Scalable approaches to the characterization of open quantum system dynamics

by

Cecilia Carolina López

Submitted to the Department of Nuclear Science and Engineering in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2009

© Massachusetts Institute of Technology 2009. All rights reserved.

Author

Department of Nuclear Science and Engineering

August 20, 2009

Certified by

David G. Cory

Professor of Nuclear Science and Engineering

Thesis Supervisor

Read by

Paola Cappellaro

Professor of Nuclear Science and Engineering

Thesis reader

Accepted by

Jacquelyn C. Yanch

Chair, Department Committee on Graduate Students
Scalable approaches to the characterization of open quantum system dynamics

by

Cecilia Carolina López

Submitted to the Department of Nuclear Science and Engineering on August 20, 2009, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Abstract

One of the biggest challenges in the physical realization of quantum information processing (QIP) is the precise control of the system. In order to achieve this, we characterize the gates, errors, and noise occurring in experimental setups. In this thesis we develop and further study characterization methods, putting particular emphasis on the scalability problem: $O(D^4)$ parameters describe the dynamics of an open quantum system of dimension $D$, thus $O(D^4)$ resources are in principle required to characterize it – which is a problem in QIP where the desired systems are large ($D = 2^n$ for $n$ qubits). We first study the fidelity decay (also called Loschmidt echo) of the system, for many steps under the progressive randomization due to a one-qubit twirl. We show how this quantity encodes useful information about the process: begin twirled. We then present a method to measure the magnitude of the multi-body correlations that scales as $O(n^n)$, when only up to $w$-body interactions are expected among the $n$ qubits. We implemented this method in a four-qubit liquid-state Nuclear Magnetic Resonance (NMR) QIP device, demonstrating its potential and feasibility. The experimental work also pointed out the need for robust procedures and the role of implementation errors, while deepening our knowledge of NMR QIP dynamics. We also report on several practical aspects of the experiment, including details on twirls using random rotations and Clifford operators. We furthermore relate this work to recent developments in the community, arriving to a more comprehensive protocol and establishing an intrinsic hierarchy of characterization algorithms. Finally, we study the many-step fidelity decay when using a flawed twirl, thus acknowledging the most realistic scenario where we have a faulty device attempting to characterize itself. Our preliminary work points towards the use of a many-step scheme that promises robust scalable tools to characterize the twirl operators themselves.

Thesis Supervisor: David G. Cory
Title: Professor of Nuclear Science and Engineering
Acknowledgments

I would like to first acknowledge my advisor Prof. David Cory for giving me the opportunity to work with him and teaching me so much about experimental physics. His amazing intuition and vast experience with NMR has been a great source of ideas and challenges, and I hope what he has taught me continues to shape and push my thinking. I joined the group because I wanted to “build bridges between theory and experiments”, and although the bridges turned out to be much longer than I expected, I owe him the progress I have made so far. I am also very grateful for his patience during our talks, without it I would not have found my ways around the research we did in these five years. Moreover, I want to thank him for his care and concern that has extended beyond the duty of research. He has supported me in having a healthy life outside the lab during these years, and also given me lots of guidance and help when finding a promising future after my PhD.

Secondly, I thank Dr. Benjamin Lévi, who has not only been my partner in crime during half of my PhD work, but also because for the bulk of knowledge he has passed to me, and more importantly for teaching and keeping me straight on my math and my logical thinking. I am very grateful for all the talks about work, physics, and science in general, that we had from the first days in the lab, over time (in Boston, Waterloo, Rome, Montreal or Bangalore!), and even now chatting on-line at odd hours Boston-Paris over his proof-reading of my thesis. It has always been a pleasure to discuss with him, and I am very grateful for his ability to explain things so clearly and soundly to me. Ben, I will never forget our struggle for the higher cause –the Pauli Matrix Liberation Front, nor your high standards for better Greek letters when running out of ways to denote $\rho$.

Thanks to Dr. Jonathan Hodges, whose role of the gloomy-grad-student-of-the-lab is only a cover for the best, most dedicated and enthusiastic combination of physicist and teacher I know. I not only have to thank him for pushing me to do more (… asking me every day if my experiments were done or sending me papers with a note “you may be interested in this”), but also for his great insight into experimental physics of all sorts. I am indeed very grateful for the energy with which he fuels the lab, the spectrometer help, his tireless will to explain how anything works, and to listen carefully and critically to any progress or problem I had. I would equally like to thank Prof. Paola Cappellaro, who has been a role-model of solid and balanced scientific career all the way. She has also spent lots of time explaining things around the lab, and listened critically to my deductions and questions. I am indebted to her for her assistance and advice at the critical points of my career: my qualifying exams, my postdoc search, and my thesis stage.

During these years I have had the fortune of having the most distracting officemate ever, Dr. Dmitry Pushin, who I also have to thank for examining my physics reasoning on several occasions (you know, he is Russian).

The Cory group has been a really cool and supportive crowd to work and share life with, for which I am very grateful: Clarice Aiello, Kevin Krsulich, Troy Borneman, Karen Lee, Sergio Valenzuela, Sarah Sheldon, Mohamed Abutaleb, Fei Yan, Jamie Yang, Michael Henry, Anatoly Dementyev, Kai Iamsunang, Matt Davidson, Sid Sinha, Peter Allen, Tim Havel, and Sekhar Ramanathan. Special thanks to Mike, for teaching me so much about the spectrometer and the lab (including secret commands and tricks); to Troy, for being my connection with the GRAPE world, dealing with the 6 $\mu$sec, and for the across-the-cluster conversations; to Kevin, for his patience and help during my computer panic attacks; and to Sekhar, for addressing my various chemistry and electronics questions, and for the permanent good vibes.
I particularly thank Karen and Clarice, for being such great friends and helping me out to take my next step after the PhD. I not only shared time in front of the spectrometer and/or the simulator with them, but also talks about science and the scientific life (and about the non-scientific life too).

I would like to thank Prof. Joseph Emerson and Dr. Marcus da Silva, for valuable discussions and explanations, that definitely improved or triggered several pieces of my work. Thanks to Prof. Lorena Viola, for many interesting discussions on a bigger frame, and for her encouragement and warmth.

Thanks also to Rachel Batista, Clare Egan and Jennifer deVries Gwinn from the NSE Department, who have certainly made life easier. Special thanks to Rachel, for her care and the flowers.

This thesis would not have been possible without the love, company, help, and support from the people that have shaped my life in Cambridge during these five years. In many ways, these are the people I have to thank the most.

From day one till today, Anuja Mahashabde and Tzu-Ching Horng have always been there for me in all possible ways, becoming just like family, through uncountable hours of talking, cooking, moving, going to restaurants, so much traveling, life adventures, tears and laughter. Thank you for opening your lives to me and making me a part of them. Thank you Tzu-Ching for never giving up on me, for many long-distance long phone calls, and for being just the way you are. Thank you Anuja for... well, for everything.

Two years down the road I met Alejandra Menchaca and Gina Escobar, who became my “Spanish-speaking friends”. I owe to them also for so many hours of talking, especially in the most difficult times. I want thank them for teaching me so much about a homeland I did not know about, for putting so many things under the right light, for the warmth, the shouting, the day trips, and the parties, and above all, the dancing. To Gini, thanks for all the traveling, down and up, and for the Lucky Star rides and our never enough New Yorker wannabe times. To Ale, thanks “for organizing”, for the full-day hanging-outs at home, and for the chipotle dip!

Thanks to Liang Sim, for being the rock you can always hold on to, my personal 911 line, my favorite proof-reader, and for being a great friend, always ready to hang out and even reader to help. Thanks to Jerry Chen, for the good times, the music, the “little unknown places”, the help, and for bringing the Bay area to Boston as much as he could. Thanks to Mridula Pore, for her warm touch, for sharing her life with me, and the unforgettable cooking lessons and sessions: I make daal almost every week. Thanks to Mira Chokshi, for our talks, her candor and wisdom, for the fun times, for lending me her family, and for cooking lessons and sessions too (mutter paneer is now my favorite entertaining dish). Thanks to Hui Deng, for her sweet touch and our talks while walking arm in arm.

And again, thanks to Dima, Ben, Paola, Karen, Sergio and Clarice, this time on the friendship side and for all kinds of help. Frankly, life in Boston would have not been the same without Dima and Ben, the parties, the dancing, the jokes, the traveling, and, of course, their fundamental insight into the male mind. Also, for the people they brought to my life: special thanks to Roman Barankov there, who has always been present at critical moments.

Almost at the end of the path, I have to thank Anna Laromaine and Inés García de la Puente, who have been very important in this special stage that is the last year, for their friendship, good times, and their support for finishing this stage and preparing for the next.

Thanks also to Erin Boyd, my NucE homework buddy, for being such great friend during
my first years at MIT. I am very grateful she opened her life and culture to me, and for all
she taught me (and also for the dancing).

On the roommate side, I have to thank Ying Yin, Nawa Sugiyama, Liang Sim and
Mridula Pore. Fate was on my side when I moved in with them, thank you for your warmth
and care.

Also I deeply thank the Mahashabde family, like I always say, my “adoptive family” in
the US.

All of you have been fundamental to my years at MIT. I will be always grateful, I would
not have finished my PhD if you had not been here. But above all, I would like to thank
the people that have been the strongest support of all.

Cristina García, the most loving and caring auntie/friend I only dreamed of finding here.
She has been the family from Argentina and the friend from the US, and she has kept me in
one piece through the tough times that passed in these years, and given me so so so many
hours of talk, joy and laughter. Diego Zocco, for how much he has helped and encouraged
me as physicist, and through life in general, from years even before MIT and the US. He
has always believed in me, giving me confidence, wisdom and unconditional support. And
finally, I thank my everlasting source of love and care, my family: my sister Laurita (“lil’
sis”), my brother Fede, my godmother Ali, my grandma Kiki, and my mom and dad.

To you I dedicate this thesis.
Contents

1 Introduction: Quantum information processors and the problem of characterizing quantum dynamics 17

2 Fidelity decay, twirling, their role in characterizing quantum dynamics, and other basic definitions 19
   2.1 The concept of fidelity .............................................. 19
   2.2 A more general framework ......................................... 21
   2.3 The concept of twirling and its relation with the fidelity ....... 22
   2.4 The generalized fidelity decay for many steps .................. 24

3 A survey of quantum processes 27
   3.1 General statement .................................................. 27
   3.2 CP vs. non-CP maps ............................................... 27
   3.3 Unital processes ................................................... 28
   3.4 Gate-dependent operations ........................................ 29
   3.5 Fluctuating ensembles ............................................. 29
       3.5.1 Dephasing Gaussian noise ................................... 30
       3.5.2 Field inhomogeneity ......................................... 31
       3.5.3 Gaussian noise in an applied field ......................... 31
       3.5.4 Error at the measurement, with Gaussian noise .......... 32
   3.6 Ensembles, incoherence and decoherence ....................... 33
   3.7 Incoherence, decoherence and memory effects ................. 35

4 The generalized fidelity decay of n qubits using one-qubit twirling 37
   4.1 General statement .................................................. 37
   4.2 Model of the map under study: Noisy channels ................. 38
   4.3 Results for specific models ...................................... 39
       4.3.1 Analytical results for $G_T$ with one-body terms only ... 40
       4.3.2 Numerical results ........................................... 44
   4.4 The initial decay rate $\gamma$ .................................... 46
   4.5 Closing remarks .................................................. 51

5 Experimental implementation of a one-step protocol to characterize spatial correlations 55
   5.1 Theoretical statement of the protocol ............................ 56
   5.2 The protocol ....................................................... 58
       5.2.1 Scalability of the method ................................. 59
   5.3 Experimental QIP in liquid-state NMR ............................ 61
5.3.1 The basics ................................................................. 61
5.3.2 Measurement ............................................................ 64
5.3.3 Simulation ................................................................. 66
5.3.4 Incoherent sum treatment of the hydrogens ....................... 67
5.3.5 rf field inhomogeneities ............................................. 69
5.4 The experiment ............................................................. 69
5.4.1 The processes under study ......................................... 70
5.4.2 The initial state preparation ....................................... 71
5.4.3 The twirl ................................................................. 72
5.4.4 The measurement ...................................................... 75
5.4.5 Negligibility of higher order multi-body terms .................. 76
5.5 Experimental results ................................................... 77
5.6 Closing remarks .......................................................... 80

6 Comparison between different protocols using twirling .......... 81
6.1 Re-stating the basics .................................................... 81
6.2 Extracting the $\chi$-coefficients ................................... 84
6.3 Comparison with the other methods: twirling in $U(2^n)$ vs. twirling in $U(D)$ . 86
6.4 What information are we extracting? ............................... 88

