i = \sum_i U_{k,i}\overline{U_{l,i}}\alpha_i\mathbb{1}_\mathcal{X}. \tag{4.17}$$

Let us define a matrix $M = (M_{l,k})_{l,k}$ where $M_{l,k} = \sum_i U_{k,i}\overline{U_{l,i}}\alpha_i$. Note, that

$$\operatorname{tr}(M) = \sum_{k,i} |U_{k,i}|^2\alpha_i = \sum_i \alpha_i > 0. \tag{4.18}$$

$(C) \implies (B)$

Let $\mathcal{F} = \mathcal{K}\left((F_k)_{k=1}^r\right)$ for some $r\in\mathbb{N}$. Define an operator $F = \sum_{i=k}^r \langle k|\otimes F_k \in \mathcal{M}(\mathbb{C}^r\otimes\mathcal{X},\mathcal{Y})$. From the assumption (C) it follows

$$F^\dagger R F = M\otimes\mathbb{1}_\mathcal{X}. \tag{4.19}$$

That implies $M\geq 0$. Take the spectral decomposition $M = U^\dagger D U$, where $U\in\mathcal{U}(\mathbb{C}^r)$ and $D\geq 0$ is a diagonal matrix. Let us define

$$A_i = \sum_k \overline{U_{i,k}}\sqrt{R}F_k. \tag{4.20}$$

Observe that $\mathcal{K}\left((\sqrt{R}F_k)_k\right) = \mathcal{K}\left((A_i)_i\right)$. We obtain

$$A_j^\dagger A_i = \sum_{k,l}\overline{U_{i,k}}U_{j,l}F_l^\dagger R F_k = \sum_{k,l}\overline{U_{i,k}}U_{j,l}M_{l,k}\mathbb{1}_\mathcal{X} = D_{j,i}\mathbb{1}_\mathcal{X}. \tag{4.21}$$

This is equivalent to $A_j^\dagger A_i \propto \delta_{ij} \mathbb{1}_\mathcal{X}$. Finally, $A_i \neq 0$ if and only if $D_{i,i} > 0$. By the fact that $M \neq 0$ we conclude the set $\{A_i : A_i \neq 0\}$ is non-empty.

(*)

Assume that $\mathcal{R} = \mathcal{K}\left((R_l)_l\right)$ and it holds (A). From the proof of implications $(A) \implies (B)$ and $(B) \implies (C)$ it follows that R can be chosen as $R = \sum_l R_l^\dagger R_l$. The relation $R_l F_k \propto \mathbb{1}_\mathcal{X}$ is a consequence of $(\mathcal{RF} \otimes \mathcal{I}_\mathcal{X})\left(|\mathbb{1}_\mathcal{X}\rangle\rangle\langle\langle\mathbb{1}_\mathcal{X}|\right) = \sum_{l,k} |R_l F_k\rangle\rangle\langle\langle R_l F_k| = p|\mathbb{1}_\mathcal{X}\rangle\rangle\langle\langle\mathbb{1}_\mathcal{X}|$ and the extremality of the point $|\mathbb{1}_\mathcal{X}\rangle\rangle\langle\langle\mathbb{1}_\mathcal{X}|$ in $\mathcal{P}(\mathcal{X} \otimes \mathcal{X})$.

(**)

If $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$, then for some $\mathcal{S} = \mathcal{K}\left((S_k)_k\right) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ and $\mathcal{R} = \mathcal{K}\left((R_l)_l\right) \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ and $p > 0$ we have

$$p|\mathbb{1}_\mathcal{X}\rangle\rangle\langle\langle\mathbb{1}_\mathcal{X}| = (\mathcal{RES} \otimes \mathcal{I}_\mathcal{X})\left(|\mathbb{1}_\mathcal{X}\rangle\rangle\langle\langle\mathbb{1}_\mathcal{X}|\right) = \sum_{l,i,k} |R_l E_i S_k\rangle\rangle\langle\langle R_l E_i S_k|. \tag{4.22}$$

From the extremality of the point $|\mathbb{1}_\mathcal{X}\rangle\rangle\langle\langle\mathbb{1}_\mathcal{X}|$ we obtain $R_l E_i S_k \propto \mathbb{1}_\mathcal{X}$. Also, as $p > 0$, there exist l_0, i_0, k_0 such that $R_{l_0} E_{i_0} S_{k_0} \neq 0$. We can take $S_* = S_{k_0}$ and $R_* = R_{l_0}$.

Now, let us assume that there exist $S_* \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ and $R_* \in \mathcal{M}(\mathcal{Y}, \mathcal{X})$ such that

$$\forall_i \quad R_* E_i S_* \propto \mathbb{1}_\mathcal{X} \tag{4.23}$$

and there exists i_0, for which it holds $R_* E_{i_0} S_* \neq 0$. There exist $q_0, q_1 > 0$ for which $\mathcal{S} = q_0 \mathcal{K}\left((S_*)\right) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ and $\mathcal{R} = q_1 \mathcal{K}\left((R_*)\right) \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$. One may note that $0 \neq \mathcal{RES} \propto \mathcal{I}_\mathcal{X}$.

□

Corollary 4.5 ([2]). *Let $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ and let $\mathcal{S} = \mathcal{K}\left((S_k)_k\right), R$ and M satisfy the condition Theorem 4.4 (C). Then, we can define a decoding operation $\mathcal{R} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ such that $\mathcal{RES} = p\mathcal{I}_\mathcal{X} \neq 0$ in the following way:*

1. *Let $M = U^\dagger D U$ be the spectral decomposition of M.*

2. *Define $A_i = \sum_{k,l} \overline{U_{i,kl}} \sqrt{R} E_l S_k$.*

3. *For each $A_i \neq 0$ define $\alpha_i :\ A_i^\dagger A_i = \alpha_i \mathbb{1}_\mathcal{X}$.*

4. *The recovery subchannel is given as $\mathcal{R} = \mathcal{K}\left(\left(\alpha_i^{-1/2} A_i^\dagger \sqrt{R}\right)_i\right)$.*

Remark 4.6. *For $\mathcal{E} = \mathcal{K}\left((E_i)_{i=1}^r\right) \in \mathcal{C}(\mathcal{Y})$ let us compare the condition from Theorem 4.4*

$$\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y}) \quad \Longleftrightarrow \quad \exists_{S,R} \forall_i \; RE_iS \propto \mathbb{1}_{\mathcal{X}}, \; \exists_{i_0} \; RE_{i_0}S \neq 0 \tag{4.24}$$

with the Knill-Laflamme conditions

$$\mathcal{E} \in \xi_1(\mathcal{X}, \mathcal{Y}) \quad \Longleftrightarrow \quad \exists_{S \neq 0} \forall_{i,j} \; S^\dagger E_j^\dagger E_i S \propto \mathbb{1}_{\mathcal{X}}. \tag{4.25}$$

The latter, is a constraint satisfaction problem with r^2 quadratic constrains $S^\dagger E_j^\dagger E_i S \propto \mathbb{1}_{\mathcal{X}}$ for the variable $S \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$, which satisfies $S \neq 0$. The parameters $E_j^\dagger E_i$ constitute a $^\dagger-$closed algebra $\operatorname{span}\left((E_j^\dagger E_i)_{i,j}\right)$, such that $\mathbb{1}_{\mathcal{Y}} \in \operatorname{span}\left((E_j^\dagger E_i)_{i,j}\right)$. In comparison, the conditions for $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$ represent a constraint satisfaction problem with r bilinear constrains $RE_iS \propto \mathbb{1}_{\mathcal{X}}$ for the variables $S \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ and $R \in \mathcal{M}(\mathcal{Y}, \mathcal{X})$. Additionally, it must hold $RE_{i_0}S \neq 0$ for some $i_0 \in \{1, \ldots, r\}$. In this problem, the parameters E_i are arbitrary operators from $\mathcal{M}(\mathcal{Y})$, which satisfy $\sum_i E_i^\dagger E_i = \mathbb{1}_{\mathcal{Y}}$.

The only property that distinguishes the tuple $(E_i)_{i=1}^r : \mathcal{K}\left((E_i)_{i=1}^r\right) \in \mathcal{C}(\mathcal{Y})$ from an arbitrary tuple $(M_i)_{i=1}^r \subset \mathcal{M}(\mathcal{Y})$ is the equality $\sum_{i=1}^r E_i^\dagger E_i = \mathbb{1}_{\mathcal{Y}}$. Nevertheless, from the perspective of our study a less restrictive property $\sum_{i=1}^r E_i^\dagger E_i > 0$ plays an equal role in studying $\xi(\mathcal{X}, \mathcal{Y})$. The following lemma clarifies this statement.

Lemma 4.7 ([2])**.** *Let $\mathcal{X}, \mathcal{Y}$ be Euclidean spaces and $\mathcal{E} = \mathcal{K}\left((E_i)\right) \in \mathcal{C}(\mathcal{Y})$. Define an invertible matrix $Q \in \mathcal{M}(\mathcal{Y})$ and an operation $\mathcal{F}(Y) = \mathcal{E}(QYQ^\dagger)$. Then, $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$ if and only if $\mathcal{F} \in \xi(\mathcal{X}, \mathcal{Y})$.*

Proof. On the one hand, from Theorem 4.4 we have

$$\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y}) \quad \Longleftrightarrow \quad \exists_{S,R} \forall_i \; RE_iS \propto \mathbb{1}_{\mathcal{X}}, \; \exists_{i_0} \; RE_{i_0}S \neq 0. \tag{4.26}$$

On the other hand, for each i it holds $RE_iS = R(E_iQ)(Q^{-1}S)$. □

4.3 Optimization of the probability of success

In general, the difficulty of finding error-correcting schemes $(\mathcal{S}, \mathcal{R})$ comes from bilinearity of the problem Eq. (4.8). Calculating the maximal probability of successful error correction $p_{\mathcal{X}}(\mathcal{E})$ defined in Eq. (4.3) is even harder task. However, if we fix an encoding operation $\mathcal{S}$ (or decoding $\mathcal{R}$), it is possible to calculate $p_{\mathcal{X}}(\mathcal{E}\mathcal{S})$ (or $p_{\mathcal{X}}(\mathcal{R}\mathcal{E})$) using SDP programming.

Lemma 4.8 ([2]). *Let $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{r-1}\right) \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$. Then, it holds*

$$p_{\mathcal{X}}(\mathcal{F}) = \max\left\{\mathrm{tr}(P) : \begin{cases} P \in \mathcal{P}(\mathbb{C}^r), \\ \mathrm{tr}_{\mathbb{C}^r}\left(R_F(P \otimes \mathbb{1}_{\mathcal{X}})\right) \leq \mathbb{1}_{\mathcal{X}}, \\ (\Pi_F \otimes \mathbb{1}_{\mathcal{X}})(P \otimes |\mathbb{1}_{\mathcal{X}}\rangle\rangle\langle\langle\mathbb{1}_{\mathcal{X}}|)(\Pi_F \otimes \mathbb{1}_{\mathcal{X}}) = P \otimes |\mathbb{1}_{\mathcal{X}}\rangle\rangle\langle\langle\mathbb{1}_{\mathcal{X}}| \end{cases}\right\}, \tag{4.27}$$

where $R_F = (FF^\dagger)^{-1}, \Pi_F = FF^{-1}$ for $F = \sum_{i=0}^{r-1} |i\rangle \otimes F_i \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$.

An optimal scheme $(\mathcal{S}, \mathcal{R}) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y}) \times s\mathcal{C}(\mathcal{X})$ which achieves the probability $p_{\mathcal{X}}(\mathcal{F})$, that is $\mathcal{R}\mathcal{F}\mathcal{S} = p_{\mathcal{X}}(\mathcal{F})\mathcal{I}_{\mathcal{X}}$, can be taken as

$$\begin{aligned} \mathcal{S}(X) &= F^{-1}(P_0 \otimes X)(F^{-1})^\dagger, \\ \mathcal{R}(X) &= X, \end{aligned} \tag{4.28}$$

where P_0 is an argument maximizing $p_{\mathcal{X}}(\mathcal{F})$ in Eq. (4.27). Additionally, if there exists another optimal scheme $(\widetilde{\mathcal{S}}, \widetilde{\mathcal{R}})$, that is $\widetilde{\mathcal{R}}\mathcal{F}\widetilde{\mathcal{S}} = p_{\mathcal{X}}(\mathcal{F})\mathcal{I}_{\mathcal{X}}$, then $\mathrm{rank}(J(\mathcal{S})) \leq \mathrm{rank}(J(\widetilde{\mathcal{S}}))$.

Moreover, if $p_{\mathcal{X}}(\mathcal{F}) > 0$, then

$$\|R_F\|_\infty^{-1} \leq p_{\mathcal{X}}(\mathcal{F}) \leq \|R_F^{-1}\|_\infty. \tag{4.29}$$

Proof. Let us fix $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{r-1}\right) \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$. Assume that for some $\widetilde{\mathcal{S}} = \mathcal{K}\left((\widetilde{S}_j)_j\right) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ and $\widetilde{\mathcal{R}} = \mathcal{K}\left((\widetilde{R_k})_k\right) \in s\mathcal{C}(\mathcal{X})$ it holds $\widetilde{\mathcal{R}}\mathcal{F}\widetilde{\mathcal{S}} = p_{\mathcal{X}}(\mathcal{F})\mathcal{I}_{\mathcal{X}} \neq 0$. From Theorem 4.4 we have $\widetilde{R}_k F_i \widetilde{S}_j \propto \mathbb{1}_{\mathcal{X}}$ and there are k_0, i_0, j_0 such that $\widetilde{R}_{k_0} F_{i_0} \widetilde{S}_{j_0} \neq 0$. Hence, for each k we have $\widetilde{R}_k \propto (F_{i_0}\widetilde{S}_{j_0})^{-1}$. That implies the map $\widetilde{\mathcal{R}}$ can be written as $\widetilde{\mathcal{R}}(X) = \widetilde{R}X\widetilde{R}^\dagger$. Now, consider another scheme $(\mathcal{S}', \mathcal{R}')$, where $\mathcal{R}'(X) = X$ and $\mathcal{S}'(X) = \widetilde{\mathcal{S}}\widetilde{\mathcal{R}}(X) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$. We get

$$\mathcal{R}'\mathcal{F}\mathcal{S}'(X) = \mathcal{F}\widetilde{\mathcal{S}}\widetilde{\mathcal{R}}(X) = \widetilde{R}^{-1}\left(\widetilde{\mathcal{R}}\mathcal{F}\widetilde{\mathcal{S}}\widetilde{\mathcal{R}}(X)\right)(\widetilde{R}^\dagger)^{-1} = p_{\mathcal{X}}(\mathcal{F})X. \tag{4.30}$$

Therefore, the scheme $(\mathcal{S}', \mathcal{R}')$ is also optimal and $\mathrm{rank}(J(\mathcal{S}')) \leq \mathrm{rank}(J(\widetilde{\mathcal{S}}))$.

From now, we can write $p_{\mathcal{X}}(\mathcal{F}) = \max\{p : \mathcal{F}\mathcal{S} = p\mathcal{I}_{\mathcal{X}}, \mathcal{S} \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})\}$. Define

$$\begin{aligned} F &= \sum_{i=0}^{r-1} |i\rangle \otimes F_i \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X}), \\ \Pi_F &= FF^{-1}, \\ R_F &= (FF^\dagger)^{-1}. \end{aligned} \tag{4.31}$$

The action of $\mathcal{F}$ may be stated as $\mathcal{F}(Y) = \mathrm{tr}_{\mathbb{C}^r}(FYF^\dagger)$. If for some $\mathcal{S} \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ it holds $\mathcal{F}\mathcal{S}(X) = \mathrm{tr}_{\mathbb{C}^r}(F\mathcal{S}(X)F^\dagger) = pX$, then according to Theorem 4.4 we have

$F\mathcal{S}(X)F^\dagger = P \otimes X$, where $P \in \mathcal{P}(\mathbb{C}^r)$. Without loss of the generality we may consider $\mathcal{S}$ such that $F^{-1}F\mathcal{S}(X)F^{-1}F = \mathcal{S}(X)$ (one can note that $\mathrm{rank}(J(\mathcal{S}))$ will not increase). Then, the equation $P \otimes X = F\mathcal{S}(X)F^\dagger$ implies

$$F^{-1}(P \otimes X)(F^{-1})^\dagger = F^{-1}F\mathcal{S}(X)F^\dagger(F^{-1})^\dagger = \mathcal{S}(X). \tag{4.32}$$

To sum up, the optimal strategy $(\mathcal{S}, \mathcal{R})$ which also minimize $\mathrm{rank}(J(\mathcal{S}))$ is of the form

$$\begin{aligned} \mathcal{S}(X) &= F^{-1}(P \otimes X)(F^{-1})^\dagger, \\ \mathcal{R}(X) &= X, \end{aligned} \tag{4.33}$$

where it holds $P \otimes X = F\mathcal{S}(X)F^\dagger = \Pi_F(P \otimes X)\Pi_F$ for all X. In that case $\mathcal{R}\mathcal{F}\mathcal{S} = p\mathcal{I}_\mathcal{X}$, where $p = \mathrm{tr}(P)$. Moreover, $\mathcal{S} \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ when $\mathrm{tr}_{\mathbb{C}^r}\left(R_F(P \otimes \mathbb{1}_\mathcal{X})\right) \leq \mathbb{1}_\mathcal{X}$. It justifies the Eq. (4.27) and the form of the optimal scheme.

Assume now that $p_\mathcal{X}(\mathcal{F}) > 0$, that is, we can find $0 \neq P \in \mathcal{P}(\mathbb{C}^r)$ satisfying $(\Pi_F \otimes \mathbb{1}_\mathcal{X})(P \otimes |\mathbb{1}_\mathcal{X}\rangle\rangle\langle\langle\mathbb{1}_\mathcal{X}|)(\Pi_F \otimes \mathbb{1}_\mathcal{X}) = P \otimes |\mathbb{1}_\mathcal{X}\rangle\rangle\langle\langle\mathbb{1}_\mathcal{X}|$ and $\mathrm{tr}_{\mathbb{C}^r}\left(R_F(P \otimes \mathbb{1}_\mathcal{X})\right) \leq \mathbb{1}_\mathcal{X}$. Let $P = p\rho$ for $\rho \in \mathcal{D}(\mathbb{C}^r)$ and define $\widetilde{P} = \left\|\mathrm{tr}_{\mathbb{C}^r}\left(R_F(\rho \otimes \mathbb{1}_\mathcal{X})\right)\right\|_\infty^{-1} \rho$. Observe that $\widetilde{P}$ also belongs to the optimization domain of $p_\mathcal{X}(\mathcal{F})$. Hence, we get

$$p_\mathcal{X}(\mathcal{F}) \geq \left\|\mathrm{tr}_{\mathbb{C}^r}\left(R_F(\rho \otimes \mathbb{1}_\mathcal{X})\right)\right\|_\infty^{-1} \geq \|R_F\|_\infty^{-1}. \tag{4.34}$$

On the other hand, it holds $\|R_F^{-1}\|_\infty^{-1}\Pi_F \leq R_F$. Hence, for any P which belongs to the optimization domain of $p_\mathcal{X}(\mathcal{F})$ it holds

$$\|R_F^{-1}\|_\infty^{-1}\mathrm{tr}(P)\mathbb{1}_\mathcal{X} \leq \mathrm{tr}_{\mathbb{C}^r}\left(R_F(P \otimes \mathbb{1}_\mathcal{X})\right) \leq \mathbb{1}_\mathcal{X}. \tag{4.35}$$

It implies that $p_\mathcal{X}(\mathcal{F}) \leq \|R_F^{-1}\|_\infty$. □

Corollary 4.9 ([2]). *Let $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{r-1}\right) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$. Then, it holds*

$$p_\mathcal{X}(\mathcal{F}) = \max\left\{\mathrm{tr}(P) : \begin{cases} P \in \mathcal{P}(\mathbb{C}^r), \\ P \otimes \mathbb{1}_\mathcal{X} \leq \widetilde{F}\widetilde{F}^\dagger, \\ (\Pi_{\widetilde{F}} \otimes \mathbb{1}_\mathcal{X})(P \otimes |\mathbb{1}_\mathcal{X}\rangle\rangle\langle\langle\mathbb{1}_\mathcal{X}|)(\Pi_{\widetilde{F}} \otimes \mathbb{1}_\mathcal{X}) = P \otimes |\mathbb{1}_\mathcal{X}\rangle\rangle\langle\langle\mathbb{1}_\mathcal{X}| \end{cases}\right\}, \tag{4.36}$$

where $\Pi_{\widetilde{F}} = \widetilde{F}\widetilde{F}^{-1}$ for $\widetilde{F} = \sum_{i=0}^{r-1} |i\rangle \otimes F_i^\dagger \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$.

An optimal scheme $(\mathcal{S}, \mathcal{R}) \in s\mathcal{C}(\mathcal{X}) \times s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ which achieves the probability $p_\mathcal{X}(\mathcal{F})$ can be taken as

$$\begin{aligned} \mathcal{S}(X) &= X, \\ \mathcal{R}(Y) &= \mathrm{tr}_{\mathbb{C}^r}\left((P_0 \otimes \mathbb{1}_\mathcal{X})(\widetilde{F}^{-1})^\dagger Y (\widetilde{F}^{-1})\right), \end{aligned} \tag{4.37}$$

where P_0 is an argument maximizing $p_\mathcal{X}(\mathcal{F})$ in Eq. (4.36).

Proof. This proof is based on the proof of Lemma 4.8. Let us define $\widetilde{F} = \sum_{i=0}^{r-1} |i\rangle \otimes F_i^\dagger \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$ and $\Pi_{\widetilde{F}} = \widetilde{F}\widetilde{F}^{-1}$ for a given $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{r-1}\right) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$. First, similarly as in Lemma 4.8 we can show that $\mathcal{S} = \mathcal{I}_\mathcal{X}$ is an optimal encoding strategy. Then, one may note that $\mathcal{R}\mathcal{F} = p\mathcal{I}_\mathcal{X}$ if and only if $\mathcal{F}^\dagger\mathcal{R}^\dagger = p\mathcal{I}_\mathcal{X}$. Therefore, an optimal decoding strategy is of the form $\mathcal{R}^\dagger(X) = \widetilde{F}^{-1}(P \otimes X)(\widetilde{F}^{-1})^\dagger$ for $P \in \mathcal{P}(\mathbb{C}^r)$, such that $P \otimes X = \Pi_{\widetilde{F}}(P \otimes X)\Pi_{\widetilde{F}}$ for any X. In that case, if $\mathcal{R}\mathcal{F}\mathcal{S} = p\mathcal{I}_\mathcal{X}$, then $p = \mathrm{tr}(P)$. Finally, we require that $\mathcal{R} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$, that is equivalent to $\mathcal{R}^\dagger(\mathbb{1}_\mathcal{X}) \le \mathbb{1}_\mathcal{Y}$, which is equivalent to $P \otimes \mathbb{1}_\mathcal{X} \le \widetilde{F}\widetilde{F}^\dagger$ due to the relation $P \otimes \mathbb{1}_\mathcal{X} = \Pi_{\widetilde{F}}(P \otimes \mathbb{1}_\mathcal{X})\Pi_{\widetilde{F}}$. □

One can note that it is possible to use a sequence of optimization procedures presented in Lemma 4.8 and Corollary 4.9 to increase the probability of successful error correction. In more detail, if we have fixed decoding operation $\mathcal{R}_0$ we run the procedure presented in Lemma 4.8 for $\mathcal{R}_0\mathcal{E}$ to calculate the encoding operation $\mathcal{S}_0$. Then, for $\mathcal{E}\mathcal{S}_0$ we run the procedure presented in Corollary 4.9 to calculate $\mathcal{R}_1$, and so on until the obtained sequence of probability values will converge.

4.4 Realization of pQEC procedure and the need for mixed state encoding

In this section, we will investigate the form of error-correcting scheme $(\mathcal{S}, \mathcal{R})$ which provides the maximal probability of successful error correction. For perfectly correctable noise channels, the encoding $\mathcal{S}$ can be realized by an isometry channel. This observation meaningfully reduces the complexity of finding error-correcting schemes – it is enough to consider a vector representation of pure states. Inspired by that, we ask if a similar behavior occurs in the probabilistic quantum error correction?

The answer to this question is surprisingly – negative. We will provide a class of noise channels $\mathcal{E}$ for which, in order to maximize the probability p of successful error correction, we need to consider a general channel realization of $\mathcal{S}$. Paraphrasing, we have to encode the initial state $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{X})$ into the mixed state $\mathcal{S}(|\psi\rangle\langle\psi|)$ to improve the probability of success.

Proposition 4.10 ([2]). *For a given channel $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$, let us fix an error-correcting scheme $(\mathcal{S}, \mathcal{R}) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y}) \times s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ such that $\mathcal{R}\mathcal{E}\mathcal{S} = p\mathcal{I}_\mathcal{X}$, for some $p > 0$. Then, the following holds:*

(A) There exist $\widetilde{\mathcal{S}} \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ and $\widetilde{\mathcal{R}} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ such that $\widetilde{\mathcal{R}}\mathcal{E}\widetilde{\mathcal{S}} = p\mathcal{I}_\mathcal{X}$.

(B) If $\mathcal{R} \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$, then there exists $\widetilde{\mathcal{S}} = \mathcal{K}\left((\widetilde{S})\right) \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ such that $\mathcal{R}\mathcal{E}\widetilde{\mathcal{S}} = \mathcal{I}_\mathcal{X}$.

(C) If $p = 1$, then there exist $\widetilde{\mathcal{S}} = \mathcal{K}\left((\widetilde{S})\right) \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ and $\widetilde{\mathcal{R}} \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$ such that $\widetilde{\mathcal{R}}\mathcal{E}\widetilde{\mathcal{S}} = \mathcal{I}_{\mathcal{X}}$.

Proof. (A)
Let $\mathcal{S} = \mathcal{K}((S_k)_k)$ and $S = \sum_k S_k^\dagger S_k \leq \mathbb{1}_{\mathcal{X}}$. Using Theorem 4.4 one can show that there exists k_0 for which $\operatorname{rank}(S_{k_0}) = \dim(\mathcal{X})$. Hence, S is invertible. Define $\widetilde{\mathcal{S}} \in \mathcal{C}(\mathcal{X}, \mathcal{Y}), \widetilde{\mathcal{R}} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ given by the equations

$$\begin{aligned} \widetilde{\mathcal{S}}(X) &= \mathcal{S}\left(S^{-1/2} X S^{-1/2}\right), \\ \widetilde{\mathcal{R}}(Y) &= S^{1/2}\mathcal{R}(Y)S^{1/2}. \end{aligned} \tag{4.38}$$

We obtain $\widetilde{\mathcal{R}}\mathcal{E}\widetilde{\mathcal{S}}(X) = S^{1/2}(\mathcal{R}\mathcal{E}\mathcal{S})\left(S^{-1/2}XS^{-1/2}\right)S^{1/2} = pX$.

(B)
Let $\mathcal{S} = \mathcal{K}((S_k)_k)$ and define $\mathcal{S}_k(X) = S_k X S_k^\dagger$. From Theorem 4.4 there exists k_0 such that $\mathcal{R}\mathcal{E}\mathcal{S}_{k_0} = p_{k_0}\mathcal{I}_{\mathcal{X}}$, for some $p_{k_0} > 0$. For any $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{X})$ it holds then

$$p_{k_0} = \operatorname{tr}\left(\mathcal{R}\mathcal{E}\mathcal{S}_{k_0}(|\psi\rangle\langle\psi|)\right) = \operatorname{tr}\left(\mathcal{S}_{k_0}(|\psi\rangle\langle\psi|)\right) = \langle\psi|S_{k_0}^\dagger S_{k_0}|\psi\rangle. \tag{4.39}$$

Hence, we get $S_{k_0}^\dagger S_{k_0} = p_{k_0}\mathbb{1}_{\mathcal{X}}$. Define $\widetilde{\mathcal{S}} = \frac{1}{p_{k_0}}\mathcal{S}_{k_0} \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ and note that $\mathcal{R}\mathcal{E}\widetilde{\mathcal{S}} = \mathcal{I}_{\mathcal{X}}$.

(C)
Let $\mathcal{S} = \mathcal{K}((S_k)_k)$ and $\mathcal{R} = \mathcal{K}((R_k)_k)$. For any $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{X})$ we have

$$1 = \operatorname{tr}(\mathcal{R}\mathcal{E}\mathcal{S}(|\psi\rangle\langle\psi|)) \leq \operatorname{tr}(\mathcal{S}(|\psi\rangle\langle\psi|)) \leq 1. \tag{4.40}$$

Therefore, for any $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{X})$ we get $\langle\psi|\left(\sum_k S_k^\dagger S_k\right)|\psi\rangle = 1$, which implies $\mathcal{S} \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$. Let $R = \sum_k R_k^\dagger R_k \leq \mathbb{1}_{\mathcal{Y}}$. Then, it holds $\operatorname{tr}\left((\mathbb{1}_{\mathcal{Y}} - R)\mathcal{E}\mathcal{S}(X)\right) = 0$. Define $\widetilde{\mathcal{R}} \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$ by the equation

$$\widetilde{\mathcal{R}}(Y) = \mathcal{R}(Y) + \operatorname{tr}\left((\mathbb{1}_{\mathcal{Y}} - R)Y\right)\rho_{\mathcal{X}}^*. \tag{4.41}$$

Observe that $\widetilde{\mathcal{R}}\mathcal{E}\mathcal{S} = \mathcal{I}_{\mathcal{X}}$. The rest of the proof follows from (B). □

We may use Proposition 4.10 (A) to state a realization of the pQEC procedure (see Fig. 4.1). For a given noise channel $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ let $(\mathcal{S}, \mathcal{R}) \in \mathcal{C}(\mathcal{X}, \mathcal{Y}) \times s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ be an error-correcting scheme for which $\mathcal{R}\mathcal{E}\mathcal{S} = p\mathcal{I}_{\mathcal{X}}$, where $p > 0$. The encoding channel $\mathcal{S}$ can be realized using the Stinespring representation given in the form $\mathcal{S}(X) = \operatorname{tr}_2\left(U_{\mathcal{S}} X U_{\mathcal{S}}^\dagger\right)$. The encoded state is then sent through $\mathcal{E}$. The decoding

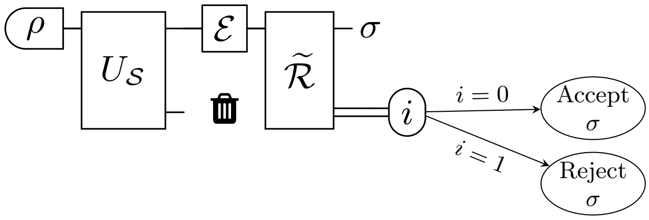


Figure 4.1: Schematic realization of the pQEC procedure for the noise channel $\mathcal{E}$.

subchannel $\mathcal{R} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ can be realized by implementing the channel $\widetilde{\mathcal{R}} \in \mathcal{C}(\mathcal{Y}, \mathcal{X} \otimes \mathbb{C}^2)$ given in the form $\widetilde{\mathcal{R}}(Y) = \mathcal{R}(Y) \otimes |0\rangle\langle 0| + \Psi(Y) \otimes |1\rangle\langle 1|$, where $\Psi \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ such that $(\mathcal{R} + \Psi) \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$. In summary, the output of the whole procedure consists of a quantum state $\sigma \in \mathcal{D}(\mathcal{X})$ and a classical label $i \in \{0, 1\}$. If the label $i = 0$ is obtained, we know that $\sigma \propto \mathcal{RES}(\rho) = p\rho$, and hence, the output state can be accepted. Otherwise, if $i = 1$, the output state $\sigma \propto \Psi\mathcal{ES}(\rho)$ should be rejected, as in general it may differ from ρ.

Now, we provide some examples which explain why (sometimes) we need to use a general channel $\mathcal{S}$ as an encoding operation (see Fig. 4.1). We start with a parametrized family of noise channels $\{\mathcal{E}_R\}_R$ for which the mixed state encoding improves the probability of successful error correction. In our example we take $\mathcal{X} = \mathbb{C}^2$ and $\mathcal{Y} = \mathbb{C}^4$. For each $R \in \mathcal{P}(\mathbb{C}^4)$ satisfying $R \leq \mathbb{1}_{\mathbb{C}^4}$ let us define a noise channel $\mathcal{E}_R \in \mathcal{C}(\mathbb{C}^4)$ given by the equation

$$\mathcal{E}_R(Y) = |0\rangle\langle 0| \otimes \operatorname{tr}_1\left(\sqrt{R} Y \sqrt{R}\right) + |1\rangle\langle 1| \otimes \operatorname{tr}\left([\mathbb{1}_{\mathbb{C}^4} - R]Y\right) \rho^*_{\mathbb{C}^2}, \tag{4.42}$$

where tr_1 is the partial trace over the first subsystem of $\mathcal{Y} = \mathbb{C}^2 \otimes \mathbb{C}^2$. For a given R we will consider:

- the optimal probability p_0 of successful error correction

$$p_0(R) \coloneqq p_{\mathbb{C}^2}(\mathcal{E}_R) = \max\left\{p : \ \mathcal{R}\mathcal{E}_R\mathcal{S} = p\mathcal{I}_{\mathbb{C}^2}, (\mathcal{S}, \mathcal{R}) \in s\mathcal{C}(\mathbb{C}^2, \mathbb{C}^4) \times s\mathcal{C}(\mathbb{C}^4, \mathbb{C}^2)\right\}, \tag{4.43}$$

- the optimal probability p_1 of successful error correction restricted to the pure state encoding:

$$p_1(R) \coloneqq \max\left\{p : \ \mathcal{R}\mathcal{E}_R\mathcal{S} = p\mathcal{I}_{\mathbb{C}^2}, \mathcal{S} = \mathcal{K}\left((S)\right) \in s\mathcal{C}(\mathbb{C}^2, \mathbb{C}^4), \mathcal{R} \in s\mathcal{C}(\mathbb{C}^4, \mathbb{C}^2)\right\}. \tag{4.44}$$

Our claim, which we will prove, is that there exists a family of operators R for which $p_0(R) > p_1(R)$.

Lemma 4.11 ([2]). *Let $R \in \mathcal{P}(\mathbb{C}^4)$ and $R \leq \mathbb{1}_{\mathbb{C}^4}$. Define Π_R as a projector on the support of R. For $\mathcal{E}_R$ defined in Eq. (4.42) we have the following simplified form of the maximization problem $p_0(R)$:*

$$p_0(R) = \max \left\{ \operatorname{tr}(P) : \begin{cases} P \in \mathcal{P}(\mathbb{C}^2), \\ \operatorname{tr}_1 \left(R^{-1}(P \otimes \mathbb{1}_{\mathbb{C}^2}) \right) \leq \mathbb{1}_{\mathbb{C}^2}, \\ \forall_{X \in \mathcal{M}(\mathbb{C}^2)} \ \Pi_R (P \otimes X) \Pi_R = P \otimes X \end{cases} \right\}. \tag{4.45}$$

An optimal scheme $(\mathcal{S}, \mathcal{R})$ which achieves the probability $p_0(R)$ can be taken as

$$\begin{aligned} \mathcal{S}(X) &= \sqrt{R}^{-1} (P_0 \otimes X) \sqrt{R}^{-1}, \\ \mathcal{R}(Y) &= \operatorname{tr}_1 \left(Y \left(|0\rangle\langle 0| \otimes \mathbb{1}_{\mathbb{C}^2} \right) \right), \end{aligned} \tag{4.46}$$

where P_0 is an argument maximizing $p_0(R)$ in Eq. (4.45). Moreover, if there exists another optimal scheme $(\widetilde{\mathcal{S}}, \widetilde{\mathcal{R}})$, then $\operatorname{rank}(J(\widetilde{\mathcal{S}})) \geq \operatorname{rank}(J(\mathcal{S}))$.

Proof. Let us investigate the form of an optimal scheme $(\mathcal{S}, \mathcal{R})$ that maximizes the probability p of successful error correction, $\mathcal{R}\mathcal{E}_R\mathcal{S} = p\mathcal{I}_{\mathbb{C}^2}$. First, according to Theorem 4.4 it must hold $\operatorname{tr}([\mathbb{1}_{\mathbb{C}^4} - R]\mathcal{S}(X))\mathcal{R}(|1\rangle\langle 1| \otimes \rho^*_{\mathbb{C}^2}) = 0$ for any $X \in \mathcal{M}(\mathbb{C}^2)$. Hence, without loss of the generality we may take $\mathcal{R}(Y) = \mathcal{R}((|0\rangle\langle 0| \otimes \mathbb{1}_{\mathbb{C}^2}) Y (|0\rangle\langle 0| \otimes \mathbb{1}_{\mathbb{C}^2}))$. $\mathcal{R}$ must be then of the form $\mathcal{R}(A \otimes B) = \operatorname{tr}(A|0\rangle\langle 0|)\, \widetilde{\mathcal{R}}(B)$, where $\widetilde{\mathcal{R}} \in s\mathcal{C}(\mathbb{C}^2)$. Consider a map $\mathcal{F} \in s\mathcal{C}(\mathbb{C}^4, \mathbb{C}^2)$ given by the equation $\mathcal{F}(Y) = \operatorname{tr}_1 \left(\sqrt{R} Y \sqrt{R} \right)$ and note that $\mathcal{R}\mathcal{E}_R\mathcal{S} = \widetilde{\mathcal{R}}\mathcal{F}\mathcal{S}$. The rest of the proof follows from Lemma 4.8. □

Let us separately consider two cases: $\operatorname{rank}(R) < 4$ and $\operatorname{rank}(R) = 4$. The first one will be discussed briefly as it will not support our claim.