7 Generalized fidelity decays
   for many steps .................................................................. 91
7.1 Generalized fidelity decays for many steps under progressive randomization . 91
7.2 A tale of two schemes ................................................... 93

8 Case studies of the generalized fidelity decays using faulty twirl operators,
   for one qubit .................................................................. 101
8.1 Physical model .............................................................. 104
8.2 Clifford gates with their angle off by a fixed amount $a$ ............ 104
8.3 Clifford gates with their direction rotated by an angle $b$
    around $\hat{z}$ ............................................................... 108
8.4 Some Clifford gates with their angle off .............................. 109
8.5 Other models ............................................................... 112
8.6 Closing remarks .......................................................... 114

9 Conclusions ..................................................................... 119
9.1 Towards efficient experimental quantum process characterization ........ 120

A Measuring probabilities by repeating experiments: Useful statistical properties . 123
A.1 The Chernoff bound ....................................................... 123
   A.1.1 Main proof ......................................................... 124
   A.1.2 Bound including both the lower and upper tails ................ 125
   A.1.3 Markov Inequality ................................................... 126
   A.1.4 Lemma I ............................................................... 126
   A.1.5 Arithmetic-Geometric Mean Inequality ......................... 127
A.2 Practical use of Chernoff Bound ....................................... 127
A.3 Mean and variance of $N$ realizations of a Bernoulli variable .......... 129
A.4 Application to the measurement of probabilities in twirling protocols .... 130
B Generalities on the Clifford twirl
B.1 Equivalence between the Clifford twirl and the Haar twirl .............. 135
B.2 Details on the Clifford gates .............................................. 136
   B.2.1 Clifford gates for one qubit ........................................ 137
B.3 Specifics on the Clifford twirl in $U(2)$ ................................. 138

C Analytical calculation of the fidelity decay after one step of a multiple
one-qubit twirl
C.1 The system ................................................................. 143
C.2 The model for the process $\Gamma$ under study ............................ 143
C.3 Randomizing the fidelity by twirling .................................... 145
C.4 Cumbersome calculations .................................................. 146
C.5 The fidelity decay after one step ........................................ 149
C.6 Characterizing a general CP map $\Gamma$ vs. characterizing a generator $G_T$ ................................. 150
C.7 Using a partial twirl in $\mathcal{H}_M$ ...................................... 150
C.8 The fidelity decay after one step with a partial twirl .................. 153
C.9 Expressions of the fidelity decay $\Delta f^{(M)}$ for a few qubits .......... 154

Bibliography .............................................................................. 156
List of Figures

2-1 Circuit representation of the twirl of a map. .......................... 23
2-2 Circuit representation of the twirl by sampling. ......................... 24
2-3 Circuit representation of the generalized fidelity over time. .... 25

3-1 The action of the map under study along a fidelity decay experiment. . . 27

4-1 Circuit representation of the algorithm with a one-qubit twirl ........... 38
4-2 Examples of the numerical calculation of $f^L(E, \rho_0, t)$ ............... 41
4-3 Fidelity decays with only one-body terms in $G_T$ .......................... 43
4-4 Examples of the numerical calculation of $f^L(E, \rho_0, t)$ including first-neighbor coupling terms in $G_T$ .................................................. 45
4-5 Examples of the numerical calculation of $f^L(E, \rho_0, t)$ including two-body terms in $G_T$ ....................................................... 45
4-6 Examples of $\zeta$, the fitting parameter for the exponential-like behavior. . 46
4-7 Examples of the decay rate $\gamma$, the fitting parameter for the initial decay. 47
4-8 Circuit representation of the algorithm focussing on a subset of $m$ qubits. 50
4-9 Example of local noise. ............................................................. 52
4-10 Example of non-local noise. ..................................................... 52

5-1 Circuit representation of the one-step protocol to measure spatial correlations. 58
5-2 The crotonic acid molecule. ..................................................... 61
5-3 Example of a pulse profile (CNOT gate for qubits 1 & 2). ................. 71
5-4 Initial state correlation (theory vs. experiment), for the pair [1, 2]. .... 72
5-5 Correlation between initial state and the implementation of the twirl without any map, to analyze the unitarity of the experimental Clifford gates (pair [2,3]). 75
5-6 Correlation between initial state and the implementation of the twirl without any map, to analyze the unitarity of the experimental Clifford gates (pair [1,4]) 76

6-1 Circuit representation of the general one-step protocol with one-qubit twirling. 83
6-2 Circuit representation a twirled $\Gamma_m$ with $\Gamma_m(\rho) = P_m \Gamma(\rho) P_m$. 87

7-1 Circuit representation of two schemes for the generalized fidelity decay over time. ................................................................. 93
7-2 Matched simulations of the generalized fidelities for a liquid-state NMR QIP device (crotonic acid in a 400MHz spectrometer). ................. 95
7-3 Inset on the simulations of the generalized fidelities for a liquid-state NMR QIP device (crotonic acid in a 400MHz spectrometer). ................. 96
Simulations of the generalized fidelities for a liquid-state NMR QIP device (crotonic acid in a 400MHz spectrometer) under the LE scheme: the perfect reversal approximation. ................................................. 97

Simulations of the generalized fidelities for a liquid-state NMR QIP device (crotonic acid in a 400MHz spectrometer), using the whole Clifford group and half of it. ................................................................. 102

Generalized fidelities using Clifford gates with a global error in their angle. .. 105

Plots of the characterizing eigenvalues (modulus and angle) when the Clifford gates have their angle off. ................................................................. 106

The LE-fidelity when the Clifford gates have their angle off, by a large amount a. ................................................................. 108

Matching of the LE-fidelity and the RT-fidelity when the Clifford gates have their angle off. ................................................................. 109

Generalized fidelities using Clifford gates with an error in their direction. . . 110

Plots of the characterizing eigenvalues (modulus and angle) when the Clifford gates $C_1$, $C_5$ and $C_9$ have their angle off. ................................................................. 111

Generalized fidelities when the Clifford gates $C_1$, $C_5$ and $C_9$ have their angle off. ................................................................. 112

The LE-fidelity when the Clifford gates $C_1$ and $C_5$ have their angle off. . . 113

Generalized fidelities using Clifford gates with a shift in $\theta n_\gamma$. .................. 114

The LE-fidelity when the Clifford gates $C_1$, $C_2$, $C_3$ and $C_4$ have their direction rotated by an angle $b$ around $\hat{z}$. ................................................................. 115

The LE-fidelity when the Clifford gates $C_1$, $C_5$ and $C_9$ have their angles off by a fluctuating small amount $x$. ................................................................. 117

Coin tossing statistics, using $N = 150$ realizations. ........................................... 128

Coin tossing statistics, using $N = 1000$ realizations. ......................................... 128

Gaussian approximation to the coin tossing statistics ............................................. 130

Statistics of the sampling of the twirl plus strong measurement, for one qubit. 132

Statistics of the sampling of the twirl plus strong measurement, for $N < |C|$ (three qubits). ................................................................. 133
List of Tables

4.1 Values for the proportionality constant $c$ between $\gamma$ and the strength $\eta^2$ arising from the fitting. ................................................................. 48

5.1 Scaling of the overhead of the cost of performing a characterization of the $\eta_I$ coefficients. ................................................................. 60

5.2 Experimental performance of the Clifford gates. .......................... 73

5.3 Results for the measured two-qubit collective coefficients for selected pairs of qubits. ................................................................. 79

B.1 Clifford gates for one qubit. ......................................................... 138

B.2 Decomposition of the Clifford gates for one qubit. ......................... 139
Chapter 1

Introduction: Quantum information processors and the problem of characterizing quantum dynamics

Quantum information processing (QIP) has become a major field of research because of its potential broad variety of applications: faster-than-classical algorithms, unique quantum processes to perform novel tasks (teleportation, quantum cryptography, etc.), efficient simulation of quantum systems, and a thought-provoking role in the exploration of quantum mechanics and its foundations. But despite of great developments, many implementation issues remain unsolved.

One of the biggest challenges in the physical realization of QIP is the precise control of the physical system. In order to achieve this, we characterize the noise and errors affecting experimental setups. Although this may seem a rather technical problem, it actually lies at the heart of quantum mechanics. Information processing and computation are physical tasks. When they are quantum, we must deal with the unique features of quantum physics, such as measurement results that are probabilistic and the possibility of having the system in a variety of states much richer than its classical counterpart.

In particular, when our system is composed of $n$ subsystems each with $d$ possible states (for example, $n$ atoms with each $d$ energy eigenstates), the system as whole can be found in $d^n$ possible states. Any process occurring in a closed system will be unitary and in turn characterized by $(d^n)^2$ real parameters (for example, the matrix elements of the Hamiltonian governing the system). Moreover, natural physical systems are very difficult to keep closed/isolated, so we must consider the dynamics of the open system. In this case, any arbitrary (not necessarily unitary) process will be described by $d^{4n}$ real parameters. We will have $d^{4n} - d^{2n}$ real parameters if we impose our process to be also trace-preserving (in principle, quantum processes need only to be trace-non-increasing). We see then that the problem of characterizing a quantum process involves a number of parameters that scales exponentially with the number of subsystems. It follows then that a strategy aiming to extract this information may involve a number of resources (experiments, measurements, or operations to be performed on the system or on the measurement data) that will be $O(d^{4n})$, thus being exponentially hard to implement. In QIP, where the subsystems are the qubits ($d = 2$: a two-level system), one is faced with the scalability problem of error characterization: the more qubits we have, the more powerful our quantum computer, but the harder
it is to characterize the performance of the QIP device. Ultimately, the aim is to achieve robust quantum process characterization methods demanding resources that scale at most polynomially in \( n \) -so-called scalable methods. And this is indeed the core of the PhD work presented here.

The first approach aiming to characterize a quantum process was given by quantum state and quantum process tomographies (QST, QPT), which allow us to fully reconstruct the dynamics occurring in the probed system after a given time step (see [1] and [2] Sec. 8.4.2). Experimental implementations of QPT have been conducted in a variety of small systems which are of interest in QIP: liquid state Nuclear Magnetic Resonance (NMR) (for example, for the quantum Fourier transform operation in [3], and for the CNOT gate in [4]), ion traps [5], linear optics [6], quantum optical lattices [7], Nitrogen-Vacancy centers [8] and superconducting qubits [9]. QPT is based on preparing the system in each of its possible states and superpositions of them (a total of \( O(2^n) \) individual states) and measuring the final state by QST after implementing the process under study. QST for one state requires \( O(2^n) \) measurements, therefore QPT requires \( O(2^{4n}) \) experiments for a system of \( n \) qubits. Thus this procedure becomes impractical beyond a few qubits.

With the advancement of QIP however, strategies for assessing the reliability of an imperfect device and counteracting errors have emerged. These are the fault-tolerance thresholds and quantum error correction codes (QECC), respectively. In particular, they have pointed out that is not necessary to know “everything” about the process occurring in a QIP device, but only certain aspects of it (for example, the fidelity of an operation or the error probabilities). Following this, in recent years the idea of getting less information at a lower cost has become a popular strategy to tackle quantum process characterization, and several works have been devoted to the subject [10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22]. Our work fits in this context: resorting on sampling and the grouping (coarse-graining) of the original \( 2^n \) parameters, we aim to obtain partial but nevertheless relevant information about the process under study in a scalable way.

The first part of this thesis introduces the concepts of fidelity decay and twirling and their relevance to quantum process characterization, and general mathematical tools and definitions (Chapter 2). Furthermore in Chapter 3 we explore different types of quantum processes that occur in the laboratory and that are the object of study.

The second part of this thesis focusses on the use of a one-qubit twirl. In Chapter 4, we study the behavior of the fidelity under the progressive randomization of the noise with time. The results on this led to the development of a protocol to characterize the spatial correlations of an arbitrary quantum process, which we experimentally tested in a liquid-state NMR device. Chapter 5 presents the protocol and is devoted to the experimental implementation and the analysis of the results. We also discuss the place of this method in relation to other schemes (Chapter 6), all utilizing a one-step twirling approach.