Corollary 4.12 ([2]). *Let us take $R \in \mathcal{P}(\mathbb{C}^4)$ such that $R \leq \mathbb{1}_{\mathbb{C}^4}$ and $\operatorname{rank}(R) < 4$. Define Π_R as a projector on the support of R. For the noise channel defined in Eq. (4.42) we have $p_0(R) = p_1(R)$. Moreover, it holds*

- *If $\operatorname{rank}(R) \leq 1$, then $p_0(R) = 0$.*
- *If $\operatorname{rank}(R) = 2, \Pi_R \neq |\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2}, |\psi\rangle \in \mathbb{C}^2$, then $p_0(R) = 0$.*
- *If $\operatorname{rank}(R) = 2, \Pi_R = |\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2}, |\psi\rangle \in \mathbb{C}^2$, then $p_0(R) = \|\operatorname{tr}_1 (R^{-1}(|\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2}))\|_\infty^{-1}$.*
- *If $\operatorname{rank}(R) = 3, \Pi_R = \mathbb{1}_{\mathbb{C}^4} - |\alpha\rangle\langle\alpha|$ and $|\alpha\rangle \in \mathbb{C}^4$ is entangled, then $p_0(R) = 0$.*
- *If $\operatorname{rank}(R) = 3, \Pi_R = \mathbb{1}_{\mathbb{C}^4} - |\psi^\perp\rangle\langle\psi^\perp| \otimes |\phi\rangle\langle\phi|$, where $|\psi^\perp\rangle, |\phi\rangle \in \mathbb{C}^2, |\psi\rangle\langle\psi| \in \mathcal{D}(\mathbb{C}^2)$, then $p_0(R) = \|\operatorname{tr}_1 (R^{-1}(|\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2}))\|_\infty^{-1}$.*

Proof. The proof is based on Lemma 4.11.

In the first case, we assume that $\text{rank}(R) \in \{0,1\}$. Then, for P satisfying $\Pi_R(P\otimes X)\Pi_R = P\otimes X$ we have

$$2\text{rank}(P) = \text{rank}(P\otimes \mathbb{1}_{\mathbb{C}^2}) = \text{rank}(\Pi_R(P\otimes \mathbb{1}_{\mathbb{C}^2})\Pi_R) \leq \text{rank}(\Pi_R) \leq 1. \tag{4.47}$$

Hence, we obtain $\text{rank}(P) \leq \frac{1}{2}$ which implies $P = 0$. In this case $p_0(R) = 0$.

In the second case, we assume that $\text{rank}(R) = 2$. Using the same argumentation for P as in the first case, we get $\text{rank}(P) \leq 1$. We can write $P = |x\rangle\langle x|$ for $|x\rangle \in \mathbb{C}^2$. Note that, if $P \neq 0$, then from the equality $\Pi_R|x,y\rangle = |x,y\rangle$ for $|y\rangle \in \mathbb{C}^2$ we get $\Pi_R = |\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2}$, for $|\psi\rangle = \frac{1}{\|x\|}|x\rangle$. Therefore, if for all $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathbb{C}^2)$ it holds $\Pi_R \neq |\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2}$, we have $p_0(R) = 0$. Otherwise, if $\Pi_R = |\psi_0\rangle\langle\psi_0| \otimes \mathbb{1}_{\mathbb{C}^2}$ for $|\psi_0\rangle\langle\psi_0| \in \mathcal{D}(\mathbb{C}^2)$, we take $P = p|\psi_0\rangle\langle\psi_0|$ for $p \geq 0$. From the assumption $p\text{tr}_1\left(R^{-1}(|\psi_0\rangle\langle\psi_0| \otimes \mathbb{1}_{\mathbb{C}^2})\right) \leq \mathbb{1}_{\mathbb{C}^2}$ we get $p_0(R) = \|\text{tr}_1\left(R^{-1}(|\psi_0\rangle\langle\psi_0| \otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1}$.

In the third case, we assume that $\text{rank}(R) = 3$. Again, P can be written in the form $P = |x\rangle\langle x|$ for $|x\rangle \in \mathbb{C}^2$. Let $\Pi_R = \mathbb{1}_{\mathbb{C}^4} - |\xi\rangle\langle\xi|$, where $|\xi\rangle\langle\xi| \in \mathcal{D}(\mathbb{C}^4)$. If $P \neq 0$, then from the equality $\Pi_R|x,y\rangle = |x,y\rangle$ for $|y\rangle \in \mathbb{C}^2$ we get $\langle\xi|x,y\rangle = 0$, for $|y\rangle \in \mathbb{C}^2$, and hence, $|\xi\rangle \propto |x^\perp, y\rangle$. Therefore, if $|\xi\rangle$ is entangled, we have $p_0(R) = 0$. Otherwise, if $\Pi_R = \mathbb{1}_{\mathbb{C}^4} - |\psi_0^\perp\rangle\langle\psi_0^\perp| \otimes |\phi_0\rangle\langle\phi_0|$ for $|\psi_0^\perp\rangle\langle\psi_0^\perp|, |\phi_0\rangle\langle\phi_0| \in \mathcal{D}(\mathbb{C}^2)$, we take $P = p|\psi_0\rangle\langle\psi_0|$ for $p \geq 0$. The assumption $p\text{tr}_1\left(R^{-1}(|\psi_0\rangle\langle\psi_0| \otimes \mathbb{1}_{\mathbb{C}^2})\right) \leq \mathbb{1}_{\mathbb{C}^2}$ implies $p_0(R) = \|\text{tr}_1\left(R^{-1}(|\psi_0\rangle\langle\psi_0| \otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1}$.

To prove that $p_0(R) = p_1(R)$ observe that in each case the argument maximizing $p_0(R)$ has the form $P_0 = p|\psi_0\rangle\langle\psi_0|$. According to Lemma 4.11 the optimal scheme $(\mathcal{S},\mathcal{R})$ can be taken as $\mathcal{S}(X) = \sqrt{R}^{-1}(P_0\otimes X)\sqrt{R}^{-1}$ and $\mathcal{R}(Y) = \text{tr}_1\left(Y\left(|0\rangle\langle 0| \otimes \mathbb{1}_{\mathbb{C}^2}\right)\right)$. As the pair $(\mathcal{S},\mathcal{R})$ belongs to the optimization domain of Eq. (4.44) we achieve the desired equality. □

In the case when the operator R is invertible, the situation is more interesting.

Corollary 4.13 ([2]). *Let us take $R \in \mathcal{P}(\mathbb{C}^4)$ such that $R \leq \mathbb{1}_{\mathbb{C}^4}$ and* $\text{rank}(R) = 4$. *For the noise channel defined in Eq.* (4.42) *we have*

$$\begin{aligned} p_0(R) &= \max\left\{\|\text{tr}_1\left(R^{-1}(\rho\otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1} :\ \rho \in \mathcal{D}(\mathbb{C}^2)\right\}, \\ p_1(R) &= \max\left\{\|\text{tr}_1\left(R^{-1}(|\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1} :\ |\psi\rangle\langle\psi| \in \mathcal{D}(\mathbb{C}^2)\right\}. \end{aligned} \tag{4.48}$$

Proof. Let us focus on $p_0(R)$ obtained in Eq. (4.45). As $\Pi_R = \mathbb{1}_{\mathbb{C}^4}$, the equation $\Pi_R(P\otimes X)\Pi_R = P\otimes X$ is always satisfied. For a given P, we can take $\rho \in \mathcal{D}(\mathbb{C}^2)$ such that $P = \text{tr}(P)\rho$. The inequality $\text{tr}(P)\text{tr}_1\left(R^{-1}(\rho\otimes \mathbb{1}_{\mathbb{C}^2})\right) \leq \mathbb{1}_{\mathbb{C}^2}$ is then equivalent to $\text{tr}(P) \leq \|\text{tr}_1\left(R^{-1}(\rho\otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1}$.

To calculate $p_1(R)$ it will be sufficient to add the constraint $\mathcal{S} = \mathcal{K}\left((S)\right)$. According to Lemma 4.11 the optimal $\mathcal{S}$ is of the form $\mathcal{S}(X) = \sqrt{R}^{-1}(P\otimes X)\sqrt{R}^{-1}$. As R is invertible, $\mathcal{S} = \mathcal{K}\left((S)\right)$ if and only if $P = |\psi\rangle\langle\psi|$ for some $|\psi\rangle \in \mathbb{C}^2$. □

Proposition 4.14 ([2]). *Let us define a unitary matrix $U \in \mathcal{U}(\mathbb{C}^4)$ which columns form the magic basis [94]*

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & i \\ 0 & i & 1 & 0 \\ 0 & i & -1 & 0 \\ 1 & 0 & 0 & -i \end{pmatrix}. \tag{4.49}$$

Let us also define a diagonal operator $D(\lambda) \coloneqq \mathrm{diag}(\lambda)$, which is parameterized by a 4−dimensional real vector $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, for which it holds $0 < \lambda_i \leq 1$. For $R = UD(\lambda)U^\dagger$ and the noise channel $\mathcal{E}_R$ defined in Eq. (4.42) we have

$$\begin{aligned} p_0(R) &= \frac{4}{\mathrm{tr}(R^{-1})}, \\ p_1(R) &= \frac{4}{\mathrm{tr}(R^{-1}) + \min\left\{\left|\frac{1}{\lambda_1} - \frac{1}{\lambda_2} - \frac{1}{\lambda_3} + \frac{1}{\lambda_4}\right|, \left|\left|\frac{1}{\lambda_1} - \frac{1}{\lambda_4}\right| - \left|\frac{1}{\lambda_2} - \frac{1}{\lambda_3}\right|\right|\right\}}. \end{aligned} \tag{4.50}$$

Proof. The proof is based on Corollary 4.13. First, we calculate $p_0(R)$. Let $|x\rangle = (x_0, x_1)^\top$. Then, we have

$$(\langle x| \otimes \mathbb{1}_{\mathbb{C}^2}) R^{-1} (|x\rangle \otimes \mathbb{1}_{\mathbb{C}^2}) = \frac{1}{2} \begin{pmatrix} \frac{|x_0|^2}{\lambda_1} + \frac{|x_1|^2}{\lambda_2} + \frac{|x_1|^2}{\lambda_3} + \frac{|x_0|^2}{\lambda_4} & \frac{x_1 \bar{x}_0}{\lambda_1} + \frac{x_0 \bar{x}_1}{\lambda_2} - \frac{x_0 \bar{x}_1}{\lambda_3} - \frac{x_1 \bar{x}_0}{\lambda_4} \\ \frac{x_0 \bar{x}_1}{\lambda_1} + \frac{x_1 \bar{x}_0}{\lambda_2} - \frac{x_1 \bar{x}_0}{\lambda_3} - \frac{x_0 \bar{x}_1}{\lambda_4} & \frac{|x_1|^2}{\lambda_1} + \frac{|x_0|^2}{\lambda_2} + \frac{|x_0|^2}{\lambda_3} + \frac{|x_1|^2}{\lambda_4} \end{pmatrix}. \tag{4.51}$$

We obtain $\mathrm{tr}\left(((\langle x| \otimes \mathbb{1}_{\mathbb{C}^2}) R^{-1} (|x\rangle \otimes \mathbb{1}_{\mathbb{C}^2})\right) = \frac{1}{2}\left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} + \frac{1}{\lambda_4}\right) \|x\|_2^2 = \frac{1}{2}\mathrm{tr}(R^{-1})\|x\|_2^2$. Hence, for any $\rho \in \mathcal{D}(\mathbb{C}^2)$ we have $\mathrm{tr}\left(R^{-1}(\rho \otimes \mathbb{1}_{\mathbb{C}^2})\right) = \frac{1}{2}\mathrm{tr}(R^{-1})$. Finally, we obtain the following upper bound

$$\|\mathrm{tr}_1\left(R^{-1}(\rho \otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1} \leq 2\left(\mathrm{tr}\left(R^{-1}(\rho \otimes \mathbb{1}_{\mathbb{C}^2})\right)\right)^{-1} = 4\left(\mathrm{tr}(R^{-1})\right)^{-1}. \tag{4.52}$$

That means, $p_0(R) \leq 4\left(\mathrm{tr}(R^{-1})\right)^{-1}$. To saturate this bound, we take the maximally mixed state $\rho = \rho^*_{\mathbb{C}^2}$ and by using Eq. (4.51) we calculate

$$\|\mathrm{tr}_1\left(R^{-1}(\rho^*_{\mathbb{C}^2} \otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1} = 2\|\mathrm{tr}_1\left(R^{-1}\right)\|_\infty^{-1} = 2\left\|\frac{1}{2}\mathrm{tr}(R^{-1})\mathbb{1}_{\mathbb{C}^2}\right\|_\infty^{-1} = 4\left(\mathrm{tr}(R^{-1})\right)^{-1}. \tag{4.53}$$

In the case of $p_1(R)$, to calculate the largest eigenvalue of $\mathrm{tr}_1\left(R^{-1}(|x\rangle\langle x| \otimes \mathbb{1}_{\mathbb{C}^2})\right)$ we use Eq. (4.51) for $|x\rangle = (|x_0|, |x_1|\alpha)^\top$, such that $|x_0|^2 + |x_1|^2 = 1$ and $|\alpha| = 1$. One may calculate that the largest eigenvalue minimized over α is given by

$$\begin{aligned} \frac{1}{4}\Bigg(\mathrm{tr}(R^{-1}) + \Bigg[&\left(\left(\frac{1}{\lambda_1} + \frac{1}{\lambda_4}\right) - \left(\frac{1}{\lambda_2} + \frac{1}{\lambda_3}\right)\right)^2 (|x_0|^2 - |x_1|^2)^2 \\ &+4\left(\left|\frac{1}{\lambda_1} - \frac{1}{\lambda_4}\right| - \left|\frac{1}{\lambda_2} - \frac{1}{\lambda_3}\right|\right)^2 |x_0|^2|x_1|^2\Bigg]^{1/2}\Bigg). \end{aligned} \tag{4.54}$$

It turns out, there are only two situations when this expression is minimized:

- For $|x_0| = 0$ and $|x_1| = 1$ (or equivalently $|x_0| = 1$ and $|x_1| = 0$), we obtain

$$\frac{1}{4}\left(\operatorname{tr}(R^{-1}) + \left|\frac{1}{\lambda_1} - \frac{1}{\lambda_2} - \frac{1}{\lambda_3} + \frac{1}{\lambda_4}\right|\right). \tag{4.55}$$

- For $|x_0| = |x_1| = \frac{1}{\sqrt{2}}$, we obtain

$$\frac{1}{4}\left(\operatorname{tr}(R^{-1}) + \left|\left|\frac{1}{\lambda_1} - \frac{1}{\lambda_4}\right| - \left|\frac{1}{\lambda_2} - \frac{1}{\lambda_3}\right|\right|\right). \tag{4.56}$$

Hence, the optimal value $p_1(R)$ equals

$$p_1(R) = \frac{4}{\operatorname{tr}(R^{-1}) + \min\left\{\left|\frac{1}{\lambda_1} - \frac{1}{\lambda_2} - \frac{1}{\lambda_3} + \frac{1}{\lambda_4}\right|, \left|\left|\frac{1}{\lambda_1} - \frac{1}{\lambda_4}\right| - \left|\frac{1}{\lambda_2} - \frac{1}{\lambda_3}\right|\right|\right\}}. \tag{4.57}$$

□

Remark 4.15. *Let R be defined as in Proposition 4.14 with parameter $\lambda = (\lambda_1, \ldots, \lambda_4) \in (0,1]^4$. Then, the pure state encoding match the mixed state encoding, $p_0(R) = p_1(R)$ if and only if*

$$\frac{1}{\lambda_1} + \frac{1}{\lambda_4} = \frac{1}{\lambda_2} + \frac{1}{\lambda_3} \vee \left|\frac{1}{\lambda_1} - \frac{1}{\lambda_4}\right| = \left|\frac{1}{\lambda_2} - \frac{1}{\lambda_3}\right|, \tag{4.58}$$

which is the $3-$dimensional subset of $(0,1]^4$. For all other values of λ we have the advantage of mixed state encoding strategy, $p_0(R) > p_1(R)$.

In an extremal case, e.g. *for $\lambda = (\frac{1}{2N}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$, $N \in \mathbb{N}$, we get $p_1(R) = \frac{1}{N+1}$ and $p_0(R) = \frac{2}{N+3}$. Especially, when $N \to \infty$ the mixed state encoding provides the advantage, $p_0(R)/p_1(R) \to 2$.*

Moreover, $\mathcal{E}_R \in \xi_1(\mathbb{C}^2, \mathbb{C}^4)$ if and only if $R = \mathbb{1}_{\mathbb{C}^4}$.

Finally, let us discuss the one-to-one relation of $p_0(R)$ with the maximum value of the output min-entropy (see for instance [95]).

Corollary 4.16. *Let us take $\Phi \in \mathcal{C}(\mathbb{C}^2)$, such that $\operatorname{rank}(J(\Phi)) = 4$ and a constant $c > 0$, such that $cJ(\Phi) \geq \mathbb{1}_{\mathbb{C}^4}$. For $R \in \mathcal{P}(\mathbb{C}^4)$ defined as $R = \left[c(\mathcal{I}_{\mathbb{C}^2} \otimes \Phi)(|\mathbb{1}_{\mathbb{C}^2}\rangle\rangle\langle\langle\mathbb{1}_{\mathbb{C}^2}|)\right]^{-1}$ and the noise channel $\mathcal{E}_R$ defined in Eq. (4.42) we have*

$$p_0(R) = \exp\left(S_{\min}(c\Phi)\right), \tag{4.59}$$

where $S_{\min}(\Psi) = \max\{-\ln(\|\Psi(\rho)\|_\infty) : \rho \in \mathcal{D}\}$ is the maximum value of the output min-entropy.

In particular, $p_0(R) > p_1(R)$ if and only if the maximally mixed state $\rho^*_{\mathbb{C}^2}$ belongs to the interior of the image of Φ taken over $\mathcal{D}(\mathbb{C}^2)$. Here, the interior is taken with respect to the subspace generated by Bloch ball [96].

Proof. The first claim follows from Corollary 4.13

$$\begin{aligned}\exp\left(S_{\min}(c\Phi)\right) &= \max\left\{\exp\left(-\ln\|c\Phi(\rho)\|_\infty\right) : \rho \in \mathcal{D}(\mathbb{C}^2)\right\} \\ &= \max\left\{\|c\Phi(\rho)\|_\infty^{-1} : \rho \in \mathcal{D}(\mathbb{C}^2)\right\} \\ &= \max\left\{\|\mathrm{tr}_1(c(\mathcal{I}_{\mathbb{C}^2} \otimes \Phi)(|\mathbb{1}_{\mathbb{C}^2}\rangle\rangle\langle\langle\mathbb{1}_{\mathbb{C}^2}|)(\rho \otimes \mathbb{1}_{\mathbb{C}^2}))\|_\infty^{-1} : \rho \in \mathcal{D}(\mathbb{C}^2)\right\} \\ &= \max\left\{\left\|\mathrm{tr}_1(R^{-1}(\rho \otimes \mathbb{1}_{\mathbb{C}^2}))\right\|_\infty^{-1} : \rho \in \mathcal{D}(\mathbb{C}^2)\right\} = p_0(R).\end{aligned} \tag{4.60}$$

From the first part it follows that to calculate $p_0(R)$ (also $p_1(R)$ – see Corollary 4.13) we need to minimize $\|\Phi(\rho)\|_\infty$ over $\rho \in \mathcal{D}(\mathbb{C}^2)$ (or $\rho = |\psi\rangle\langle\psi| \in \mathcal{D}(\mathbb{C}^2)$ for $p_1(R)$). Recall, that the image of Φ is an ellipse submerged into Bloch ball [96]. Moreover, the points in the interior of this set are the image of mixed states and the border is the image taken over pure states (sometimes mixed states are also transferred into border but they do not add new points to the border). Hence, if $\rho^*_{\mathbb{C}^2}$ belongs to the interior, then $p_0(R) > p_1(R)$. On the other hand, if $\rho^*_{\mathbb{C}^2}$ is not in the interior ($\rho^*_{\mathbb{C}^2}$ may be located in the border or outside the image) then the function $\rho \mapsto \|\Phi(\rho)\|_\infty$ is minimized for some pure state ρ. □

4.5 Advantage of pQEC procedure

Perfectly correctable noise channels constitute only a small subset of probabilistically correctable ones. This behavior will be the object of our investigation in this section. We begin our analysis with some basic observations concerning $\xi(\mathcal{X},\mathcal{Y})$ and $\xi_1(\mathcal{X},\mathcal{Y})$ (see Eq. (4.2)).

Proposition 4.17 ([2]). *For any $\mathcal{X}$, $\mathcal{Y}$ we have the following properties:*

(A) $\xi_1(\mathcal{X},\mathcal{Y}) \subset \xi(\mathcal{X},\mathcal{Y})$,

(B) If $\dim(\mathcal{X}) > \dim(\mathcal{Y})$, *then* $\xi(\mathcal{X},\mathcal{Y}) = \emptyset$,

(C) If $\dim(\mathcal{X}) \leq \dim(\mathcal{Y})$, *then* $\xi_1(\mathcal{X},\mathcal{Y}) \neq \emptyset$,

(D) If $\dim(\mathcal{X}) = \dim(\mathcal{Y})$, *then* $\xi_1(\mathcal{X},\mathcal{Y}) = \xi(\mathcal{X},\mathcal{Y})$.

Proof. (*D*)
Let us take $\mathcal{E} = \mathcal{K}\left((E_i)_i\right) \in \xi(\mathcal{X},\mathcal{Y})$. From Theorem 4.4 there exist $S_* \in \mathcal{M}(\mathcal{X},\mathcal{Y})$ and $R_* \in \mathcal{M}(\mathcal{Y},\mathcal{X})$ such that $R_* E_i S_* \propto \mathbb{1}_\mathcal{X}$, and there exists i_0 for which it holds $R_* E_{i_0} S_* \neq 0$. It implies that R_* and S_* are invertible, so for all i we have $E_i \propto R_*^{-1} S_*^{-1}$. Hence, $\mathrm{rank}(J(\mathcal{E})) = 1$, so we can write $\mathcal{E}(X) = EXE^\dagger$, for $E \in \mathcal{U}(\mathcal{X})$. By taking $\mathcal{R} = \mathcal{I}_\mathcal{X}$ and $\mathcal{S} = \mathcal{E}^\dagger$ we get $\mathcal{E} \in \xi_1(\mathcal{X},\mathcal{Y})$.

□

We see that if $\dim(\mathcal{X}) = \dim(\mathcal{Y})$, then there is no need to consider the pQEC procedure. The situation changes if we encode the initial information into a larger space, $\dim(\mathcal{Y}) > \dim(\mathcal{X})$. In the following proposition, we will show that $\xi_1(\mathcal{X},\mathcal{Y}) \subsetneq \xi(\mathcal{X},\mathcal{Y})$ for $\dim(\mathcal{Y}) > \dim(\mathcal{X})$.

Proposition 4.18 ([2]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be Euclidean spaces for which* $\dim(\mathcal{X}) < \dim(\mathcal{Y})$. *Then, the set $\xi_1(\mathcal{X},\mathcal{Y})$ is a nowhere dense subset of $\xi(\mathcal{X},\mathcal{Y})$.*

Proof. First, we will prove that $\xi_1(\mathcal{X},\mathcal{Y})$ is a closed set. Define a sequence $(\mathcal{E}_n)_{n\in\mathbb{N}} \subset \xi_1(\mathcal{X},\mathcal{Y})$ that converges to $\mathcal{E} = \lim_{n\to\infty} \mathcal{E}_n \in \mathcal{C}(\mathcal{Y})$. From Proposition 4.10 there exist two sequences $(\mathcal{S}_n)_{n\in\mathbb{N}} \subset \mathcal{C}(\mathcal{X},\mathcal{Y})$ and $(\mathcal{R}_n)_{n\in\mathbb{N}} \subset \mathcal{C}(\mathcal{Y},\mathcal{X})$ such that $\mathcal{R}_n\mathcal{E}_n\mathcal{S}_n = \mathcal{I}_\mathcal{X}$ for $n \in \mathbb{N}$. Both sets $\mathcal{C}(\mathcal{X},\mathcal{Y})$ and $\mathcal{C}(\mathcal{Y},\mathcal{X})$ are compact, so there exists a subsequence $(n_k)_{k\in\mathbb{N}}$, such that $(\mathcal{S}_{n_k})_{k\in\mathbb{N}}$, $(\mathcal{R}_{n_k})_{k\in\mathbb{N}}$ converge to some $\mathcal{S} \in \mathcal{C}(\mathcal{X},\mathcal{Y}), \mathcal{R} \in \mathcal{C}(\mathcal{Y},\mathcal{X})$, respectively. Hence, we obtain $\mathcal{R}\mathcal{E}\mathcal{S} = \lim_{k\to\infty} \mathcal{R}_{n_k}\mathcal{E}_{n_k}\mathcal{S}_{n_k} = \mathcal{I}_\mathcal{X}$. That ends this part of the proof.

To show that $\xi_1(\mathcal{X},\mathcal{Y})$ is nowhere dense in $\xi(\mathcal{X},\mathcal{Y})$, it is enough to prove $\operatorname{int}(\xi_1(\mathcal{X},\mathcal{Y})) = \emptyset$. Therefore, for any $\mathcal{E} \in \xi_1(\mathcal{X},\mathcal{Y})$ we will construct a sequence of channels $(\mathcal{E}_n)_{n\in\mathbb{N}} \subset \mathcal{C}(\mathcal{Y})$ that converges to $\mathcal{E}$ and for which $\mathcal{E}_n \in \xi(\mathcal{X},\mathcal{Y})$, and $\mathcal{E}_n \notin \xi_1(\mathcal{X},\mathcal{Y})$, for $n \in \mathbb{N}$.

Fix $\mathcal{E} \in \xi_1(\mathcal{X},\mathcal{Y})$. From Proposition 4.10 there exist $\mathcal{S} = \mathcal{K}((S)) \in \mathcal{C}(\mathcal{X},\mathcal{Y})$ and $\mathcal{R} \in \mathcal{C}(\mathcal{Y},\mathcal{X})$ such that $\mathcal{R}\mathcal{E}\mathcal{S} = \mathcal{I}_\mathcal{X}$. From Theorem 4.4 we have

$$\mathcal{E}\mathcal{S} = \mathcal{K}((A_i)_i): \quad A_i \neq 0, A_j^\dagger A_i \propto \delta_{ij}\mathbb{1}_\mathcal{X}. \tag{4.61}$$

As $\dim(\mathcal{X}) < \dim(\mathcal{Y})$, there exists $|y\rangle\langle y| \in \mathcal{D}(\mathcal{Y})$ such that $\langle y|A_1 = 0$. Let us define a sequence of channels $\mathcal{E}_n \in \mathcal{C}(\mathcal{Y})$ given by

$$\mathcal{E}_n(Y) = \frac{n}{n+1}\mathcal{E}(Y) + \frac{\operatorname{tr}(Y)}{n+1}|y\rangle\langle y|. \tag{4.62}$$

One can note that $\lim_{n\to\infty} \mathcal{E}_n = \mathcal{E}$. We take $\mathcal{S}_n = \mathcal{S}$ and $\mathcal{R}_n = \mathcal{K}\left((A_1^\dagger)\right)$ for $n \in \mathbb{N}$ and obtain

$$\mathcal{R}_n\mathcal{E}_n\mathcal{S}_n(X) = \frac{n}{n+1}A_1^\dagger\mathcal{E}\mathcal{S}(X)A_1 = \frac{n}{n+1}\|A_1\|_\infty^4 X. \tag{4.63}$$

As $A_1 \neq 0$, it follows that $\mathcal{E}_n \in \xi(\mathcal{X},\mathcal{Y})$. Now, for each $n \in \mathbb{N}$, let $\widetilde{\mathcal{S}}_n \in \mathcal{C}(\mathcal{X},\mathcal{Y})$ and $\widetilde{\mathcal{R}}_n \in s\mathcal{C}(\mathcal{Y},\mathcal{X})$ be arbitrary maps satisfying $0 \neq \widetilde{\mathcal{R}}_n\mathcal{E}_n\widetilde{\mathcal{S}}_n \propto \mathcal{I}_\mathcal{X}$. It holds that $\widetilde{\mathcal{R}}_n(|y\rangle\langle y|) = 0$. Finally, for any $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{X})$ we have

$$\operatorname{tr}\left(\widetilde{\mathcal{R}}_n\mathcal{E}_n\widetilde{\mathcal{S}}_n(|\psi\rangle\langle\psi|)\right) = \frac{n}{n+1}\operatorname{tr}\left(\widetilde{\mathcal{R}}_n\mathcal{E}\widetilde{\mathcal{S}}_n(|\psi\rangle\langle\psi|)\right) \leq \frac{n}{n+1}. \tag{4.64}$$

Hence, we obtain $\mathcal{E}_n \notin \xi_1(\mathcal{X},\mathcal{Y})$. □

4.5.1 Choi rank of correctable noise channels

The intensity of a noise channel $\mathcal{E}$ can be connected with its Choi rank $r = \operatorname{rank}(J(\mathcal{E}))$. Given $\mathcal{E}$ in the Stinespring form, the Choi rank describes the dimension of an environment system which unitarily interacts with the encoded information. If the interaction is the weakest ($r = 1$) we deal with unitary noise channels, which are always perfectly correctable. The strongest interaction ($r = \dim(\mathcal{Y})^2$) is a property of noise channels that are difficult to correct. For example, the maximally depolarizing channel $\mathcal{E}(Y) = \operatorname{tr}(Y)\rho^*_{\mathcal{Y}}$, which can not be corrected, has the maximal Choi rank.

Proposition 4.19 ([2]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be some Euclidean spaces such that* $\dim(\mathcal{Y}) \geq \dim(\mathcal{X})$. *The following relations hold:*

$$\begin{aligned} &(A) \quad \max\{\operatorname{rank}(J(\mathcal{E})) : \mathcal{E} \in \xi_1(\mathcal{X},\mathcal{Y})\} &&= \dim(\mathcal{Y})^2 - \dim(\mathcal{Y})\dim(\mathcal{X}) + \left\lfloor \tfrac{\dim(\mathcal{Y})}{\dim(\mathcal{X})} \right\rfloor, \\ &(B) \quad \max\{\operatorname{rank}(J(\mathcal{E})) : \mathcal{E} \in \xi(\mathcal{X},\mathcal{Y})\} &&= \dim(\mathcal{Y})^2 - \dim(\mathcal{X})^2 + 1. \end{aligned} \tag{4.65}$$

Proof. Let us define $d = \dim(\mathcal{X})$, $s = \dim(\mathcal{Y})$ and $k = \left\lfloor \frac{s}{d} \right\rfloor$.

(A)
Take $\mathcal{E} = \mathcal{K}\left((E_i)_{i=1}^r\right) \in \xi_1(\mathcal{X},\mathcal{Y})$, where $r = \operatorname{rank}(J(\mathcal{E}))$. From Proposition 4.10 there exist $\mathcal{S} = \mathcal{K}((S)) \in \mathcal{C}(\mathcal{X},\mathcal{Y})$ and $\mathcal{R} \in \mathcal{C}(\mathcal{Y},\mathcal{X})$ such that $\mathcal{R}\mathcal{E}\mathcal{S} = \mathcal{I}_{\mathcal{X}}$. According to Theorem 4.4 it holds

$$\mathcal{K}\left((E_i S)_{i=1}^r\right) = \mathcal{K}\left((A_i)_{i=1}^{r'}\right): \quad A_i \neq 0, A_j^\dagger A_i \propto \delta_{ij}\mathbb{1}_{\mathcal{X}}. \tag{4.66}$$

If $r' < r$, then let us define $A_i = 0$ for $i = r'+1,\ldots,r$. As $\mathcal{K}\left((A_i)_{i=1}^{r'}\right) = \mathcal{K}\left((A_i)_{i=1}^r\right)$, there exists a Kraus decomposition $\mathcal{E} = \mathcal{K}\left((E'_i)_{i=1}^r\right)$ such that $A_i = E'_i S$ for each $i \leq r$. For $A_i \neq 0$ images of A_i are orthogonal and $\operatorname{rank}(A_i) = d$. Hence, $r'd \leq s$ which is equivalent to $r' \leq k$. For $i > r'$ it holds that $\left(\mathbb{1}_{\mathcal{Y}} \otimes S^\top\right)|E'_i\rangle\rangle = 0$. Note that the Kraus operators E'_i are linearly independent and it holds

$$\dim(\ker(\mathbb{1}_{\mathcal{Y}} \otimes S^\top)) = s^2 - \operatorname{rank}(\mathbb{1}_{\mathcal{Y}} \otimes S^\top) = s^2 - \operatorname{rank}(\mathbb{1}_{\mathcal{Y}})\operatorname{rank}(S) = s^2 - sd. \tag{4.67}$$

Therefore, we get $r - r' = \dim(\operatorname{span}(E'_i, i > r')) \leq \dim(\ker(\mathbb{1}_{\mathcal{Y}} \otimes S^\top)) = s^2 - sd$ and eventually $r \leq s^2 - sd + k$. To saturate this bound, let us define $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ given by

$$\mathcal{E}(Y) = \sum_{i=0}^{k-1} E_i Y E_i^\dagger + \operatorname{tr}\left((\mathbb{1}_{\mathcal{Y}} - \Pi)Y\right)\rho^*_{\mathcal{Y}}, \tag{4.68}$$

where

$$E_i = \frac{1}{\sqrt{k}} \sum_{j=0}^{d-1} |j + id\rangle\langle j| \in \mathcal{M}(\mathcal{Y}), \qquad \text{for } i = 0, \ldots, k-1,$$
$$\Pi = \sum_{j=0}^{d-1} |j\rangle\langle j| \in \mathcal{P}(\mathcal{Y}). \tag{4.69}$$

Note that $\Pi = \sum_{i=0}^{k-1} E_i^\dagger E_i$ and $(\mathbb{1}_\mathcal{Y} \otimes \Pi)|E_i\rangle\rangle = |E_i\rangle\rangle$. Therefore, we obtain

$$\operatorname{rank}(J(\mathcal{E})) = \operatorname{rank}\left(\sum_{i=0}^{k-1} |E_i\rangle\rangle\langle\langle E_i| + \rho_\mathcal{Y}^* \otimes (\mathbb{1}_\mathcal{Y} - \Pi)\right) = \operatorname{rank}\left(\sum_{i=0}^{k-1} |E_i\rangle\rangle\langle\langle E_i|\right) + \operatorname{rank}\left(\rho_\mathcal{Y}^* \otimes (\mathbb{1}_\mathcal{Y} - \Pi)\right) = s^2 - sd + k. \tag{4.70}$$

Finally, let us define $\mathcal{S} = \mathcal{K}\left((S)\right) \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$, where $S = \sum_{j=0}^{d-1} |j\rangle_\mathcal{Y}\langle j|_\mathcal{X}$, and $\mathcal{R} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ given by $\mathcal{R}(Y) = kS^\dagger \left(\sum_{i=0}^{k-1} E_i^\dagger Y E_i\right) S$. We can observe that $\mathcal{R}\mathcal{E}\mathcal{S} = \mathcal{I}_\mathcal{X}$, so $\mathcal{E} \in \xi_1(\mathcal{X}, \mathcal{Y})$.