Liquid-state NMR has long served as a test-bed for the exploration and development of techniques to improve the control and performance of QIP devices. In this case, the experimental work of Chapter 5 showed us that no protocol is realistic unless it deals with errors in its own implementation. That is why the third part of this thesis (Chapters 7 and 8) focusses on a bigger picture for using fidelity decays as a tool to characterize quantum processes, taking from Chapter 4 and the work in [13, 14, 17] to consider now a many-step fidelity decay in a more realistic scenario where we face the fact that we are using a faulty device to characterize itself. Finally, in Chapter 9, we present our conclusions.
Chapter 2

Fidelity decay, twirling, their role in characterizing quantum dynamics, and other basic definitions

2.1 The concept of fidelity

The fidelity, also called Loschmidt echo, is defined as follows. Consider that the system under study is in a pure state $|\psi\rangle$ in a Hilbert space $\mathcal{H}$, and let it evolve under some evolution operator (propagator) $\tilde{K}$ but later reverse this motion, so that the system is theoretically back to its initial state. If during the evolution of the system there was a perturbation, so $\tilde{K}$ was implemented instead of $K$, the overlap between the final and the initial state is

$$f(K^{-1} \circ \tilde{K}, \psi) = |\langle \psi | K^\dagger \tilde{K} | \psi \rangle|^2 = \langle \psi | K^\dagger \tilde{K} | \psi \rangle \langle \psi | \tilde{K}^\dagger K | \psi \rangle$$

(2.1)

This is the fidelity $f$. This concept originated in the early days of statistical mechanics from discussions about irreversibility between Josef Loschmidt and Ludwig Boltzmann. It was later introduced to quantum mechanics as a tool to study chaos in quantum systems. A good introduction to the topic, including its history, can be found in [23].

If there is no perturbation, $f = 1$, otherwise, $f < 1$, therefore also the name fidelity decay. The concept of fidelity however has been applied in general to measures of similarity between two states (in this case, between $\tilde{K}|\psi\rangle$ and $K|\psi\rangle$). For example, in [2] Sec. 9.2.2, we have $f(|\phi\rangle, |\varphi\rangle) = |\langle \phi | \varphi \rangle|$, or in [24, 25] $f(|\phi\rangle, |\varphi\rangle) = |\langle \phi | \varphi \rangle|^2$. For this reason, the term Loschmidt echo is probably more accurate. Nevertheless, in the QIP community the name fidelity has prevailed and that is what will be used here - following eq. (2.1).

In QIP, $K$ would be for example a computational gate, and it is of interest to take an average of (2.1) over all possible pure initial states:

$$F(K^{-1} \circ \tilde{K}) = \int d\psi f(K^{-1} \circ \tilde{K}, \psi) = \int d\psi |\langle \psi | K^\dagger \tilde{K} | \psi \rangle|^2$$

(2.2)

The integral is over the uniform (Haar) measure $d\psi$ on (pure) state space, normalized so that $\int d\psi = 1$. Even beyond QIP, the average (2.2) is known to encode information about the operation $K^{-1} \circ \tilde{K}$ (see for example [26]). In the context of QIP, $F$ is called the average fidelity, and characterizes the average performance of a faulty $\tilde{K}^\dagger K$ over the possible pure

19
states of the system. In the most general case, a faulty implementation of $K^\dagger K$ will not be the composition of two unitary operators $K^{-1} \circ K$ as we denoted, but actually a more general non-unitary process, a quantum map\(^5\) $\Gamma$:

$$
\Gamma(\rho) = \sum_k A_k \rho A_k^\dagger
$$

(2.3a)

$$
\sum_k A_k^\dagger A_k = \mathbb{I}
$$

(2.3b)

which is the typical Kraus operator-sum representation of a trace-preserving, hermiticity-preserving, completely positive (CP) map. Alternative, we could use a representation with continuous parameters $\vec{\eta}$

$$
\Gamma(\rho) = \int P(\vec{\eta}) E(\vec{\eta}) \rho E^\dagger(\vec{\eta}) d\vec{\eta}
$$

(2.4a)

$$
\int P(\vec{\eta}) E^\dagger(\vec{\eta}) E(\vec{\eta}) d\vec{\eta} = \mathbb{I}
$$

(2.4b)

where $P(\vec{\eta})$ must be a real nonnegative distribution for both representations (2.3) and (2.4) to be equivalent. Eq. (2.3) is the standard form to describe non-unitary CP processes. However, (2.4) is better suited to describing non-unitary dynamics arising from stochastic Hamiltonians (when there is an underlying static distribution of Hamiltonians). This representation will be particularly useful in Chapter 4. Moreover, if $P$ is a real but not necessarily non-negative distribution, eqs. (2.4) describe the most general open quantum process — as recently shown in [27]. This accommodates also for non-CP maps, which typically arise from correlations of the initial state of the system with other external degrees of freedom\(^6\), or from the inversion of CP maps that yield a non-CP operation (see for example, the phase-flip map example in [28]). Examples of non-CP maps of practical importance can be found in [29]. Experimental examples of non-CP dynamics have been reported in [3, 8]. We will refer to $P$ as a distribution (not necessarily a “probability distribution”), keeping in mind that it can take negative values.

We now re-express the average fidelity as

$$
F(\Gamma) = \int d\psi \langle \psi | \Gamma(|\psi\rangle\langle \psi|) |\psi\rangle
$$

(2.5)

The original idea was that $\Gamma$ is a map which ideally is expected to be the identity map (for example, a noisy channel through which quantum information is transmitted). In that case, $F < 1$ is an indication that we are facing a faulty realization of $\Gamma$. The performance of a gate is then analyzed by composing it with its inverse, which is the idea behind the concept of an echo like in eq. (2.1) [30, 31]. When analyzing the composition of a gate with its inverse, $F$ has been sometimes called *average gate fidelity* [32]. Eq. (2.5) is well defined in cases where $\Gamma$ is not ideally the identity map, but in that case we need to redefine what value of $F$ signals a perfect implementation and address the issue of whether such value can be obtained if and only if there are no spurious processes. So till Chapter 5 we

---

\(^5\)By map we mean a process that takes states to states like $\Gamma : \mathcal{H} \to \mathcal{H}$ independently of the state to which is applied.

\(^6\)The concept of external degrees of freedom will be further explained in Chapter 3.
will restrict ourselves to characterizing only noise affecting an identity map. This includes noisy channels, gate+gate reversal studies (particularly meaningful with gates that give the identity when applied twice, like the CNOT gate - which is of paramount importance in QIP), and do-nothing gates (which are of relevance in systems where the Hamiltonian of the system cannot be turned off, so it requires a degree of control to refocus its natural evolution -like in liquid-state NMR).

2.2 A more general framework

Consider that our system consists of \( n \) qubits, thus \( D = 2^n \) is its dimension. We move from the Hilbert space \( \mathcal{H}_D \) to the Liouville space \( \mathcal{L}_{D \times D} \) where we can easily account for non-pure (mixed) states represented by density matrices \( \rho \). Moreover, an arbitrary \( D \times D \) matrix \( \rho \) in \( \mathcal{L}_{D \times D} \) (acting on \( \mathcal{H}_D \)) can be represented by a vector \( |\rho\rangle \) resulting from stacking the columns of \( \rho \) in left-to-right order on top of one another to get a single column vector of dimension \( D^2 \). The superoperator matrix of size \( D^2 \times D^2 \) arising due to a CP map \( \Gamma \) will be

\[
\hat{\Gamma} = \sum_k A_k \otimes A_k \quad \text{or} \quad \hat{\Gamma} = \int P(\vec{\eta})E(\vec{\eta}) \otimes E(\vec{\eta})d\vec{\eta}
\]  

where the overline denotes the complex conjugate. We can of course choose a different column-representation of the matrix elements of \( \rho \) and build \( \hat{\Gamma} \) accordingly, but we will follow this particular framework, which was neatly presented in [33]. We denote the columnized states as \(|\ldots\rangle \rangle\). In the case of pure states, \(|\psi\rangle\rangle\) will be the vector state corresponding to a pure vector state \(|\psi\rangle\rangle\) in the Hilbert space, which in turn corresponds to a density matrix \( \rho_\psi = |\psi\rangle\langle\psi| \) in the Liouville space. The superoperator matrix acting on these columnized states in \( \mathcal{L}_{D \times D} \) will be denoted with a hat symbol on top: \( \hat{\Gamma} \). Whenever we work with the matrix states in \( \mathcal{L}_{D \times D} \), we will use a bold type that will require an input argument, for example: \( \rho(t+1) = \Gamma_\prime(\rho(t)) \). We keep the bare symbol \( \Gamma \) to refer to the physical map in general.

The average fidelity is now

\[
F(\Gamma) = \int d\psi \langle\psi|\hat{\Gamma}|\psi\rangle
\]  

where \(|\psi\rangle\rangle\) is a state vector in the Liouville space corresponding to pure states only. It is of interest to analyze the performance of \( \hat{\Gamma} \) over any state (not only pure states),

\[
F_e(\Gamma) = \frac{1}{D^2} \sum_{l=1}^{D^2} \langle\rho_l|\hat{\Gamma}|\rho_l\rangle = \frac{1}{D^2} \sum_{l,l' = 1}^{D} \langle l'|\Gamma(l|l)\langle l'|l\rangle\rangle = \frac{\text{Tr}[\hat{\Gamma}]}{D^2}
\]  

where the last expression is a result in [34]. Notice the \( \rho_l \) must be a set of states that form an operator basis for \( \mathcal{H}_D \), as pointed out in [32]. \( F_e \) is usually referred as the entanglement fidelity [32, 35], although it is called gate fidelity in [34]. Nevertheless, there is a relation between the average fidelity \( F \) and the entanglement fidelity \( F_e \) [36, 32]

\[
F(\Gamma) = \frac{DF_e(\Gamma) + 1}{D + 1}
\]  

(2.9)
which allows us to calculate one in terms of the other.

Determining any of these two quantities would require, in principle, non-scalable resources. We could either perform quantum process tomography on $\Gamma$ to obtain all of the matrix elements of $\Gamma$ and then calculate $\text{Tr}[\hat{\Gamma}]$, or we could prepare a convenient set of $D^2$ pure initial states $\rho_i = |\psi_i\rangle\langle\psi_i|$ that form an operator basis for $\mathcal{H}_D$, apply $\Gamma$, and measure the overlap between the initial and final states. This last strategy, introduced by Nielsen in [32] and requiring $O(D^2)$ experiments, represents a great improvement over the $O(D^4)$ experiments required for QPT. Nielsen’s method is not far from QPT, but just sorts out the measurement smartly so that full QST can be avoided granted that we are only interested in $F$, $F_e$. Although the method is still non-scalable, it shows that if we are aiming to obtain partial information (in this case, the average fidelity $F$ instead of a full characterization of the map), it is possible to reduce the cost. It also shows, however, that the mere reduction of the targeted parameters (in this case drastically from $D^4 - D^2$ to 1) does not necessarily translate to an equivalent reduction in the costs, nor it is trivial to direct our strategy towards the targeted parameters.

The average fidelity $F$ or the entanglement fidelity $F_e$ (either also named sometimes gate fidelity) is the quantity that is widely used to characterize the performance of a gate in quantum computation. When preparing an experiment or doing numerical simulations, this is the typical quantity that measures the “goodness” of a potential implementation. And when an experimental realization is presented, either $F$ or $F_e$ assess the degree of success. For example, in the previous chapter we gave some examples of experimental QPT: in [3, 4, 5, 8, 9], either the average fidelity or the entanglement fidelity (considering a perfect reversal) are reported along as an indicator of success.

### 2.3 The concept of twirling and its relation with the fidelity

After the work by Nielsen, it was noticed by Emerson et al. in [13] that the integral over the Haar measure $d\hat{\psi}$ on state space in (2.7) can be replaced by the Haar measure $dU$ in the space of $D \times D$ unitary matrices (this is, in $U(D)$),

$$F(\Gamma) = \int d\hat{\psi} \langle \psi | \hat{\Gamma} | \psi \rangle = \int dU \langle \psi_0 | \hat{U}^{-1} \hat{\Gamma} \hat{U} | \psi_0 \rangle = \int dU f(U^{-1} \circ \Gamma \circ U, \psi_0)$$  \hfill (2.10)

$$= \langle \psi_0 | \left( \int dU \hat{U}^{-1} \hat{\Gamma} \hat{U} \right) | \psi_0 \rangle = f(\Gamma_T, \psi_0)$$  \hfill (2.11)

where $\psi_0$ is an arbitrary but fixed pure initial state. The action of wrapping $\Gamma$ with this Haar average, giving $\hat{\Gamma}_T = \int dU \hat{U}^{-1} \hat{\Gamma} \hat{U}$ is called a *twirl* of $\Gamma$, a terminology adopted from [37, 38]. Originally, the concept of twirling appeared in [37] to average states; we will use the term here exclusively to twirl maps.

This idea is illustrated in Fig. 2-1. We can say that $f(\Gamma_T, \psi_0)$ is a fidelity decay under the randomization of the noise.

As suggested in [13], the integral in (2.11) can be replaced by a sampling over the (infinite) set of unitary random matrices (rotations) $R_s$ distributed according to the invariant
Haar measure [39],

\[ F(\Gamma) = \int dU f(U^{-1} \circ \Gamma \circ U, \psi_0) \rightarrow \frac{1}{N} \sum_{s=1}^{N} f(R_{s}^{-1} \circ \Gamma \circ R_s, \psi_0) \quad (2.12) \]

It was then stated in [13] that each measurement of \( N \) realizations of \( f(U_{s}^{-1} \circ \Gamma \circ U_s, \psi_0) \) would approach the average \( F(\Gamma) \) as \( O(1/\sqrt{DN}) \). The strategy of approximating the integral with a sampling of the \( R_s \) was further formalized in [14] by using the Chernoff bound (see Appendix Secs. A.1 and A.2), as we will explain in a moment.

Shortly after the idea of sampling over random rotations emerged, the work by Dankert et al. [15] established that the above integral over the Haar measure on the unitary group \( U(D) \) is equivalent to a finite sum over the operations \( C_s \) belonging to the Clifford group:

\[ F(\Gamma) = \int dU f(U^{-1} \circ \Gamma \circ U, \psi_0) = \frac{1}{|C|} \sum_{s=1}^{|C|} f(C_{s}^{-1} \circ \Gamma \circ C_s, \psi_0) \quad (2.13) \]

(for more information on the Clifford operators, refer to Appendix B). Unfortunately, the size of the Clifford group \(|C| \) scales as \( O(\exp(n^2)) \), so the exact realization of the Clifford twirl is non-scalable. Again, we must resort on sampling, this time sampling over the set of Clifford gates.

We will then implement the twirl in Fig. 2.1(b) by sampling over a pool of operators \( U_s \), a pool whose size is either infinite (when using random rotations) or grows exponentially with \( n \) at least. This strategy is depicted in Fig. 2.2. The experimental approach would be to prepare the system in a convenient \( \psi_0 \) of our choice, draw a twirl operator \( U_s \) randomly from the pool, apply it to the system, apply or allow the action of the map \( \Gamma \) under study, reverse the previous operator, and measure the projection of the resulting state \( |\rho_s\rangle = U_s^{-1} \tilde{\Gamma} U_s |\psi_0\rangle \) onto \( |\psi_0\rangle \), that is \( f(U_{s}^{-1} \circ \Gamma \circ U_s, \psi_0) = \langle \rho_s |\psi_0\rangle \).

The sampling makes this approach scalable, as we explain as follows. If we are interested in measuring the probability of finding the system in the state \( \psi_0 \), this outcome will be a random binary variable \( x_s = f(U_{s}^{-1} \circ \Gamma \circ U_s, \psi_0) \) whose mean is, following eq. (2.12), \( \mathbb{E}[x_s] = F(\Gamma) \). The average of many realizations, \( \bar{x} = \sum_{s=1}^{N} x_s/N \), will tend to \( F(\Gamma) \) and will be retrieved with a standard error \( \sigma \leq 1/\sqrt{N} \) (following the Central Limit Theorem), so for a desired precision \( \epsilon \) we must require \( N \geq \epsilon^{-2} \). On the other hand, the Chernoff bound tells us that for a desired precision \( \epsilon \) and an error probability \( \delta \), we must have \( N \geq \log(2/\delta)/(2\epsilon^2) \). These two conditions do not compete with each other, since one is a bound to the variance of the distribution (in the \( N \rightarrow \infty \) limit), while the other is a bound to the error probability (which is a closed result valid for arbitrary \( N \)). Since we want to fulfill
both conditions, we must observe that the Chernoff bound imposes a stronger requirement on \( N \) when the desired error probability is \( \delta < 2e^{-2} \approx 0.27 \). Since we want \( \delta << 1 \), by fulfilling the Chernoff bound with \( \delta < 0.2 \) we are already guaranteed that the dispersion of the mean value will be \( \sigma \leq 1/\sqrt{N} \) in the large \( N \) limit.

In any case, \( N \) is independent of the size of the system, thus ensuring the scalability of the procedure (unless, of course, the targeted probabilities are expected to be of the \( O(1/\sqrt{D}) \), which would be the extreme case where \( \Gamma \) accounts for a completely random channel). The idea of sampling is key to the scalability of the twirling methods, but nevertheless these are known results from probability theory. We present the statistics of the sampling in more detail, in Appendix A.

For this scheme to be efficient, however, the implementation of each the \( U_s \) must be scalable. This is indeed the case with the Clifford operators, since any Clifford in \( U(2^n) \) can be implemented using \( O(n^2) \) one-qubit and two-qubit gates (see Appendix Sec. B.2) therefore each can be implemented efficiently.

On the other hand, the exact implementation of each random rotation operator \( R_s \) in \( U(2^n) \) scales exponentially with \( n \) [39, 40]. A possible solution to this problem is the use of pseudo-random operators [40, 41], which are operations that can be implemented efficiently while retaining features of the truly random operators. Such strategy calls for the use of \( m \) one-qubit and two-qubit operations, with \( m \) scaling no faster than \( poly(n) \).

### 2.4 The generalized fidelity decay for many steps

Following the spirit of the Loschmidt echo, we may wonder what happens if we let the system evolve for longer before measuring its overlap with the initial state. In the Loschmidt echo literature (for example [23, 30, 31, 42, 43, 44]) we find that the dynamics of the system reveals itself also in the time-dependence of the fidelity decay. Beyond the first step, which retrieves the average fidelity and the entanglement fidelity, we may find other signatures of the quantum process under study.

We observe two possible ways of generalizing the fidelity decay from one step to many steps, depicted in Fig. 2-3. In Fig. 2-3(a) we see the algorithm in a Loschmidt echo fashion: the system is allowed to evolve for \( t \) steps affected by a perturbation \( \Gamma \), and then its evolution is perfectly reversed. In Fig. 2-3(b) we just repeat the twirling step \( t \) times. But thanks to
the unitarily invariance of the twirl (that is: \( \hat{V}^{-1} \hat{\Gamma}^T \hat{V} = \hat{\Gamma}^T \) for any unitary map \( \hat{V} \)), both schemes are equivalent. This equivalence will be further studied (and proved) in Chapter 7.

(a) \( \rho_0 \left\{ \begin{array}{c} U = \Gamma \end{array} \right\}_T \)

(b) \( \rho_0 \left\{ \begin{array}{c} U = \Gamma \end{array} \right\}_T \)

Figure 2-3: Circuit representation of the generalization of the fidelity decay over time. (a) The algorithm seen as a “Loschmidt echo”. (b) Equivalent algorithm, seen as a “repeated twirl”. The equivalence of both schemes is due to the average over \( U \), when we take them as unitary operators invariant under the Haar measure.

These two schemes produce what we call a randomized fidelity decay: we are studying the map \( \Gamma \) but after randomizing it with the twirl, progressively, in many steps.

Formally, we define these randomized fidelities as

\[
\begin{align*}
    f^{LE}(\Gamma, t, \rho_0) &= \int dU_t \ldots \int dU_1 \langle \rho_0 | \hat{U}_1^{-1} \ldots \hat{U}_t^{-1} \hat{\Gamma}_t \hat{U}_t \ldots \hat{\Gamma}_1 \hat{U}_1 | \rho_0 \rangle \quad (2.14) \\
    f^{RT}(\Gamma, t, \rho_0) &= \int dU_t \ldots \int dU_1 \langle \rho_0 | \hat{U}_1^{-1} \hat{\Gamma}_1 \hat{U}_1 \ldots \hat{U}_1^{-1} \hat{\Gamma}_1 \hat{U}_1 | \rho_0 \rangle \quad (2.15) \\
    f^{RT}(\Gamma, t, \rho_0) &= f^{LE}(\Gamma, t, \rho_0) \quad (2.16)
\end{align*}
\]

where we have allowed for a potential dependence on the initial state \( \rho_0 \), and also we leave the possibility of having different maps \( \Gamma \) at different steps. \( RT \) stands for repeated twirl, while \( LE \) stands for Loschmidt Echo. We are implicitly using a measure for similarity between two mixed states \( \rho_A \) and \( \rho_B \) given by \( \langle \rho_A | \rho_B \rangle = \langle \rho_B | \rho_A \rangle = \text{Tr}[\rho_A \rho_B] \). This quantity does not satisfy in general the desired properties for a measure of fidelity between two states [25], but it does so when one of the states is pure, and it will prove to be a measurable quantity that will retrieve useful information about the quantum process under study, using an initial state \( \rho_0 \) that could be mixed in principle.

Unlike the Loschmidt echo, where there is only one variable \( t \), here we have two time scales: the time \( T \) required for one time step (i.e., the typical time required to implement \( \hat{U}_1^{-1} \hat{\Gamma} \hat{U}_1 \)); and the number of steps we repeat - which is the discrete variable we will address as “time” \( t \) (although \( t \) will be a dimensionless variable: \( t \) steps will take a time \( Tt \)).

The Loschmidt echo decay exhibits in general traces of both the evolution and the perturbation. In our case, “the evolution” will be a twirl, more specifically, a twirl where the average is over the Haar measure (or equivalently a Clifford twirl, as explained above). As we will see, different types of twirl generate different decays, but defining features of the decay are given by the map \( \Gamma \) under study.

We have already established that the twirl in \( U(D) \) retrieves the average fidelity, eq. 2.11. In Chapters 4 and 5 we will study the fidelity decay using a different form of twirl: a one-qubit twirl. That is, a twirl that is invariant over the Haar measure in \( U(2) \) for each qubit, not in \( U(D) \). Our aim is to explore what information can be extracted in this case. On the other hand, in Chapters 7 and 8 we will come back to the twirl in \( U(D) \).
But before anything else, in the next Chapter we will discuss in more detail the possible scenarios for the map $\Gamma$ under study.
Chapter 3

A survey of quantum processes

3.1 General statement

The generalized fidelity decay over time probes the quantum dynamics of the system in many ways. Describing the dynamics as in eqs. (2.4), we can say that at each time there is an operator $E$ acting on the system. The protocol is to sample the twirl with $N$ realizations, as in Fig. 2-2, repeat one step $t$ times, as in Fig. 2-3. The evolution of the system is then broken into $N \times t$ events where $E$ acts over the system, as represented in Fig. 3-1.

![Figure 3-1: Circuit representation of the action of the map $\Gamma$ under study, along a fidelity decay experiment where the twirl is sampled $N$ times and the evolution goes for $t$ steps. The action of $\Gamma$ is represented by the action of an operator $E$ and a distribution $P$, in the way described by eqs. (2.4). The two time-scales given by $T$ and $t \ T$ are also depicted.](image-url)

3.2 CP vs. non-CP maps

As mentioned in Chapter 2, when describing a quantum map in terms of $E(\eta)$ and a distribution $P(\eta)$ (eqs. 2.4), we consider the most general quantum dynamics. This includes non-CP maps when $P(\eta)$ is allowed to take negative values. However, there are some caveats. A non-CP map $\Gamma_-$ acting on a valid density matrix $\rho$ (i.e., a matrix that is hermitian and positive with $\text{Tr}[\rho] = 1$) can give rise to a final density matrix $\rho' = \Gamma_-(\rho)$ that may be nonpositive. Such states are nonphysical in the sense that they could retrieve negative probabilities of measuring the system in a particular state. In particular, for $\rho = |\psi\rangle\langle\psi|$, we
could have $f = \langle \rho | \rho' \rangle = \langle \psi | \rho' | \psi \rangle < 0$, and we can no longer claim that the fidelity decay will be indeed a decay, $f \leq 1$.

Further work in the theory of open quantum system dynamics is in order at this point: although non-CP processes are expected to occur on general grounds (see [27, 29] and references therein), it is not yet clear what are the consequences of allowing arbitrary non-positive dynamics on the system we want to observe and measure.

In this thesis we will keep the general description given by an arbitrary operator $E$ and a real distribution $P$ that could take negative values. However we restrict ourselves to processes where the reduced dynamics yield a final state whose projection onto the initial state is a positive quantity. Furthermore, in this Chapter, where we gather concepts about open quantum systems, we refer in principle to CP maps, which are the maps that constitute the lore of open quantum systems dynamics so far\footnote{See a simple discussion, for example, in [2] Sec. 8.5. A more elaborated discussion can be found in [45, 46, 47].}

Let’s go back to Fig. 3-1. Each $E$ is parameterized by coefficients $\tilde{\eta}$, that in turn may fluctuate obeying a distribution $P(\tilde{\eta})$. If the coefficients are constant, then $P$ is a delta function. In what follows we consider various scenarios according to the variation of the coefficients, determined by the nature of the system under study. In the most general scenario, for two steps of one realization we have that the resulting map $\Gamma$ is

$$\hat{\Gamma}(s_1, s_2, \tau_1, \tau_2) = \int \int P(s_1, \tau_1)(\eta_1) \hat{E}_{s_1, \tau_1}(\eta_1) P(s_2, \tau_2)(\eta_2) \hat{E}_{s_2, \tau_2}(\eta_2) d\eta_1 d\eta_2 \quad (3.1)$$

where we have denoted $\hat{E} = E \otimes E$, and for simplicity we consider a single parameter $\eta$ at each step (recall that in principle we have $D^2$ complex parameters $\tilde{\eta}$ for every operator $E$). When $\hat{E}_{s_1, \tau_1} = \hat{E}_{s_2, \tau_2} = \hat{E}_\tau$, and $P(s_1, \tau_1) = P(s_2, \tau_2) = P_\tau$ we say that the map under study is gate-independent. That is, it does not depend on the twirl operator associated to that event.