(B)
Take $\mathcal{E} = \mathcal{K}\left((E_i)_{i=1}^r\right) \in \xi(\mathcal{X}, \mathcal{Y})$, where $r = \operatorname{rank}(J(\mathcal{E}))$. According to Theorem 4.4 there exist $S_* \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ and $R_* \in \mathcal{M}(\mathcal{Y}, \mathcal{X})$ such that $R_* E_i S_* \propto \mathbb{1}_\mathcal{X}$, and there exists i_0 for which it holds $R_* E_{i_0} S_* \neq 0$. We may assume that $\|R_*\|_\infty \leq 1$ and $\|S_*\|_\infty \leq 1$. Hence, according to Theorem 4.4 we get

$$\mathcal{K}\left(\left(\sqrt{R_*^\dagger R_*} E_i S_*\right)_{i=1}^r\right) = \mathcal{K}\left((A_i)_{i=1}^{r'}\right) : \quad A_i \neq 0, A_j^\dagger A_i \propto \delta_{ij} \mathbb{1}_\mathcal{X}. \tag{4.71}$$

If $r' < r$, then let us define $A_i = 0$ for $i = r'+1, \ldots, r$. As $\mathcal{K}\left((A_i)_{i=1}^{r'}\right) = \mathcal{K}\left((A_i)_{i=1}^r\right)$, there exists a Kraus decomposition $\mathcal{E} = \mathcal{K}\left((E_i')_{i=1}^r\right)$ such that $A_i = \sqrt{R_*^\dagger R_*} E_i' S_*$ for each $i \leq r$. Let Π be the projector on the support of $R_*^\dagger R_*$. Observe that $\operatorname{rank}(\Pi) = d$. Then, for each $i \leq r$ we have $\Pi A_i = A_i$ and for $i \leq r'$ we have $\operatorname{rank}(A_i) = d$. The relation $A_j^\dagger A_i \propto \delta_{ij} \mathbb{1}_\mathcal{X}$ implies that there exists exactly one $A_i \neq 0$, hence $r' = 1$. For $i > 1$ we have $\left(\sqrt{R_*^\dagger R_*} \otimes S_*^\top\right) |E_i'\rangle\rangle = 0$. Note that the Kraus operators E_i' are linearly independent and it holds

$$\dim\left(\ker\left(\sqrt{R_*^\dagger R_*} \otimes S_*^\top\right)\right) = s^2 - \operatorname{rank}\left(\sqrt{R_*^\dagger R_*} \otimes S_*^\top\right) = s^2 - d^2. \tag{4.72}$$

Therefore, we obtain $r-1 = \dim(\mathrm{span}(E'_i, i>1)) \leq \dim\left(\ker\left(\sqrt{R_*^\dagger R_*} \otimes S_*^\top\right)\right) = s^2 - d^2$ and eventually $r \leq s^2 - d^2 + 1$. To saturate this bound, we define $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ given by

$$\mathcal{E}(Y) = \frac{\Pi Y \Pi + \operatorname{tr}(\Pi Y)(\mathbb{1}_\mathcal{Y} - \Pi)}{s-d+1} + \operatorname{tr}((\mathbb{1}_\mathcal{Y} - \Pi)Y)\,\rho_\mathcal{Y}^*, \tag{4.73}$$

where $\Pi = \sum_{j=0}^{d-1} |j\rangle\langle j| \in \mathcal{P}(\mathcal{Y})$. Note, that

$$\begin{aligned}\operatorname{rank}(J(\mathcal{E})) &= \operatorname{rank}\left(\frac{1}{s-d+1}(|\Pi\rangle\rangle\langle\langle\Pi| + (\mathbb{1}_\mathcal{Y} - \Pi)\otimes\Pi) + \rho_\mathcal{Y}^* \otimes (\mathbb{1}_\mathcal{Y} - \Pi)\right)\\ &= \operatorname{rank}(|\Pi\rangle\rangle\langle\langle\Pi|) + \operatorname{rank}((\mathbb{1}_\mathcal{Y} - \Pi)\otimes\Pi) + \operatorname{rank}(\rho_\mathcal{Y}^* \otimes (\mathbb{1}_\mathcal{Y} - \Pi))\\ &= s^2 - d^2 + 1.\end{aligned} \tag{4.74}$$

Define $\mathcal{S} = \mathcal{K}((S)) \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$, where $S = \sum_{j=0}^{d-1} |j\rangle_\mathcal{Y}\langle j|_\mathcal{X}$ and $\mathcal{R} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ given by $\mathcal{R}(Y) = S^\dagger Y S$. We can observe that $\mathcal{R}\mathcal{E}\mathcal{S} = \frac{\mathcal{I}_\mathcal{X}}{s-d+1}$, so $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$. □

In Proposition 4.17 we showed that if $\dim(\mathcal{X}) = \dim(\mathcal{Y})$, then the the pQEC procedure gives us no advantage. Indeed, the only reversible noise channels, in this case, are unitary noise channels, that is channels with the Choi rank equal to one. We can ask, what is the maximum value of $r \in \mathbb{N}$, such that all noise channels which Choi rank is less or equal r, can be corrected perfectly or probabilistically, respectively. Formally speaking, for any $\mathcal{X}$ and $\mathcal{Y}$ we define the following quantities:

$$\begin{aligned} r_1(\mathcal{X}, \mathcal{Y}) &:= \max\left\{r \in \mathbb{N} : \quad \forall_{\mathcal{E}\in\mathcal{C}(\mathcal{Y})}\ \operatorname{rank}(J(\mathcal{E})) \leq r \implies \mathcal{E} \in \xi_1(\mathcal{X}, \mathcal{Y})\right\},\\ r(\mathcal{X}, \mathcal{Y}) &:= \max\left\{r \in \mathbb{N} : \quad \forall_{\mathcal{E}\in\mathcal{C}(\mathcal{Y})}\ \operatorname{rank}(J(\mathcal{E})) \leq r \implies \mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})\right\}.\end{aligned} \tag{4.75}$$

In the next sections we will investigate the behavior of r_1 and r. It would be useful to note the monotonicity of $r(\mathcal{X}, \mathcal{Y})$ w.r.t. the dimension of $\mathcal{Y}$, which follows from Lemma 4.7. Let $\mathcal{Y}, \mathcal{Y}'$ be such Euclidean spaces that $\dim(\mathcal{Y}) \leq \dim(\mathcal{Y}')$. Take $\mathcal{E} = \mathcal{K}((E_i)_i) \in \mathcal{C}(\mathcal{Y}')$. There exist two projectors $\Pi_1, \Pi_2 \in \mathcal{P}(\mathcal{Y}')$, such that $\operatorname{rank}(\Pi_1) = \operatorname{rank}(\Pi_2) = \dim(\mathcal{Y})$ and for $\mathcal{F} = \mathcal{K}((\Pi_2 E_i \Pi_1)_i)$ we have $\operatorname{rank}(\operatorname{tr}_1(J(\mathcal{F}))) = \dim(\mathcal{Y})$. To construct Π_2 observe that

$$\int \operatorname{tr}(\mathcal{E}^\dagger(U\Pi_0 U^\dagger))dU = \operatorname{tr}\left(\mathcal{E}^\dagger\left(\frac{\dim(\mathcal{Y})}{\dim(\mathcal{Y}')}\mathbb{1}_{\mathcal{Y}'}\right)\right) = \dim(\mathcal{Y}), \tag{4.76}$$

where $\operatorname{rank}(\Pi_0) = \dim(\mathcal{Y})$ and the integral is taken over Haar-distributed unitary matrices $U \in \mathcal{U}(\mathcal{Y}')$. As $\mathcal{E}^\dagger(U\Pi_0 U^\dagger) \leq \mathbb{1}_{\mathcal{Y}'}$ there exists Π_2 that $\operatorname{rank}(\mathcal{E}^\dagger(\Pi_2)) \geq \dim(\mathcal{Y})$. The construction of Π_1 follows naturally.

Hence, if there exists a scheme $(\mathcal{S}, \mathcal{R})$ such that $0 \neq \mathcal{R}\mathcal{F}\mathcal{S} \propto \mathcal{I}_\mathcal{X}$, then $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y}')$, which gives

$$r(\mathcal{X}, \mathcal{Y}) \leq r(\mathcal{X}, \mathcal{Y}'). \tag{4.77}$$

4.5.2 From bi-linear to linear problem

Let us now consider a particular class of noise channels $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ satisfying $\operatorname{rank}(\mathcal{E}(\mathbb{1}_\mathcal{Y})) = \dim(\mathcal{X})$. For each such channel $\mathcal{E} = \mathcal{K}\left((E_i)_i\right)$, we may consider an associated channel $\mathcal{F} = \mathcal{K}\left((V^\dagger E_i)_i\right) \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$, where $V \in \mathcal{U}(\mathcal{X}, \mathcal{Y})$ is an isometry operator with the image on the support of $\mathcal{E}(\mathbb{1}_\mathcal{Y})$. It turns out that for $\mathcal{F} \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$ we can easily exploit Lemma 4.8.

Corollary 4.20 ([2]). *Let $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{r-1}\right) \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$. Define $\Pi_F = FF^{-1}$, where $F = \sum_{i=0}^{r-1} |i\rangle \otimes F_i \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$. Then, it holds $p_\mathcal{X}(\mathcal{F}) \in \{0, 1\}$. Moreover, $\mathcal{F}$ is perfectly correctable for $\mathcal{X}$ if and only if there exists $0 \neq |\psi\rangle \in \mathbb{C}^r$ such that $(\Pi_F \otimes \mathbb{1}_\mathcal{X})(|\psi\rangle \otimes |\mathbb{1}_\mathcal{X}\rangle\rangle) = |\psi\rangle \otimes |\mathbb{1}_\mathcal{X}\rangle\rangle$.*

Proof. Let us assume that for a given $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{r-1}\right) \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$ there exists error-correcting scheme $(\mathcal{S}, \mathcal{R}) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y}) \times s\mathcal{C}(\mathcal{X})$ such that $\mathcal{R}\mathcal{F}\mathcal{S} = p\mathcal{I}_\mathcal{X} \neq 0$. From Lemma 4.8, we may take $\mathcal{R} = \mathcal{I}_\mathcal{X}$. Hence, from Proposition 4.10 we have $p_\mathcal{X}(\mathcal{F}) = 1$. Now, from Lemma 4.8 we know that $p_\mathcal{X}(\mathcal{F}) > 0$ if and only if there exists $0 \neq P \in \mathcal{P}(\mathbb{C}^r)$ such that $(\Pi_F \otimes \mathbb{1}_\mathcal{X})(P \otimes |\mathbb{1}_\mathcal{X}\rangle\rangle\langle\langle\mathbb{1}_\mathcal{X}|)(\Pi_F \otimes \mathbb{1}_\mathcal{X}) = P \otimes |\mathbb{1}_\mathcal{X}\rangle\rangle\langle\langle\mathbb{1}_\mathcal{X}|$. This condition is equivalent to $(\Pi_F \otimes \mathbb{1}_\mathcal{X})(|\psi\rangle \otimes |\mathbb{1}_\mathcal{X}\rangle\rangle) = |\psi\rangle \otimes |\mathbb{1}_\mathcal{X}\rangle\rangle$ for some $0 \neq |\psi\rangle \in \mathbb{C}^r$. □

Proposition 4.21 ([2]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be some complex Euclidean spaces and $\dim(\mathcal{X}) \leq \dim(\mathcal{Y})$.*

(A) If $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ is a noise channel such that $\operatorname{rank}(\mathcal{E}(\mathbb{1}_\mathcal{Y})) = \dim(\mathcal{X})$ and $\operatorname{rank}(J(\mathcal{E})) < \frac{\dim(\mathcal{Y})\dim(\mathcal{X})}{\dim(\mathcal{X})^2-1}$, then $\mathcal{E} \in \xi_1(\mathcal{X}, \mathcal{Y})$.

(B) There exists a noise channel $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ such that $\operatorname{rank}(\mathcal{E}(\mathbb{1}_\mathcal{Y})) = \dim(\mathcal{X})$ and $\operatorname{rank}(J(\mathcal{E})) \geq \frac{\dim(\mathcal{Y})\dim(\mathcal{X})}{\dim(\mathcal{X})^2-1}$, for which we have $\mathcal{E} \notin \xi(\mathcal{X}, \mathcal{Y})$.

Proof. (*A*)
Let us take $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ and denote $r = \operatorname{rank}(J(\mathcal{E}))$. Assume that $\operatorname{rank}(\mathcal{E}(\mathbb{1}_\mathcal{Y})) = \dim(\mathcal{X})$ and $r < \frac{\dim(\mathcal{Y})\dim(\mathcal{X})}{\dim(\mathcal{X})^2-1}$. Consider an associated to $\mathcal{E}$ channel $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{r-1}\right) \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$. Define $\Pi_F = FF^{-1}$, where $F = \sum_{i=0}^{r-1} |i\rangle \otimes F_i \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$. Observe that $\dim(\ker\left((\mathbb{1}_{\mathbb{C}^r\otimes\mathcal{X}} - \Pi_F) \otimes \mathbb{1}_\mathcal{X}\right)) = \dim(\mathcal{Y})\dim(\mathcal{X})$ and $\dim(\operatorname{span}(|\psi\rangle \otimes |\mathbb{1}_\mathcal{X}\rangle\rangle : |\psi\rangle \in \mathbb{C}^r)) = r$. Therefore, as $\dim(\mathcal{Y})\dim(\mathcal{X}) + r > r\dim(\mathcal{X})^2$ there exists $0 \neq |\psi\rangle \in \mathbb{C}^r$, such that $(\Pi_F \otimes \mathbb{1}_\mathcal{X})(|\psi\rangle \otimes |\mathbb{1}_\mathcal{X}\rangle\rangle) = |\psi\rangle \otimes |\mathbb{1}_\mathcal{X}\rangle\rangle$. It follows from Corollary 4.20 that $\mathcal{E} \in \xi_1(\mathcal{X}, \mathcal{Y})$.

(*B*)
Let us take $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{r-1}\right) \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$ defined as in the part (*A*) of the proof. We have that $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$ if and only if there exists $S \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ such that $F_i S = c_i \mathbb{1}_\mathcal{X}$ and $c_{i_0} \neq 0$ for some i_0. Let $F = \sum_{i=0}^{r-1} |i\rangle \otimes F_i \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$ and

$|c\rangle = \sum_{i=0}^{r-1} c_i|i\rangle$. Hence, $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$ if and only if it holds $FS = |c\rangle \otimes \mathbb{1}_{\mathcal{X}} \neq 0$. This is equivalent to

$$\mathcal{E} \notin \xi(\mathcal{X}, \mathcal{Y}) \iff ((F \otimes \mathbb{1}_{\mathcal{X}})|S\rangle\rangle = |c\rangle \otimes |\mathbb{1}_{\mathcal{X}}\rangle\rangle \implies |S\rangle\rangle = 0). \tag{4.78}$$

Therefore, in this proof, we will construct appropriate operator F. Formally, the operator F should be an isometry operator, but by Lemma 4.7, it is enough to define F such that $\operatorname{rank}(F) = \dim(\mathcal{Y})$.

Let $d = \dim(\mathcal{X}), s = \dim(\mathcal{Y})$ and fix $r \in \mathbb{N}$, such that $r \geq \frac{sd}{d^2-1}$. We start with the case $s = kd$ for $k \in \mathbb{N}$. Consider the decomposition $F = \sum_{i=0}^{r-1} |i\rangle \otimes F_i$, where $F_i \in \mathcal{M}(\mathcal{Y}, \mathcal{X})$. For $i = 0, \ldots, k-1$ we define

$$F_i = \langle i| \otimes \mathbb{1}_{\mathcal{X}}. \tag{4.79}$$

Let $\left\{\mathbb{1}_{\mathcal{X}}, (M_j)_{j=0}^{d^2-2}\right\} \subset \mathcal{M}(\mathcal{X})$ be a basis of $\mathcal{M}(\mathcal{X})$. For each $i = k, \ldots, r-1$ we define

$$F_i = \sum_{j=0}^{d^2-2} \delta(j + (i-k)(d^2-1) < k)\langle j + (i-k)(d^2-1)| \otimes M_j. \tag{4.80}$$

Observe, that $\operatorname{rank}(F) = s$. Let us take S which satisfies $F_i S \propto \mathbb{1}_{\mathcal{X}}$ for each i. Basing on the equations with indices $i = 0, \ldots, k-1$ we get $S = |c\rangle \otimes \mathbb{1}_{\mathcal{X}}$ for some $|c\rangle = \sum_{j=0}^{k-1} c_j|j\rangle$. Note, that if for any $i = k, \ldots, r-1$ it holds $F_i S \propto \mathbb{1}_{\mathcal{X}}$, then $c_j = 0$ for each $j = (i-k)(d^2-1), \ldots, d^2-2+(i-k)(d^2-1)$. From the assumption $r \geq \frac{sd}{d^2-1}$ we have $(r-k)(d^2-1) \geq k$, hence, all entries c_j are zeroed. It implies $S = 0$.

The case $s = kd + l$ for $l = 1, \ldots, d-1$ is more technically engaging than the previous case but it is based on the same idea. It will be only briefly discussed. For $i = 0, \ldots, k-1$ we can define F_i similarly as in the previous case, that is $F_i \sim \langle i| \otimes \mathbb{1}_{\mathcal{X}}$. The operator F_k has a special form, $F_k \sim (\langle k| \otimes \sum_{j=0}^{l-1} |j\rangle\langle j|) + N$, where the image of N is contained in $\operatorname{span}(|j\rangle : j \geq l)$. Here, the operator S which satisfy $F_i S \propto \mathbb{1}_{\mathcal{X}}$ has the form $S \sim |c\rangle \otimes \mathbb{1}_{\mathcal{X}}$ for some $|c\rangle = \sum_{j=0}^{k} c_j|j\rangle$. We can choose N such that $d(d-l)$ entries c_j will be zeroed if $F_k S \propto \mathbb{1}_{\mathcal{X}}$. Finally, operators F_i for $i = k+1, \ldots, r-1$ has the analogous form as Eq. (4.80) – each nullify (d^2-1) entries. In total, the number of entries c_j which can be zeroed is not less than $k+1$. Indeed, it holds

$$d(d-l) + (r-k-1)(d^2-1) \geq k+1. \tag{4.81}$$

Therefore, $S = 0$, which ends the proof. □

4.5.3 Schur noise channels

In this subsection, we restrict our attention to a particular family of noise channels whose Kraus operators are diagonal in the computational basis. In the literature, these channels are referred to as Schur channels [62, Theorem 4.19]. We use them to study an upper bound for $r(\mathcal{X},\mathcal{Y})$ and $r_1(\mathcal{X},\mathcal{Y})$.

Lemma 4.22 ([2]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be Euclidean spaces such that $\dim(\mathcal{Y}) \geq \dim(\mathcal{X})$. Then, there exists a Schur channel $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ such that $\operatorname{rank}(J(\mathcal{E})) = \left\lceil \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1} \right\rceil$ and $\mathcal{E} \notin \xi(\mathcal{X},\mathcal{Y})$. Moreover, there exists a Schur channel $\mathcal{F} \in \mathcal{C}(\mathcal{Y})$ such that $\operatorname{rank}(J(\mathcal{F})) = \left\lceil \sqrt{\left\lceil \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1} \right\rceil} \right\rceil$ and $\mathcal{F} \notin \xi_1(\mathcal{X},\mathcal{Y})$. Especially, that implies*

$$\begin{aligned} r(\mathcal{X},\mathcal{Y}) &< \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1}, \\ r_1(\mathcal{X},\mathcal{Y}) &< \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1}}. \end{aligned} \tag{4.82}$$

Proof. Let $d = \dim(\mathcal{X})$, $s = \dim(\mathcal{Y})$ and $s = k(d-1) - w$, where $k = \left\lceil \frac{s}{d-1} \right\rceil$ and $w \in \{0,\dots,d-2\}$. First, we will show that $r(\mathcal{X},\mathcal{Y}) < k$. Define a Schur channel $\mathcal{E} = \mathcal{K}\left((E_i)_{i=0}^{k-1}\right) \in \mathcal{C}(\mathcal{Y})$ given by

$$\begin{aligned} E_i &= \sum_{j=0}^{d-2} |j+(d-1)i\rangle\langle j+(d-1)i|, \quad i = 0,\dots,k-2, \\ E_{k-1} &= \sum_{j=0}^{d-2-w} |j+(d-1)(k-1)\rangle\langle j+(d-1)(k-1)|. \end{aligned} \tag{4.83}$$

Observe that $\operatorname{rank}(J(\mathcal{E})) = k$. From Theorem 4.4 we know that $\mathcal{E} \in \xi(\mathcal{X},\mathcal{Y})$ if and only if there exist $S_* \in \mathcal{M}(\mathcal{X},\mathcal{Y})$ and $R_* \in \mathcal{M}(\mathcal{Y},\mathcal{X})$, such that $R_* E_i S_* \propto \mathbb{1}_{\mathcal{X}}$ for all i and there exists i_0 for which it holds $R_* E_{i_0} S_* \neq 0$. As $\operatorname{rank}(E_i) \leq d-1$, if we have $R_* E_i S_* \propto \mathbb{1}_{\mathcal{X}}$, then $R_* E_i S_* = 0$ for all i. That implies $\mathcal{E} \notin \xi(\mathcal{X},\mathcal{Y})$.

Now, let us define $l = \left\lceil \sqrt{k} \right\rceil$. We will prove that $r_1(\mathcal{X},\mathcal{Y}) < l$. Due to the relation $\operatorname{span}_{\mathbb{C}}\left(|\psi\rangle\langle\psi| : |\psi\rangle\langle\psi| \in \mathcal{D}(\mathbb{C}^l)\right) = \mathcal{M}(\mathbb{C}^l)$, we may define unit vectors $|\psi_a\rangle$, for $a = 0,\dots,l^2-1$, such that $\operatorname{span}_{\mathbb{C}}\left(\{|\psi_a\rangle\langle\psi_a|\}_a\right) = \mathcal{M}(\mathbb{C}^l)$. Let us define $F_i \in \mathcal{M}(\mathcal{Y})$ for $i = 0,\dots,l-1$ given by

$$F_i = \sum_{a=0}^{k-1} \langle\psi_a|i\rangle E_a, \tag{4.84}$$

for E_a defined in Eq. (4.83). Observe that F_i are linearly independent. We have that

$$\sum_{i=0}^{l-1} F_i^\dagger F_i = \sum_{i=0}^{l-1}\sum_{a,b=0}^{k-1} \langle i|\psi_b\rangle\langle\psi_a|i\rangle E_b^\dagger E_a = \sum_{i=0}^{l-1}\sum_{a=0}^{k-1}\langle i|\psi_a\rangle\langle\psi_a|i\rangle E_a = \sum_{a=0}^{k-1} E_a = \mathbb{1}_\mathcal{Y}. \tag{4.85}$$

Now, we introduce a Schur channel $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{l-1}\right) \in \mathcal{C}(\mathcal{Y})$. Assume indirectly that $\mathcal{F} \in \xi_1(\mathcal{X},\mathcal{Y})$. Then, according to Proposition 4.10 and Theorem 4.4 there exists $S \in \mathcal{M}(\mathcal{X},\mathcal{Y})$, which satisfies $S^\dagger S = \mathbb{1}_\mathcal{X}$ and $M \in \mathcal{M}(\mathbb{C}^l)$, such that $S^\dagger F_j^\dagger F_i S = M_{j,i}\mathbb{1}_\mathcal{X}$. Therefore, we get

$$\begin{aligned} M\otimes\mathbb{1}_\mathcal{X} &= \sum_{j,i}|j\rangle\langle i|\otimes S^\dagger F_j^\dagger F_i S = (\mathbb{1}\otimes S^\dagger)\sum_{j,i}\left(|j\rangle\langle i|\otimes\sum_{a=0}^{k-1}\langle j|\psi_a\rangle\langle\psi_a|i\rangle E_a\right)(\mathbb{1}\otimes S) \\ &= \sum_{a=0}^{k-1}|\psi_a\rangle\langle\psi_a|\otimes S^\dagger E_a S. \end{aligned} \tag{4.86}$$

For each $a = 0,\ldots,k-1$ we can use Gram-Schmidt orthogonalization to define X_a, such that $\mathrm{tr}(X_a|\psi_a\rangle\langle\psi_a|) \neq 0$ and $\mathrm{tr}(X_a|\psi_b\rangle\langle\psi_b|) = 0$ whenever $a \neq b$. Hence, we obtain $\mathrm{tr}(X_a M)\mathbb{1}_\mathcal{X} = \mathrm{tr}(X_a|\psi_a\rangle\langle\psi_a|)S^\dagger E_a S$. As $\mathrm{rank}(E_a) \leq d-1$ we get $S^\dagger E_a S = 0$ for all a. It implies that $0 = \sum_{a=0}^{k-1} S^\dagger E_a S = S^\dagger S = \mathbb{1}_\mathcal{X}$, which gives the contradiction. That means $\mathcal{F} \notin \xi_1(\mathcal{X},\mathcal{Y})$. It is enough to observe that $r_1(\mathcal{X},\mathcal{Y}) < \mathrm{rank}(J(\mathcal{F})) = l$. □

The bounds obtained in Lemma 4.22 are asymptotically tight for Schur noise channels with $\dim(\mathcal{Y}) \to \infty$. To prove the tightness of the bound for perfectly correctable noise channels, we may use the construction provided in [61]. Hence, if we take a Schur channel $\mathcal{E} = \mathcal{K}((E_i)_i) \in \mathcal{C}(\mathcal{Y})$, such that $\mathrm{rank}(J(\mathcal{E})) \approx \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1}}$, we obtain $\mathcal{E} \in \xi_1(\mathcal{X},\mathcal{Y})$. In the following proposition we will prove the tightness for probabilistically correctable Schur noise channels.

Proposition 4.23 ([2]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be Euclidean spaces and* $\dim(\mathcal{X}) \leq \dim(\mathcal{Y})$. *For any Schur channels $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$, such that* $(\dim(\mathcal{X})-1)\mathrm{rank}(J(\mathcal{E})) < \dim(\mathcal{Y})$, *it holds $\mathcal{E} \in \xi(\mathcal{X},\mathcal{Y})$.*

Proof. Let $\Delta \in \mathcal{C}(\mathcal{Y})$ be the maximally dephasing channel, that is $\Delta(Y) = \sum_i |i\rangle\langle i|Y|i\rangle\langle i|$. Let us fix r such that $(\dim(\mathcal{X})-1)r < \dim(\mathcal{Y})$. We will show that if $\mathcal{E} = \mathcal{K}((E_i)) \in \mathcal{C}(\mathcal{Y})$, such that $E_i = \Delta(E_i)$ for each i and $\mathrm{rank}(J(\mathcal{E})) \leq r$, then $\mathcal{E} \in \xi(\mathcal{X},\mathcal{Y})$. Observe that the thesis is true in two particular situations:

- For $\dim(\mathcal{X}) = 1$ and $\dim(\mathcal{Y}) \geq 1$.

- For $r = 1$ and $\dim(\mathcal{Y}) \geq \dim(\mathcal{X})$.

Let us take $\mathcal{E} = \mathcal{K}((E_i)) \in \mathcal{C}(\mathcal{Y})$, such that $\operatorname{rank}(J(\mathcal{E})) \leq r$ and $E_i = \Delta(E_i)$ for each i. We may assume that $\operatorname{rank}(J(\mathcal{E})) = r$. Therefore, there exists a projector $\Pi \in \mathcal{P}(\mathcal{Y})$, such that $\operatorname{rank}(\Pi) = r$ and $\Delta(\Pi) = \Pi$, and for which the operators $\Pi E_i \Pi$ are linearly independent. Let us consider the map $\mathcal{F} = \mathcal{K}\left((\Pi^\perp E_i \Pi^\perp)_{i=1}^r\right)$. Define $\mathcal{X}' = \mathbb{C}^{\dim(\mathcal{X})-1}$. By the recurrence and Theorem 4.4 for $\mathcal{F}$ there exist $S'_* \in \mathcal{M}(\mathcal{X}', \mathcal{Y})$ and $R'_* \in \mathcal{M}(\mathcal{Y}, \mathcal{X}')$, such that $R'_* \Pi^\perp E_i \Pi^\perp S'_* = c_i \mathbb{1}_{\mathcal{X}'}$ and $c_{i_0} \neq 0$ for some i_0. Let $|s\rangle \in \mathcal{C}(\mathcal{Y})$ be the flat superposition. As $\Pi E_i \Pi$ are diagonal and linearly independent, there exists the vector $|r\rangle$ such that $\langle r|\Pi E_i \Pi|s\rangle = c_i$. We may define an encoding operator S_* by adding a column $\Pi|s\rangle$ to the operator $\Pi^\perp S'_*$. In the same manner, we may construct R_* by adding a row $\langle r|\Pi$ to the operator $R'_* \Pi^\perp$. It is easy to check that S_*, R_* satisfy Theorem 4.4, so $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$. □

In the case of Schur channels we have a clear separation between probabilistically and perfectly correctable noise channels.

Corollary 4.24 ([2]). *Let $\mathcal{Y}$ be an Euclidean space such that $\dim(\mathcal{Y}) \geq 2$ and let $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ be a Schur channel. Then, $\mathcal{E} \in \xi(\mathbb{C}^2, \mathcal{Y})$ if and only if $\dim(\mathcal{Y}) > \operatorname{rank}(J(\mathcal{E}))$. Moreover, if $\mathcal{E} \in \xi(\mathbb{C}^2, \mathcal{Y})$ then $p_{\mathbb{C}^2}(\mathcal{E}) \geq \frac{1}{\operatorname{rank}(J(\mathcal{E}))^2}$.*

Proof. Let $r = \operatorname{rank}(J(\mathcal{E}))$. If $\dim(\mathcal{Y}) > r$, then from Proposition 4.23 if follows $\mathcal{E} \in \xi(\mathbb{C}^2, \mathcal{Y})$. Assume now that $\dim(\mathcal{Y}) = r$. Let $\mathcal{E} = \mathcal{K}((E_i)_{i=1}^r)$ and define $M = (M_{i,j})_{i,j=1,\ldots,r}$ such that $\operatorname{rank}(M) = r$. Fix $S \in \mathcal{M}(\mathbb{C}^2, \mathcal{Y})$ and $R \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^2)$ and observe that the following conditions are equivalent:

- For all i it holds $RE_iS = c_i \mathbb{1}_{\mathbb{C}^2}$ and $c_{i_0} \neq 0$ for some i_0.
- For all i it holds $R \sum_j M_{i,j} E_j S = d_i \mathbb{1}_{\mathbb{C}^2}$ and $d_{i_0} \neq 0$ for some i_0.

Since $\mathcal{E}$ is a Schur channel we can take M such that for all i it holds $\sum_j M_{i,j} E_j = |i\rangle\langle i|$. It implies that $\mathcal{E} \notin \xi(\mathbb{C}^2, \mathcal{Y})$.

Now, we will prove that $p_{\mathbb{C}^2}(\mathcal{E}) \geq \frac{1}{\operatorname{rank}(J(\mathcal{E}))^2}$ for $\dim(\mathcal{Y}) > r$. It is enough to show this inequality for $\mathcal{Y} = \mathbb{C}^{r+1}$. Let us fix a Schur channel $\mathcal{E} = \mathcal{K}\left((E_i)_{i=0}^{r-1}\right) \in \mathcal{C}(\mathbb{C}^{r+1})$. For $i \in \{0, \ldots, r\}$ define

$$|x_i\rangle \coloneqq \sum_{j=0}^{r-1} (E_j)_{i,i} |j\rangle. \tag{4.87}$$

Observe that $\langle x_i|x_i\rangle = 1$ for all i. First we will show that there exist $i_0 \in \{0, \ldots, r\}$ and a vector $|v_{i_0}\rangle = ((v_{i_0})_i)_{i \neq i_0}$ such that $|x_{i_0}\rangle = \sum_{i \neq i_0} (v_{i_0})_i |x_i\rangle$ and $\langle v_{i_0}|v_{i_0}\rangle \leq r$. Naturally, this statement is true for $r = 1$. By induction, we assume that this statement is true for $r - 1$ and we will show that it implies its validity for r. In the first case, assume that there is $i_0 \in \{0, \ldots, r\}$ such that vectors $|x_i\rangle$ for $i \neq i_0$ are linearly dependent. That means, we have r vectors $|x_i\rangle, i \neq i_0$ which belong to a some subspace $\mathbb{C}^{r-1}$. We may use the induction step to prove the correctness of our statement. In the second case, we assume that for all i_0 the vectors $|x_i\rangle$ for $i \neq i_0$ are linearly independent. That means, for each $i_0 \in \{0, \ldots, r\}$ the vector $|x_{i_0}\rangle$ can be uniquely expressed as a linear combination of $|x_i\rangle$ for $i \neq i_0$ with coefficients $((v_{i_0})_i)_{i \neq i_0}$. Let us define $Q = \sum_{i=0}^r |x_i\rangle\langle x_i|$ and $Q_i = Q - |x_i\rangle\langle x_i| > 0$. One can show that $\langle v_i|v_i\rangle = \langle x_i|Q_i^{-1}|x_i\rangle$ for all i. We obtain

$$
\begin{aligned}
\langle v_i|v_i\rangle &= \mathrm{tr}(Q_i^{-1}|x_i\rangle\langle x_i|) = \mathrm{tr}(Q_i^{-1}(Q - Q_i)) = \mathrm{tr}(\sqrt{Q}Q_i^{-1}\sqrt{Q}) - r \\
&= \mathrm{tr}\left(\left(\sqrt{Q}^{-1}(Q - |x_i\rangle\langle x_i|)\sqrt{Q}^{-1}\right)^{-1}\right) - r \\
&= \mathrm{tr}\left(\left(\mathbb{1}_{\mathbb{C}^r} - \sqrt{Q}^{-1}|x_i\rangle\langle x_i|\sqrt{Q}^{-1}\right)^{-1}\right) - r = \frac{\langle x_i|Q^{-1}|x_i\rangle}{1 - \langle x_i|Q^{-1}|x_i\rangle}.
\end{aligned}
\tag{4.88}
$$

On the other hand, we have $\sum_{i=0}^r \langle x_i|Q^{-1}|x_i\rangle = \mathrm{tr}(Q^{-1}Q) = r$. There exists $i_0 \in \{0, \ldots, r\}$ such that $\langle x_{i_0}|Q^{-1}|x_{i_0}\rangle \leq \frac{r}{r+1}$ and hence, $\langle v_{i_0}|v_{i_0}\rangle \leq r$ which ends the proof of our statement.

Now, without loss of generality we assume that there is a vector $|v\rangle = \sum_{i=1}^r v_i|i\rangle$ such that $|x_0\rangle = \sum_{i=1}^r v_i|x_i\rangle$ and $\langle v|v\rangle \leq r$. Define $c = \max(1, \|v\|_2)$ and let us take

$$
\begin{aligned}
R &= \frac{1}{\sqrt{rc}}|0\rangle\langle 0| + \frac{1}{\sqrt{r}}|1\rangle\left(\sum_{i=1}^r \langle i|\right) \in \mathcal{M}(\mathbb{C}^{r+1}, \mathbb{C}^2), \\
S &= |0\rangle\langle 0| + \frac{1}{c}|v\rangle\langle 1| \in \mathcal{M}(\mathbb{C}^2, \mathbb{C}^{r+1}).
\end{aligned}
\tag{4.89}
$$

Observe that $\|S\|_\infty, \|R\|_\infty \leq 1$ and for any $j \in \{0, \ldots, r-1\}$ we have $RE_jS = \frac{(E_j)_{0,0}}{\sqrt{rc}}\mathbb{1}_{\mathbb{C}^2}$. Eventually, we obtain

$$
p_{\mathbb{C}^2}(\mathcal{E}) \geq \sum_{j=0}^{r-1} \frac{|(E_j)_{0,0}|^2}{rc^2} = \frac{1}{rc^2} \geq \frac{1}{r^2}. \tag{4.90}
$$

□

4.5.4 Correctable noise channels with bounded Choi rank

In this section we will study the behavior of $r(\mathcal{X}, \mathcal{Y})$ and $r_1(\mathcal{X}, \mathcal{Y})$. Lower and upper bounds for both quantities will be summarized in the following theorem.

Theorem 4.25 ([2]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be some Euclidean spaces such that* $\dim(\mathcal{Y}) \geq \dim(\mathcal{X})$. *Then, we have*

$$\max\left\{r \in \mathbb{N} : \dim(\mathcal{X}) \leq \frac{\left\lceil \frac{\dim(\mathcal{Y})}{r^2} \right\rceil + r^2}{r^2+1}\right\} \leq r_1(\mathcal{X}, \mathcal{Y}) \leq \left\lceil \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X}) - 1}} \right\rceil - 1,$$
$$\left\lceil \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X}) - 1}} \right\rceil - 1 \leq r(\mathcal{X}, \mathcal{Y}) < \frac{\dim(\mathcal{Y})\dim(\mathcal{X})}{\dim(\mathcal{X})^2 - 1}. \tag{4.91}$$

Proof. The inequalities $r_1(\mathcal{X}, \mathcal{Y}) \leq \left\lceil \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1}} \right\rceil - 1$ and $r(\mathcal{X}, \mathcal{Y}) < \frac{\dim(\mathcal{Y})\dim(\mathcal{X})}{\dim(\mathcal{X})^2-1}$ follow from Lemma 4.22 and Proposition 4.21, respectively.