In eq. (3.1) we have allowed the latter step to be correlated to the former through $P_{s_2, \tau_2}$ depending on both $\eta_1$, $\eta_2$. When $P_{s_2, \tau_2}(\eta_2 | \eta_1) = P_{s_1, \tau_1}(\eta_2)$, we say that the successive steps are uncorrelated. Alternatively, we could have a strong correlation of the form $P_{s_2, \tau_2}(\eta_2 | \eta_1) = \delta(\eta_1 - \eta_2)$. Moreover, a correlated process may be characterized by a correlation time $\tau_c$ that indicates for how long (for how many steps) the successive steps will remain correlated. Two simple settings are given by: events that are uncorrelated from one step to the other ($\tau_c < T$), and events that remain correlated for the length of one experiment ($t$ steps) but uncorrelated from the previous one (so different realizations can be regarded as uncorrelated). If the order of magnitude of $\tau_c > T$ is known, this can be enforced by leaving a time much longer than $\tau_c$ between different realizations.

We discuss now in more detail some salient cases given this framework.

### 3.3 Unital processes

These processes take the identity operator to itself. This means that the maximally mixed state $\mathbb{I}/D$ is an eigenstate of the process: $\Gamma(\mathbb{I}/D) = \mathbb{I}/D$. A process will be unital if and only if $\int P(s, \tau)(\tilde{\eta}) E_{s, \tau}(\tilde{\eta}) E_{s, \tau}^\dagger(\tilde{\eta}) d\tilde{\eta} = \mathbb{I}$.

A typical case is when each $E_{s, \tau}$ is itself a unitary operator. In this case there is a nice
interpretation for $\Gamma$: $E$ represents the action of a Hamiltonian during a time $T$ characterized by parameters $\vec{\eta}$ that fluctuate according to $P(\vec{\eta})$. This type of processes will be considered in Chapter 4.

Notice that unital processes are not necessarily unitary. By non-unitary we mean a process that although preserve the trace of $\rho$ (which is associated with the conservation of probability), it need not preserve its purity, namely $\text{Tr}[\rho^2]$. For example, the depolarizing channel characterized by a probability $p$

$$\Gamma(\rho) = p \frac{I}{D} + (1 - p)\rho$$

is unital but not unitary.

### 3.4 Gate-dependent operations

Imagine that the operator $E$ is correlated to the particular twirl operator $U$ that is applied at that moment. We say then we have a gate-dependence, and at a given step $\tau$, each realization $s$ will carry a different map. Since we average over the many realizations $s$, we observe the averaged action of $E_{s,\tau}$: $\sum_s E_{s,\tau}\rho E_{s,\tau}^\dagger$. This resulting map may also have fluctuating parameters, in which case we will be observing the net effect of the map

$$\Gamma_\tau(\rho) = \int \sum_s P_{s,\tau}(\vec{\eta}) E_{s,\tau}(\vec{\eta})\rho E_{s,\tau}^\dagger(\vec{\eta})d\vec{\eta}$$

Typically, for operations that are purely gate-dependent, $E_{s,\tau}$ will be correlated to a $U_s$ independently of the particular time at which it was applied and independently of the previous steps, thus $E_{s,\tau} = E_s$. This type of dynamics will be considered in Chapters 7 and 8.

On the other hand, successive applications of this gate-dependent map may be correlated one from the other. We will briefly address this case in Chapter 8.

The main feature of gate-dependent maps is that the results for a perfect twirl no longer apply, since we are now twirling a map that depends on the twist.

### 3.5 Fluctuating ensembles

It may be the case that the fluctuations in $E_{s,\tau}(\vec{\eta})$ are due to an ensemble average in space or time that is independent of the particular realization. There is, for example, an ensemble average in space in liquid-state NMR QIP, where the state $\rho$ of a single qubit actually represents the state of many systems: the nuclei in the sample.

On the other hand, some physical systems like NV centers or superconducting qubits, characterized by a low contrast measurement (either by principle or due to current technical limitations), require “ensemble measurement”: the repetition of the same experiment many times in order to obtain the result of a measurement. This opposes to the concept of “single-shot” (high contrast) measurement, where this outcome is definite after one experimental implementation - this leaving aside the averages that may be necessary to record an expectation value.
This must not be confused with the repetition of a single experiment to obtain the probability of an outcome, as it is required when using strong measurements: even if we are guaranteed that a two-level system is in $|0\rangle$ or in $|1\rangle$ exclusively (with probability 1), systems with low contrast measurement require many realizations to obtain the result - while systems with strong (single-shot, high contrast) measurement (such as ion traps) require only one.

When performing these averages, $E$ may present fluctuations across the ensemble (whether in time or in space). This kind of processes will be adequately described by a distribution $P$. If $P$ does not depend on $s$ but simply we have that the fluctuation happens as many realizations are implemented, the resulting dynamics is of this type. We have thus non-unitary dynamics arising due to the existence of an ensemble average. Let’s see some examples, that illustrate various scenarios of fluctuating ensembles and will later help us discuss some general aspects of non-unitary processes.

### 3.5.1 Dephasing Gaussian noise

A simple picture of this type of process can be obtained by studying the effect of $E$ for only one qubit, which is simple enough to handle analytically, with $E(\theta) = \exp(-i(\theta/2)\sigma_z)$. In order to reflect the fluctuating nature of the process, we study the evolution of the state averaged over different realizations of $\theta$.

We take the dynamics of $E$ to be fast enough so there is no correlation between different steps, that is: the fluctuations of $\theta$ are uncorrelated after a time $O(T)$. We are assuming then that $E_{s,\tau}(\theta) = E(\theta)$ and $P_{s,\tau}(\theta) = P(\theta)$ for all $s, \tau$. As we average different realizations $s$ with a fluctuation that is independent of $s$, we get a non-unitary map from the originally unitary operators $E$. We have

$$
\rho(t) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} E(\theta_1) \cdots E(\theta_1) P(0) E(\theta_1) \cdots E(\theta_1) P(\theta_1) \cdots P(\theta_1) d\theta_1 \cdots d\theta_t
$$

(3.4)

where, for example, $P$ is a Gaussian distribution centered in $\alpha$ with a deviation $\sigma$:

$$
P(\theta) = \frac{\exp(-\frac{(\theta-\alpha)^2}{2\sigma^2})}{\sqrt{2\pi\sigma^2}}
$$

(3.5)

If we start with a general initial state $\rho(0) = (I + \sigma_x q_x + \sigma_y q_y + \sigma_z q_z)/2$, we obtain

$$
\rho(t) = \frac{I}{2} + \frac{\sigma_z}{2} q_z + \frac{\sigma_x}{2} (q_x \cos(\alpha t) - q_y \sin(\alpha t)) + \frac{\sigma_y}{2} (q_x \sin(\alpha t) + q_y \cos(\alpha t)) \exp(-\delta(t))
$$

(3.6)

where $\delta(t) = \sigma^2 t/2$. This is an exponential decrease of the transverse polarization of the qubit (together with a rotation), leaving the longitudinal polarization unchanged. This suppression of the transverse polarization is usually referred to as dephasing. This non-
unitary process has the following Kraus representation

\[
\rho(t) = M_1(t)\rho(0)M_1^\dagger(t) + M_2(t)\rho(0)M_2^\dagger(t) \tag{3.7a}
\]

\[
M_1(t) = \left(\frac{1 + e^{-\delta(t)}}{2}\right)^{1/2} \mathbb{1} e^{-i(\alpha t/2)\sigma_z} \tag{3.7b}
\]

\[
M_2(t) = \left(\frac{1 - e^{-\delta(t)}}{2}\right)^{1/2} \sigma_z e^{-i(\alpha t/2)\sigma_z} \tag{3.7c}
\]

This clearly shows that the process is composed of two parts: a rotation (unitary operation) around the \( \hat{z} \) axis (since the values of \( \theta \) are distributed centered at \( \alpha \)) and a phase-flip channel with probability \( (1 - e^{-\delta})/2 \leq 0.5 \) of phase-flipping the qubit.

### 3.5.2 Field inhomogeneity

Consider an NMR sample consisting of a spin-1/2 nucleus. We apply to this spin a magnetic field along the \( \hat{z} \) direction, but due to inhomogeneities in the field, its magnitude has a dependence along the \( z \)-axis, where the sample extends from \(-L/2\) to \(L/2\). We can approximate the variation as a simple linear dependence so the Hamiltonian of the system is \( H(z) = (\omega_0 + Gz)\sigma_z/2 \), assuming that \( B_0/(GL) >> 1 \). We perform the NMR measurement choosing not to distinguish the \( z \) variable, and collect the net magnetization from the sample as a whole. The implicit distribution for the values of \( z \) in this case is \( \{P(z) = 1/L \text{ if } |z| < L/2; P(z) = 0 \text{ otherwise}\} \). Thus the resulting map on the one-qubit ensemble is

\[
\rho(t) = \int_{-\infty}^{\infty} (E(z))^t \rho(0)(E^t(z))t P(z)dz = \int_{-L/2}^{L/2} (E(z))^t \rho(0)(E^t(z))t \frac{dz}{L}
\]

where \( E(z) = \exp(-iH(z)T) \). We find then a Kraus decomposition of the form

\[
\rho(t) = M_1(t)\rho(0)M_1^\dagger(t) + M_2(t)\rho(0)M_2^\dagger(t) \tag{3.8a}
\]

\[
M_1(t) = \left(\frac{1 + \text{sinc}(GLTt/2)}{2}\right)^{1/2} \mathbb{1} e^{-i(\omega_0 Tt/2)\sigma_z} \tag{3.8b}
\]

\[
M_2(t) = \left(\frac{1 - \text{sinc}(GLTt/2)}{2}\right)^{1/2} \sigma_z e^{-i(\omega_0 Tt/2)\sigma_z} \tag{3.8c}
\]

which is the same dynamics discussed for eqs. 3.7, with a different time-dependence.

### 3.5.3 Gaussian noise in an applied field

As in Section 3.5.1, let’s consider \( E(\theta) = \exp(-i(\theta/2)\sigma_z) \), an operator resulting from the application of a field along the \( \hat{z} \) direction with magnitude \( \omega = \theta/T \). We consider that \( \theta \) fluctuates according to a Gaussian distribution (3.5). This case is very similar to the previous one in Sec. 3.5.2, but now the underlying distribution will be Gaussian, like in 3.5.1. But unlike in Sec. 3.5.1, consecutive steps are correlated since we take \( \omega \) to be stable for the duration \( O(Tt) \) of one experiment, and fluctuate from one realization to the other. In

\^[3]\text{We have taken } \hbar = 1, \text{ as in the rest of this thesis.}
terms of $P$, the value of $\theta$ at each step will be the same and given by the same distribution, which is equivalent to a correlation of the form $P(\theta_1, \ldots, \theta_t) = P(\theta)\delta(\theta_1 - \theta) \ldots \delta(\theta_t - \theta)$, thus

$$
\rho(t) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} E(\theta_1) \ldots E(\theta_t) \rho(0) E(\theta_t)^\dagger E(\theta_1)^\dagger P(\theta_1, \ldots, \theta_t) \, d\theta_1 \ldots d\theta_t
= \int_{-\infty}^{\infty} \exp(-i\theta/2)t\sigma_z)\rho(0)\exp(+i\theta/2)t\sigma_z) \ldots P(\theta) \, d\theta
$$

This yields a Kraus form exactly as in (3.7), but with $\delta(t) = \sigma^2 t^2/2$. The decay is slower with the short-time correlated processes, and we can interpret this as that the rapid change of the short-time correlation makes it overall less harmful.