To show the lower-bound for $r_1(\mathcal{X}, \mathcal{Y})$ we use the result from [61]. The authors showed that the noise channel $\mathcal{E} = \mathcal{K}\left((E_i)_{i=1}^r\right) \in \mathcal{C}(\mathcal{Y})$ satisfying

$$(D+1)(\dim(\mathcal{X}) - 1) + 1 \leq \left\lceil \frac{\dim(\mathcal{Y})}{D} \right\rceil, \tag{4.92}$$

where $D = \dim(\mathrm{span}(E_j^\dagger E_i : i, j = 1, \ldots, r))$ is perfectly correctable for $\mathcal{X}$, that is $\mathcal{E} \in \xi_1(\mathcal{X}, \mathcal{Y})$. Let us observe that $D \leq \mathrm{rank}(J(\mathcal{E}))^2$ and there are quantum channels satisfying $D = \mathrm{rank}(J(\mathcal{E}))^2$ (we may take $\mathcal{E}$ as a random quantum channel sampled according to Definition 3.1 and use Proposition 3.12). Note that due to the upper-bound on $r_1(\mathcal{X}, \mathcal{Y})$ we consider channels with the Choi rank not greater than $\dim(\mathcal{Y})$, hence, $D = \mathrm{rank}(J(\mathcal{E}))^2$ is achievable. As a result, if for $r \in \mathbb{N}$ we have $(r^2+1)(\dim(\mathcal{X}) - 1) + 1 \leq \left\lceil \frac{\dim(\mathcal{Y})}{r^2} \right\rceil$, then for all $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$, $\mathrm{rank}(J(\mathcal{E})) \leq r$, it holds $\mathcal{E} \in \xi_1(\mathcal{X}, \mathcal{Y})$. Therefore,

$$r_1(\mathcal{X}, \mathcal{Y}) \geq \max\left\{r \in \mathbb{N} : \dim(\mathcal{X}) \leq \frac{\left\lceil \frac{\dim(\mathcal{Y})}{r^2} \right\rceil + r^2}{r^2+1}\right\}. \tag{4.93}$$

Now, we will show that $\left\lceil \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1}} \right\rceil - 1 \leq r(\mathcal{X}, \mathcal{Y})$. Take arbitrary $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ such that $\mathrm{rank}(J(\mathcal{E}))^2(\dim(\mathcal{X}) - 1) < \dim(\mathcal{Y})$. We will show $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$. Let us denote $r = \mathrm{rank}(J(\mathcal{E}))$. Consider a Kraus representation $\mathcal{E} = \mathcal{K}\left((E_j)_{j=1}^r\right)$ and define the following set

$$A = \left\{s \in \mathbb{N} :\ \exists_{\Pi_s \in \mathcal{P}(\mathcal{Y})}\ \Pi_s = \Pi_s^2, \mathrm{rank}(\Pi_s) = s, \mathrm{rank}(\mathcal{E}^\dagger(\Pi_s)) = \dim(\mathcal{Y})\right\}. \tag{4.94}$$

Observe that $\dim(\mathcal{Y}) \in A$ and if some $s \in A$, then $sr \geq \dim(\mathcal{Y})$. Define $s_0 = \min(A)$ and consider a corresponding projector $\Pi_{s_0} \in \mathcal{P}(\mathcal{Y})$, such that $\mathrm{rank}(\Pi_{s_0}) = s_0$

and $\text{rank}(\mathcal{E}^\dagger(\Pi_{s_0})) = \dim(\mathcal{Y})$. Let us take a orthonormal collection of vectors $|v_i\rangle$, where $i = 1, \ldots, s_0$ for which we have $\Pi_{s_0} = \sum_{i=1}^{s_0} |v_i\rangle\langle v_i|$. From the assumption $s_0 = \min(A)$, for any i we get $\text{rank}(\mathcal{E}^\dagger(\Pi_{s_0} - |v_i\rangle\langle v_i|)) < \dim(\mathcal{Y})$. Therefore, we may define vectors $\mathcal{Y} \ni |w_i\rangle \neq 0$ such that $\mathcal{E}^\dagger(\Pi_{s_0} - |v_i\rangle\langle v_i|)|w_i\rangle = 0$. Observe that for each i, there exists E_j for which $\langle v_i|E_j|w_i\rangle \neq 0$. Let us define $F_j = (\langle v_a|E_j|w_b\rangle)_{a,b=1,\ldots,s_0}$ for $j = 1, \ldots, r$. Note, that F_j are diagonal operators and it holds $\sum_j F_j^\dagger F_j > 0$. From $r^2(\dim(\mathcal{X}) - 1) < \dim(\mathcal{Y})$ and $s_0 r \geq \dim(\mathcal{Y})$ we have

$$r(\dim(\mathcal{X}) - 1) < \frac{\dim(\mathcal{Y})}{r} \leq s_0. \tag{4.95}$$

Utilizing Proposition 4.23, Lemma 4.7 and Theorem 4.4 there exist $S_* \in \mathcal{M}(\mathcal{X}, \mathbb{C}^{s_0})$ and $R_* \in \mathcal{M}(\mathbb{C}^{s_0}, \mathcal{X})$, such that $R_* F_j S_* \propto \mathbb{1}_\mathcal{X}$ and there exists j_0, for which it holds $R_* F_{j_0} S_* \neq 0$. That implies $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$. □

For now, we will calculate explicitly $r(\mathcal{X}, \mathcal{Y})$ and $r_1(\mathcal{X}, \mathcal{Y})$ for $\mathcal{X} = \mathbb{C}^2$ and $\mathcal{Y} = \mathbb{C}^3, \mathbb{C}^4$.

Proposition 4.26 ([2])**.** *For all* $\mathcal{E} \in \mathcal{C}(\mathbb{C}^4)$ *satisfying* $\text{rank}(J(\mathcal{E})) \leq 2$ *we have* $\mathcal{E} \in \xi(\mathbb{C}^2, \mathbb{C}^4)$.

Proof. Let us fix $\mathcal{E} = \mathcal{K}((E_0, E_1)) \in \mathcal{C}(\mathbb{C}^4)$. From the equality $E_0^\dagger E_0 + E_1^\dagger E_1 = \mathbb{1}_{\mathbb{C}^4}$ we may write the singular decomposition of E_0, E_1 in the form: $E_0 = U_0 D_0 V$ and $E_1 = U_1 D_1 V$, where $U_0, U_1, V \in \mathcal{U}(\mathbb{C}^4)$ and $D_0, D_1 \in \mathcal{P}(\mathbb{C}^4)$ are diagonal operators satisfying $D_0^2 + D_1^2 = \mathbb{1}_{\mathbb{C}^4}$. In order to show that $\mathcal{E} \in \xi(\mathbb{C}^2, \mathbb{C}^4)$ we will use Theorem 4.4. We will prove that there exist $S_* \in \mathcal{M}(\mathbb{C}^2, \mathbb{C}^4)$ and $R_* \in \mathcal{M}(\mathbb{C}^4, \mathbb{C}^2)$, such that $R_* E_0 S_* = c_0 \mathbb{1}_{\mathbb{C}^2}$, $R_* E_1 S_* = c_1 \mathbb{1}_{\mathbb{C}^2}$ for some $c_0, c_1 \in \mathbb{C}$ satisfying $(c_0, c_1) \neq (0, 0)$. Let us introduce the following notation

$$\begin{aligned} |x_i\rangle &= (D_0)_{ii} U_0 |i\rangle, \quad i = 0, \ldots, 3, \\ |y_i\rangle &= (D_1)_{ii} U_1 |i\rangle, \quad i = 0, \ldots, 3. \end{aligned} \tag{4.96}$$

Note that vectors $|x_i\rangle$ are orthogonal (the same holds for $|y_i\rangle$) and for each $i = 0, \ldots, 3$ we have $|x_i\rangle \neq 0$ or $|y_i\rangle \neq 0$. We may write S_* and R_* in the following form

$$\begin{aligned} S_* &= V^\dagger(|S_0\rangle\langle 0| + |S_1\rangle\langle 1|), \\ R_* &= |0\rangle\langle R_0| + |1\rangle\langle R_1|, \end{aligned} \tag{4.97}$$

for some vectors $|S_0\rangle, |S_1\rangle, |R_0\rangle, |R_1\rangle \in \mathbb{C}^4$. The rest of the proof will be divided into three cases.

In the first case, we assume there exists $i_3 \in \{0, \ldots, 3\}$ such that vectors $|x_{i_3}\rangle, |y_{i_3}\rangle$ are linearly independent. Define indices $i_0, i_1, i_2 \in \{0, \ldots, 3\}$ as the remaining labels, such that $\{i_0,\dots,i_3\}$ covers the whole set $\{0,\dots,3\}$. Let $(a_0,a_1,a_2)^\top \in \mathbb{C}^3$ be a normalized vector orthogonal to vectors $(\langle y_{i_3}|x_{i_0}\rangle, \langle y_{i_3}|x_{i_1}\rangle, \langle y_{i_3}|x_{i_2}\rangle)^\dagger$ and $(\langle x_{i_3}|y_{i_0}\rangle, \langle x_{i_3}|y_{i_1}\rangle, \langle x_{i_3}|y_{i_2}\rangle)^\dagger$. Take $|S_1\rangle = |i_3\rangle$ and $|S_0\rangle = a_0|i_0\rangle + a_1|i_1\rangle + a_2|i_2\rangle$. Define $|x\rangle = a_0|x_{i_0}\rangle + a_1|x_{i_1}\rangle + a_2|x_{i_2}\rangle$ and $|y\rangle = a_0|y_{i_0}\rangle + a_1|y_{i_1}\rangle + a_2|y_{i_2}\rangle$. We obtain

$$\begin{aligned} E_0 S_* &= |x\rangle\langle 0| + |x_{i_3}\rangle\langle 1|, \\ E_1 S_* &= |y\rangle\langle 0| + |y_{i_3}\rangle\langle 1|. \end{aligned} \tag{4.98}$$

It is not hard to observe that $|x\rangle \neq 0$ or $|y\rangle \neq 0$. If $|x\rangle \neq 0$, take $|R_0\rangle = |x\rangle$, else take $|R_0\rangle = |y\rangle$. As the vectors $|x_{i_3}\rangle, |y_{i_3}\rangle$ are linearly independent we may define

$$(b_0,b_1)^\top := \begin{pmatrix} \langle x_{i_3}|x_{i_3}\rangle & \langle y_{i_3}|x_{i_3}\rangle \\ \langle x_{i_3}|y_{i_3}\rangle & \langle y_{i_3}|y_{i_3}\rangle \end{pmatrix}^{-1} (\langle R_0|x\rangle, \langle R_0|y\rangle)^\top. \tag{4.99}$$

Take $|R_1\rangle = \bar{b}_0|x_{i_3}\rangle + \bar{b}_1|y_{i_3}\rangle$. Finally, we may check that it holds

$$\begin{aligned} R_* E_0 S_* &= (|0\rangle\langle R_0| + |1\rangle\langle R_1|)(|x\rangle\langle 0| + |x_{i_3}\rangle\langle 1|) = \langle R_0|x\rangle \mathbb{1}_{\mathbb{C}^2}, \\ R_* E_1 S_* &= (|0\rangle\langle R_0| + |1\rangle\langle R_1|)(|y\rangle\langle 0| + |y_{i_3}\rangle\langle 1|) = \langle R_0|y\rangle \mathbb{1}_{\mathbb{C}^2}. \end{aligned} \tag{4.100}$$

In the second case, we assume that there exists a pair of vectors $|y_{i_0}\rangle, |y_{i_1}\rangle$ for $i_0 \neq i_1$, such that $|y_{i_0}\rangle = |y_{i_1}\rangle = 0$. Then, the vectors $|x_{i_0}\rangle, |x_{i_1}\rangle$ are orthonormal. We simply define $|S_0\rangle = |i_0\rangle, |S_1\rangle = |i_1\rangle$, $|R_0\rangle = |x_{i_0}\rangle$ and $|R_1\rangle = |x_{i_1}\rangle$. One can calculate that $R_* E_0 S_* = \mathbb{1}_{\mathbb{C}^2}$ and $R_* E_1 S_* = 0$.

In the third case, for all $i \in \{0,\dots,3\}$ vectors $|x_i\rangle, |y_i\rangle$ are not linearly independent and there is at most one zero vector $|y_{i_3}\rangle$ for some $i_3 \in \{0,\dots,3\}$. Define indices $i_0,i_1,i_2 \in \{0,\dots,3\}$ as the remaining labels, such that $\{i_0,\dots,i_3\}$ covers the whole set $\{0,\dots,3\}$. Define the matrix

$$M = \begin{pmatrix} \langle y_{i_0}|x_{i_0}\rangle & \langle y_{i_1}|x_{i_1}\rangle & \langle y_{i_2}|x_{i_2}\rangle \\ \langle y_{i_0}|y_{i_0}\rangle & \langle y_{i_1}|y_{i_1}\rangle & \langle y_{i_2}|y_{i_2}\rangle \end{pmatrix}. \tag{4.101}$$

In the first sub-case we assume that $\mathrm{rank}(M) = 1$. Define $b = \frac{\langle y_{i_1}|y_{i_1}\rangle}{\langle y_{i_0}|y_{i_0}\rangle}$. We can take $|S_0\rangle = |i_0\rangle$, $|S_1\rangle = |i_1\rangle$, $|R_0\rangle = |y_{i_0}\rangle$ and $|R_1\rangle = \frac{1}{b}|y_{i_1}\rangle$. One can calculate that $R_* E_0 S_* = \langle y_{i_0}|x_{i_0}\rangle \mathbb{1}_{\mathbb{C}^2}$ and $R_* E_1 S_* = \langle y_{i_0}|y_{i_0}\rangle \mathbb{1}_{\mathbb{C}^2}$.

In the second sub-case we assume that $\mathrm{rank}(M) = 2$. Define indices $j_1, j_2 \in \{0,1,2\}$, such that

$$\mathrm{rank}\begin{pmatrix} M_{0,j_1} & M_{0,j_2} \\ M_{1,j_1} & M_{1,j_2} \end{pmatrix} = 2. \tag{4.102}$$

Define $j_0 \in \{0,1,2\}$ as the remaining label, such that $\{j_0,j_1,j_2\}$ covers the whole set $\{0,1,2\}$. Take $|S_0\rangle = |i_{j_0}\rangle$, $|R_0\rangle = |y_{i_{j_0}}\rangle$ and define

$$(b_1,b_2)^\top := \begin{pmatrix} \langle y_{i_{j_1}}|x_{i_{j_1}}\rangle & \langle y_{i_{j_2}}|x_{i_{j_2}}\rangle \\ \langle y_{i_{j_1}}|y_{i_{j_1}}\rangle & \langle y_{i_{j_2}}|y_{i_{j_2}}\rangle \end{pmatrix}^{-1} (\langle y_{i_{j_0}}|x_{i_{j_0}}\rangle, \langle y_{i_{j_0}}|y_{i_{j_0}}\rangle)^\top. \tag{4.103}$$

We may take $|S_1\rangle = |i_{j_1}\rangle + |i_{j_2}\rangle$ and $|R_1\rangle = \bar{b}_1|y_{i_{j_1}}\rangle + \bar{b}_2|y_{i_{j_2}}\rangle$. Direct calculations reveal that $R_* E_0 S_* = \langle y_{i_{j_0}}|x_{i_{j_0}}\rangle \mathbb{1}_{\mathbb{C}^2}$ and $R_* E_1 S_* = \langle y_{i_{j_0}}|y_{i_{j_0}}\rangle \mathbb{1}_{\mathbb{C}^2}$. □

By Theorem 4.25 and Proposition 4.26 we get the following advantage of the pQEC protocol for $\mathcal{X} = \mathbb{C}^2$ and $\mathcal{Y} = \mathbb{C}^4$.

Corollary 4.27 ([2]). *For $\mathcal{X} = \mathbb{C}^2$ and $\mathcal{Y} = \mathbb{C}^4$ we have*

$$r_1(\mathcal{X}, \mathcal{Y}) < r(\mathcal{X}, \mathcal{Y}). \tag{4.104}$$

In particular, it holds

$$\begin{array}{ll} r_1(\mathbb{C}^2, \mathbb{C}^3) = 1 & \quad r(\mathbb{C}^2, \mathbb{C}^3) = 1 \\ r_1(\mathbb{C}^2, \mathbb{C}^4) = 1 & \quad r(\mathbb{C}^2, \mathbb{C}^4) = 2 \end{array} \tag{4.105}$$

Example of pQEC qubit code

Consider the following scenario. You have a task to transfer a given qubit state $\rho \in \mathcal{D}(\mathbb{C}^2)$ through a quantum communication line represented by a noise channel $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ of the form $\mathcal{E}(Y) = \operatorname{tr}_2\left(U(Y \otimes |\psi\rangle\langle\psi|)U^\dagger\right)$, where $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathbb{C}^2)$ and $U \in \mathcal{U}(\mathcal{Y} \otimes \mathbb{C}^2)$. We ask, what is the minimal size of the communication line, $\dim(\mathcal{Y})$, which is large enough to recover the state ρ with the pQEC procedure? To answer this question, observe that the channel $\mathcal{E}$ satisfies $\operatorname{rank}(J(\mathcal{E})) \leq 2$. In Proposition 4.26 we noticed that such channels are probabilistically correctable for a given input space $\mathbb{C}^2$, if $\dim(\mathcal{Y}) = 4$. Therefore, to correctly transfer a qubit state through $\mathcal{E}$, we may define an error-correcting scheme with only two physical qubits.

It is worth mentioning that some channels $\mathcal{E} \in \mathcal{C}(\mathbb{C}^4)$ which satisfy $\operatorname{rank}(J(\mathcal{E})) = 2$ are perfectly correctable for a space $\mathbb{C}^2$. If $\mathcal{E}$ is not an extreme point in the set of all channels $\mathcal{C}(\mathbb{C}^4)$, then $\mathcal{E}$ is mixed-unitary channel of the form $\mathcal{E}(Y) = pUYU^\dagger + (1-p)VYV^\dagger$ for some $U, V \in \mathcal{U}(\mathbb{C}^4)$ and $p \in (0,1)$ [97]. In that case, it was shown in [98] that $\mathcal{E} \in \xi_1(\mathbb{C}^2, \mathbb{C}^4)$. Nevertheless, if we consider a random channel $\mathcal{E} \in \mathcal{C}(\mathbb{C}^4)$ defined according to Definition 3.1 we will see that almost surely it is an extremal channel (see Proposition 3.12). What is more, by Theorem 4.29 we even know that $\mathcal{P}\left(\mathcal{E} \in \xi_1(\mathbb{C}^2, \mathbb{C}^4)\right) < 1$. A particular example of a Schur channel which is not perfectly correctable for $\mathbb{C}^2$ was constructed in the proof of Lemma 4.22. In fact, it follows from this construction that almost all Schur channels are not perfectly correctable.

We provide the following the pQEC procedure based on Proposition 4.26 to probabilistically correct any $\mathcal{E} \in \mathcal{C}(\mathbb{C}^4)$ which satisfy $\operatorname{rank}(J(\mathcal{E})) = 2$.

Algorithm 28: Probabilistic QEC qubit code [2]

Input: $\mathcal{E} \in \mathcal{C}(\mathbb{C}^4)$ such that $\operatorname{rank}(J(\mathcal{E})) \leq 2$.
Output: The pQEC procedure with success probability $p > 0$.

Let $\mathcal{E} = \mathcal{K}\left((E_0, E_1)\right)$.
Define $S_* \in \mathcal{M}(\mathbb{C}^2, \mathbb{C}^4)$ and $R_* \in \mathcal{M}(\mathbb{C}^4, \mathbb{C}^2)$, such that $R_* E_0 S_* \propto \mathbb{1}_{\mathbb{C}^2}$, $R_* E_1 S_* \propto \mathbb{1}_{\mathbb{C}^2}$ and $R_* E_0 S_* \neq 0 \vee R_* E_1 S_* \neq 0$ according to the proof of Proposition 4.26.
Define

$$
\begin{aligned}
Q &= S_*^\dagger S_*, \\
S &= S_* Q^{-1/2}, \\
R &= \frac{Q^{1/2} R_*}{\|Q^{1/2} R_*\|_\infty}.
\end{aligned}
$$

Calculate $p \in (0, 1]$, such that $R\left(\mathcal{E}\left(SXS^\dagger\right)\right) R^\dagger = pX$ for any $X \in \mathcal{M}(\mathbb{C}^2)$.
Define $U_S \in \mathcal{U}(\mathbb{C}^4)$ which satisfies $U_S(\mathbb{1}_{\mathbb{C}^2} \otimes |0\rangle) = S$.
Let $R = \sigma_1 |z_1\rangle\langle t_1| + \sigma_2 |z_2\rangle\langle t_2|$ be the singular value decomposition of R. Define $U_R \in \mathcal{U}(\mathbb{C}^4)$ which satisfies

$$
\begin{aligned}
U_R|t_1\rangle &= |0, 0\rangle, \\
U_R|t_2\rangle &= |1, 0\rangle.
\end{aligned}
$$

Define $R' = R U_R^\dagger (\mathbb{1}_{\mathbb{C}^2} \otimes |0\rangle)$.
Define $V_R \in \mathcal{U}(\mathbb{C}^4)$ which satisfies $(\mathbb{1}_{\mathbb{C}^2} \otimes \langle 0|) V_R (\mathbb{1}_{\mathbb{C}^2} \otimes |0\rangle) = R'$.
Run the QEC procedure presented in Figure 4.2 for $|\psi\rangle, U_S, U_R, V_R$.
Let σ_{exp} be the output state of the procedure presented in Figure 4.2. Use the post-processing of the measurements' output (i, j) according to the following table:

Labels	**Probability**	**Status**	**Action**	**Result**
$(i, j) = (0, 0)$	p	QEC succeeded	Accept σ_{exp}	$\sigma_{\text{exp}} = \|\psi\rangle\langle\psi\|$
$(i, j) \neq (0, 0)$	$1 - p$	QEC failed	Reject σ_{exp}	$\sigma_{\text{exp}}\ ?\ \|\psi\rangle\langle\psi\|$

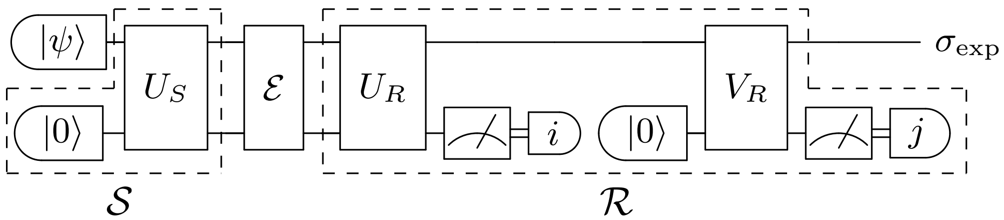


Figure 4.2: The circuit representing the pQEC procedure. We have access to two physical qubits. The first qubit is in the state $|\psi\rangle$. This state will be encoded. The second state we set equal to $|0\rangle$. We implement the two-qubit encoding unitary operator U_S. Then, the encoded state, $U_S(|\psi\rangle \otimes |0\rangle)$, is affected by the noise channel $\mathcal{E}$. After that, we start the decoding procedure. We implement the two-qubit unitary rotation U_R. We measure the second qubit in the standard basis and obtain a classical label $i \in \{0, 1\}$. We prepare a third qubit in the state $|0\rangle$ and implement a two-qubit unitary rotation V_R. We measure the third qubit in the standard basis and obtain a classical label $j \in \{0, 1\}$. If $(i, j) = (0, 0)$ we accept the output state, otherwise, we reject it and request resend. It is worth mentioning that this circuit can be implemented on currently available gate-model quantum computers such as Rigetti or IonQ using three qubits. It is also possible to implement this circuit by using only two qubits and mid-circuit measurement provided by IBMQ.

4.5.5 Random noise channels

In this section, we will show the advantage of the pQEC procedure for randomly generated noise channels. We will follow the procedure of sampling quantum channels considered in Section 3.1.

Theorem 4.29 ([2]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be some Euclidean spaces such that* $\dim(\mathcal{Y}) \geq \dim(\mathcal{X})$ *and $r \in \mathbb{N}$ be a parameter. For a random quantum channel $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ sampled according to the measure $\mu^{Kraus}_{\mathcal{Y},\mathcal{Y};r}$ defined in Definition 3.2 it holds:*

$$\begin{aligned} r < \frac{\dim(\mathcal{X})\dim(\mathcal{Y})}{\dim(\mathcal{X})^2 - 1} &\implies \mathcal{P}\left(\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})\right) = 1, \\ \mathcal{P}\left(\mathcal{E} \in \xi_1(\mathcal{X}, \mathcal{Y})\right) = 1 &\implies r < \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X}) - 1}}. \end{aligned} \tag{4.106}$$

Proof. Following the construction in Definition 3.2, for $r \in \mathbb{N}$ satisfying $r < \frac{\dim(\mathcal{X})\dim(\mathcal{Y})}{\dim(\mathcal{X})^2-1}$, let $(G_i)_{i=1}^r \subset \mathcal{M}(\mathcal{Y})$ be a tuple of random and independent Ginibre matrices and $Q = \sum_{i=1}^r G_i^\dagger G_i$. Define the projector $\Pi = \sum_{i=0}^{\dim(\mathcal{X})-1} |i\rangle\langle i|$ and consider the set

$$A = \left\{ (G_i)_{i=1}^r : \operatorname{rank}(Q) = \dim(\mathcal{Y}), \operatorname{rank}\left(\sum_{i=1}^r G_i^\dagger \Pi G_i \right) = \min\{r \dim(\mathcal{X}), \dim(\mathcal{Y})\} \right\}. \tag{4.107}$$

One can observe that $\mathcal{P}((G_i)_{i=1}^r \in A) = 1$. Let $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ be a random channel defined as $\mathcal{E} = \mathcal{K}\left((G_i Q^{-1/2})_{i=1}^r\right)$ for $(G_i)_{i=1}^r \in A$. Define $S = Q^{1/2}\tilde{S}$ for $\tilde{S} \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ and $R = \tilde{R}\Pi$ for $\tilde{R} \in \mathcal{M}(\mathcal{Y}, \mathcal{X})$. We obtain $RG_iQ^{-1/2}S = \tilde{R}\Pi G_i \tilde{S}$. Utilizing Lemma 4.7, Proposition 4.21 and Theorem 4.4 for $\tilde{\mathcal{E}} = \mathcal{K}\left((\Pi G_i)_{i=1}^r\right)$, there exist $\tilde{S}, \tilde{R}$, such that $\tilde{R}\Pi G_i \tilde{S} \propto \mathbb{1}_\mathcal{X}$ and $\tilde{R}\Pi G_{i_0}\tilde{S} \neq 0$ for some i_0. Hence, $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$.

Now, for a given $r \in \mathbb{N}$ let us define $B = \{\mathcal{E} \in \mu_{\mathcal{Y},\mathcal{Y};r}^{Kraus} : \ \mathcal{E} \in \xi_1(\mathcal{X}, \mathcal{Y})\}$. From the assumption $\mathcal{P}(B) = 1$, we obtain that B is a dense subset of $\{\mathcal{E} \in \mathcal{C}(\mathcal{Y}) : \operatorname{rank}(J(\mathcal{E})) \leq r\}$. Imitating the proof of Proposition 4.18, we get that if $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ and $\operatorname{rank}(J(\mathcal{E})) \leq r$, then $\mathcal{E} \in \xi_1(\mathcal{X}, \mathcal{Y})$. That implies $r \leq r_1(\mathcal{X}, \mathcal{Y})$. By using Lemma 4.22 we obtain the desired inequality. □

Corollary 4.30 ([2]). *Let $\mathcal{E} = \mathcal{K}\left((E_i)_{i=0}^{r-1}\right) \in \mathcal{C}(\mathcal{Y})$ be a random quantum channel defined according to Definition 3.2 and assume that $r \leq \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})}$. Define a sequence $V_1, V_2, \ldots \in \mathcal{U}(\mathcal{X}, \mathcal{Y})$ of random isometry matrices sampled according to the Haar measure. Let $R_{F_n} = (F_n F_n^\dagger)^{-1}$ for $F_n = \sum_{i=0}^{r-1} |i\rangle \otimes V_n^\dagger E_i \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$. Then, almost surely it holds*

$$\begin{aligned} p_\mathcal{X}(\mathcal{E}) &\geq \sup_{n\in\mathbb{N}} \max\left\{\operatorname{tr}(P) : P \in \mathcal{P}(\mathbb{C}^r), \operatorname{tr}_{\mathbb{C}^r}\left(R_{F_n}(P \otimes \mathbb{1}_\mathcal{X})\right) \leq \mathbb{1}_\mathcal{X}\right\} \\ &\geq \max\left\{\lambda_{\min}((\mathbb{1}_{\mathbb{C}^r} \otimes V^\dagger) E E^\dagger (\mathbb{1}_{\mathbb{C}^r} \otimes V)) : \ V \in \mathcal{U}(\mathcal{X}, \mathcal{Y})\right\}, \end{aligned} \tag{4.108}$$

where $\lambda_{\min}$ is the smallest eigenvalue and $E = \sum_{i=0}^{r-1} |i\rangle \otimes E_i$.

Proof. For $r \in \mathbb{N}$ satisfies $r \dim(\mathcal{X}) \leq \dim(\mathcal{Y})$, let $(G_i)_{i=0}^{r-1} \subset \mathcal{M}(\mathcal{Y})$ be a tuple of random and independent Ginibre matrices and $Q = \sum_{i=0}^{r-1} G_i^\dagger G_i$. Define a sequence $V_1, V_2, \ldots \in \mathcal{U}(\mathcal{X}, \mathcal{Y})$ of random isometry matrices sampled according to the Haar measure. Consider the following sets

$$\begin{aligned} A_0 &= \{\left((G_i)_{i=0}^{r-1}, (V_n)_{n\in\mathbb{N}}\right) : \operatorname{rank}(Q) = \dim(\mathcal{Y})\}, \\ A_m &= \{\left((G_i)_{i=0}^{r-1}, (V_n)_{n\in\mathbb{N}}\right) : \operatorname{rank}\left(\sum_{i=0}^{r-1} G_i^\dagger V_m V_m^\dagger G_i\right) = r \dim(\mathcal{X})\}. \end{aligned} \tag{4.109}$$

One can observe that $\mathcal{P}(\bigcap_{n\geq 0} A_n) = 1$. For $n \in \mathbb{N}$ let $R_{F_n} = (F_n F_n^\dagger)^{-1}, \Pi_{F_n} = F_n F_n^{-1}$, where $F_n = \sum_{i=0}^{r-1} |i\rangle \otimes V_n^\dagger E_i \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$. Utilizing Lemma 4.8 we obtain

$$p_{\mathcal{X}}(\mathcal{E}_r) \geq \sup \left\{ \operatorname{tr}(P) : \begin{cases} n \in \mathbb{N}, \\ P \in \mathcal{P}(\mathbb{C}^r), \\ \operatorname{tr}_{\mathbb{C}^r} \left(R_{F_n}(P \otimes \mathbb{1}_{\mathcal{X}}) \right) \leq \mathbb{1}_{\mathcal{X}}, \\ (\Pi_{F_n} \otimes \mathbb{1}_{\mathcal{X}})(P \otimes |\mathbb{1}_{\mathcal{X}}\rangle\rangle\langle\langle \mathbb{1}_{\mathcal{X}}|)(\Pi_{F_n} \otimes \mathbb{1}_{\mathcal{X}}) = P \otimes |\mathbb{1}_{\mathcal{X}}\rangle\rangle\langle\langle \mathbb{1}_{\mathcal{X}}| \end{cases} \right\}. \tag{4.110}$$

It holds that $F_n^\dagger F_n = Q^{-1/2} \sum_{i=0}^{r-1} G_i^\dagger V_n V_n^\dagger G_i Q^{-1/2}$. Hence, almost surely $\operatorname{rank}(F_n) = r \dim(\mathcal{X})$ which implies that $\Pi_{F_n} = \mathbb{1}_{\mathbb{C}^r \otimes \mathcal{X}}$. Therefore, we get

$$p_{\mathcal{X}}(\mathcal{E}_r) \geq \sup_{n \in \mathbb{N}} \max \left\{ \operatorname{tr}(P) : P \in \mathcal{P}(\mathbb{C}^r), \operatorname{tr}_{\mathbb{C}^r} \left(R_{F_n}(P \otimes \mathbb{1}_{\mathcal{X}}) \right) \leq \mathbb{1}_{\mathcal{X}} \right\}. \tag{4.111}$$

To prove the second inequality note that for each $n \in \mathbb{N}$ the value of max is positive. By Lemma 4.8 and the fact that $R_{F_n} > 0$ for each $n \in \mathbb{N}$, we get

$$\begin{aligned} &\sup_{n \in \mathbb{N}} \max \left\{ \operatorname{tr}(P) : P \in \mathcal{P}(\mathbb{C}^r), \operatorname{tr}_{\mathbb{C}^r} \left(R_{F_n}(P \otimes \mathbb{1}_{\mathcal{X}}) \right) \leq \mathbb{1}_{\mathcal{X}} \right\} \\ \geq &\sup_{n \in \mathbb{N}} \|R_{F_n}\|_\infty^{-1} = \sup_{n \in \mathbb{N}} \|((\mathbb{1}_{\mathbb{C}^r} \otimes V_n^\dagger) E E^\dagger (\mathbb{1}_{\mathbb{C}^r} \otimes V_n))^{-1}\|_\infty^{-1} \\ = &\sup_{n \in \mathbb{N}} \lambda_{\min}((\mathbb{1}_{\mathbb{C}^r} \otimes V_n^\dagger) E E^\dagger (\mathbb{1}_{\mathbb{C}^r} \otimes V_n)) \\ = &\max \left\{ \lambda_{\min}((\mathbb{1}_{\mathbb{C}^r} \otimes V^\dagger) E E^\dagger (\mathbb{1}_{\mathbb{C}^r} \otimes V)) : \ V \in \mathcal{U}(\mathcal{X}, \mathcal{Y}) \right\}, \end{aligned} \tag{4.112}$$

where in the last equality we used the fact that the subset $\{V_n : n \in \mathbb{N}\}$ is almost surely dense in the set $\mathcal{U}(\mathcal{X}, \mathcal{Y})$ and the fact that $\lambda_{\min}$ is a continuous function. □

4.6 Numerical examples

In this section we numerically investigate the effectiveness of the pQEC codes construction provided in Corollary 4.30. We check the behavior of a lower bound for $p_{\mathcal{X}}(\mathcal{E})$, where $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ is a random quantum channel sampled according to the measure $\mu_{\mathcal{Y},\mathcal{Y};r}^{Kraus}$ defined in Definition 3.2.

For a given tuple (y, x, r), where $x = \dim(\mathcal{X})$ and $y = \dim(\mathcal{Y})$ we will sample M random channels $\mathcal{E} \in \mu_{\mathcal{Y},\mathcal{Y};r}^{Kraus}$. To meet the assumptions of Corollary 4.30 we consider (y, x, r) such that $rx \leq y$. For each random channel N Haar-random isometry matrices $V_n \in \mathcal{U}(\mathcal{X}, \mathcal{Y})$, where $n = 1, \ldots, N$, will be sampled. By using the SDP programming we then calculate

$$p = \max_{n=1,\ldots,N} \max \left\{ \operatorname{tr}(P) : P \in \mathcal{P}(\mathbb{C}^r), \operatorname{tr}_{\mathbb{C}^r} \left(R_{F_n}(P \otimes \mathbb{1}_{\mathcal{X}}) \right) \leq \mathbb{1}_{\mathcal{X}} \right\} \tag{4.113}$$

with a precision $\epsilon = 10^{-5}$. We plot the results as the following histograms with fixed parameters $N = 30$ and $M = 300$.

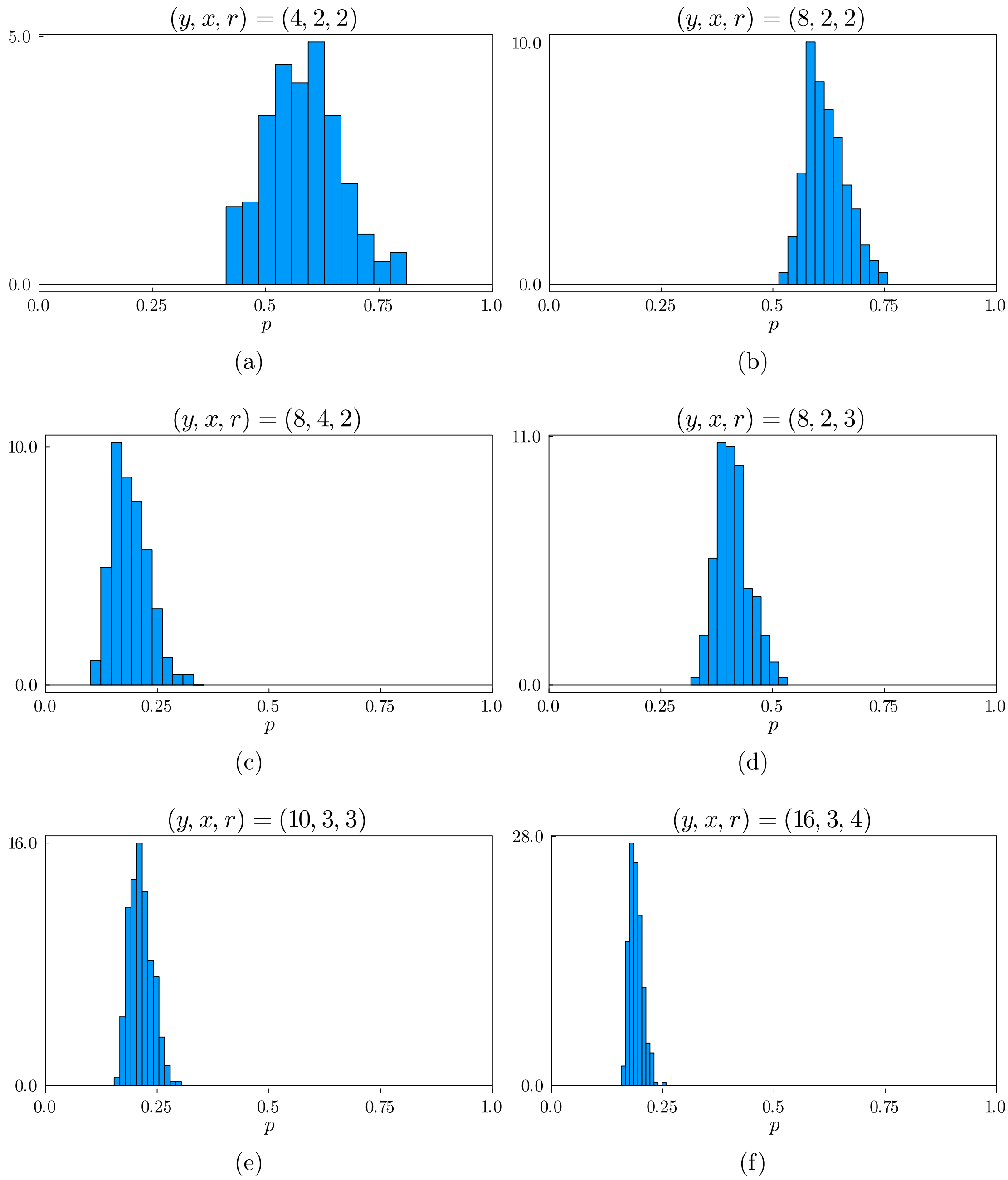


Figure 4.3: The estimation of probability density function of the probability of successful error correction p defined as in Eq. (4.113).

4.7 Generalization

In the last section of this chapter, we raise a subject of correcting not only single noise channel, but correcting all noise channels from given family. This approach is useful for studying multi-qubit noises, see for example [99, Chapter 10].

Let us denote by Υ an arbitrary family of noise subchannels, that is $\Upsilon \subset s\mathcal{C}(\mathcal{Y})$. We ask if there exists error-correcting scheme $(\mathcal{S}, \mathcal{R})$, such that for all $\mathcal{E} \in \Upsilon$ we have $\mathcal{R}\mathcal{E}\mathcal{S} = p_{\mathcal{E}}\mathcal{I}_{\mathcal{X}}$, for some $p_{\mathcal{E}} \geq 0$. Note, that $p_{\mathcal{E}}$ may differ for different $\mathcal{E}$, hence, we shall introduce a quantity to "globally" control the effectiveness of $(\mathcal{S}, \mathcal{R})$. We propose the following approach.

Without loss of the generality we may assume that Υ is convex. Let μ be some probability measure defined on the set Υ, such that the support of μ is equal to Υ. The choice of μ is arbitrary. As an example, we can take μ as the flat measure, representing the maximal uncertainty in the process of probing random noise channels $\mathcal{E}$ from Υ. We assume that noise $\mathcal{E} \in \Upsilon$ are probed according to μ. We will say, that the scheme $(\mathcal{S}, \mathcal{R})$ will be a valid error-correcting scheme for Υ and μ if in average, the probability of successful error correction is non zero, that is

$$\int_{\Upsilon} p_{\mathcal{E}}\mu(d\mathcal{E}) > 0. \tag{4.114}$$

Proposition 4.31 ([2]). *Let $\Upsilon \subset s\mathcal{C}(\mathcal{Y})$ be a nonempty and convex family of noise subchannels. Define μ to be a probability measure defined on Υ and assume that the support of μ is equal to Υ. Let us define the average noise subchannel of Υ with respect to μ*

$$\bar{\mathcal{E}} = \int_{\Upsilon} \mathcal{E}\mu(d\mathcal{E}). \tag{4.115}$$

We fix $(\mathcal{S}, \mathcal{R}) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y}) \times s\mathcal{C}(\mathcal{Y}, \mathcal{X})$. The following conditions are equivalent:

(A) For each $\mathcal{E} \in \Upsilon$ there exists $p_{\mathcal{E}} \geq 0$ such that $\mathcal{R}\mathcal{E}\mathcal{S} = p_{\mathcal{E}}\mathcal{I}_{\mathcal{X}}$ and $\int_{\Upsilon} p_{\mathcal{E}}\mu(d\mathcal{E}) > 0$.

(B) It holds that $0 \neq \mathcal{R}\bar{\mathcal{E}}\mathcal{S} \propto \mathcal{I}_{\mathcal{X}}$.

Proof. $(A) \implies (B)$
This implication is straightforward.
$(B) \implies (A)$
Let us assume that $\mathcal{R}\bar{\mathcal{E}}\mathcal{S} = p\mathcal{I}_{\mathcal{X}}$ for $p > 0$. There exists a k dimensional affine subspace L such that $\Upsilon \subset$ L and $\mathrm{int}(\Upsilon) \neq \emptyset$. Take an arbitrary $\mathcal{E}_0 \in \Upsilon$. There exist $\mathcal{E}_1, \ldots, \mathcal{E}_k \in \Upsilon$ such that convex hull of points $\mathcal{E}_0, \ldots, \mathcal{E}_k$ is a k-dimensional simplex Δ_k. For any state $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{X})$ it holds

$$p|\psi\rangle\langle\psi| = \mathcal{R}\bar{\mathcal{E}}\mathcal{S}(|\psi\rangle\langle\psi|) = \int_{\Upsilon} \mathcal{R}\mathcal{E}\mathcal{S}(|\psi\rangle\langle\psi|)\mu(d\mathcal{E}) \geq \int_{\Delta_k} \mathcal{R}\mathcal{E}\mathcal{S}(|\psi\rangle\langle\psi|)\mu(d\mathcal{E}). \tag{4.116}$$

Inside Δ_k each $\mathcal{E}$ can be uniquely represented as $\sum_{i=0}^k q_i(\mathcal{E})\mathcal{E}_i$, where $(q_i(\mathcal{E}))_{i=0}^k$ is a probability vector which depends on $\mathcal{E}$. Hence,

$$p|\psi\rangle\langle\psi| \geq \sum_{i=0}^{k} \int_{\Delta_k} q_i(\mathcal{E})\mathcal{R}\mathcal{E}_i\mathcal{S}(|\psi\rangle\langle\psi|)\mu(d\mathcal{E}) \geq \left(\int_{\Delta_k} q_0(\mathcal{E})\mu(d\mathcal{E})\right)\mathcal{R}\mathcal{E}_0\mathcal{S}(|\psi\rangle\langle\psi|). \tag{4.117}$$

There exists ϵ small ball B_ϵ around $\mathcal{E}_0$, such that for each $\mathcal{E} \in B_\epsilon \cap \Delta_k$ it holds $q_0(\mathcal{E}) \geq \frac{1}{2}$. Hence, $\int_{\Delta_k} q_0(\mathcal{E})\mu(d\mathcal{E}) \geq \frac{1}{2}\mu\left(B_\epsilon \cap \Delta_k\right) > 0$, where in the last inequality we used the fact that the support of μ is equal to Υ. Therefore, it holds that for any $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{X})$ we have $\mathcal{R}\mathcal{E}_0\mathcal{S}(|\psi\rangle\langle\psi|) \propto |\psi\rangle\langle\psi|$ and from Lemma 4.1 there exists $p_{\mathcal{E}_0} \geq 0$ such that $\mathcal{R}\mathcal{E}_0\mathcal{S} = p_{\mathcal{E}_0}\mathcal{I}_{\mathcal{X}}$. The instant relation $\int_\Upsilon p_\mathcal{E}\mu(d\mathcal{E}) = p > 0$ ends the proof. $\square$

Chapter 5

Application of pQEC procedure to approximate QEC

This chapter includes unpublished, author results.

5.1 Motivation

In the real case scenario, it is impossible to correct the given noise channel perfectly. Even an ϵ-small fluctuation can affect perfectly correctable channel, *e.g.* $\mathcal{I}_\mathcal{Y}$, and make it not correctable, *e.g.* transform it into the channel $(1-\epsilon)\mathcal{I}_\mathcal{Y}+\epsilon\Phi_*$. Therefore, it is crucial to consider approximate quantum error-correcting codes [37–39]. The accuracy of QEC codes can be measured in many ways, although, in this chapter, we will focus on the average channel (entanglement) fidelity defined as [40, 41]

$$F_{\text{avg}}(\Phi) := \frac{1}{\dim(\mathcal{X})^2}\langle\langle \mathbb{1}_\mathcal{X}|J(\Phi)|\mathbb{1}_\mathcal{X}\rangle\rangle. \tag{5.1}$$

In this set-up, the goal is to find the best encoding strategy $\mathcal{S} \in \mathcal{C}(\mathcal{X},\mathcal{Y})$ and decoding strategy $\mathcal{R} \in \mathcal{C}(\mathcal{Y},\mathcal{X})$ for a given $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$, such that $F_{\text{avg}}(\mathcal{RES})$ is maximized. If $\mathcal{E} \in \xi_1(\mathcal{X},\mathcal{Y})$, then it is possible to achieve $F_{\text{avg}}(\mathcal{RES}) = 1$.

In the probabilistic approach, the ability to post-process unwanted output states leads us to the multi-objective optimization. On one hand we should minimize the distance between $\Phi = \mathcal{RES} \neq 0$ and $\mathcal{I}_\mathcal{X}$. On the other hand, we should maximize the probability of successful error correction. In this chapter, we will focus on the following quantities:

- the conditional average channel fidelity defined as [44]

$$F_{|\text{avg}}(\Phi) := \frac{1}{\dim(\mathcal{X})}\langle\langle \mathbb{1}_\mathcal{X}|\frac{J(\Phi)}{\text{tr}J(\Phi)}|\mathbb{1}_\mathcal{X}\rangle\rangle, \tag{5.2}$$

- the average probability of success:

$$p_{\text{avg}}(\Phi) := \text{tr}(\Phi(\rho^*_{\mathcal{X}})). \tag{5.3}$$

Note that, if $\Phi \in \mathcal{C}(\mathcal{X})$, then $F_{\text{avg}}(\Phi) = F_{|\text{avg}}(\Phi)$. Below we provide some properties of the introduced measures, that justify their choice.

Proposition 5.1. *Let $0 \neq \Phi \in s\mathcal{C}(\mathcal{X})$. It holds that*

(A) $\left\| \frac{\Phi}{p_{\text{avg}}(\Phi)} - \mathcal{I}_{\mathcal{X}} \right\|_{\diamond} \leq 2\dim(\mathcal{X})\sqrt{1 - F_{|\text{avg}}(\Phi)}$.

(B) For any $\rho \in \mathcal{D}(\mathcal{X})$ we have $|\text{tr}\Phi(\rho) - p_{\text{avg}}(\Phi)| \leq 2\dim(\mathcal{X})\sqrt{1 - F_{|\text{avg}}(\Phi)}p_{\text{avg}}(\Phi)$.

(C) For any $\rho \in \mathcal{D}(\mathcal{X} \otimes \mathcal{Z})$ it holds $\left\| \frac{(\Phi \otimes \mathcal{I}_{\mathcal{Z}})(\rho)}{\text{tr}(\Phi \otimes \mathcal{I}_{\mathcal{Z}})(\rho)} - \rho \right\|_1 \leq 4\dim(\mathcal{X})\sqrt{1 - F_{|\text{avg}}(\Phi)}$.

Proof. (A) By Fuchs-van de Graaf inequality [62] and properties of the diamond norm, for each $\rho \in \mathcal{D}(\mathcal{X})$ we have

$$\begin{aligned} &\left\| (\mathbb{1}_{\mathcal{X}} \otimes \sqrt{\rho^\top})(J(\Phi)/p_{\text{avg}}(\Phi) - J(\mathcal{I}_{\mathcal{X}}))(\mathbb{1}_{\mathcal{X}} \otimes \sqrt{\rho^\top}) \right\|_1 \\ \leq &\dim(\mathcal{X}) \left\| \frac{J(\Phi)}{\text{tr}J(\Phi)} - \frac{J(\mathcal{I}_{\mathcal{X}})}{\dim(\mathcal{X})} \right\|_1 \leq 2\dim(\mathcal{X})\sqrt{1 - F_{|\text{avg}}(\Phi)}. \end{aligned} \tag{5.4}$$

(B) By the data processing inequality and point (A) for each $\rho \in \mathcal{D}(\mathcal{X})$ we have

$$\begin{aligned} \left| \frac{\text{tr}\Phi(\rho)}{p_{\text{avg}}(\Phi)} - 1 \right| &\leq \left\| (\mathbb{1}_{\mathcal{X}} \otimes \sqrt{\rho^\top})(J(\Phi)/p_{\text{avg}}(\Phi) - J(\mathcal{I}_{\mathcal{X}}))(\mathbb{1}_{\mathcal{X}} \otimes \sqrt{\rho^\top}) \right\|_1 \\ &\leq 2\dim(\mathcal{X})\sqrt{1 - F_{|\text{avg}}(\Phi)}. \end{aligned} \tag{5.5}$$

(C) By points (A) and (B) for each $\rho \in \mathcal{D}(\mathcal{X} \otimes \mathcal{Z})$ we have

$$\begin{aligned} &\left\| \frac{(\Phi \otimes \mathcal{I}_{\mathcal{Z}})(\rho)}{\text{tr}(\Phi \otimes \mathcal{I}_{\mathcal{Z}})(\rho)} - \rho \right\|_1 \leq \left\| \frac{\Phi}{p_{\text{avg}}(\Phi)} - \mathcal{I}_{\mathcal{X}} \right\|_{\diamond} + \left| \frac{\text{tr}\Phi(\text{tr}_{\mathcal{Z}}(\rho))}{p_{\text{avg}}(\Phi)} - 1 \right| \\ \leq &4\dim(\mathcal{X})\sqrt{1 - F_{|\text{avg}}(\Phi)}. \end{aligned} \tag{5.6}$$

□

In Chapter 4 we observed that reducing the probability of success allows us to correct errors generated by $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$, such that $\mathcal{E} \notin \xi_1(\mathcal{X}, \mathcal{Y})$. However, the comparison between QEC and pQEC procedures provided there, was qualitative, not quantitative. There are channels $\mathcal{E}$ such that $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$ and $\mathcal{E} \notin \xi_1(\mathcal{X}, \mathcal{Y})$, but we do not know yet if we can construct a good approximate quantum error correction code, *e.g.* a deterministic scheme $(\mathcal{S}, \mathcal{R})$, which will provide $F_{\text{avg}}(\mathcal{R}\mathcal{E}\mathcal{S}) \geq 1 - \epsilon$. The goal of this chapter is to compare QEC and pQEC procedures by the means of introduced fidelity measures.

Example 1

We are given 4 qubits. Each qubit is independently affected by the noise $\mathcal{E}(\rho) = (1-p)\rho + \frac{p}{3}(\sigma_x\rho\sigma_x + \sigma_y\rho\sigma_y + \sigma_z\rho\sigma_z)$, where $p \in [0,1]$ and $\sigma_x, \sigma_y, \sigma_z$ are the Pauli operators. Our goal is to encode and successfully decode one qubit of information. As there is no $[[4,1,3]]_2$ code due to Singleton bound [100], any deterministic error-correcting strategy achieves fidelity $F_{\text{avg}}(\mathcal{R}\mathcal{E}^{\otimes 4}\mathcal{S}) \leq 1 - cp$, where $c > 0$ is some constant which depends on $(\mathcal{S}, \mathcal{R})$.

Let us define $[[4,1,2]]_2$ code, *e.g.* defined by the codewords $|0_S\rangle = \frac{1}{2}|\mathbb{1}_2\rangle\rangle|\mathbb{1}_2\rangle\rangle$, $|1_S\rangle = \frac{1}{2}|\sigma_x\rangle\rangle|\sigma_x\rangle\rangle$, where $\mathcal{S} = \mathcal{K}\left((|0_S\rangle\langle 0| + |1_S\rangle\langle 1|)\right)$. The decoding $\mathcal{R}$ is given as $\mathcal{R} = \mathcal{S}^\dagger$. This probabilistic error-correcting strategy achieves $p_{\text{avg}}(\mathcal{R}\mathcal{E}^{\otimes 4}\mathcal{S}) \geq (1-p)^4 \geq 1 - 4p$ and $F_{|\text{avg}}(\mathcal{R}\mathcal{E}^{\otimes 4}\mathcal{S}) \geq \frac{1}{4}\frac{4(1-p)^4}{1-4p(1-p)^3} \geq 1 - 11p^2$.

The introduced example indicates that it is possible to significantly increase the fidelity by lowering the probability of success. In fact, there is a trade-off between $F_{|\text{avg}}(\mathcal{R}\mathcal{E}\mathcal{S})$ and $p_{\text{avg}}(\mathcal{R}\mathcal{E}\mathcal{S})$; the high value of $p_{\text{avg}}(\mathcal{R}\mathcal{E}\mathcal{S})$ will decrease the value of $F_{|\text{avg}}(\mathcal{R}\mathcal{E}\mathcal{S})$ and vice versa. The best possible error-correcting schemes $(\mathcal{S}, \mathcal{R})$ form a Pareto front as presented in Figure 5.1.

5.2 Noise model and problem formulation

In this chapter we consider noise quantum channels of the form

$$\mathcal{E}_p = (1-p)\mathcal{I}_{\mathcal{Y}} + p\mathcal{E}, \tag{5.7}$$

where $p \in [0,1]$ and $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$. We assume that p is an unknown parameter, that is, the encoding operation $\mathcal{S} \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ and decoding operation $\mathcal{R} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ should be independent of p. We will provide a construction of error-correcting schemes $(\mathcal{S}, \mathcal{R})$, which achieve a high value of $F_{\text{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$ for the deterministic decoding $\mathcal{R} \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$, and $F_{|\text{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$ for the probabilistic decoding $\mathcal{R} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$. Additionally, we require that if no error occurred ($p = 0$), the scheme $(\mathcal{S}, \mathcal{R})$ will perfectly recover the initial information with some positive probability of success, that is

$$\mathcal{R}\mathcal{S} = q\mathcal{I}_{\mathcal{X}} \neq 0. \tag{5.8}$$

Lowering the value of q may improve the fidelity even if the channel $\mathcal{E}$ is mixed unitary and self-adjoint.

Example 2

Let us consider $\mathcal{X} = \mathbb{C}^2, \mathcal{Y} = \mathbb{C}^3$ and $\mathcal{E}_p \in \mathcal{C}(\mathcal{Y})$

$$\mathcal{E}_p(Y) = (1-p)H_0 Y H_0 + p\left(\frac{1}{2}H_1 Y H_1 + \frac{1}{2}H_2 Y H_2\right), \tag{5.9}$$

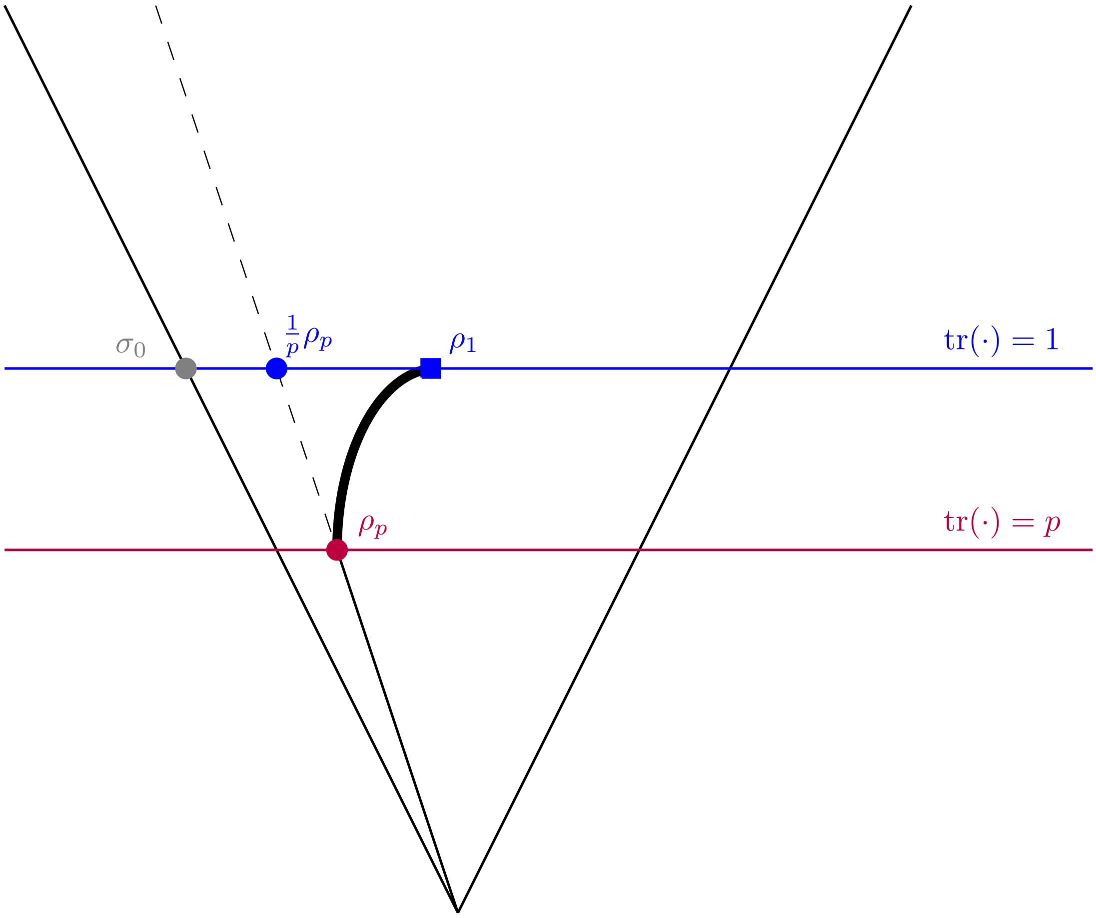


Figure 5.1: A visualization of a Pareto front (thick line) of outputs of possible error-correcting schemes $(\mathcal{S}, \mathcal{R})$, calculated for the input state σ_0. The state $\sigma_0 \in \mathcal{D}(\mathcal{X})$ is encoded into $\mathcal{S}(\sigma_0)$ and then sent through a noise channel $\mathcal{E}$. At one end of the Pareto front, the best deterministic recovery operation $\mathcal{R}$ transfers $\mathcal{ES}(\sigma_0)$ to $\mathcal{RES}(\sigma_0) = \rho_1$. At the other end of the Pareto front, the best probabilistic recovery operation (maximizing $F_{|\mathrm{avg}}$) transfers $\mathcal{ES}(\sigma_0)$ to $\mathcal{RES}(\sigma_0) = \rho_p$ with the probability of success $\mathrm{tr}(\rho_p) = p$. As we can see, σ_0 is closer to the post-processed state $\frac{1}{p}\rho_p$, than ρ_1.

where

$$H_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, H_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \tag{5.10}$$

First, let us take $\mathcal{S} = \mathcal{K}((S))$ and $\mathcal{R} = \mathcal{S}^\dagger$, where $S \in \mathcal{U}(\mathcal{X}, \mathcal{Y})$. Due to Theorem 4.4, for each p we have $\mathcal{R}\mathcal{E}_p\mathcal{S} \propto \mathcal{I}_\mathcal{X}$ if and only if $S^\dagger H_i S \propto 1\!\!1_\mathcal{X}$ for $i = 0, 1, 2$. The latter holds if $S^\dagger(H_0 - H_1)S \propto 1\!\!1_\mathcal{X}$ and $S^\dagger(H_0 + H_1)S \propto 1\!\!1_\mathcal{X}$, which means $S = \begin{pmatrix} U \\ 0 \end{pmatrix}$, where $U \in \mathcal{U}(\mathcal{X})$. That implies $S^\dagger H_2 S = U^\dagger|1\rangle\langle 1|U \propto 1\!\!1_\mathcal{X}$, which is impossible. Hence, there is no scheme $(\mathcal{S}, \mathcal{S}^\dagger)$ such that $\mathcal{S}^\dagger\mathcal{E}_p\mathcal{S} \propto \mathcal{I}_\mathcal{X}$.

Now, let us take different encoding and decoding bases, *e.g.* $S = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $R = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ and define $\mathcal{S} = \mathcal{K}((S))$, $\mathcal{R} = \mathcal{K}((R))$. For each p it holds $\mathcal{R}\mathcal{E}_p\mathcal{S} = \frac{1}{2}\mathcal{I}_\mathcal{X}$. Alternatively, the best deterministic decoding strategy $\widetilde{\mathcal{R}}$ for a given $\mathcal{S}$ satisfying $\widetilde{\mathcal{R}}\mathcal{S} = \mathcal{I}_\mathcal{X}$, is equal to $\widetilde{\mathcal{R}}(Y) = S^\dagger Y S + |0\rangle\langle 0|\langle 2|Y|2\rangle$. For each p it holds $\widetilde{\mathcal{R}}\mathcal{E}_p\mathcal{S} = (1 - \frac{1}{2}p)\mathcal{I}_\mathcal{X} + \frac{1}{2}p\sum_{i=0}^{1}\mathcal{K}((|i\rangle\langle i|))$. Hence, the Pareto front for $\mathcal{E}_p\mathcal{S}$ contains two extremal points:

- the point maximizing the probability of success $(p_{\text{avg}}, F_{|\text{avg}}) = (1, 1 - p/4)$,
- the point maximizing the fidelity $(p_{\text{avg}}, F_{|\text{avg}}) = (1/2, 1)$.

5.3 Construction of approximate codes

In this section we provide a numerically efficient algorithm that constructs approximate error-correcting schemes for a given noise quantum channels $\mathcal{E}_p$ defined as in Eq. (5.7) and a given input space $\mathcal{X}$. The output of this procedure will consist of an encoding channel $\mathcal{S} = \mathcal{K}((S)) \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ and a decoding operation: deterministic $\mathcal{R}_1 \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$, probabilistic $\mathcal{R}_0 = \mathcal{K}((R)) \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$, such that $\mathcal{R}_1\mathcal{S} = \mathcal{I}_\mathcal{X}$ and $0 \neq \mathcal{R}_0\mathcal{S} \propto \mathcal{I}_\mathcal{X}$. The construction of $(\mathcal{S}, \mathcal{R}_1)$ and $(\mathcal{S}, \mathcal{R}_0)$ is focused on maximizing $F_{\text{avg}}(\mathcal{R}_1\mathcal{E}_p\mathcal{S})$ and $F_{|\text{avg}}(\mathcal{R}_0\mathcal{E}_p\mathcal{S})$, respectively. This procedure is based on the proof of Theorem 4.25 and the proof of Proposition 4.23.

5.3.1 Encoding $\mathcal{S}$

First, we construct $\mathcal{S} = \mathcal{K}((S))$, where $S \in \mathcal{U}(\mathcal{X}, \mathcal{Y})$. We assume, that we are given the Kraus decomposition of $\mathcal{E}$ (not necessarily a canonical decomposition), that is $\mathcal{E} = \mathcal{K}\left((E_i)_{i=1}^{\text{rank}(J(\mathcal{E}))}\right)$.

Definition of an operation $\mathcal{F}$

In the first step, we sort the Kraus operators of $\mathcal{E}$ in a such way that

$$\mathrm{tr}(E_i^\dagger E_i) \geq \mathrm{tr}(E_{i+1}^\dagger E_{i+1}), \tag{5.11}$$

for $i = 1, \ldots, \mathrm{rank}(J(\mathcal{E})) - 1$. The quality of the construction will increase if we consider a canonical decomposition of $\mathcal{E}$. Next, let us fix $r' \in \mathbb{N}$, such that $r'^2(\dim(\mathcal{X}) - 1) < \dim(\mathcal{Y}) \leq (r' + 1)^2(\dim(\mathcal{X}) - 1)$ and

$$r = \min(r', \mathrm{rank}(J(\mathcal{E})) + 1). \tag{5.12}$$

Define $\mathcal{F} = \mathcal{K}\left((F_i)_{i=1}^r\right)$, where $F_1 = \mathbb{1}_\mathcal{Y}, F_2 = E_1, \ldots, F_r = E_{r-1}$. Observe that $\mathrm{rank}(\mathcal{F}^\dagger(\mathbb{1}_\mathcal{Y})) = \dim(\mathcal{Y})$.

Initialization of random decoding

According to Theorem 4.25 and Lemma 4.7 it holds $\mathcal{F} \in \xi(\mathcal{X}, \mathcal{Y})$. Following the proof of Theorem 4.25 we define

$$A = \left\{P \in \mathcal{P}(\mathcal{Y}) : \mathrm{rank}(\mathcal{F}^\dagger(P)) = \dim(\mathcal{Y})\right\}. \tag{5.13}$$

The set A is non-empty. We need to find $P_0 \in A$, such that for all $P \in \mathcal{P}(\mathcal{Y})$ satisfying $\mathrm{rank}(P) < \mathrm{rank}(P_0)$ it holds $\mathrm{rank}(\mathcal{F}^\dagger(P)) < \dim(\mathcal{Y})$. We will find P_0 indirectly by sampling Ginibre matrices and utilizing the following lemma.

Lemma 5.2 ([2]). *Let $\mathcal{F} \in \mathcal{C}(\mathcal{Y})$. Assume that there exists $P_0 \in \mathcal{P}(\mathcal{Y})$ such that* $\mathrm{rank}(\mathcal{F}^\dagger(P_0)) = \dim(\mathcal{Y})$. *Let us consider a random complex Ginibre matrix $G \in \mathcal{M}(\mathbb{C}^{\mathrm{rank}(P_0)}, \mathcal{Y})$ Then, almost surely it holds* $\mathrm{rank}(\mathcal{F}^\dagger(GG^\dagger)) = \dim(\mathcal{Y})$.

Proof. This lemma follows from the observation, that $P(\det(\mathbb{1}_\mathcal{Y} + cM) = 0) = 0$ for any matrix $M \in \mathcal{M}(\mathcal{Y})$ and an absolutely continuous random variable $c \in \mathbb{R}$. □

Let $s = \left\lceil \frac{\dim(\mathcal{Y})}{r} \right\rceil \geq 2$, $G_s \in \mathcal{M}(\mathbb{C}^s, \mathcal{Y})$ be a complex Ginibre matrix. If $\mathrm{rank}(\mathcal{F}^\dagger(G_s G_s^\dagger)) = \dim(\mathcal{Y})$, we have found $P_0 = G_s G_s^\dagger$ and we can proceed to the next step carrying the random decoding matrix G_s. Otherwise, we increase $s \leftarrow s + 1$ and define $G_s \leftarrow \begin{pmatrix} G_s & v \end{pmatrix}$, where $v \in \mathcal{Y}$ is a random complex Gaussian vector. We repeat the process until we will find the appropriate G_s, that is G_s satisfying $\mathrm{rank}(\mathcal{F}^\dagger(G_s G_s^\dagger)) = \dim(\mathcal{Y})$.

Diagonalization of $\mathcal{F}$

For each $i = 1, \ldots, s$ we define $\Pi_i = \left(\mathcal{F}^\dagger(G_s(\mathbb{1}_{\mathbb{C}^s} - |i\rangle\langle i|)G_s^\dagger)\right)^0$. Then, we take

$$|w_i\rangle = (\mathbb{1}_\mathcal{Y} - \Pi_i)G_s|i\rangle \tag{5.14}$$

and define $W_s = \sum_{i=1}^s |w_i\rangle\langle i| \in \mathcal{M}(\mathbb{C}^s, \mathcal{Y})$. Let $D_b = G_s^\dagger F_b W_s \in \mathcal{M}(\mathbb{C}^s)$. The matrices D_b are by the construction diagonal. Moreover, by the definition of G_s it holds that $\operatorname{rank}(\Pi_i) < \dim(\mathcal{Y})$ for any i. In addition to that, as G_s is a complex Ginibre matrix, the projector Π_i and the Gaussian vector $G_s|i\rangle$ are independent random variables. Therefore, for $i = 1, \ldots, s$ almost surely it holds $|w_i\rangle \neq 0$ and as a consequence

$$\operatorname{rank}(D_1) = \operatorname{rank}(G_s^\dagger F_1 W_s) = \operatorname{rank}(G_s^\dagger W_s) = s. \tag{5.15}$$

The assumptions of Proposition 4.23 are satisfied and we can proceed to the final step.

Definition of S

According to the proof of Proposition 4.23 an encoding operation for $\mathcal{K}\left((D_b)_{b=1}^r\right)$ can be taken as an appropriate binary matrix, such that in each row there is at most one non-zero element. Hence, in this step of construction, we will define $\mathrm{IDX}_j \subset \{1, \ldots, s\}$ for $j = 1, \ldots, \dim(\mathcal{X})$, such that $\mathrm{IDX}_{j_1} \cap \mathrm{IDX}_{j_2} = \emptyset$ for $j_1 \neq j_2$. Let us define $M = (M_{a,b})_{\substack{a=1,\ldots,s\\ b=1,\ldots,r}} \in \mathcal{M}(\mathbb{C}^r, \mathbb{C}^s)$, where $M_{a,b} = (D_b)_{a,a}$. For each $j = 1, \ldots, \dim(\mathcal{X})$ do the following: consider

$$\texttt{temp\_rank} = \operatorname{rank}\left((M_{a,b})_{\substack{a\in\{1,\ldots,s\}\setminus\bigcup_{k=1}^{j-1}\mathrm{IDX}_k\\ b=1,\ldots,r}}\right). \tag{5.16}$$

Then, find `temp_rank` linearly independent vectors among $(M_{z_i,b})_{b=1,\ldots,r}$, where $z_i \in \{1, \ldots, s\}\setminus\bigcup_{k=1}^{j-1}\mathrm{IDX}_k$ and $i = 1, \ldots, \texttt{temp\_rank}$. Eventually, define $\mathrm{IDX_j} = \{z_1, \ldots, z_{\texttt{temp\_rank}}\}$.

We update $\mathrm{IDX}_{\dim(\mathcal{X})} \leftarrow \{z\}$, where $z \in \mathrm{IDX}_{\dim(\mathcal{X})}$ and introduce $S \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ defined as

$$S|j\rangle = \sum_{i\in\mathrm{IDX}_j} W_s|i\rangle, \tag{5.17}$$

where $j = 1, \ldots, \dim(\mathcal{X})$. As $\operatorname{rank}(W_s) = s$ it holds that $\operatorname{rank}(S) = \dim(\mathcal{X})$. To satisfy the assumption $S \in \mathcal{U}(\mathcal{X}, \mathcal{Y})$, we normalize S according to

$$S \leftarrow S(S^\dagger S)^{-1/2}. \tag{5.18}$$

5.3.2 Decoding $\mathcal{R}$

Deterministic

For $\mathcal{S} = \mathcal{K}((S)) \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$, we will construct $\mathcal{R}_1 \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$, such that $\mathcal{R}_1\mathcal{S} = \mathcal{I}_\mathcal{X}$ and the value of $F_{\text{avg}}(\mathcal{R}_1\mathcal{E}\mathcal{S})$ is relatively high. To make sure that $\mathcal{R}_1\mathcal{S} = \mathcal{I}_\mathcal{X}$ it must hold

$$\mathcal{R}_1(Y) = \mathcal{S}^\dagger(Y) + \mathcal{R}(\Pi Y \Pi), \tag{5.19}$$

where $\mathcal{R} \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$ and $\Pi = \mathbb{1}_\mathcal{Y} - SS^\dagger$. For $\mathcal{F}(X) = \Pi(\mathcal{E}\mathcal{S}(X))\Pi$ we define $\mathcal{R}$ to be the Petz recovery map [101,102]

$$\mathcal{R}(Y) = \mathcal{F}^\dagger(\mathcal{F}(\mathbb{1}_\mathcal{X})^{-1/2} Y \mathcal{F}(\mathbb{1}_\mathcal{X})^{-1/2}). \tag{5.20}$$

Note, that $\mathcal{R} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ may be not be trace-preserving, but adding the part $\operatorname{tr}\left((\mathbb{1}_\mathcal{Y} - (\mathcal{F}(\mathbb{1}_\mathcal{X}))^0)Y\right)\rho^*_\mathcal{X}$ to $\mathcal{R}$ will not change the value of $F_{\text{avg}}(\mathcal{R}_1\mathcal{E}_p\mathcal{S})$ for any $p \in [0,1]$. Eventually, we define

$$\mathcal{R}_1(Y) = \mathcal{S}^\dagger(Y) + \mathcal{F}^\dagger(\mathcal{F}(\mathbb{1}_\mathcal{X})^{-1/2} Y \mathcal{F}(\mathbb{1}_\mathcal{X})^{-1/2}). \tag{5.21}$$

Probabilistic

For $\mathcal{S} = \mathcal{K}((S)) \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$, we will construct $\mathcal{R}_0 \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$, such that $0 \neq \mathcal{R}_0\mathcal{S} \propto \mathcal{I}_\mathcal{X}$ and the value of $F_{|\text{avg}}(\mathcal{R}_0\mathcal{E}_p\mathcal{S})$ is relatively high for all $p \in [0,1]$.