### 3.5.4 Error at the measurement, with Gaussian noise

We consider a different situation now. We will analyze the effect of a map that represents an error in the measurement combined with an ensemble over time. The repeated measurements occur during a time $T$, in which $N = T/(Tt)$ individual experiments with $t$ steps and total net time $Tt$ each have been implemented. The experiment is such that the system does not evolve (or its evolution is perfectly refocused), but an error occurs at the end, in the measurement step. So we have that $\rho(t) = \rho(0)$, but at the moment of measuring the property $A$, the actual outcome is $\text{Tr}(\{\rho(t) \{E(\theta_t)AE(\theta_t)\}],$ with $A$ either an observable (when measuring mean values) or a projector operator (when measuring probabilities). $E(\theta_t)$ represents an error in the measurement. In this toy example we take $E(\theta_t) = \exp(-i\theta/2)t\sigma_z)$. This is representative of a process arising due to an ensemble average that does not accommodate potential correlations between successive steps, because it only happens at the end of the evolution, during the measurement. However, there could be in principle a correlation between different realizations.

We take the fluctuations of $\theta_t$ to be governed by a Gaussian distribution that becomes broader as the equipment is being used (in a random walk fashion). At each realization $s$, $P(\theta_t, s) = \exp(-\frac{\theta^2}{2\sigma_0^2})/\sigma_0 \sqrt{t/2\pi}$. To compare this case with the previous ones, consider

$$
\text{Tr}[\rho(t)E(\theta_t)AE(\theta_t)] = \text{Tr}[\rho'(t)A],
$$

so now we can analyze the evolution of $\rho'(t)$ and then consider a perfect measurement. Thus we have

$$
\rho'(t) = \frac{1}{N} \sum_{s=1}^{N} \int_{-\infty}^{\infty} E(\theta_t)^\dagger \rho(0) E(\theta_t) P(\theta_t, s) \, d\theta_t
$$

which gives a dynamics with two Kraus operators

$$
M_1(t) = \left(\frac{1 + \epsilon(t)}{2}\right)^{1/2} \mathbf{1}, \quad (3.9a)
$$

$$
M_2(t) = \left(\frac{1 - \epsilon(t)}{2}\right)^{1/2} \sigma_z, \quad (3.9b)
$$

with $\epsilon(t) = \frac{1 - \exp(-\sigma_0^2 N^2 t^2/2)}{N(\exp(\sigma_0^2 N^2 t^2/2) - 1)}$. Thus a phase-flipping channel with phase-flip probability $\frac{1 - \epsilon(t)}{2}$.

We notice that if $P$ was independent of the realization (independent of $s$) we would simply have a constant phase-flip probability of $(1 - e^{-\sigma^2/2})/2$. Increasing the number of
realizations \((N \to \infty)\) will not suppress the error in the measurement. However it can suppress the effect of the correlation between different realizations.

### 3.6 Ensembles, incoherence and decoherence

A non-unitary process, which we describe in terms of an operator \(E\) and a distribution \(P\), can also be seen as the action of a unitary operator \(U\) over a bigger Hilbert space that contains the system under study and an extra space usually referred to as environment or bath [48, 27]. This extra space represents degrees of freedom that are unknown, or that we cannot (or decide not to) control and/or observe, that belong either to the system itself or to other physical systems coupled to it. We must consider then the initial state of the system and then environment, \(\rho_{SE}\), and the evolution given by the action of \(U\). The dynamics solely on the system (tracing out the environment) results in a non-unitary map \(\Gamma\) on the system, initially in \(\rho_S = \text{Tr}_E[\rho_{SE}]\).

Beyond this abstract but powerful picture, we must remember that the environment is not necessarily an external object interacting with the system, but sometimes internal degrees of freedom that have been left aside.

Examples of the latter are relaxation mechanisms in liquid-state NMR, where we have a two-level system representing an ensemble of \(O(10^{21})\) nuclei contained in a macroscopic sample. Since the sample is in a liquid state, the molecules undergo rapid molecular motions of large amplitude and random character at a much faster rate than we can observe or control, so certain intermolecular and intramolecular dynamics (such as the dipolar interaction between distant molecules) must be considered “on average”. The residual effect are then NMR relaxation processes ([49] Chap. VIII).

On the other hand, we can consider the “relaxation” (radiation) process for an excited atom. Spontaneous emission is actually a property of the atom-vacuum system, not of the isolated atom. It is the result of the presence of an environment. Moreover, if we place the atom between mirrors or in a cavity, spontaneous emission can be manipulated, even inhibited or enhanced [50]. This clearly shows the presence of an external system (the cavity), which may now also play a role as a part of the system under consideration, or again be considered as an environment.

In these examples, the extra (external, internal) degrees of freedom are responsible for well-known unitary dynamics. This is how processes that are non-unitary are generally associated with a picture of an environment interacting with the system of interest. Even if in many cases a such picture is too abstract, it is a generally widespread idea and very useful in many occasions. For example, the typical relaxation mechanisms for a qubit (any two-level system) are described in terms of “\(T_1\)” and “\(T_2\)” relaxation processes. \(T_1\) relaxation, also called dissipation, is responsible for the change of the diagonal terms of the density matrix \(\rho\) describing the state (in the basis of the eigenstates of the Hamiltonian governing the system). On the other hand, \(T_2\) relaxation, also called dephasing (cf. Sec. 3.5.1), is responsible for the suppression of the off-diagonal terms. These two ubiquitous processes are described extensively in textbooks and in the literature as either the action of fluctuating inaccessible parameters or due to the coupling with an environment; see for example [2] Secs. 8.3.5 & 8.3.6.

Non-unitary processes arising from the presence of an environment coupled to the system are usually called \textit{decoherent} processes. Decoherence was originally a concept associated with
the transition from quantum to classical behavior [51], or in general with the evolution of pure states into mixed states ([52] Sec. 3.2). It is generally associated with irreversibility, since a stand-alone superoperator $\hat{\Gamma}$ is invertible only if it is unitary ([52] Sec. 3.2.1).

However, we must acknowledge that non-unitary dynamics that remain correlated for periods of time longer that the time required to implement a map $\hat{\Gamma}$ can in principle be reversed by refocusing the action of the map, if we have the ability to control the system at that time scale. Going back to eq. (3.1), if we can control our system in order to implement an operation $E$, we can engineer $E_2$ so $E_2 = E_1^{-1}$. If the correlation between steps is such that $P_2(\eta_2|\eta_1) = \delta(\eta_1 - \eta_2)$, it is possible then to reverse the action of the first step $\hat{\Gamma}_1 = \int P_1(\eta_1)E_1(\eta_1)$. However, if the dynamics at each step are completely uncorrelated, so $P_2(\eta_2|\eta_1) = P_2(\eta_2)$, we will have two uncorrelated maps $\hat{\Gamma}_1$ and $\hat{\Gamma}_2$, and we will be unable to refocus the evolution. Notice that the distribution $P$ could be constant in the sense that $P_2$ could be the same than $P_1$, but the key is whether the variation of $\eta_2$ is correlated to the variation of $\eta_1$. Processes that exhibit correlations over successive steps allowing the experimenter to reverse non-unitary dynamics are called incoherent: as explained in [54], if the typical modulation frequency of the parameters that represent our command over the system is larger than the variation of the parameters in the quantum process under observation, time-dependent control fields often will give us the power to affect and revert undesired dynamics. This is will be mainly the case for processes exhibiting some incoherence along a given experiment of length $O(Tt)$. It will not trivially be the case if the incoherence arises across different realizations in time, as with ensemble measurement.

We notice at this point the role of the correlation time $\tau_c$. For example, it could happen that the dynamics are correlated only for a time that, although larger than the $T$ required to implement a gate, is shorter than the time required to complete an experiment. Another case is when many experiments must be performed consecutively and $\tau_c$ keeps them correlated, which may not be a desirable feature. In conclusion, the average over the distribution $P$ can be of very different nature (as different as the variety of experimental setups in experimental QIP), and the type of correlations between elements of the ensemble can retrieve very different behaviors.

As in the example of Sec. 3.5.1, the existence of an ensemble averaging unitary evolutions, either in space or in time, cannot grant the existence of what we call incoherent behavior. The existence of an ensemble average in our experiment does not imply automatically incoherence, which we rather reserve to indicate correlations over different time steps. As discussed in [54], many processes that are typically considered as decoherent can be also stated in terms of an average of unitary operators: the distinction between decoherence and incoherence remains practical and depends primarily on the correlation time and the time scale of the variation of the experimental control parameters.

Referring again to the picture of having the system coupled to an abstract environment, we have that any non-unitary process implies a flow of information that leaves the system towards the environment, which is also seen as feature of irreversibility. In this scenario, decoherent processes take place for example when the state of the environment is refreshed at each step. The relevance of resetting the state of the environment (or equivalently having an environment of infinite size) was illustrated for example in [53].

If on the contrary the environment is continuously evolving coupled to the system (a more realistic scenario), the system+environment picture can give rise to incoherent dynamics. For example, the environment could have been originally in a classical mixture of eigenstates
of the joint evolution $\mathcal{U}$ during a time $T$, like $\int p(\epsilon)|\epsilon\rangle\langle\epsilon|d\epsilon$. In that case, the incoherence will arise from the weighted average of $\langle\epsilon|\mathcal{U}|\epsilon\rangle^T$ - the net evolution operator acting over the system. In such scenario we have again the possibility of refocusing the evolution of the system, depending of course on our degree of control, our knowledge of the full $\mathcal{U}$, the existence of a drift Hamiltonian that may cause the evolution of the environment, etc.

### 3.7 Incoherence, decoherence and memory effects

As we compare the two models in Secs. 3.5.1 and 3.5.3 we see little difference, however they illustrate a very important concept. We can calculate the superoperator $\hat{\Gamma}$ in the Zeeman basis for one qubit $\{|0\rangle, |1\rangle\}$, as described in Sec. 2.2. The matrix elements of the superoperator are then given in the basis $\{|0\rangle, |1\rangle\}$. We obtain

$$\hat{\Gamma}(t) = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & e^{i\alpha} e^{-\delta(t)} & 0 & 0 \\
0 & 0 & 0 & e^{-i\alpha} e^{-\delta(t)} \\
0 & 0 & 0 & 1
\end{pmatrix}$$

In the example of Sec. 3.5.1, which is a short-time correlated process, we have that

$$\hat{\Gamma}(t)\hat{\Gamma}(t') = \hat{\Gamma}(t + t')$$

thanks to the fact that $\delta(t)$ is linear in $t$. That is not the case for the long-correlated (incoherent) process in Sec. 3.5.3, given the quadratic time dependence in $\delta(t)$: the net map over many steps is not the composition of the individual steps. We can easily see also the same behavior in the field inhomogeneity example in Sec. 3.5.2. On the other hand, in Sec. 3.5.4 the memory effect arises because the maps per realization are correlated in time but across the ensemble average. Nevertheless, in this toy example the effect of this correlation will be suppressed if $N >> 1/(\sigma^2_t)$. If, as we mentioned in the end, there were no correlation between realizations, we would have a net memory-less non-unitary process.

In general, we can make the correlations over times of $O(Tt)$ responsible for these memory effects, a signature of non-Markovianity.

It is an open topic whether correlations over different events/steps, irreversibility and non-Markovianity are necessarily linked. Also, considering now the presence of the twist in between the events of Fig. 3-1, it is arguable whether there are maps that exhibit some correlation between different events that do not have a gate-dependence. The twist operators are the result of the action of an external Hamiltonian $\mathcal{H}$ that the experimenter controls. If there is some fluctuation in the parameters $\vec{\eta}$ characterizing the evolution of the system, that probably arises from some stochastic dynamics that is present at the level of the control Hamiltonian, thus affecting the implementation of a gate $U$ in a gate-dependent way: since $U = \mathcal{T}[e^{-i\int_0^t \mathcal{H} \, dt}]$, it is very unlikely that we can factor out the fluctuating piece of $\mathcal{H}$ independently of $U$. It seems that the only case where we could have incoherent gate-independent processes is when actually the map $\hat{\Gamma}$ under study is incoherent and relatively much stronger than any spurious dynamics in the implementation of the twist operators $U_s$.

---

1. $T$ represents the Dyson time-ordering operator specifying how the integral should take into account the non-commutativity of $\mathcal{H}$ at different times.
Closing this Chapter, we emphasize that classifying the possible processes that may occur in a quantum system is a very ambitious task, and indeed there is active research in this area. But although we refrained from developing a strong categorization, in this Chapter we have listed some salient features that will assist us in describing the quantum processes that we will study in this thesis. Summarizing, these are:

- unitarity
- unitality
- existence of correlations between different steps: i.e., correlations over a time scale $O(T)$ vs. correlations over a time $O(Tt)$
- existence of correlations between different realizations
- gate-dependence of the map with the particular twirl operator associated with it
Chapter 4

The generalized fidelity decay of $n$ qubits using one-qubit twirling

Motivated by the equivalence between the Haar state-twirl and the Haar matrix-twirl (2.10), the first application of twirling by sampling to study the fidelity decay [13] used a twirl in $U(2^n)$. The natural question afterwards was: what could be gained by implementing a weaker twirl, in particular, a twirl in $U(2)^{\otimes n}$?