This construction has two parameters: $q \in (0,1)$ and $\epsilon > 0$. First, we can choose arbitrary $q \in (0,1)$ and define $\mathcal{F} = \mathcal{E}_q\mathcal{S}$. For relatively small values of $q \simeq 0$, this procedure will return $\mathcal{R}_0$ such that $p_{\text{avg}}(\mathcal{R}_0\mathcal{S})$ is high, but $F_{|\text{avg}}(\mathcal{R}_0\mathcal{E}_p\mathcal{S})$ tends to be small for $p \simeq 1$. Analogously, if we take $q \simeq 1$, then $p_{\text{avg}}(\mathcal{R}_0\mathcal{S})$ will be small, but $F_{|\text{avg}}(\mathcal{R}_0\mathcal{E}_p\mathcal{S})$ will be greater for $p \simeq 1$.

Second, we can choose $\epsilon > 0$. As there are noise channels, such that the optimal $\mathcal{R}_0$ satisfies $\mathcal{R}_0\mathcal{S} = 0$, we need ϵ to determine which values of $p_{\text{avg}}(\mathcal{R}_0\mathcal{S})$ are acceptable and which of them are in the machine epsilon's range.

Let $\Pi_s = \frac{1}{\dim(\mathcal{X})}|S\rangle\rangle\langle\langle S| + (\mathbb{1}_\mathcal{Y} - SS^\dagger) \otimes \mathbb{1}_\mathcal{X}$. Let $|v_0\rangle$ be the eigenvector of the largest eigenvalue of

$$\left(\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S\right)^{-1/2} J(\mathcal{F}) \left(\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S\right)^{-1/2}. \tag{5.22}$$

Define $R \in \mathcal{M}(\mathcal{Y}, \mathcal{X})$ by the equation

$$|R^\dagger\rangle\rangle = \left(\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S\right)^{-1/2} |v_0\rangle. \tag{5.23}$$

If $\left|\langle\langle S|R^\dagger\rangle\rangle\right| < \epsilon$, then use the substitution $|R^\dagger\rangle\rangle \leftarrow (1-\epsilon)|R^\dagger\rangle\rangle + \epsilon|S\rangle\rangle$. Normalize R according to $R \leftarrow \frac{R}{\|R\|_\infty}$ and define $\mathcal{R}_0 = \mathcal{K}((R))$.

5.3.3 Algorithm

Let us summarize the provided construction as an algorithm:

Algorithm 3: Construction of approximate codes based on Theorem 4.25

Input: $\mathcal{E} = \mathcal{K}\left((E_i)_{i=1}^{\mathrm{rank}(J(\mathcal{E}))}\right) \in \mathcal{C}(\mathcal{Y})$, vector space $\mathcal{X}$.

Output: $\mathcal{S} = \mathcal{K}((S)) \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$, $\mathcal{R}_1 \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$ and $\mathcal{R}_0 \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$: $\mathcal{R}_1\mathcal{S} = \mathcal{I}_\mathcal{X}$, $0 \neq \mathcal{R}_0\mathcal{S} \propto \mathcal{I}_\mathcal{X}$ and the values $F_{\mathrm{avg}}(\mathcal{R}_1\mathcal{E}_p\mathcal{S})$, and $F_{|\mathrm{avg}}(\mathcal{R}_0\mathcal{E}_p\mathcal{S})$ are relatively high, for $\mathcal{E}_p = (1-p)\mathcal{I}_\mathcal{Y} + p\mathcal{E}$, $p \in [0,1]$.

Sort Kraus operators $\mathcal{E} = \mathcal{K}\left((E_i)_{i=1}^{\dim(\mathcal{Y})^2}\right)$: such that $\mathrm{tr}(E_i^\dagger E_i) \geq \mathrm{tr}(E_{i+1}^\dagger E_{i+1})$ and $E_i = 0$ for $i > \mathrm{rank}(J(\mathcal{E}))$.
Define $r = \left\lceil \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1}} \right\rceil$.
Define $\mathcal{F} = \mathcal{K}((F_i)_{i=1}^r) : F_1 = \mathbb{1}_\mathcal{Y}, F_2 = E_1, \ldots, F_r = E_{r-1}$.
for $s = \left\lceil \frac{\dim(\mathcal{Y})}{r} \right\rceil$ **to** $\dim(\mathcal{Y})$ **do**
  Define $G_s \in \mathcal{M}(\mathbb{C}^s, \mathcal{Y})$ - complex Ginibre matrix.
  if $\mathrm{rank}(\mathcal{F}^\dagger(G_sG_s^\dagger)) = \dim(\mathcal{Y})$ **then**
    break
  end
end
for $i = 1$ **to** s **do**
  Define $\Pi_i = \left[\mathcal{F}^\dagger(G_s(\mathbb{1}_{\mathbb{C}^s} - |i\rangle\langle i|)G_s^\dagger)\right]^0$.
  Define $|w_i\rangle = (\mathbb{1}_\mathcal{Y} - \Pi_i)G_s|i\rangle$.
end
Let $W_s = \sum_{i=1}^s |w_i\rangle\langle i|$ and $M = (M_{a,b})_{\substack{a=1,\ldots,s\\ b=1,\ldots,r}}$, where $M_{a,b} = \langle a|G_s^\dagger F_b W_s|a\rangle$.
for $j = 1$ **to** $\dim(\mathcal{X})$ **do**
  Define `temp_rank` = $\mathrm{rank}\left((M_{a,b})_{\substack{a\in\{1,\ldots,s\}\setminus\bigcup_{k=1}^{j-1}\mathrm{IDX}_k\\ b=1,\ldots,r}}\right)$.
  Find `temp_rank` linearly independent vectors among $(M_{z_i,b})_{b=1,\ldots,r}$, where $z_i \in \{1,\ldots,s\}\setminus\bigcup_{k=1}^{j-1}\mathrm{IDX}_k$ and $i = 1,\ldots,$ `temp_rank`.
  Define $\mathrm{IDX_j} = \{z_1, \ldots, z_{\texttt{temp\_rank}}\}$.
end
Update $\mathrm{IDX}_{\dim(\mathcal{X})} \leftarrow \{z\}$, where $z \in \mathrm{IDX}_{\dim(\mathcal{X})}$.
Define $S = \sum_{j=1}^{\dim(\mathcal{X})} \sum_{i\in\mathrm{IDX}_j} W_s|i\rangle\langle j|$.
Let $S \leftarrow S(S^\dagger S)^{-1/2}$ and define $\mathcal{S} = \mathcal{K}((S))$.
Let $P_1 = \mathbb{1}_\mathcal{Y} - SS^\dagger$ and $\Phi_1(X) = P_1\mathcal{E}\mathcal{S}(X)P_1$.
Define $\mathcal{R}_1(Y) = \mathcal{S}^\dagger(Y) + \Phi_1^\dagger(\Phi_1(\mathbb{1}_\mathcal{X})^{-1/2} Y \Phi_1(\mathbb{1}_\mathcal{X})^{-1/2})$.
Fix $q \in (0,1)$ and $\epsilon > 0$ and define $\Phi_0 = \mathcal{E}_q\mathcal{S}$.
Let $P_0 = \frac{1}{\dim(\mathcal{X})}|S\rangle\rangle\langle\langle S| + (\mathbb{1}_\mathcal{Y} - SS^\dagger) \otimes \mathbb{1}_\mathcal{X}$.
Let $|v_0\rangle$ be the eigenvector of the largest eigenvalue of $\left(P_0(\Phi_0(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})P_0\right)^{-1/2} J(\Phi_0) \left(P_0(\Phi_0(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})P_0\right)^{-1/2}$.
Define $R \in \mathcal{M}(\mathcal{Y}, \mathcal{X})$ according to $|R^\dagger\rangle\rangle = \left(P_0(\Phi_0(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})P_0\right)^{-1/2} |v_0\rangle$.
if $|\langle\langle S|R^\dagger\rangle\rangle| < \epsilon$ **then**
  $|R^\dagger\rangle\rangle \leftarrow (1-\epsilon)|R^\dagger\rangle\rangle + \epsilon|S\rangle\rangle$.
end
Define $\mathcal{R}_0 = \mathcal{K}\left(\left(\frac{R}{\|R\|_\infty}\right)\right)$.
Result: $\mathcal{S}, \mathcal{R}_1, \mathcal{R}_0$

5.3.4 Complexity

Let us discuss briefly the complexity of provided construction. Let $y = \dim(\mathcal{Y})$. To sort Kraus operators of $\mathcal{E}$ we need $\mathcal{O}(y^4)$ operations. We can improve efficiency of the construction by considering sorted Kraus operations from a canonical decomposition. However, it has the complexity $\mathcal{O}(y^6)$. Finding the matrix G_s has the worst-case complexity of $\mathcal{O}(y^5)$. The same holds for the matrix W_s. Finding the partition of M can be roughly upper bounded by $\mathcal{O}(y^3)$. Finally, to find $\mathcal{R}_1$ we need $\mathcal{O}(y^5)$ operations and to find $\mathcal{R}_0$ we need $\mathcal{O}(y^6)$ operations in the worst-case scenario.

For comparison, computing $\mathcal{E}(\rho)$, where $\mathcal{E}$ is given in the Kraus decomposition, requires $\mathcal{O}(y^5)$ operations in the worst-case scenario.

5.3.5 Comments

The presented construction of approximate codes returns in general suboptimal schemes. Finding the optimal error-correcting scheme $(\mathcal{S}, \mathcal{R})$ is in fact numerically demanding problem - the optimization domain of F_{avg} is more or less isomorphic to the set of separable states. The best approach so far to find the optimal solution is to utilize SDP hierarchy and de Finetti theorem [62] as presented in [103].

Nevertheless, the performance of our construction can be improved. One may repeat the procedure of sampling Ginibre matrices G_s and take the best result. It was observed [40] that random codes, especially for high-dimensional $\mathcal{Y}$ are highly suitable for QEC. Note, that our construction of $\mathcal{S}$ is not purely random, but is tuned for noise channels $\mathcal{E}$ with low Choi rank. In particular, if $(\text{rank}(J(\mathcal{E})) + 1)^2(\dim(\mathcal{X}) - 1) < \dim(\mathcal{Y})$, then the proposed scheme $(\mathcal{S}, \mathcal{R}_0)$ achieves $F_{|\text{avg}}(\mathcal{R}_0\mathcal{E}_p\mathcal{S}) = 1$ for all p and satisfies $0 \neq \mathcal{R}_0\mathcal{S} \propto \mathcal{I}_{\mathcal{X}}$.

The performance of our construction can also be improved by iteratively optimizing the decoding $\mathcal{R}$ for a fixed $\mathcal{S}$ and optimizing the encoding $\mathcal{S}$ for a fixed $\mathcal{R}$ [104–106]. More specifically, let us focus our attention on the function $F_{\text{avg}}(\mathcal{R}\mathcal{E}\mathcal{S})$. Let $\mathcal{S}^{(1)} = \mathcal{S}$ be the encoding provided by our construction. We can calculate the best decoding operation $\mathcal{R}^{(1)}$, that is maximizing $\mathcal{R} \mapsto F_{\text{avg}}(\mathcal{R}\mathcal{E}\mathcal{S}^{(1)})$, by using SDP programming [105]. Then, we may find the best encoding operation $\mathcal{S}^{(2)}$, that is maximizing $\mathcal{S} \mapsto F_{\text{avg}}(\mathcal{R}^{(1)}\mathcal{E}\mathcal{S})$, by using SDP programming [105]. We repeat this process until we reach the locally optimal error-correcting scheme.

Finally, we may ask if for a given $\mathcal{S}$ the construction of $\mathcal{R}_0$ and $\mathcal{R}_1$ is justified. The deterministic decoding $\mathcal{R}_1$ was defined by using Petz recovery map. It is known, that this decoding operation is generally suboptimal, but guarantees relatively high fidelity [101]. Its greatest advantage is simple and computationally efficient construction. For the same reasons, we could use decoding operations defined by iterative algorithm EigQER [107].

For low-dimensional spaces $\mathcal{Y}$ we can find the best recovery $\mathcal{R}_1$ in reasonable time by using SDP programming and the following proposition.

Proposition 5.4 ([108]). *Let $\mathcal{E}_p = (1-p)\mathcal{I}_\mathcal{Y} + p\mathcal{E}$, where $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ and $p \in [0,1]$. For a given encoding operation $\mathcal{S} = \mathcal{K}((S)) \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$, the decoding operation $\mathcal{R}_1 \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$ which maximizes $F_{\text{avg}}(\mathcal{R}_1\mathcal{E}_p\mathcal{S})$ for any $p \in [0,1]$ and satisfies $\mathcal{R}_1\mathcal{S} = \mathcal{I}_\mathcal{X}$ is given as*

$$\mathcal{R}_1(Y) = \mathcal{S}^\dagger(Y) + \mathcal{R}(\Pi Y \Pi), \tag{5.24}$$

where $\Pi = \mathbb{1}_\mathcal{Y} - SS^\dagger$ and $\mathcal{R}$ is calculated via SDP optimization

$$J(\mathcal{R}^\dagger) = \operatorname{argmax}\left\{\operatorname{tr}(QJ(\mathcal{F})) : Q \in \mathcal{P}(\mathcal{Y}\otimes\mathcal{X}), \operatorname{tr}_\mathcal{X}(Q) = \mathbb{1}_\mathcal{Y}\right\}, \tag{5.25}$$

where $\mathcal{F}(X) = \Pi(\mathcal{E}\mathcal{S}(X))\Pi$.

Proof. To satisfy $\mathcal{R}_1\mathcal{S} = \mathcal{I}_\mathcal{X}$ we need to consider $\mathcal{R}_1 \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$ in the form $\mathcal{R}_1(Y) = \mathcal{S}^\dagger(Y) + \mathcal{R}(\Pi Y \Pi)$, where $\mathcal{R} \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$. Then, we obtain

$$F_{\text{avg}}(\mathcal{R}_1\mathcal{E}_p\mathcal{S}) = (1-p) + pF_{\text{avg}}(\mathcal{R}_1\mathcal{E}\mathcal{S}) = (1-p) + pF_{\text{avg}}(\mathcal{S}^\dagger\mathcal{E}\mathcal{S}) + pF_{\text{avg}}(\mathcal{R}\mathcal{F}). \tag{5.26}$$

Therefore, we need to optimize $F_{\text{avg}}(\mathcal{R}\mathcal{F})$ over quantum channels $\mathcal{R} \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$. It holds

$$\begin{aligned} &\max\left\{F_{\text{avg}}(\mathcal{R}\mathcal{F}) : \mathcal{R} \in \mathcal{C}(\mathcal{Y}, \mathcal{X})\right\} = \max\left\{\frac{1}{\dim(\mathcal{X})^2}\operatorname{tr}(J(\mathcal{R}^\dagger)J(\mathcal{F})) : \mathcal{R} \in \mathcal{C}(\mathcal{Y}, \mathcal{X})\right\} \\ &= \max\left\{\frac{1}{\dim(\mathcal{X})^2}\operatorname{tr}(QJ(\mathcal{F})) : Q \in \mathcal{P}(\mathcal{Y}\otimes\mathcal{X}), \operatorname{tr}_\mathcal{X}(Q) = \mathbb{1}_\mathcal{Y}\right\}. \end{aligned} \tag{5.27}$$

□

The construction of $\mathcal{R}_0$ is not as straightforward as $\mathcal{R}_1$. First, it may happen that the best probabilistic decoding $\mathcal{R}_0$, maximizing $\mathcal{R} \mapsto F_{|\text{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$, satisfies $\mathcal{R}_0\mathcal{S} = 0$. Hence, $\mathcal{R}_0$ can be optimal only up to some level of tolerance $\epsilon > 0$. Second, unlike $\mathcal{R}_1$, there may not exist universal $\mathcal{R}_0$, which is the best for all noise channels $\mathcal{E}_p$, where $p \in [0,1]$.

The construction of $\mathcal{R}_0$ provided in our procedure is optimal for $\mathcal{E}_q$, where $q \in (0,1)$ is the parameter of this procedure. We justify this claim in the next proposition.

Proposition 5.5. *Let $\mathcal{E}_q = (1-q)\mathcal{I}_\mathcal{Y} + q\mathcal{E}$, where $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ and $q < 1$. For a given encoding operation $\mathcal{S} = \mathcal{K}((S)) \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ define $\mathcal{F} = \mathcal{E}_q\mathcal{S}$. Then, it holds*

$$\begin{aligned} &\sup\left\{F_{|\text{avg}}(\mathcal{R}\mathcal{F}) : 0 \neq \mathcal{R}\mathcal{S} \propto \mathcal{I}_\mathcal{X}, \mathcal{R} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})\right\} \\ &= \frac{1}{\dim(\mathcal{X})}\left\|(\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X})\otimes\mathbb{1}_\mathcal{X})\Pi_S)^{-1/2} J(\mathcal{F}) (\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X})\otimes\mathbb{1}_\mathcal{X})\Pi_S)^{-1/2}\right\|_\infty, \end{aligned} \tag{5.28}$$

where $\Pi_S = \frac{1}{\dim(\mathcal{X})}|S\rangle\rangle\langle\langle S| + (\mathbb{1}_\mathcal{Y} - SS^\dagger) \otimes \mathbb{1}_\mathcal{X}$.

The sequence of decoding operations $(\mathcal{R}_n)_{n\in\mathbb{N}} \subset s\mathcal{C}(\mathcal{Y},\mathcal{X})$ *that in the limit are optimal is defined by a sequence* $\left(\mathcal{K}\left(\left(\frac{R_n}{\|R_n\|_\infty}\right)\right)\right)_{n\in\mathbb{N}}$, *where*

$$|R_n^\dagger\rangle\rangle = (1 - 1/n) \left(\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S\right)^{-1/2} |v_0\rangle + 1/n|S\rangle\rangle \tag{5.29}$$

and $|v_0\rangle$ *is the eigenvector of the leading eigenvalue of*

$$\left(\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S\right)^{-1/2} J(\mathcal{F}) \left(\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S\right)^{-1/2}. \tag{5.30}$$

Proof. Let us fix $\mathcal{R} \in s\mathcal{C}(\mathcal{Y},\mathcal{X})$ such that $0 \neq \mathcal{R}\mathcal{S} \propto \mathcal{I}_\mathcal{X}$. In particular, that implies $\mathcal{R}\mathcal{F} \neq 0$ and we can calculate $F_{|\text{avg}}(\mathcal{R}\mathcal{F})$. Take a Kraus decomposition $\mathcal{R} = \mathcal{K}((R_i)_i)$ for which $R_i S \neq 0$ for all i. Let us consider $\mathcal{R}_i = \mathcal{K}((R_i))$ for i satisfying $\mathcal{R}_i\mathcal{F} \neq 0$. For any $a_i \geq 0$ and $b_i > 0$ it holds $\frac{\sum_i a_i}{\sum_i b_i} \leq \max_i \frac{a_i}{b_i}$. Hence,

$$F_{|\text{avg}}(\mathcal{R}\mathcal{F}) \leq \max_i F_{|\text{avg}}(\mathcal{R}_i\mathcal{F}). \tag{5.31}$$

From now, we consider the probabilistic decoding operation of the form $\mathcal{R} = \mathcal{K}((R)) \in s\mathcal{C}(\mathcal{Y},\mathcal{X})$. Let $\Pi_S = \frac{1}{\dim(\mathcal{X})}|S\rangle\rangle\langle\langle S| + (\mathbb{1}_\mathcal{Y} - SS^\dagger) \otimes \mathbb{1}_\mathcal{X}$. It holds that $RS \propto \mathbb{1}_\mathcal{X}$ if and only if $\Pi_S|R^\dagger\rangle\rangle = |R^\dagger\rangle\rangle$. Therefore,

$$\begin{aligned}
&\sup\left\{F_{|\text{avg}}(\mathcal{R}\mathcal{F}) : 0 \neq \mathcal{R}\mathcal{S} \propto \mathcal{I}_\mathcal{X}\right\}\\
=&\sup\left\{\frac{\langle\langle R^\dagger|J(\mathcal{F})|R^\dagger\rangle\rangle}{\dim(\mathcal{X})\mathrm{tr}\left((R \otimes \mathbb{1}_\mathcal{X})J(\mathcal{F})(R^\dagger \otimes \mathbb{1}_\mathcal{X})\right)} : 0 \neq RS \propto \mathbb{1}_\mathcal{X}\right\}\\
=&\frac{1}{\dim(\mathcal{X})}\sup\left\{\frac{\langle\langle R^\dagger|\Pi_S J(\mathcal{F})\Pi_S|R^\dagger\rangle\rangle}{\langle\langle R^\dagger|\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S|R^\dagger\rangle\rangle} : \langle\langle S|R^\dagger\rangle\rangle \neq 0\right\},
\end{aligned} \tag{5.32}$$

where we used the substitution $|R^\dagger\rangle\rangle \leftarrow \Pi_S|R^\dagger\rangle\rangle$. Note that

$$(1-p)|S\rangle\rangle\langle\langle S| \leq \Pi_S J(\mathcal{F})\Pi_S \leq \dim(\mathcal{X})\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S. \tag{5.33}$$

Without loss of the generality we may assume that $\left(\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S\right)^0 |R^\dagger\rangle\rangle = |R^\dagger\rangle\rangle$. Next, we substitute $|R^\dagger\rangle\rangle \leftarrow \left(\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S\right)^{-1/2} |v\rangle$, where $\|v\|_2 = 1$. The objective function in our optimization problem is now equal to

$$|v\rangle \mapsto \langle v| \left(\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S\right)^{-1/2} J(\mathcal{F}) \left(\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S\right)^{-1/2} |v\rangle \tag{5.34}$$

and it is upper bounded by

$$\left\|\left(\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S\right)^{-1/2} J(\mathcal{F}) \left(\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S\right)^{-1/2}\right\|_\infty. \tag{5.35}$$

Taking into account all substitutions, in the limit, we can achieve this value by considering a sequence of decoding matrices $\left(\frac{R_n}{\|R_n\|_\infty}\right)_{n\in\mathbb{N}}$, where

$$|R_n^\dagger\rangle\rangle = (1 - 1/n)\left(\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S\right)^{-1/2} |v_0\rangle + 1/n|S\rangle\rangle \tag{5.36}$$

and $|v_0\rangle$ is the eigenvector of the leading eigenvalue of $\left(\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S\right)^{-1/2}$ $J(\mathcal{F})\left(\Pi_S(\mathcal{F}(\mathbb{1}_\mathcal{X}) \otimes \mathbb{1}_\mathcal{X})\Pi_S\right)^{-1/2}$. □

5.4 Numerical results for random quantum channels

In the last section we will numerically investigate the effectiveness of the construction provided in Section 5.3. The noise channel will be defined as in Section 5.2, that is $\mathcal{E}_p = (1-p)\mathcal{I}_\mathcal{Y} + p\mathcal{E}$ for $p \in [0,1]$ and $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$. The channel $\mathcal{E}$ will be sampled according to the measure $\mu^{Kraus}_{\mathcal{Y},\mathcal{Y};r}$ defined in Definition 3.2. For each $\mathcal{E}$ we will use Algorithm 3 to generate an encoding operation $\mathcal{S} \in \mathcal{C}(\mathcal{X},\mathcal{Y})$, a deterministic decoding operation $\mathcal{R}_1 \in \mathcal{C}(\mathcal{Y},\mathcal{X})$ and a probabilistic decoding operation $\mathcal{R}_0 \in s\mathcal{C}(\mathcal{Y},\mathcal{X})$. To define $\mathcal{R}_0$ we will fix the parameter $\epsilon = 10^{-3}$ and try different values of the second parameter $q \in (0,1)$. To distinguish the decoding operations depending on q we use the notation $\mathcal{R}_{0,q}$. For low-dimensional $\mathcal{Y}$ we will also calculate $\mathcal{R}_{1,\mathrm{sdp}}$ which is the deterministic decoding operation defined according to Proposition 5.4. We will check the behavior of $p \mapsto F_{|\mathrm{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$, where $\mathcal{R} \in \{\mathcal{R}_1, \mathcal{R}_{1,\mathrm{sdp}}, \mathcal{R}_{0,q}\}$ and also $p \mapsto p_{\mathrm{avg}}(\mathcal{R}_{0,q}\mathcal{E}_p\mathcal{S})$. Moreover, we will define a deterministic scheme $(\mathcal{S}_*, \mathcal{R}_*)$ that depends only on $\mathcal{X}$ and $\mathcal{Y}$, of the form: $\mathcal{S}_* = \mathcal{K}((S_*)) \in \mathcal{C}(\mathcal{X},\mathcal{Y})$, where $S_* = \sum_{i=1}^{\dim(\mathcal{X})} |i\rangle_\mathcal{Y}\langle i|_\mathcal{X}$ and $\mathcal{R}_*(Y) = \mathcal{S}_*^\dagger(Y) + \mathrm{tr}\left((\mathbb{1}_\mathcal{Y} - S_*S_*^\dagger)Y\right)\rho^*_\mathcal{X}$. This scheme will provide a reference point for the fidelity function.

All simulations are described by a tuple of parameters: (y, x, r, k), where $y = \dim(\mathcal{Y})$ is the dimension of the target system, $x = \dim(\mathcal{X})$ is the dimension of the code, $r = \mathrm{rank}(J(\mathcal{E}))$ is the expected Choi rank of random quantum channel $\mathcal{E} \in \mu^{Kraus}_{\mathcal{Y},\mathcal{Y};r}$ and k is the number of generated random channels $\mathcal{E}$. The calculations are performed with precision at least 10^{-5}. We used the `Julia` programming language along with quantum package `QuantumInformation.jl` [109] and SDP optimization via SCS solver [110, 111]. The code used to generate the following plots is available on GitHub [112].

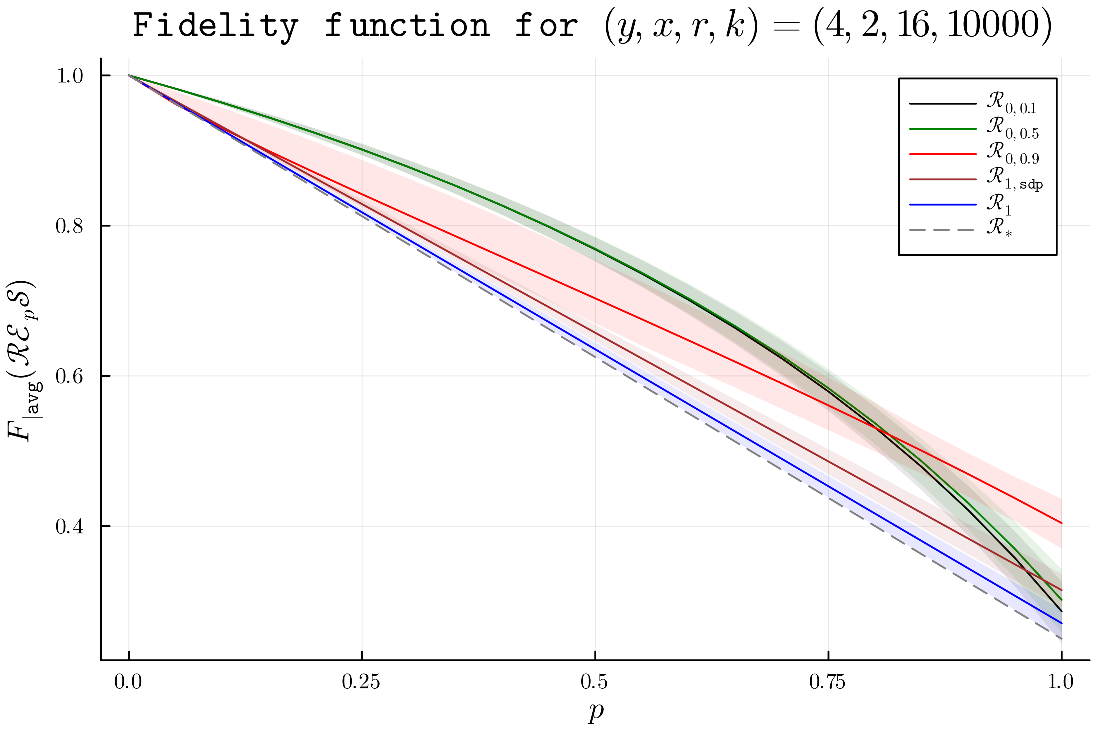
Let us analyze the results from Fig. 5.2. First, we can observe that probabilistic decoding in general increases the fidelity. However, there are regions where it does not - depending on q. In this simulation we used $q = 0.1, 0.5, 0.9$. The plots confirm that if q is small then we achieve greater probability of success and better fidelity for small values of p. Although, for $p \simeq 1$ it may happen that deterministic decoding, especially the optimal one $\mathcal{R}_{1,\text{sdp}}$, performs better than $\mathcal{R}_{0,0.1}$. On the other hand, taking $q = 0.9$ we see that $\mathcal{R}_{0,0.9}$ is the best strategy for very noisy channel $\mathcal{E}_p$. Surprisingly, when $p \to 0$, the decoding $\mathcal{R}_{0,0.9}$ is worse even than the reference strategy $\mathcal{R}_*$. To enhance this effect we take more extremal values of q (see Fig. 5.3).

In the next example (Fig. 5.4) we increased the ration y/x and decreased the Choi rank of $\mathcal{E}$. We can observe that our construction significantly increases the fidelity (above the reference line). Moreover, we see that difference between $\mathcal{R}_{1,\text{sdp}}$ and $\mathcal{R}_1$ is not that clear. Finally, the probabilistic decoding $\mathcal{R}_{0,0.3}$ and $\mathcal{R}_{0,0.6}$ beats the deterministic one in whole range $p \in [0,1]$. However, the strategy $\mathcal{R}_{0,0.9}$ is worse than the deterministic one for small p.

Let us check if our construction can indeed generate codes achieving the fidelity equal one for all p, that is $0 \neq \mathcal{R}_{0,q}\mathcal{E}_p\mathcal{S} \propto \mathcal{I}_\mathcal{X}$. By the construction and Theorem 4.25 it should happen if $(r+1)^2(x-1) < y$. We confirmed this numerically in Fig. 5.5. Two things are worth noting: there is significant drop in the probability of success; there is almost no difference between the best decoding operation and the Petz recovery map.

We may go even further and ask: If $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ is random quantum channel, then is it possible to generate codes that $0 \neq \mathcal{R}_{0,q}\mathcal{E}_p\mathcal{S} \propto \mathcal{I}_\mathcal{X}$, under the assumption of Theorem 4.29, that is $(r+1) < \frac{xy}{x^2-1}$ (or at least under the assumption of Corollary 4.30, $(r+1) \leq \frac{y}{x}$). We give the positive answer in both cases - see Fig. 5.6 and Fig. 5.7. Moreover, the answer does not depend on the value of q. Unfortunately, this construction does not return perfect deterministic schemes, that is schemes $(\mathcal{R}_{1,\text{sdp}}, \mathcal{S})$ such that $\mathcal{R}_{1,\text{sdp}}\mathcal{E}_p\mathcal{S} = \mathcal{I}_\mathcal{X}$, even if according to Theorem 4.25 they should exist. We provide an example of such situation in Fig. 5.8. Note that the performance of $\mathcal{R}_1$ and $\mathcal{R}_{1,\text{sdp}}$ matches.

In the final examples we compare the performance of $\mathcal{R}_{0,q}$ and $\mathcal{R}_1$ for high-dimensional $\mathcal{Y}$: see Fig. 5.9, Fig. 5.10 and Fig. 5.11.



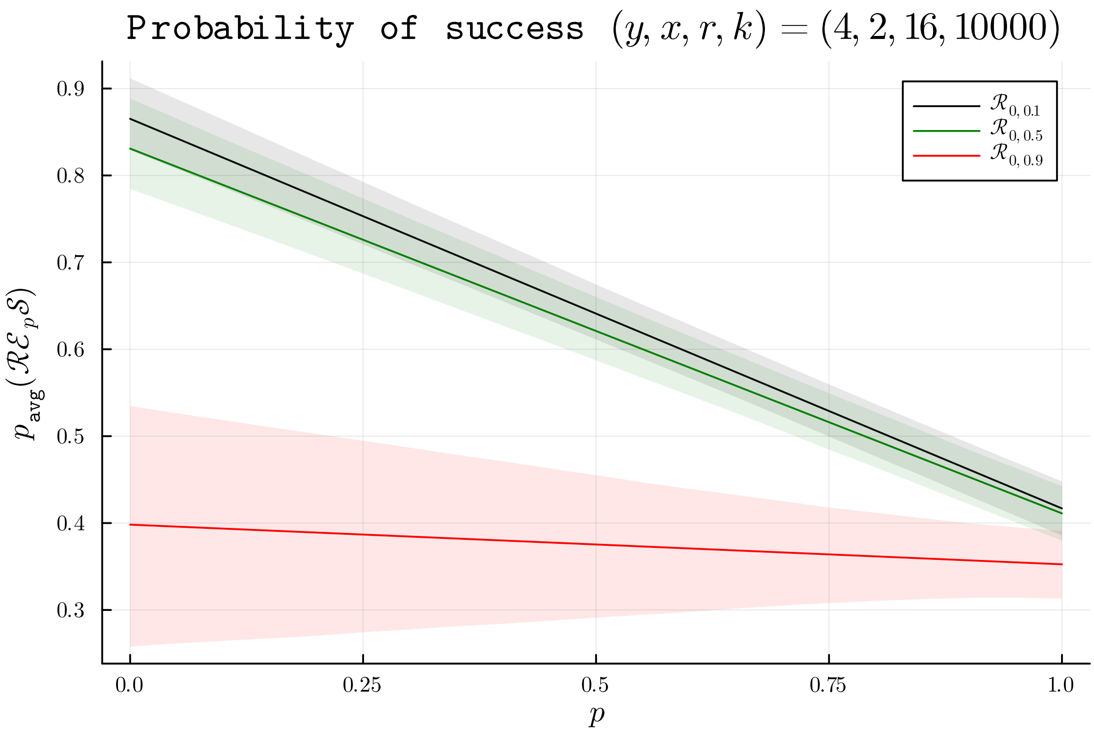


Figure 5.2: First plot: the mean value of $p \mapsto F_{|\text{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$ with the interquartile range calculated for $\mathcal{R} \in \{\mathcal{R}_1, \mathcal{R}_{1,\text{sdp}}, \mathcal{R}_{0,q}\}$. The dashed line represents $p \mapsto F_{|\text{avg}}(\mathcal{R}_*\mathcal{E}_p\mathcal{S}_*)$. Second plot: the mean value of $p \mapsto p_{\text{avg}}(\mathcal{R}_{0,q}\mathcal{E}_p\mathcal{S})$ with the interquartile range.

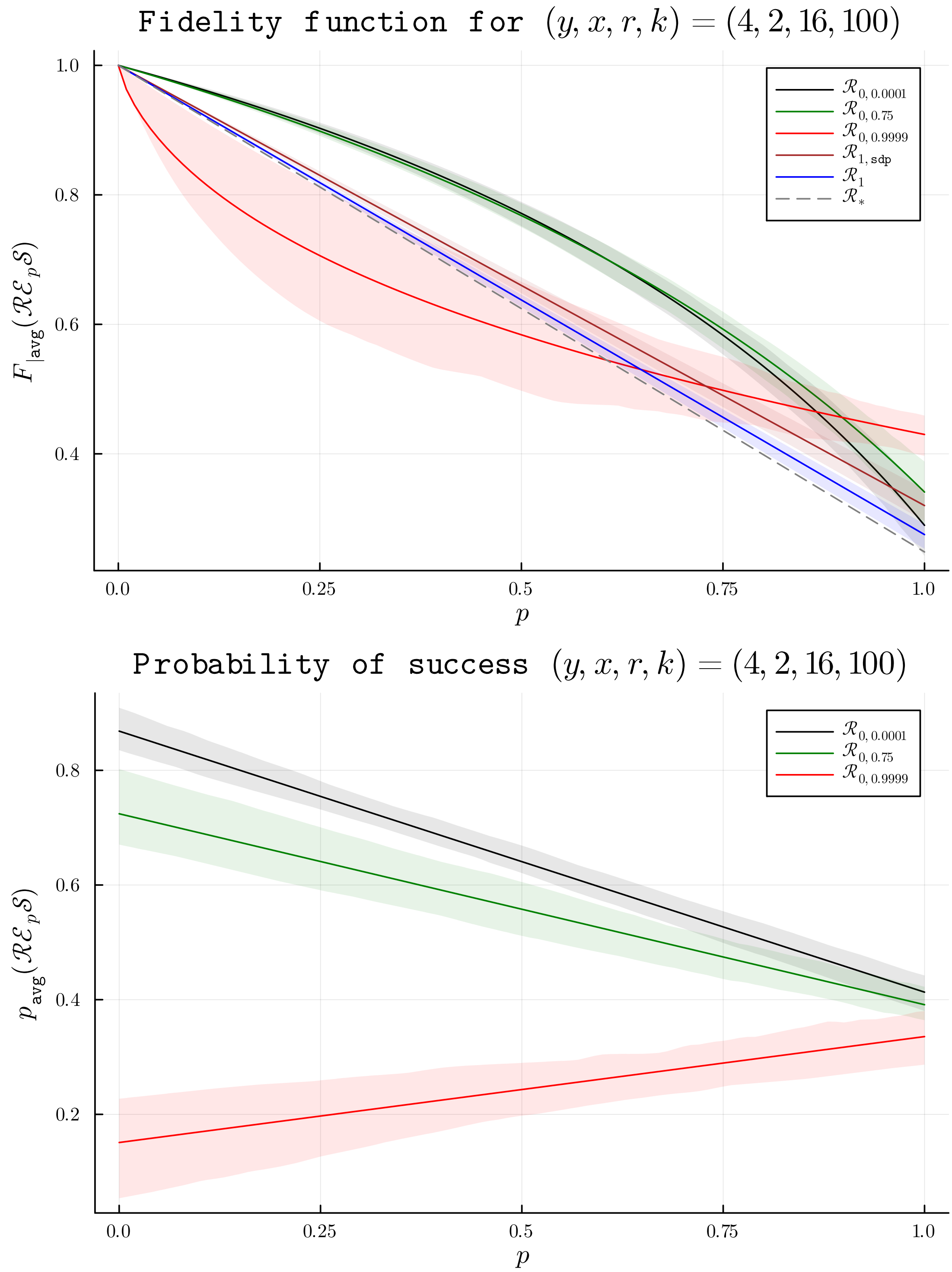


Figure 5.3: First plot: the mean value of $p \mapsto F_{|\text{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$ with the interquartile range calculated for $\mathcal{R} \in \{\mathcal{R}_1, \mathcal{R}_{1,\text{sdp}}, \mathcal{R}_{0,q}\}$. The dashed line represents $p \mapsto F_{|\text{avg}}(\mathcal{R}_*\mathcal{E}_p\mathcal{S}_*)$. Second plot: the mean value of $p \mapsto p_{\text{avg}}(\mathcal{R}_{0,q}\mathcal{E}_p\mathcal{S})$ with the interquartile range.

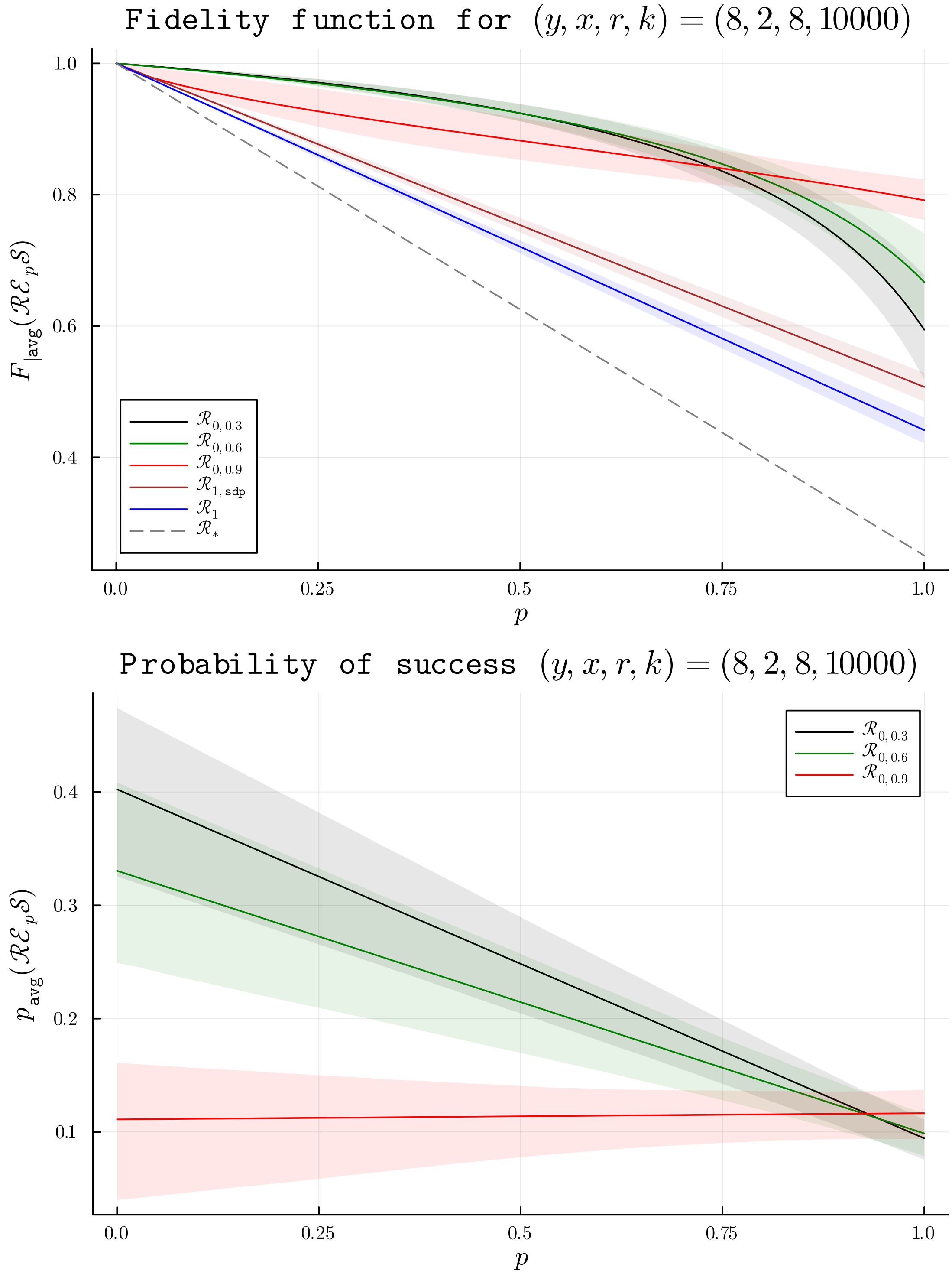


Figure 5.4: First plot: the mean value of $p \mapsto F_{|\mathrm{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$ with the interquartile range calculated for $\mathcal{R} \in \{\mathcal{R}_1, \mathcal{R}_{1,\mathrm{sdp}}, \mathcal{R}_{0,q}\}$. The dashed line represents $p \mapsto F_{|\mathrm{avg}}(\mathcal{R}_*\mathcal{E}_p\mathcal{S}_*)$. Second plot: the mean value of $p \mapsto p_{\mathrm{avg}}(\mathcal{R}_{0,q}\mathcal{E}_p\mathcal{S})$ with the interquartile range.

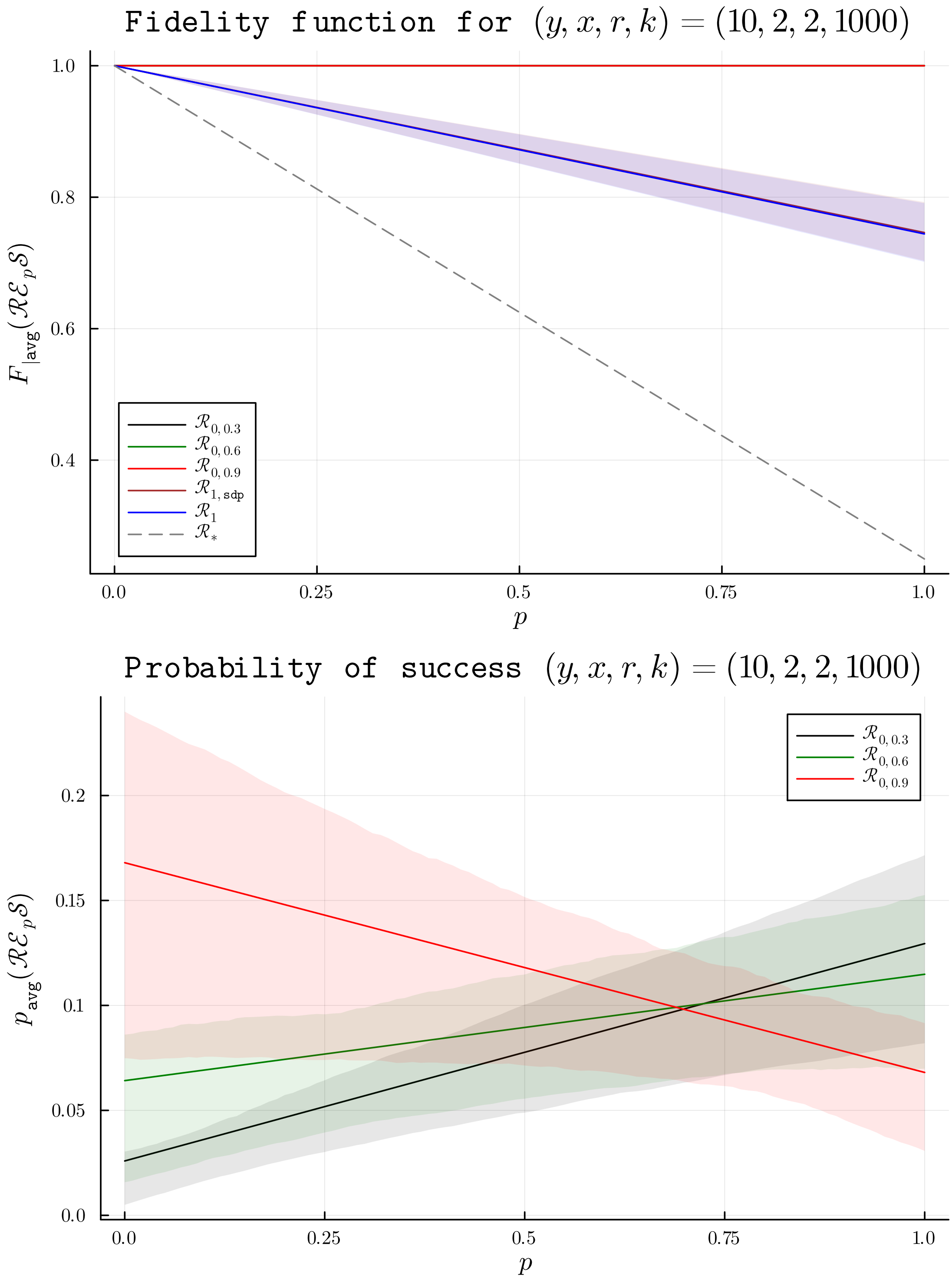


Figure 5.5: First plot: the mean value of $p \mapsto F_{|\text{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$ with the interquartile range calculated for $\mathcal{R} \in \{\mathcal{R}_1, \mathcal{R}_{1,\text{sdp}}, \mathcal{R}_{0,q}\}$. The dashed line represents $p \mapsto F_{|\text{avg}}(\mathcal{R}_*\mathcal{E}_p\mathcal{S}_*)$. Second plot: the mean value of $p \mapsto p_{\text{avg}}(\mathcal{R}_{0,q}\mathcal{E}_p\mathcal{S})$ with the interquartile range.

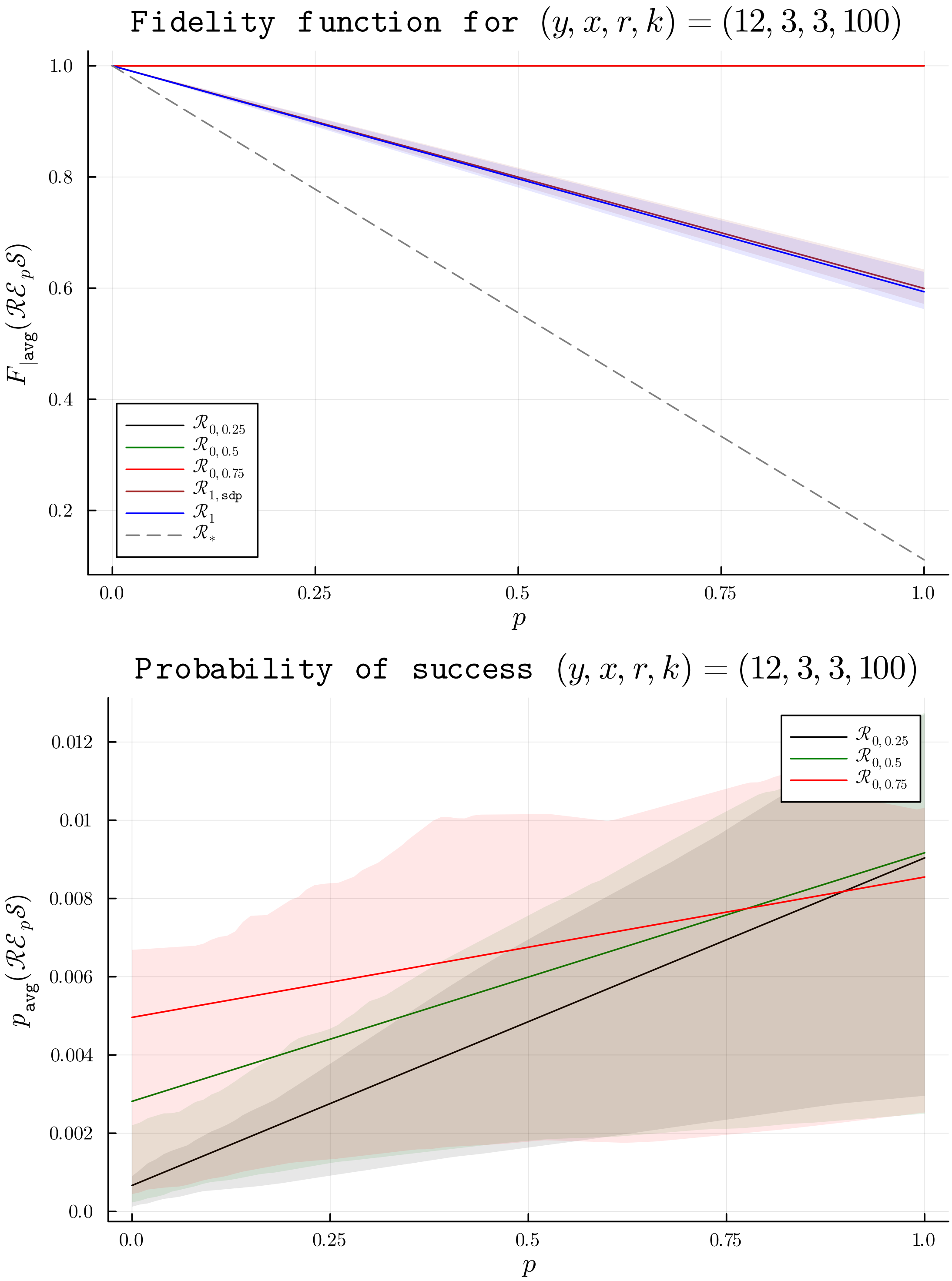
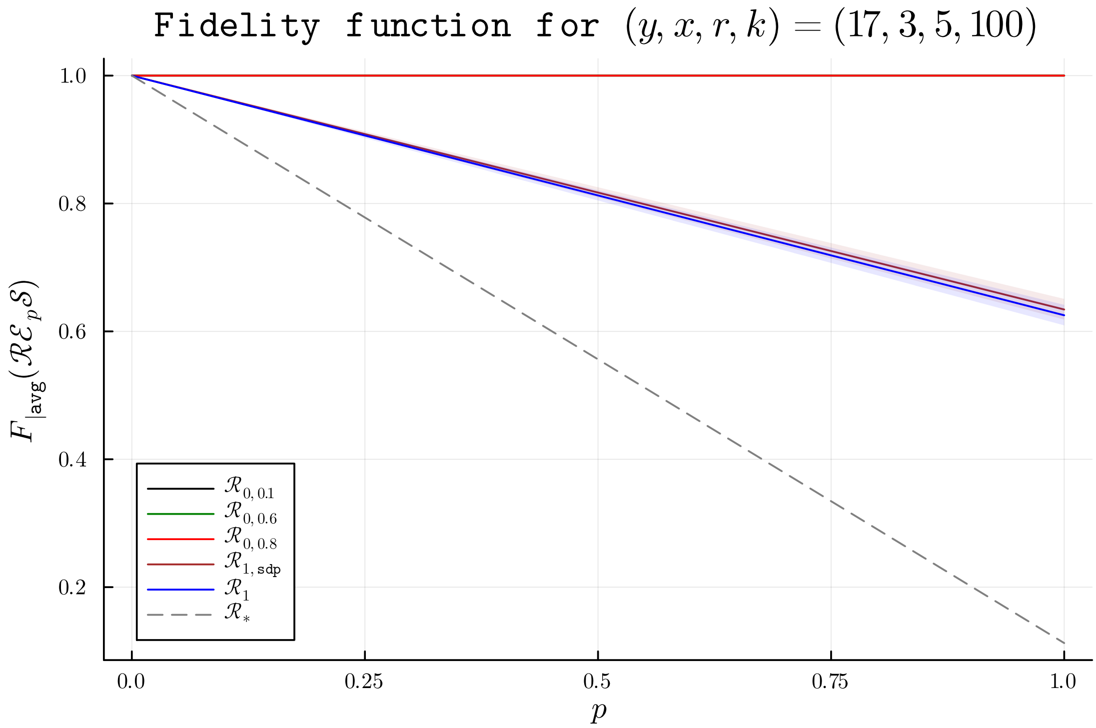


Figure 5.6: First plot: the mean value of $p \mapsto F_{|\text{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$ with the interquartile range calculated for $\mathcal{R} \in \{\mathcal{R}_1, \mathcal{R}_{1,\text{sdp}}, \mathcal{R}_{0,q}\}$. The dashed line represents $p \mapsto F_{|\text{avg}}(\mathcal{R}_*\mathcal{E}_p\mathcal{S}_*)$. Second plot: the mean value of $p \mapsto p_{\text{avg}}(\mathcal{R}_{0,q}\mathcal{E}_p\mathcal{S})$ with the interquartile range.



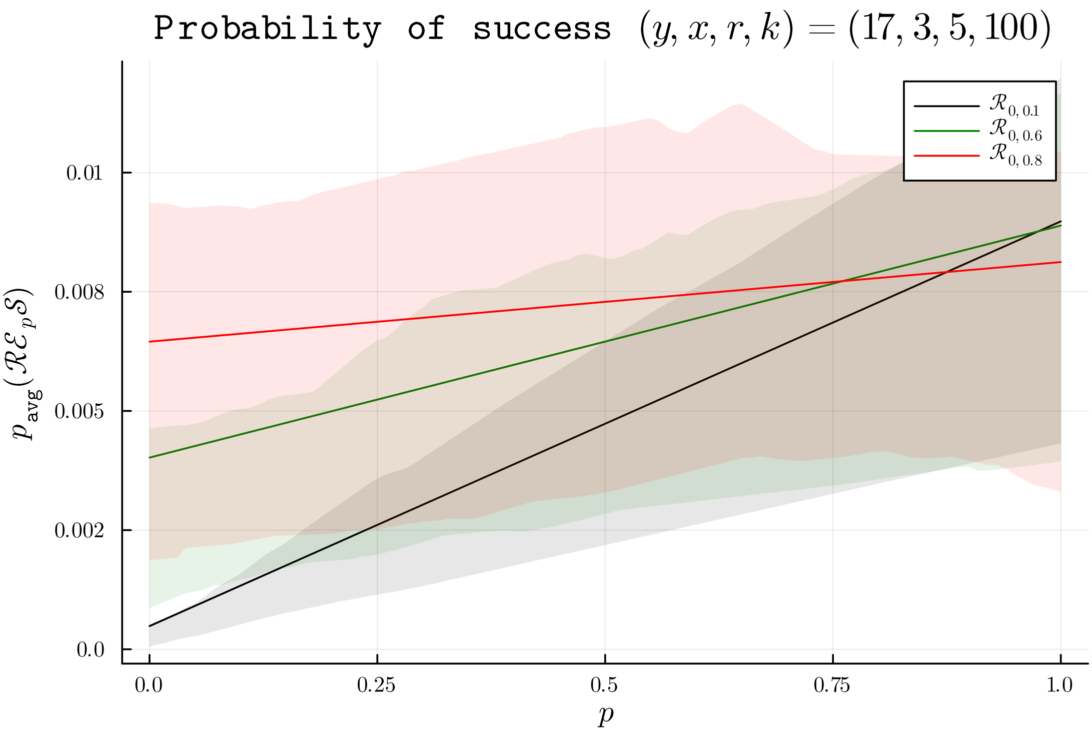
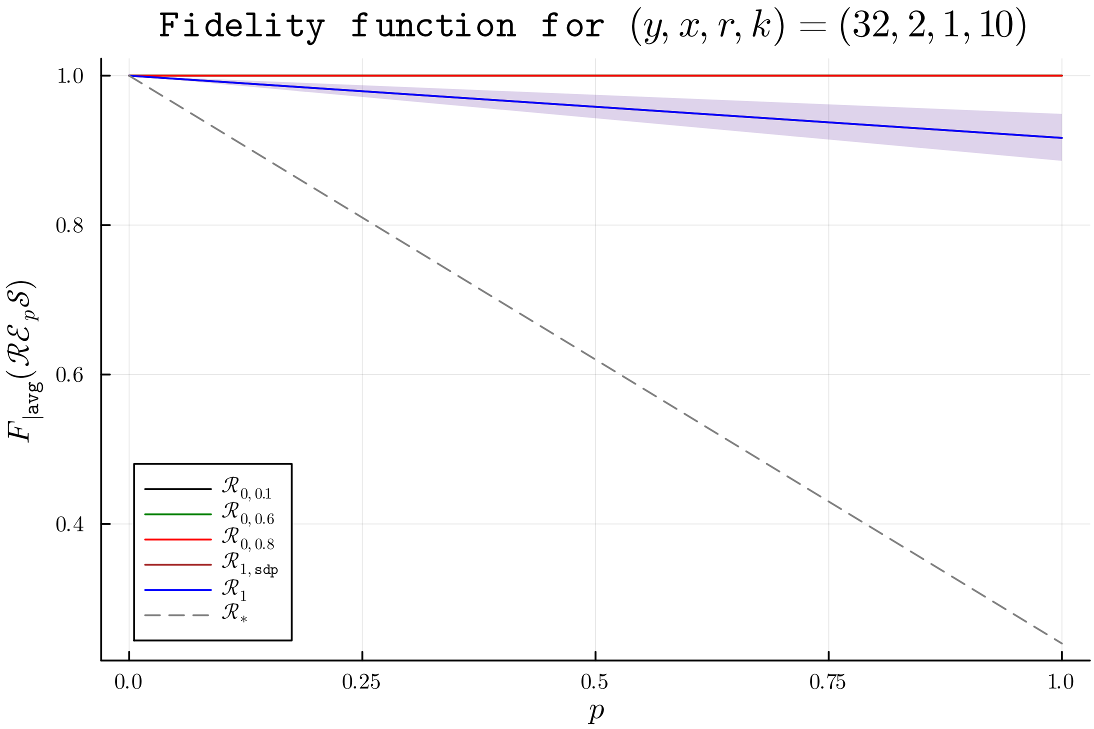


Figure 5.7: First plot: the mean value of $p \mapsto F_{|\mathrm{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$ with the interquartile range calculated for $\mathcal{R} \in \{\mathcal{R}_1, \mathcal{R}_{1,\mathrm{sdp}}, \mathcal{R}_{0,q}\}$. The dashed line represents $p \mapsto F_{|\mathrm{avg}}(\mathcal{R}_*\mathcal{E}_p\mathcal{S}_*)$. Second plot: the mean value of $p \mapsto p_{\mathrm{avg}}(\mathcal{R}_{0,q}\mathcal{E}_p\mathcal{S})$ with the interquartile range.



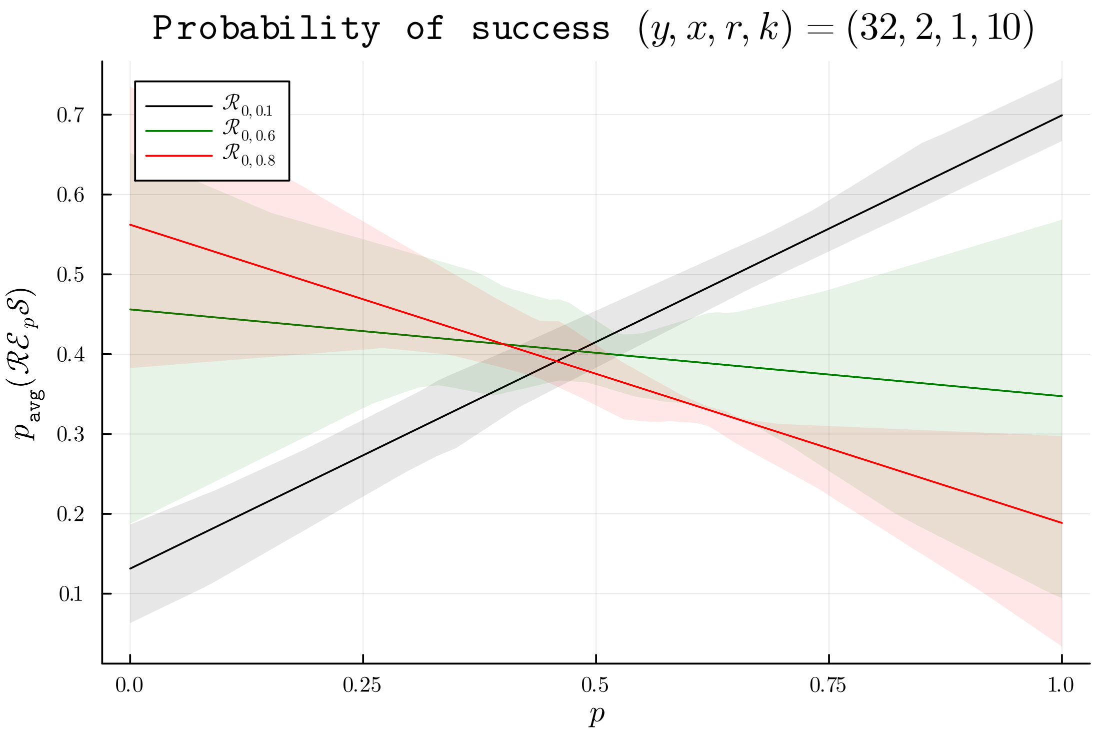


Figure 5.8: First plot: the mean value of $p \mapsto F_{|\text{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$ with the interquartile range calculated for $\mathcal{R} \in \{\mathcal{R}_1, \mathcal{R}_{1,\text{sdp}}, \mathcal{R}_{0,q}\}$. The dashed line represents $p \mapsto F_{|\text{avg}}(\mathcal{R}_*\mathcal{E}_p\mathcal{S}_*)$. Second plot: the mean value of $p \mapsto p_{\text{avg}}(\mathcal{R}_{0,q}\mathcal{E}_p\mathcal{S})$ with the interquartile range.

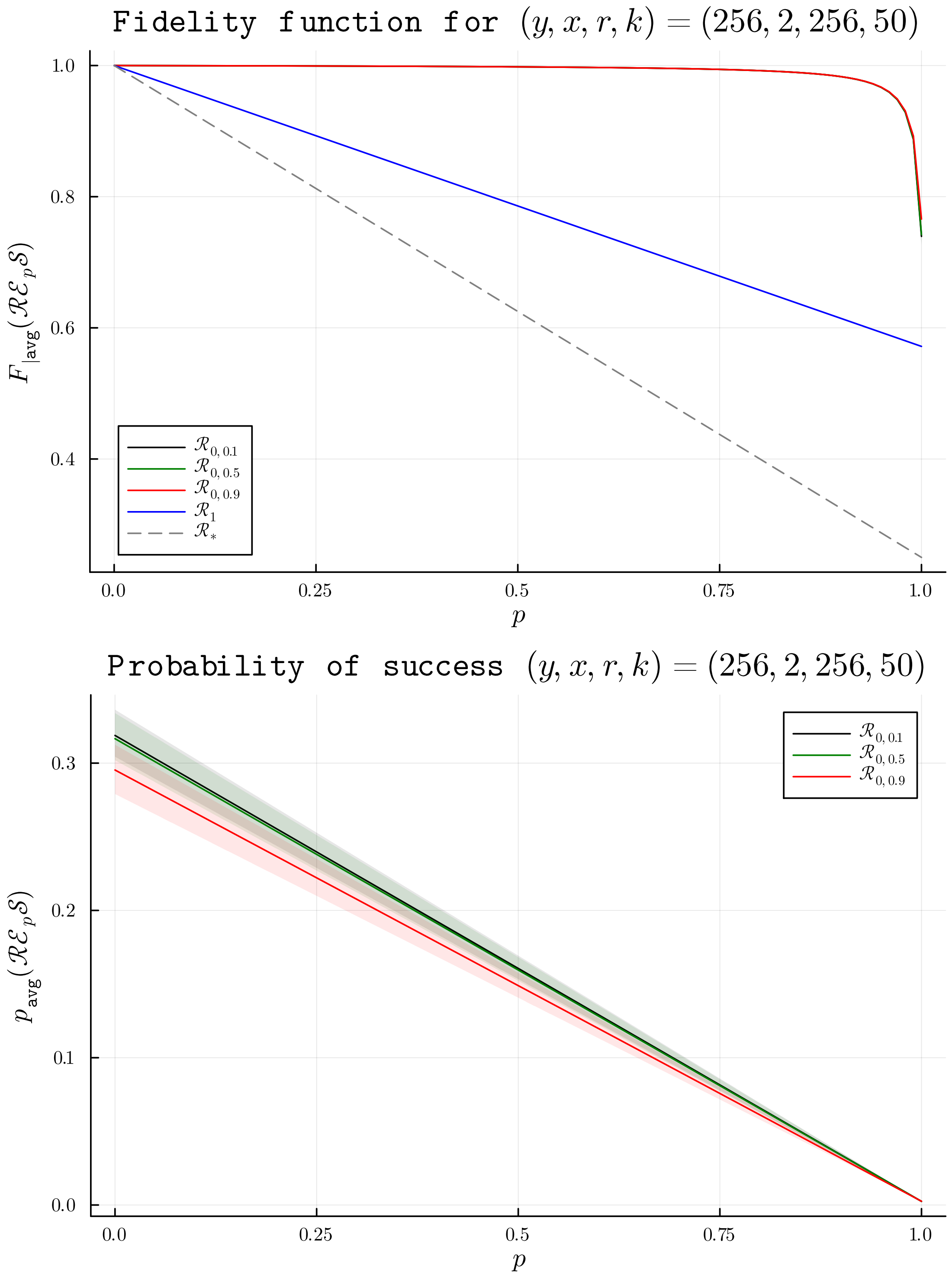
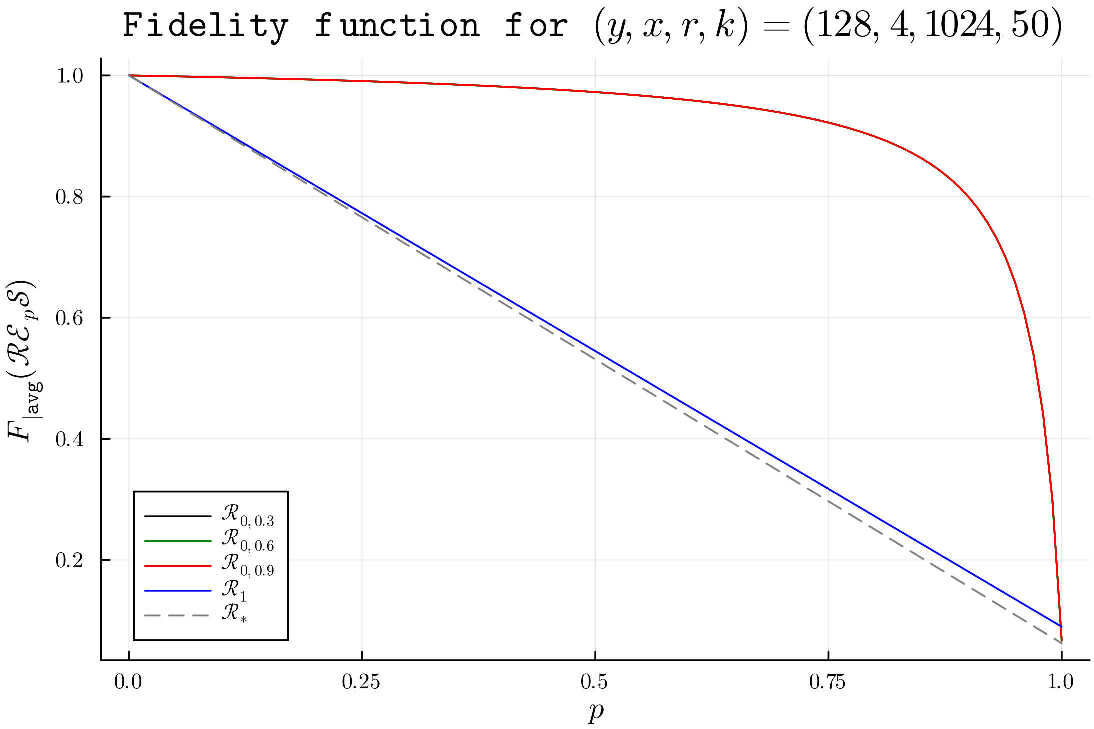


Figure 5.9: First plot: the mean value of $p \mapsto F_{|\mathrm{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$ with the interquartile range calculated for $\mathcal{R} \in \{\mathcal{R}_1, \mathcal{R}_{0,q}\}$. The dashed line represents $p \mapsto F_{|\mathrm{avg}}(\mathcal{R}_*\mathcal{E}_p\mathcal{S}_*)$. Second plot: the mean value of $p \mapsto p_{\mathrm{avg}}(\mathcal{R}_{0,q}\mathcal{E}_p\mathcal{S})$ with the interquartile range.



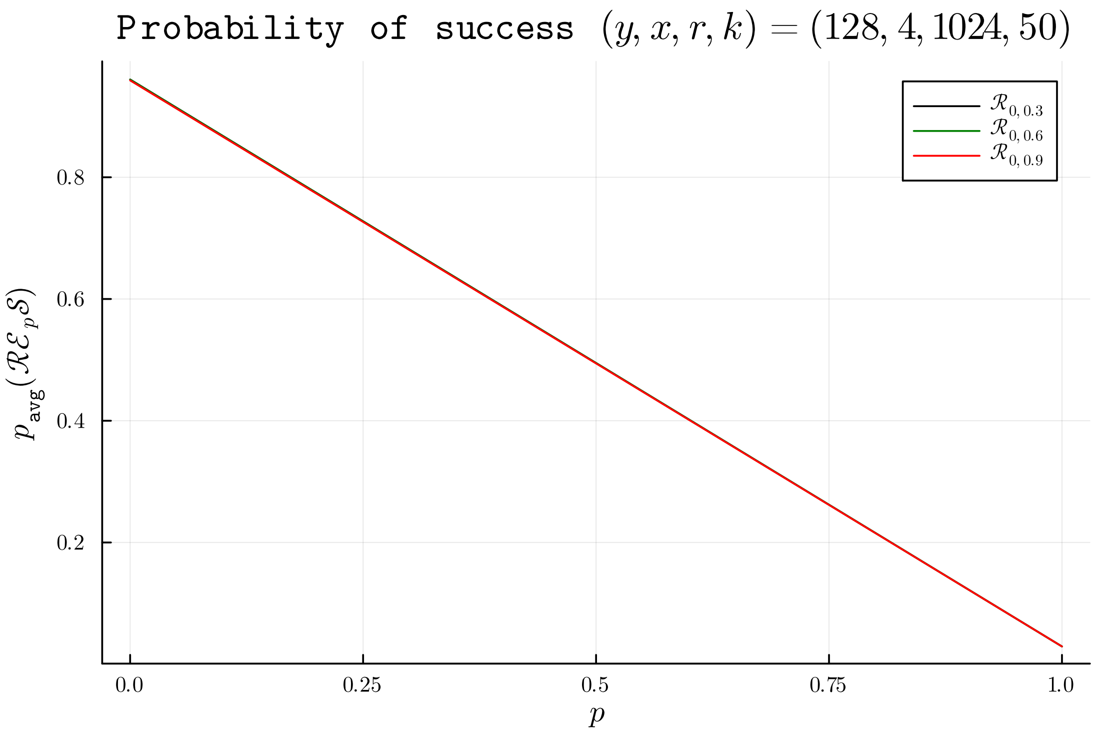
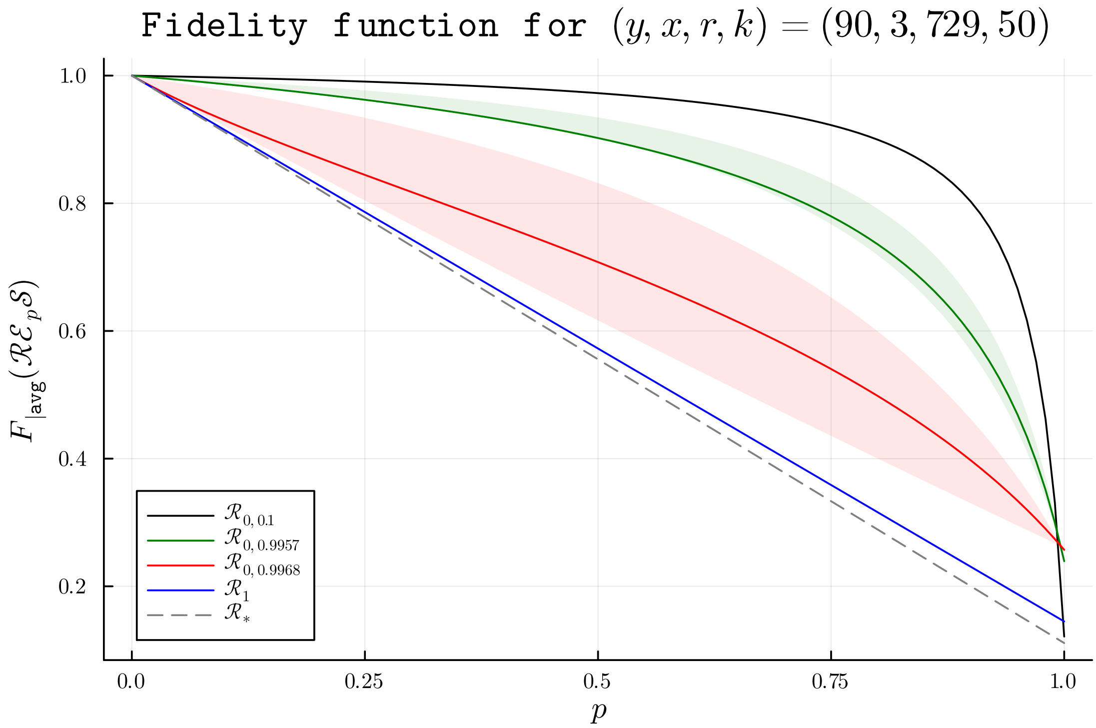


Figure 5.10: First plot: the mean value of $p \mapsto F_{|\mathrm{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$ with the interquartile range calculated for $\mathcal{R} \in \{\mathcal{R}_1, \mathcal{R}_{0,q}\}$. The dashed line represents $p \mapsto F_{|\mathrm{avg}}(\mathcal{R}_*\mathcal{E}_p\mathcal{S}_*)$. Second plot: the mean value of $p \mapsto p_{\mathrm{avg}}(\mathcal{R}_{0,q}\mathcal{E}_p\mathcal{S})$ with the interquartile range.