In this Chapter we will study the generalized fidelity decay for $n$ qubits each going under a one-qubit randomization of a quantum process $\Gamma$. In Sec. 4.2 we will describe the model for the map $\Gamma$ under study, which will be a unital noisy channel. In Sec. 4.3 we will show analytical and numerical results for specific noise models, which exhibit a weak noise regime where the fidelity decay has an exponential-like behavior. This will allow us to define in Sec. 4.4 an initial decay rate $\gamma$ which will prove to be a fruitful quantity to characterize the structure of the noise.

4.1 General statement

We will take, for the calculations in this Chapter, the scheme depicted in Fig. 2-3(a), and we will use a twirl in terms of a Haar average, as it was originally presented in [14]. We will consider map described as in eqs. (2.4), that is: by means of operator $E$ that may vary according to a distribution $P$. In this Chapter, $E$ may fluctuate over realizations but does not depend on the particular realization $s$, thus

$$
\langle \langle \rho(R,E,t)|\rho_0 \rangle \rangle = \text{Tr}[\rho(R,E,t)\rho_0]
$$

(4.1a)

$$
= \text{Tr} \left[ \left( R_t^\dagger \ldots R_1^\dagger \right) \left( E_t R_t \ldots E_1 R_1 \right) \rho_0 \left( R_t^\dagger E_t^\dagger \ldots R_1^\dagger E_1^\dagger \right) \left( R_t \ldots R_1 \right) \rho_0 \right]
$$

(4.1b)

$$
f^{LE}(\Gamma,\rho_0,t) = \int \int P(\vec{\eta}) \, \text{Tr}[\rho(R,E(\vec{\eta}),t)\rho_0] \, dR \, d\vec{\eta}
$$

(4.1c)

$$
= \left\langle \int \text{Tr}[\rho(R,E(\vec{\eta}),t)\rho_0] \, dR \right\rangle_P
$$

(4.1d)

where we have defined $\langle \ldots \rangle_P = \int P(\vec{\eta}) \ldots d\vec{\eta}$. In an experiment, the average over the distribution $P$ will occur as we sample the twirl, that is, it will be implemented as we
implement the average over the Haar measure.

We take the random operators drawn at a realization $s$ at time $\tau$ to be $R_{s,\tau} = R_{s,\tau}^{(1)} \otimes R_{s,\tau}^{(2)} \otimes \ldots R_{s,\tau}^{(n)}$, where $R_{s,\tau}^{(j)}$ is a random rotation of the $j$-th qubit. The resulting algorithm is depicted on Fig. 4-1. This is a one-qubit twirl for each of the $n$ qubits: a net twirl in $U(2)^\otimes n$.

The $R_{s,\tau}^{(j)}$ are drawn uniformly from $SU(2)$ with respect to the invariant Haar measure (a global phase is irrelevant, therefore we work in $SU(2)$ rather than in $U(2)$). Their expression in the Zeeman basis is [39]

$$R_{s,\tau}^{(j)} = \begin{pmatrix}
\cos(\phi_{j,s,\tau})e^{i\psi_{j,s,\tau}} & \sin(\phi_{j,s,\tau})e^{i\chi_{j,s,\tau}} \\
-\sin(\phi_{j,s,\tau})e^{-i\chi_{j,s,\tau}} & \cos(\phi_{j,s,\tau})e^{-i\psi_{j,s,\tau}}
\end{pmatrix}$$

(4.2)

with $\psi$ and $\chi$ drawn uniformly from the interval $[0, 2\pi)$, and $\phi = \arcsin(\sqrt{\xi})$ with $\xi$ uniformly distributed in $[0, 1]$. Notice that these random operators can be efficiently implemented, since they are single-qubit operations with an efficient gate decomposition (Solovay-Kitaev algorithm; see [55]). In this respect, we are not affected by the scalability issue that arises in the implementation of random operators in $U(D)$ with arbitrary $D$, which may lead to the use of pseudo-random operators [40, 41].

$$\rho_0 \rightarrow \begin{pmatrix}
\rho_{(1)} & & & \\
& \Gamma & & \\
& & \rho_{(n)} & \\
& & & \rho_{(n)\dagger}
\end{pmatrix} \rightarrow \rho(R, E, t)$$

Figure 4-1: Circuit representation of the algorithm, eqs. (4.1), choosing individual rotations as the twirl operators.

### 4.2 Model of the map under study: Noisy channels

In this Chapter, the process under analysis will be a noisy channel represented by a propagator $E = \exp(-iG_T)$, where $G_T$ is the noise generator. As indicated in eqs. (2.4), the map $\Gamma$ under study is

$$\Gamma(\rho) = \int P(\tilde{\eta}) \exp(-iG_T(\tilde{\eta}))\rho \exp(-iG_T(\tilde{\eta}))d\tilde{\eta}$$

(4.3)

We already see that the form of eqs. (4.1) does not allow for a gate-dependence in $E$ as the fluctuations are independent of the particular realization $s$.

We assume we have relatively good control of the system under study, so the random rotations can be implemented with sufficient accuracy (so we can treat them analytically as perfect) and then the dynamics of interest is only present in the operators $E_t$. We consider
\[ E \] to be a deviation from \( \mathbb{I} \) of the form

\[ E = \exp(-iG_\tau) = \exp \left( -i \sum_{l=1}^{D^2-1} \eta^l P_l \right) \]

(4.4)

Without losing generality, the generator is decomposed in the (generalized) Pauli operators (also called Product Operator basis [56]),

\[ P_l = \bigotimes_{j=1}^n P_l^{(j)} \]

(5.5)

where each \( P_l^{(j)} \) is an operator in the space of qubit \( j \) and is either a Pauli matrix \( \{ \sigma_x, \sigma_y, \sigma_z \} \) or the identity \( \mathbb{I} \) for the \( j \)-th qubit. Collectively, \( \{ \mathbb{I}, \sigma_x, \sigma_y, \sigma_z \} \) form the Pauli group for one qubit. We will assign the identity \( \mathbb{I} \) in \( U(D) \) to \( l = 0 \), which is out of the sum in eq. (4.4). This means that at least one factor in each \( P_l \), \( l > 0 \) is a Pauli matrix and therefore they are traceless. Notice that \( P_l^\dagger = P_l \) and \( \text{Tr}[P_l P_l^\dagger] = D \delta_{l,l'} \).

When necessary, we shall denote the coefficients \( \eta^l \) in more detail as \( \eta^{p,q,...} \), where \( j > k > \ldots \) label qubits, and \( p,q,\ldots = x,y,z \). Therefore \( \eta^{p,q,...} \) labels a term in eq. (4.4) that is a product of \( \sigma_p \) for qubit \( j \), \( \sigma_q \) for qubit \( k \), etc., and \( \mathbb{I} \) for the qubits absent in the subscript. Notice that the number of qubits in the subscript gives the Pauli weight\(^3\) of that particular \( P_l \). Therefore, the one-body terms (Pauli weight 1) go with coefficients \( \eta^x_j, \eta^y_j, \eta^z_j \), two-body terms (Pauli weight 2) are \( \eta^{p,q}_{j,k} \), etc. To avoid double counting of multi-body terms, the labeling of the qubits must obey \( j < k < \ldots \) and so on. For example, the Pauli operator \( \sigma_x^{(1)} \sigma_z^{(3)} \sigma_y^{(4)} \) has Pauli weight 3, and the corresponding coefficient is \( \eta_{x,z,y} \).

We will work with real coefficients \( \eta^l \), restraining \( E \) to be unitary and therefore restraining \( \Gamma \) to be unital (i.e., a noise channel that maps the identity \( \mathbb{I} \) to itself). We can regard \( E \) to be a residual operator resulting from the action of \( \Gamma \) during a time \( T \) through an effective Hamiltonian \( G_\tau / T \), for the time step under consideration. In turn, the coefficients in \( G_\tau \) are allowed to fluctuate according to a distribution \( P(\tilde{\eta}) \).

### 4.3 Results for specific models

The properties of \( G_\tau \) give rise to a variety of noise models. We will distinguish three major cases depending on the fluctuation of \( E \):

- **U** - Unitary: \( E \) may change from one step to the other, but does not fluctuate over realizations neither during a time step (so \( P \) is a delta function).

- **SC** - Short-time correlated: \( E \) fluctuates during a time step and/or over realizations but the variations are not correlated from one step to the other (as in the toy model in Sec. 3.5.1).

- **LC** - Long-times correlated: \( E \) fluctuates and the variations are coherently correlated along a realization \( s \) requiring a time \( Tt \): the time required to obtain \( \rho(R,E,t) \). Similar to the toy model in Sec. 3.5.3.

\(^3\)The Pauli weight is defined as the number of non-identity factors in a generalized Pauli operator: [57] p. 13; [2] p. 467.
In types LC and SC, the change of $E$ will be given by a change of the coefficients $\eta$ in $G_T$; these are randomly drawn according to the distribution $P(\tilde{\eta})$. Unitary noise gives unitary processes, while the other two give rise to non-unitary ones. We study $f^{LE}(\Gamma, \rho_0, t)$ for some specific cases, including only one-body terms (analytical and numerical results) and one-body and two-body terms (numerically). The two distributions considered were:

- a constant distribution $(\eta^p_0 = \alpha, \eta^{p,q}_{j,k} = \beta \forall j, p, k, q)$
- each $\eta^p_0$ (resp. $\eta^{p,q}_{j,k}$) randomly drawn from a Gaussian distribution with mean value $\alpha (\beta)$ and standard deviation $\sigma_\alpha (\sigma_\beta)$.

We will refer to the coefficients $\eta$ or to $\alpha, \beta, \sigma_\alpha, \sigma_\beta$ collectively as the “noise strength” $\eta$, and the powers of $\eta$ will include any monomial combination of degree equal to the given power.

As we report in more detail in the next Sections, our analytical and numerical calculations found that the decay behavior of $f^{LE}(\Gamma, \rho_0, t)$ is a shifted exponential. In particular we observe:

- Linear initial decay: $f^{LE}(\Gamma, \rho_0, t) \approx f_0 (1 - \gamma t)$ for $t$ sufficiently small and with $f_0 = \text{Tr}[\rho_0^2]$ (we shall call $\gamma$ the initial decay rate);
- Constant long-time limit: $f^{LE}(\Gamma, \rho_0, t) \to 1/D$ for $t \to \infty$.

The scales “$t$ sufficiently small” or $t \to \infty$ are set by the value of the parameters $\eta$. Our numerical calculations ranged up to a noise strength of 0.4. For higher strength, the saturation value $1/D$ is reached in only a few steps and not much can be extracted from this fidelity decay. An example to illustrate these calculations is given in Fig. 4-2.

We now present our analytical and numerical results in more detail. The analytical results were derived with mathematical tools developed in [94, 95], a good presentation thereof can be found in [96].

### 4.3.1 Analytical results for $G_T$ with one-body terms only

Analytical expressions of the generalized fidelity decay can be obtained when $G_T$ has one-body terms only and $\rho_0$ is a separable initial state:

$$E_t = \bigotimes_{j=1}^n E_{t}^{(j)}, \quad E_{t}^{(j)} = \exp \left( -i \eta^{(j)}_{n,j,t} \right)$$  \hspace{1cm} (4.6)

(the actual directions $\hat{n}_{j,t}$ are irrelevant) and we have

$$\rho_0 = \rho_0^{(1)} \otimes \rho_0^{(2)} \otimes \ldots \otimes \rho_0^{(n)}$$  \hspace{1cm} (4.7)

Under these conditions the fidelity of the whole system is the product of the fidelities for each qubit:

$$f^{LE}(\Gamma, \rho_0, t) = \prod_{j=1}^n \left( f^{(j)}(E, \rho_0, t) \right)_P$$  \hspace{1cm} (4.8)
Figure 4-2: (semi-log) Examples of the numerical calculation of $f^{LE}(\Gamma, \rho_0, t)$ with $N = 100$ realizations for 8 qubits, initially all in the $|0\rangle$ state. Here we took $G_T$ with one-body terms only ($\Box$, $\blacktriangle$, $\bigcirc$) and with one-body and all the two-body terms ($\blacklozenge$, $\blacklozenge$, $\blacklozenge$).

$\Box$: type LC, $P$ Gaussian with $\alpha = 0$, $\sigma_\alpha = 0.05$.
$\blacktriangle$: type U, $P$ constant with $\alpha = 0.05$.
$\bigcirc$: type SC, $P$ Gaussian with $\alpha = 0$, $\sigma_\alpha = 0.05$.
$\blacklozenge$: type LC, $P$ Gaussian with $\alpha = \beta = 0$, $\sigma_\alpha = \sigma_\beta = 0.05$.
$\blacklozenge$: type U, $P$ constant with $\alpha = \beta = 0.05$.
$\bigodot$: type SC, $P$ Gaussian with $\alpha = \beta = 0$, $\sigma_\alpha = \sigma_\beta = 0.05$.