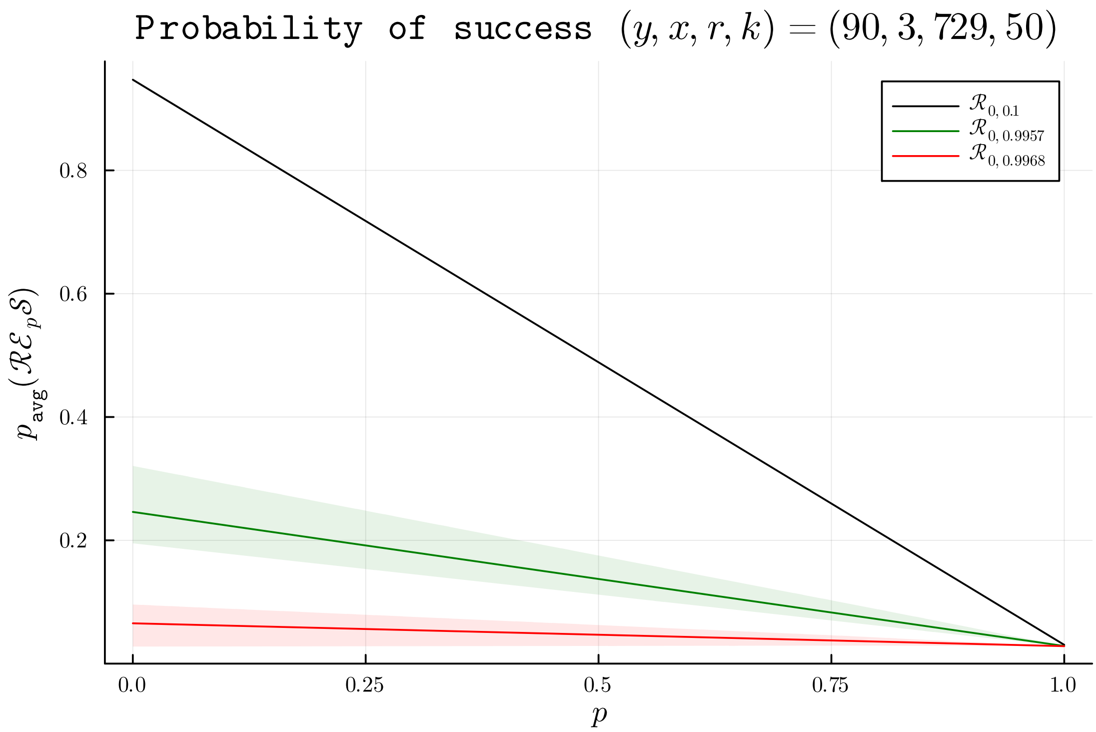


Figure 5.11: First plot: the mean value of $p \mapsto F_{|\mathrm{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$ with the interquartile range calculated for $\mathcal{R} \in \{\mathcal{R}_1, \mathcal{R}_{0,q}\}$. The dashed line represents $p \mapsto F_{|\mathrm{avg}}(\mathcal{R}_*\mathcal{E}_p\mathcal{S}_*)$. Second plot: the mean value of $p \mapsto p_{\mathrm{avg}}(\mathcal{R}_{0,q}\mathcal{E}_p\mathcal{S})$ with the interquartile range.

Chapter 6

Summary and discussion

The main focus of this dissertation was the development of pQEC codes for general noise channels. The conducted analysis showed that probabilistic codes are suitable for very noisy quantum systems. The results of this thesis confirm the formulated **Hypothesis**:

The usage of probabilistic quantum error correction codes can improve the quality of quantum systems disturbed by general noise channels.

In more detail, we generalized the Kill-Laflamme theorem. We established the necessary and sufficient conditions to check if a given noise channel is probabilistically correctable, Theorem 4.4. We used these conditions to show that the pQEC codes, in a comparison to the QEC codes, are able to correct noise channels from a larger set, Proposition 4.18 and Proposition 4.19. A clear separation between probabilistic and deterministic codes was observed for: Schur noise channels, Lemma 4.22 and Proposition 4.23, also for noise channels of Choi rank $\mathrm{rank}(J(\mathcal{E})) = 2$, Proposition 4.26, and finally for random quantum channels, Theorem 4.29. Each of these examples indicated a trade-off between the probability of successful error correction and the quality of the code.

We proved that noise channels with bounded Choi rank, $\mathrm{rank}(J(\mathcal{E}))^2(\dim(\mathcal{X}) - 1) < \dim(\mathcal{Y})$, are probabilistically correctable, Theorem 4.25. The proof of this theorem was constructive and led us to the construction of approximate pQEC codes, Algorithm 3. This procedure was numerically efficient, having the complexity $\mathcal{O}(\dim(\mathcal{Y})^6)$ in the worst-case scenario.

We used Algorithm 3 to compare the effectiveness of deterministic and probabilistic decoding operations. We numerically checked the quality of our procedure by using randomly sampled quantum noise channels. In particular, for high-dimensional vector spaces $\mathcal{Y}$ and noise channels $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ with the expected Choi rank behaving like $\mathrm{rank}(J(\mathcal{E})) = \mathcal{O}\left(\frac{\dim(\mathcal{Y})^2}{\dim(\mathcal{X})^2}\right)$, we observed the separation between probabilistic and deterministic error-correcting schemes, Fig. 5.10 and Fig. 5.11.

The numerical investigation of pQEC codes was possible due to advancement in the methods of generating random quantum channels. In this dissertation, we showed the equivalence of sampling techniques based on different representations of quantum channels, Proposition 3.4 and Proposition 3.5. In particular, the method based on the Kraus representation happened to be the most suitable for numerical simulation, providing appropriate diversity of sampled channels and having a simple implementation. Additionally, in Section 3.1.6 and Section 3.1.7 we generalized techniques of generating quantum channels and showed how to effectively generate random quantum subchannels, quantum instruments and quantum super-channels, and how to obtain the uniform measure. All the presented methods are structured and may be developed further to cover all higher-order quantum operations, like quantum deterministic networks, quantum probabilistic networks and quantum testers.

This dissertation sets future research directions concerning approximate pQEC codes and numerical simulation:

- A parameterized method of generating random channels: The random ensemble $\mu^{Kraus}_{\mathcal{Y},\mathcal{Y};\dim(\mathcal{Y})^2}$ provides the flat measure in the set $\mathcal{C}(\mathcal{Y})$. However, for a high-dimensional $\mathcal{Y}$ and small sample size k (see Section 5.4) the effect of the measure concentration occurs, *e.g.* Theorem 3.26. Due to that effect, in the numerical investigation in Section 5.4 we took into account different ratios of $\operatorname{rank}(J(\mathcal{E}))/\dim(\mathcal{Y})$. This effect also justifies why we need to set the theoretical background for parameterized methods, for example:

  For a given parameters $p_i > 0$ define $\mathcal{E} = \mathcal{K}\left((E_i)_i\right)$, where $E_i = p_i G_i Q^{-1/2}$, $Q = \sum_i p_i^2 G_i^\dagger G_i$ and G_i are complex Ginibre matrices.

- Improving the performance of Algorithm 3: We may ask how to chose wisely the value of the parameter $q \in (0, 1)$ such that a probabilistic decoding provides better fidelity than a deterministic for $\mathcal{E}_p$ in the whole range $p \in [0, 1]$, see Fig. 5.3.

  In some ranges of r/y there is a problem with the probability of success, see Fig. 5.6. The algorithm returns a scheme $(\mathcal{S}, \mathcal{R}_0)$ such that $\mathcal{R}_0\mathcal{E}_p\mathcal{S} = c_p\mathcal{I}_\mathcal{X}$, but c_p is low. The question is how to increase the value of c_p without using optimization methods, such as SDP programming, Corollary 4.9.

  We can examine how the fidelity varies if we re-sample $(\mathcal{S}, \mathcal{R}_{0,q})$ for a fixed noise channel $\mathcal{E}_p$. In Fig. 6.1 we showed this effect for $k = 100$ random codes.

  In the same way, we may investigate if alternating the optimization of $\mathcal{S}$ and $\mathcal{R}_0$ will increase the fidelity significantly. To do so, we may utilize

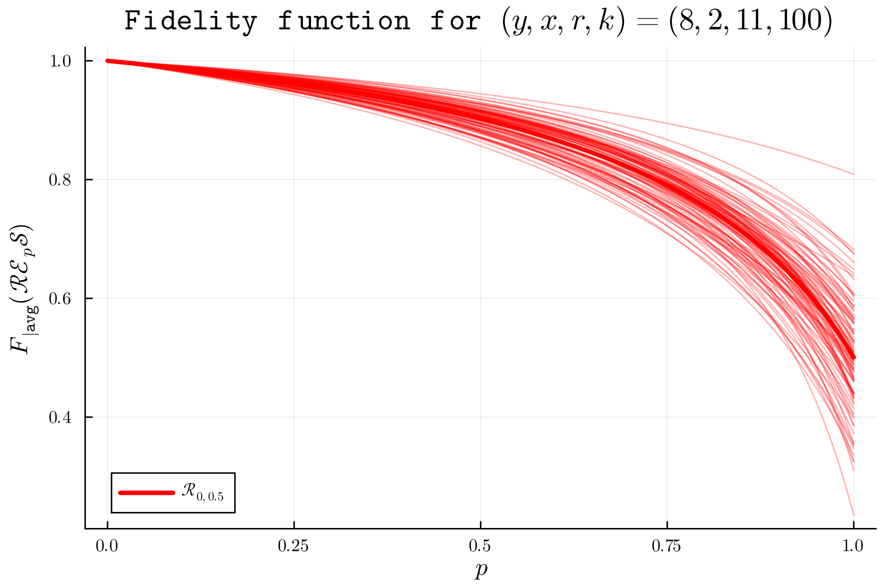


Figure 6.1: The function $p \mapsto F_{|\text{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$ calculated for a fixed channel $\mathcal{E}_p$ and $k = 100$ random schemes $(\mathcal{S}, \mathcal{R}_{0,0.5})$ defined as in Algorithm 3.

Proposition 5.5. In Fig. 6.2 we showed the result of such optimization for the fixed $\mathcal{E}_p$.

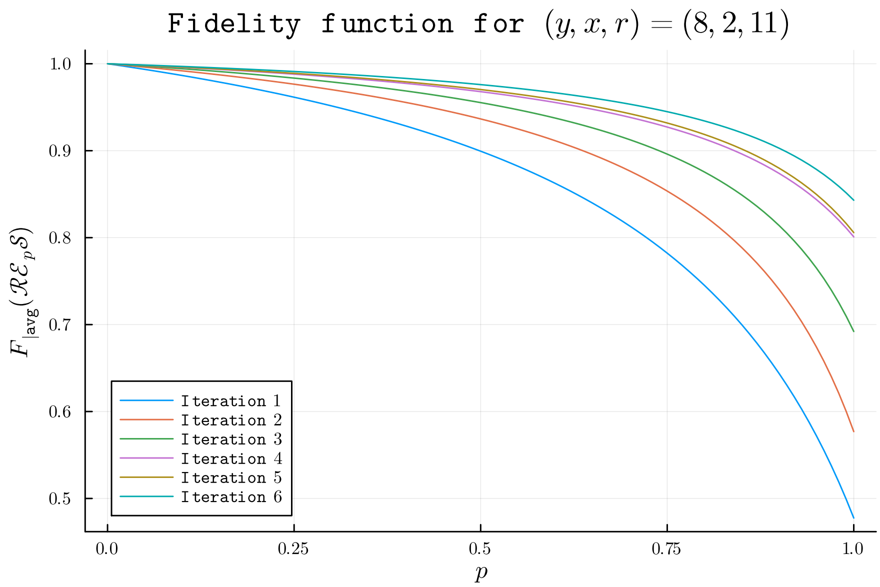


Figure 6.2: The function $p \mapsto F_{|\text{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$ calculated for a fixed channel $\mathcal{E}_p$ and 6 schemes $(\mathcal{S}, \mathcal{R}_0)$ defined by alternating the optimization of $\mathcal{S}$ and $\mathcal{R}_0$.

- Checking the performance of Algorithm 3 for tensor product of noise channels: For a qubit depolarizing quantum channel $\Phi_p(X) = (1-p)X + \frac{p}{3}(\sigma_x X \sigma_x + \sigma_y X \sigma_y + \sigma_z X \sigma_z)$ let us define a noise channel on 5 qubits $\mathcal{E}_p = \Phi_p^{\otimes 5}$. In Figure 6.3 we compare the fidelity achieved by [[5, 1, 3]] code [113] with the probabilistic code given by Algorithm 3 and alternating the optimization of $\mathcal{S}$ and $\mathcal{R}_0$.

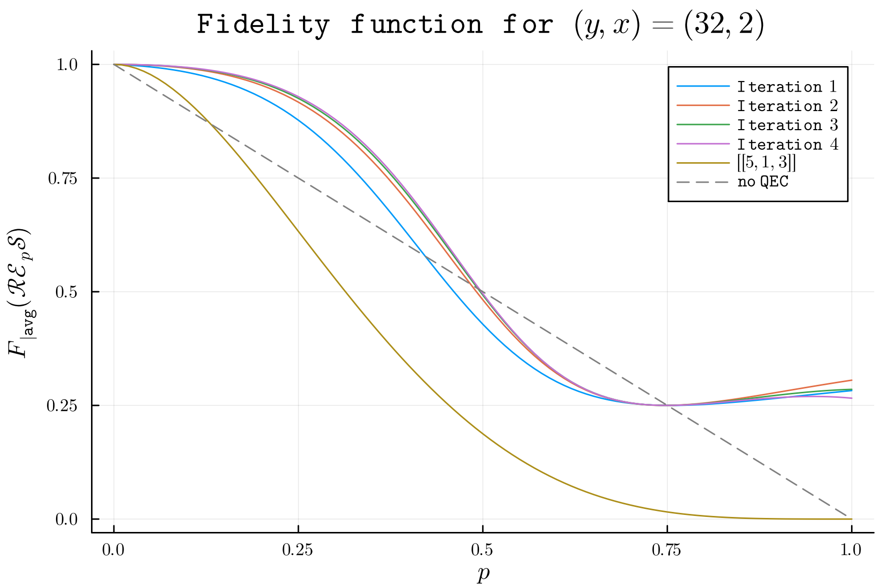


Figure 6.3: The function $p \mapsto F_{|\text{avg}}(\mathcal{R}\mathcal{E}_p\mathcal{S})$ calculated for a channel $\mathcal{E}_p = \Phi_p^{\otimes 5}$. The probabilistic schemes $(\mathcal{S}, \mathcal{R}_0)$ are defined by alternating the optimization of $\mathcal{S}$ and $\mathcal{R}_0$.

In the similar way, it would be interesting to check the performance of our construction for a composition of general noise channels, for example: $\mathcal{E}_p \circ \mathcal{F}_p$, $\mathcal{F}_p \circ \mathcal{E}_p$ or $\mathcal{E}_p \circ \cdots \circ \mathcal{E}_p$, where $\mathcal{E}_p = (1-p)\mathcal{I}_\mathcal{Y} + p\mathcal{E}$ and $\mathcal{F}_p = (1-p)\mathcal{I}_\mathcal{Y} + p\mathcal{F}$.

Bibliography

[1] R. Kukulski, I. Nechita, Ł. Pawela, Z. Puchała, and K. Życzkowski, "Generating random quantum channels," *Journal of Mathematical Physics*, vol. 62, no. 6, p. 062201, 2021.

[2] R. Kukulski, Ł. Pawela, and Z. Puchała, "On the probabilistic quantum error correction," *IEEE Transactions on Information Theory*, 2023.

[3] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[4] R. S. Ingarden, "Quantum information theory," *Reports on Mathematical Physics*, vol. 10, no. 1, pp. 43–72, 1976.

[5] J. Preskill, "Quantum computing in the nisq era and beyond," *Quantum*, vol. 2, p. 79, 2018.

[6] `https://www.ibm.com/quantum` The web resource at `https://www.ibm.com/quantum`. Accessed on 2023-02-18.

[7] `https://www.rigetti.com/` The web resource at `https://www.rigetti.com/`. Accessed on 2023-02-18.

[8] `https://www.xanadu.ai/` The web resource at `https://www.xanadu.ai/`. Accessed on 2023-02-18.

[9] `https://ionq.com/` The web resource at `https://ionq.com/`. Accessed on 2023-02-18.

[10] J. Preskill, "Quantum computing 40 years later," *arXiv preprint arXiv:2106.10522*, 2021.

[11] D. A. Lidar and T. A. Brun, *Quantum error correction.* Cambridge University Press, 2013.

[12] E. Knill and R. Laflamme, "Theory of quantum error-correcting codes," *Physical Review A*, vol. 55, no. 2, p. 900, 1997.

[13] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek, "Perfect quantum error correcting code," *Physical Review Letters*, vol. 77, no. 1, p. 198, 1996.

[14] P. W. Shor, "Scheme for reducing decoherence in quantum computer memory," *Physical Review A*, vol. 52, no. 4, p. R2493, 1995.

[15] A. R. Calderbank and P. W. Shor, "Good quantum error-correcting codes exist," *Physical Review A*, vol. 54, no. 2, p. 1098, 1996.

[16] A. M. Steane, "Simple quantum error-correcting codes," *Physical Review A*, vol. 54, no. 6, p. 4741, 1996.

[17] A. M. Steane, "Error correcting codes in quantum theory," *Physical Review Letters*, vol. 77, no. 5, p. 793, 1996.

[18] A. Steane, "Multiple-particle interference and quantum error correction," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 452, no. 1954, pp. 2551–2577, 1996.

[19] M. Grassl, W. Geiselmann, and T. Beth, "Quantum Reed-Solomon codes," in *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes: 13th International Symposium, AAECC-13 Honolulu, Hawaii, USA, November 15–19, 1999 Proceedings 13*, pp. 231–244, Springer, 1999.

[20] D. Gottesman, "Class of quantum error-correcting codes saturating the quantum hamming bound," *Physical Review A*, vol. 54, no. 3, p. 1862, 1996.

[21] D. Gottesman, *Stabilizer codes and quantum error correction.* Ph.D. dissertation, California Institute of Technology, 1997.

[22] T. Jackson, M. Grassl, and B. Zeng, "Codeword stabilized quantum codes for asymmetric channels," in *2016 IEEE International Symposium on Information Theory (ISIT)*, pp. 2264–2268, IEEE, 2016.

[23] M. F. Ezerman, S. Ling, B. Özkaya, and P. Solé, "Good stabilizer codes from quasi-cyclic codes over f 4 and f 9," in *2019 IEEE International Symposium on Information Theory (ISIT)*, pp. 2898–2902, IEEE, 2019.

[24] H. Bombin and M. Martin-Delgado, "Topological quantum error correction with optimal encoding rate," *Physical Review A*, vol. 73, no. 6, p. 062303, 2006.

[25] P. Andreasson, J. Johansson, S. Liljestrand, and M. Granath, "Quantum error correction for the toric code using deep reinforcement learning," *Quantum*, vol. 3, p. 183, 2019.

[26] J. P. Bonilla-Ataides, D. K. Tuckett, S. D. Bartlett, S. T. Flammia, and B. J. Brown, "The xzzx surface code," *arXiv preprint arXiv:2009.07851*, 2020.

[27] M. Grassl, T. Beth, and M. Roetteler, "On optimal quantum codes," *International Journal of Quantum Information*, vol. 2, no. 01, pp. 55–64, 2004.

[28] M. Grassl and M. Rötteler, "Quantum mds codes over small fields," in *2015 IEEE International Symposium on Information Theory (ISIT)*, pp. 1104–1108, IEEE, 2015.

[29] F. Huber and M. Grassl, "Quantum codes of maximal distance and highly entangled subspaces," *Quantum*, vol. 4, p. 284, 2020.

[30] D. J. MacKay, G. Mitchison, and P. L. McFadden, "Sparse-graph codes for quantum error correction," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2315–2330, 2004.

[31] D. Gottesman, "Fault-tolerant quantum computation with constant overhead," *arXiv preprint arXiv:1310.2984*, 2013.

[32] N. P. Breuckmann and J. N. Eberhardt, "Quantum low-density parity-check codes," *PRX Quantum*, vol. 2, no. 4, p. 040101, 2021.

[33] D. Gottesman, "Quantum error correction and fault-tolerance," *Quantum Information Processing: From Theory to Experiment*, vol. 199, p. 159, 2006.

[34] T. Brun, I. Devetak, and M.-H. Hsieh, "Correcting quantum errors with entanglement," *Science*, vol. 314, no. 5798, pp. 436–439, 2006.

[35] T. A. Brun, I. Devetak, and M.-H. Hsieh, "Catalytic quantum error correction," *IEEE Transactions on Information Theory*, vol. 60, no. 6, pp. 3073–3089, 2014.

[36] M. Grassl, "Entanglement-assisted quantum communication beating the quantum singleton bound," *Physical Review A*, vol. 103, no. 2, p. L020601, 2021.

[37] D. W. Leung, M. A. Nielsen, I. L. Chuang, and Y. Yamamoto, "Approximate quantum error correction can lead to better codes," *Physical Review A*, vol. 56, no. 4, p. 2567, 1997.

[38] B. Schumacher and M. D. Westmoreland, "Approximate quantum error correction," *Quantum Information Processing*, vol. 1, pp. 5–12, 2002.

[39] C. Bény and O. Oreshkov, "General conditions for approximate quantum error correction and near-optimal recovery channels," *Physical Review Letters*, vol. 104, no. 12, p. 120501, 2010.

[40] R. Klesse, "Approximate quantum error correction, random codes, and quantum channel capacity," *Physical Review A*, vol. 75, no. 6, p. 062315, 2007.

[41] J. Li, "Some progress on quantum error correction for discrete and continuous error models," *IEEE Access*, vol. 8, pp. 46998–47012, 2020.

[42] D.-S. Wang, G. Zhu, C. Okay, and R. Laflamme, "Quasi-exact quantum computation," *Physical Review Research*, vol. 2, no. 3, p. 033116, 2020.

[43] D.-S. Wang, Y.-J. Wang, N. Cao, B. Zeng, and R. Laflamme, "Theory of quasi-exact fault-tolerant quantum computing and valence-bond-solid codes," *arXiv preprint arXiv:2105.14777*, 2021.

[44] H. Barnum, M. A. Nielsen, and B. Schumacher, "Information transmission through a noisy quantum channel," *Physical Review A*, vol. 57, no. 6, p. 4153, 1998.

[45] N. Yamamoto, "Exact solution for the max-min quantum error recovery problem," in *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pp. 1433–1438, IEEE, 2009.

[46] Y.-J. Sun, M. Xu, and Y. Deng, "An optimal quantum error-correcting procedure using quantifier elimination," *Quantum Information Processing*, vol. 20, no. 5, p. 170, 2021.

[47] A. E. Ashikhmin, A. M. Barg, E. Knill, and S. N. Litsyn, "Quantum error detection. I. Statement of the problem," *IEEE Transactions on Information Theory*, vol. 46, no. 3, pp. 778–788, 2000.

[48] A. E. Ashikhmin, A. M. Barg, E. Knill, and S. N. Litsyn, "Quantum error detection. II. Bounds," *IEEE Transactions on Information Theory*, vol. 46, no. 3, pp. 789–800, 2000.

[49] N. M. Linke, M. Gutierrez, K. A. Landsman, C. Figgatt, S. Debnath, K. R. Brown, and C. Monroe, "Fault-tolerant quantum error detection," *Science Advances*, vol. 3, no. 10, p. e1701074, 2017.

[50] M. Koashi and M. Ueda, "Reversing measurement and probabilistic quantum error correction," *Physical Review Letters*, vol. 82, no. 12, p. 2598, 1999.

[51] J. Fern and J. Terilla, "Probabilistic quantum error correction," *arXiv preprint quant-ph/0209058*, 2002.

[52] K. Wang, X. Zhao, and T. Yu, "Environment-assisted quantum state restoration via weak measurements," *Physical Review A*, vol. 89, no. 4, p. 042320, 2014.

[53] P. Barberis-Blostein, D. Norris, L. Orozco, and H. Carmichael, "From quantum feedback to probabilistic error correction: manipulation of quantum beats in cavity QED," *New Journal of Physics*, vol. 12, no. 2, p. 023002, 2010.

[54] X. Xiao and Y.-L. Li, "Protecting qutrit-qutrit entanglement by weak measurement and reversal," *The European Physical Journal D*, vol. 67, no. 10, pp. 1–7, 2013.

[55] S.-C. Wang, Z.-W. Yu, W.-J. Zou, and X.-B. Wang, "Protecting quantum states from decoherence of finite temperature using weak measurement," *Physical Review A*, vol. 89, no. 2, p. 022318, 2014.

[56] A. Scott, "Probabilities of failure for quantum error correction," *Quantum Information Processing*, vol. 4, no. 5, pp. 399–431, 2005.

[57] K. Kasai, M. Hagiwara, H. Imai, and K. Sakaniwa, "Quantum error correction beyond the bounded distance decoding limit," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 1223–1230, 2011.

[58] L.-M. Duan and G.-C. Guo, "Probabilistic cloning and identification of linearly independent quantum states," *Physical Review Letters*, vol. 80, no. 22, p. 4999, 1998.

[59] M. Sedlák, A. Bisio, and M. Ziman, "Optimal probabilistic storage and retrieval of unitary channels," *Physical Review Letters*, vol. 122, no. 17, p. 170502, 2019.

[60] Z. Puchała, Ł. Pawela, A. Krawiec, R. Kukulski, and M. Oszmaniec, "Multiple-shot and unambiguous discrimination of von Neumann measurements," *Quantum*, vol. 5, p. 425, 2021.

[61] E. Knill, R. Laflamme, and L. Viola, "Theory of quantum error correction for general noise," *Physical Review Letters*, vol. 84, no. 11, p. 2525, 2000.

[62] J. Watrous, *The theory of quantum information.* Cambridge University Press, 2018.

[63] W. Bruzda, V. Cappellini, H.-J. Sommers, and K. Zyczkowski, "Random quantum operations," *Physics Letters. A*, vol. 373, no. 3, pp. 320–324, 2009.

[64] M. A. Nielsen and I. L. Chuang, "Quantum computation and quantum information," 2010.

[65] I. Bengtsson and K. Życzkowski, *Geometry of quantum states: an introduction to quantum entanglement.* Cambridge University Press, 2017.

[66] K.-E. Hellwig and K. Kraus, "Pure operations and measurements," *Communications in Mathematical Physics*, vol. 11, no. 3, pp. 214–220, 1969.

[67] A. Y. Kitaev, "Quantum computations: algorithms and error correction," *Russian Mathematical Surveys*, vol. 52, no. 6, p. 1191, 1997.

[68] J. Watrous, "Semidefinite programs for completely bounded norms," *arXiv preprint arXiv:0901.4709*, 2009.

[69] M. L. Mehta, *Random matrices.* Elsevier, 2004.

[70] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An introduction to random matrices.* No. 118, Cambridge University Press, 2010.

[71] J. A. Mingo and R. Speicher, *Free probability and random matrices*, vol. 35. Springer, 2017.

[72] J. Ginibre, "Statistical ensembles of complex, quaternion, and real matrices," *Journal of Mathematical Physics*, vol. 6, no. 3, pp. 440–449, 1965.

[73] L. Zhang, "Dirac delta function of matrix argument," *International Journal of Theoretical Physics*, vol. 60, no. 7, pp. 2445–2472, 2021.

[74] K. Zyczkowski and H.-J. Sommers, "Truncations of random unitary matrices," *Journal of Physics A: Mathematical and General*, vol. 33, no. 10, p. 2045, 2000.

[75] M. Pozniak, K. Zyczkowski, and M. Kus, "Composed ensembles of random unitary matrices," *Journal of Physics A: Mathematical and General*, vol. 31, no. 3, p. 1059, 1998.

[76] F. Mezzadri, "How to generate random matrices from the classical compact groups," *Notices of the American Mathematical Society*, vol. 54, no. 5, pp. 592–604, 2007.

[77] R. Duan, H. Wu, and R. Zhou, "Faster matrix multiplication via asymmetric hashing," *arXiv preprint arXiv:2210.10173*, 2022.

[78] G. Gour, "Comparison of quantum channels by superchannels," *IEEE Transactions on Information Theory*, vol. 65, no. 9, pp. 5880–5904, 2019.

[79] A. Bisio, G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Quantum networks: general theory and applications," *arXiv preprint arXiv:1601.04864*, 2016.

[80] H.-J. Sommers and K. Życzkowski, "Statistical properties of random density matrices," *Journal of Physics A: Mathematical and General*, vol. 37, no. 35, p. 8457, 2004.

[81] K. Zyczkowski and H.-J. Sommers, "Induced measures in the space of mixed quantum states," *Journal of Physics A: Mathematical and General*, vol. 34, no. 35, p. 7111, 2001.

[82] J. Wallman, C. Granade, R. Harper, and S. T. Flammia, "Estimating the coherence of noise," *New Journal of Physics*, vol. 17, no. 11, p. 113020, 2015.

[83] B. Collins, "Moments and cumulants of polynomial random variables on unitarygroups, the itzykson-zuber integral, and free probability," *International Mathematics Research Notices*, vol. 2003, no. 17, pp. 953–982, 2003.

[84] B. Collins and I. Nechita, "Random quantum channels i: graphical calculus and the bell state phenomenon," *Communications in Mathematical Physics*, vol. 297, no. 2, pp. 345–370, 2010.

[85] Z. Puchała and J. A. Miszczak, "Symbolic integration with respect to the haar measure on the unitary group," *arXiv preprint arXiv:1109.4244*, 2011.

[86] L. Devroye, "Sample-based non-uniform random variate generation," in *Proceedings of the 18th conference on Winter simulation*, pp. 260–265, 1986.

[87] B. Collins and P. Śniady, "Integration with respect to the haar measure on unitary, orthogonal and symplectic group," *Communications in Mathematical Physics*, vol. 264, no. 3, pp. 773–795, 2006.

[88] K. Korzekwa, S. Czachórski, Z. Puchała, and K. Życzkowski, "Coherifying quantum channels," *New Journal of Physics*, vol. 20, no. 4, p. 043028, 2018.

[89] I. Nechita, Z. Puchała, Ł. Pawela, and K. Życzkowski, "Almost all quantum channels are equidistant," *Journal of Mathematical Physics*, vol. 59, no. 5, p. 052201, 2018.

[90] Z. Puchała, Ł. Pawela, and K. Życzkowski, "Distinguishability of generic quantum states," *Physical Review A*, vol. 93, no. 6, p. 062112, 2016.

[91] M. Fukuda, R. König, and I. Nechita, "Rtni—a symbolic integrator for haar-random tensor networks," *Journal of Physics A: Mathematical and Theoretical*, vol. 52, no. 42, p. 425303, 2019.

[92] B. Collins, I. Nechita, and D. Ye, "The absolute positive partial transpose property for random induced states," *Random Matrices: Theory and Applications*, vol. 1, no. 03, p. 1250002, 2012.

[93] D. W. Kribs, R. Laflamme, D. Poulin, and M. Lesosky, "Operator quantum error correction," *arXiv preprint quant-ph/0504189*, 2005.

[94] S. Hill and W. K. Wootters, "Entanglement of a pair of quantum bits," *Physical Review Letters*, vol. 78, no. 26, p. 5022, 1997.

[95] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, "On quantum Rényi entropies: A new generalization and some properties," *Journal of Mathematical Physics*, vol. 54, no. 12, p. 122203, 2013.

[96] D. Braun, O. Giraud, I. Nechita, C. Pellegrini, and M. Žnidarič, "A universal set of qubit quantum channels," *Journal of Physics A: Mathematical and Theoretical*, vol. 47, no. 13, p. 135302, 2014.

[97] M. Girard, D. Leung, J. Levick, C.-K. Li, V. Paulsen, Y. T. Poon, and J. Watrous, "On the mixed-unitary rank of quantum channels," *Communications in Mathematical Physics*, pp. 1–33, 2022.

[98] M.-D. Choi, D. W. Kribs, and K. Życzkowski, "Quantum error correcting codes from the compression formalism," *Reports on Mathematical Physics*, vol. 58, no. 1, pp. 77–91, 2006.

[99] M. A. Nielsen and I. Chuang, "Quantum computation and quantum information," 2002.

[100] M. Grassl, F. Huber, and A. Winter, "Entropic proofs of singleton bounds for quantum error-correcting codes," *IEEE Transactions on Information Theory*, vol. 68, no. 6, pp. 3942–3950, 2022.

[101] H. K. Ng and P. Mandayam, "Simple approach to approximate quantum error correction based on the transpose channel," *Physical Review A*, vol. 81, no. 6, p. 062342, 2010.

[102] S. Beigi, N. Datta, and F. Leditzky, "Decoding quantum information via the petz recovery map," *Journal of Mathematical Physics*, vol. 57, no. 8, p. 082203, 2016.

[103] M. Berta, F. Borderi, O. Fawzi, and V. B. Scholz, "Semidefinite programming hierarchies for constrained bilinear optimization," *Mathematical Programming*, pp. 1–49, 2021.

[104] M. Reimpell and R. F. Werner, "Iterative optimization of quantum error correcting codes," *Physical Review Letters*, vol. 94, no. 8, p. 080501, 2005.

[105] R. L. Kosut and D. A. Lidar, "Quantum error correction via convex optimization," *Quantum Information Processing*, vol. 8, pp. 443–459, 2009.

[106] S. Taghavi, R. L. Kosut, and D. A. Lidar, "Channel-optimized quantum error correction," *IEEE Transactions on Information Theory*, vol. 56, no. 3, pp. 1461–1473, 2010.

[107] A. S. Fletcher, P. W. Shor, and M. Z. Win, "Structured near-optimal channel-adapted quantum error correction," *Physical Review A*, vol. 77, no. 1, p. 012320, 2008.

[108] A. S. Fletcher, P. W. Shor, and M. Z. Win, "Optimum quantum error recovery using semidefinite programming," *Physical Review A*, vol. 75, no. 1, p. 012338, 2007.

[109] P. Gawron, D. Kurzyk, and Ł. Pawela, "QuantumInformation.jl—a julia package for numerical computation in quantum information theory," *PLOS ONE*, vol. 13, p. e0209358, dec 2018.

[110] B. O'Donoghue, E. Chu, N. Parikh, and S. Boyd, "Conic optimization via operator splitting and homogeneous self-dual embedding," *Journal of Optimization Theory and Applications*, vol. 169, pp. 1042–1068, June 2016.

[111] B. O'Donoghue, E. Chu, N. Parikh, and S. Boyd, "SCS: Splitting conic solver, version 3.2.1." `https://github.com/cvxgrp/scs`, Nov. 2021.

[112] `https://github.com/rkukulski/phd_code`. Permanent link to code/repository, Accessed: 2023-04-27.

[113] C. Cafaro and S. Mancini, "Quantum stabilizer codes for correlated and asymmetric depolarizing errors," *Physical Review A*, vol. 82, no. 1, p. 012306, 2010.