We observe that

$$f^{(j)}(R_t^{(j)}, \ldots, R_1^{(j)}, E, \rho_0) = \text{Tr} \left[ \left( R_t^{(j)\dagger} \ldots R_1^{(j)\dagger} \right) \left( E_t^{(j)} R_t^{(j)} \ldots E_1^{(j)} R_1^{(j)} \right) \rho_0^{(j)} \right]$$

$$f^{(j)}(E, \rho_0, t) = \int \ldots \int f^{(j)}(R_t^{(j)}, \ldots, R_1^{(j)}, E, \rho_0) \, dR_t^{(j)} \ldots dR_1^{(j)}$$

where each integral is an average over the normalized Haar measure on $U(2)$. Eq. (4.10) is a polynomial function of $R_t^{(j)}$ and $R_t^{(j)\dagger}$. General results for computing this kind of averages in $U(D)$ are presented in [94, 95, 96]. Here we state the following particular results

$$\int \text{Tr}[ARB R^{\dagger}] \, dR = \frac{1}{2} \text{Tr}[A]\text{Tr}[B]$$

$$\int \text{Tr}[\rho R^{\dagger} A R \rho R^{\dagger} B R] \, dR = \frac{1}{3} \text{Tr}[AB] \left( 1 - \frac{\text{Tr}[\rho^2]}{2} \right) + \frac{1}{3} \text{Tr}[A]\text{Tr}[B] \left( \text{Tr}[\rho^2] - \frac{1}{2} \right)$$

where all the operators belong to $\mathcal{H}_2$ and we have used $\text{Tr}[\rho] = 1$. Applying these formulas,

$$f^{(j)}(E, \rho_0, t) = f^{(j)}(E, \rho_0, t - 1) \frac{\left| \text{Tr}_j[E_t^{(j)}] \right|^2 - 1}{3} + \frac{2}{3} - \frac{\left| \text{Tr}_j[E_t^{(j)}] \right|^2}{6}$$

41
This shows that $f^{LE}(\Gamma, \rho_0, t)$ only depends on the fidelity at a previous time $t - 1$, thus exhibiting at this point an intrinsic exponential-type decay. At the same time, this also shows that the precise decay law is not a simple exponential. Even for constant Unitary noise, where $\text{Tr}_j[E^{(j)}_t]$ is the same at all times, eq. (4.14) gives a shifted exponential for $f^{(j)}(E, \rho_0, t)$. And then eq. (4.8) indicates that the decay is a product of those shifted exponentials.

From eq. (4.14) it is possible to compute a closed expression for $f^{LE}(\Gamma, \rho_0, t)$ in several cases, accounting also for the time variation of the coefficients in $E_t$ given by $P$. For unitary processes with a constant distribution $P(\eta_1, \ldots, \eta_t) = \prod_{j=1}^n \delta(\eta_{j,1} - \alpha) \ldots \delta(\eta_{j,t} - \alpha)$ we have

$$\langle f^{(j)}(E, \rho_0, t) \rangle_P = \frac{1}{2} + \left( f_0^{(j)} - \frac{1}{2} \right) \exp(-\lambda t)$$

(4.15)

$$f^{LE}(\Gamma, \rho_0, t) = \frac{1}{2^n} \left( 1 + \left( 2f_0^{(j)} - 1 \right) e^{-\lambda t} \right)^n$$

(4.16)

where

$$\lambda = -\ln \left( \frac{4\cos^2(\alpha) - 1}{3} \right) \approx \frac{4\alpha^2}{3} + O(\alpha^4)$$

(4.17)

and $f_0^{(j)} = \text{Tr}_j[\rho_0^{(j)}]$. To be precise, (4.17) is valid for $\alpha < \pi/6 \approx 0.52$ - the range in which $\lambda$ is real. Otherwise $\langle f^{(j)}(E, \rho_0, t) \rangle_P$ oscillates (a “step-by-step” decaying oscillation, that is, with a global $(-1)^t$ factor). For unitary processes with a non-uniform distribution like $P(\eta_t) = \delta(\eta_t - \alpha_j)$ the result is very similar: for each qubit, the $\lambda$ in (4.15) must be replaced by the respective $\lambda_j$, and eq. (4.16) will composed of $n$ factors with different $\lambda$’s.

We also have closed expressions for the scenarios with correlated fluctuations, including the additional averages over a Gaussian distribution $P_G(y, \alpha, \sigma_\alpha) = \exp(-\frac{(y-\alpha)^2}{2\sigma_\alpha^2})/\sigma_\alpha \sqrt{2\pi}$ of coefficients and already assuming $\alpha, \sigma_\alpha << 1$.

For long-time correlated processes,

$$P(\eta_1, \ldots, \eta_t) = \prod_{j=1}^n P_G(\eta_{j,1}, \alpha_j, \sigma_{\alpha,j}) \delta(\eta_{j,2} - \eta_{j,1}) \ldots \delta(\eta_{j,t} - \eta_{j,1})$$

(4.18)

$$\langle f^{(j)}(E, \rho_0, t) \rangle_P \approx \frac{1}{2} + \left( f_0^{(j)} - \frac{1}{2} \right) \frac{\exp(-a_j(t))}{\sqrt{1 + \frac{\pi}{3} \sigma_{\alpha,j}^2 t}}$$

(4.19)

$$a_j(t) = \frac{\frac{4}{3} \sigma_{\alpha,j}^2 t}{1 + \frac{\pi}{3} \sigma_{\alpha,j}^2 t}$$

(4.20)

For noise with short correlation time:

$$P(\eta_1, \ldots, \eta_t) = \prod_{j=1}^n P_G(\eta_{j,1}, \alpha_j, \sigma_{\alpha,j}) \ldots P_G(\eta_{j,t}, \alpha_j, \sigma_{\alpha,j})$$

(4.21)

$$\langle f^{(j)}(E, \rho_0, t) \rangle_P \approx \frac{1}{2} + \left( f_0^{(j)} - \frac{1}{2} \right) \exp(-\xi_j t)$$

(4.22)

$$\xi_j = -\ln \left( \frac{1 + 2\cos(2\alpha_j) \exp(-2\sigma_{\alpha,j}^2)}{3} \right)$$

(4.23)
Fig. 4-3 shows some examples of numerical calculations together with the theoretical result, exhibiting perfect agreement.

\[
\begin{align*}
\text{Figure 4-3: (semi-log) Examples of the numerical calculation of } & f^{LE}(\Gamma, \rho_0, t) \text{ with only one-body terms in } G_T. \\
& N = 100 \text{ realizations, 8 qubits initially all in the } |0\rangle \text{ state. The fitting shown is given by the corresponding theoretical result (see text). With } \\
& \alpha_j = \alpha \text{ and } \sigma_{\alpha,j} = \sigma_\alpha \text{ for } j = 1, \ldots, 8: \\
\square & \text{ type LC, } P \text{ Gaussian with } \alpha = 0, \sigma_\alpha = 0.08. \\
\bullet & \text{ type SC, } P \text{ Gaussian with } \alpha = 0, \sigma_\alpha = 0.08. \\
\bigcirc & \text{ type U, } P \text{ constant with } \alpha = 0.08. \\
\blacktriangle & \text{ type SC, } P \text{ Gaussian with } \alpha = 0.08, \sigma_\alpha = 0.04. \\
\ast & \text{ type LC, } P \text{ Gaussian with } \alpha = 0.08, \sigma_\alpha = 0.08.
\end{align*}
\]

The exponential-like behavior, at least for weak noise strength, can be shown in more general grounds as follows. We take eq. (4.14) and re-state it as

\[
f^{(j)}(E, \rho_0, t) = f^{(j)}(E, \rho_0, t-1)p_{j,t} + q_{j,t} \tag{4.24}
\]

with \( p_{j,t} = (4\cos^2(\eta_{j,t}) - 1)/3 \) and \( q_{j,t} = 2\sin^2(\eta_{j,t})/3 \). We analyze the two correlated noise models we proposed. For SC noise, each step gets averaged over the corresponding distribution \( P \), so

\[
\left\langle f^{(j)}(E, \rho_0, t) \right\rangle_P = \frac{1}{2} \left( 1 + (2f_0^{(j)} - 1) (\langle p_j \rangle_P)^t \right) \text{ for SC noise} \tag{4.25}
\]

which gives a shifted exponential for each qubit, and then a product of them for the whole fidelity decay - independently of the details of the distribution. On the other hand, for the LC noise, we average over the many steps, so we consider \( p_{j,t} = p_j \) and \( q_{j,t} = q_j \) and multiply all the steps, so for the whole evolution we have

\[
f^{(j)}(E, \rho_0, t) = \frac{1}{2} \left( 1 + (2f_0^{(j)} - 1)p_j^t \right) \tag{4.26}
\]
Then we take the average as
\[
\left\langle f^{(j)}(E, \rho_0, t) \right\rangle_p = \frac{1}{2} \left( 1 + (2f_0^{(j)} - 1) \left\langle p_j^0 \right\rangle_p \right) \quad \text{for LC noise} \quad (4.27)
\]

In general the behavior of \( \left\langle p_j^0 \right\rangle_p \) cannot be anticipated. In particular, for negative \( p_j \) (which happens for strong noise, with \( \eta \) above some threshold) we will have a step-by-step oscillation. But if the noise is weak enough, we can approximate \( (p(\eta))^t = (4\cos^2(\eta) - 1)^t / 3^t \approx 1 - 4\eta^2 t / 3 \approx \exp(-4\eta^2 t / 3) \). In this case \( \left\langle p_j^0 \right\rangle_p \approx 1 - 4 \left\langle p_j^0 \right\rangle_p / 3 \approx \exp(-4 \left\langle p_j^0 \right\rangle_p t / 3) \) and we get an exponential-like behavior.

If we also allow some correlation between the different qubits, so the product (4.8) must be taken before averaging over \( P \), we will get also an exponential-like behavior, since in general we will be dealing with
\[
\left\langle \prod_{j=1}^{n} f^{(j)}(E, \rho_0, t) \right\rangle_p = \frac{1}{2^n} \left( 1 + (2f_0^{(j)} - 1) \left\langle p_j^0 \right\rangle_p \right)^n = \sum_{k=0}^{n} \left( \begin{array}{c} n \\ k \end{array} \right) \left\langle p_j^k \right\rangle_p (2f_0^{(j)} - 1)^k
\]

and again for weak noise \( \left\langle p_j^k \right\rangle_p \) will give an exponential-like behavior in \( tk \). Thus the whole fidelity will be a sum of exponentials.

### 4.3.2 Numerical results

When multi-body terms are present in \( G_T \), the non-separability of the \( E \) prevents us from getting closed results for the generalized fidelity decay. We studied numerically the case with only one-body and two-body terms in \( G_T \), obtaining a linear initial decay and a saturation value for long times, as mentioned in Sec. 4.3. Figs. 4-4 and 4-5 show some examples of the numerical calculation of \( f^{LE}(\Gamma, \rho_0, t) \) and a curve-fitting following the formula
\[
f^{LE}(\Gamma, \rho_0, t) = e^{-\zeta t} \left( f_0 - \frac{1}{D} \right) + \frac{1}{D} \quad (4.29)
\]

where \( \zeta \) is the only fitting parameter. As before, \( f_0 = \text{Tr}[\rho_0^2] \). We do not expect an exact agreement with this formula; we chose a shifted exponential as it is the simplest exponential-type decay that interpolates the initial and long time behavior. Also this is the exact expression for the fidelity decay when the random operators are rotations in \( U(D) \) (cf. [13]).

In practice, the fitting is simply a linear fitting of \( \log(f^{LE}(\Gamma, \rho_0, t) - 1/N) \). Of course we must be careful with the values of \( f^{LE}(\Gamma, \rho_0, t) \) close to 0 (since in this range numerically we will have null and negative values of the fidelity); thus we just use the points with \( f^{LE} \) higher than a certain cut-off value \( f_{co} \). Notice that for fast decays, \( f_{co} \) should be low enough to include a sufficiently large number of points to fit.

For Fig. 4-4 we took the system to be in an initial pure state, and we examine different \( P \) distributions assuming that only one-body and first neighbor couplings are present in \( G_T \).

For Fig. 4-5 we took the system to have initially some qubits in the maximally mixed state \( I/2 \). We chose a particular \( P \) distribution (long-time correlated noise with coefficients drawn from a Gaussian distribution centered around 0) and we worked with different noise strengths, including a different set of non-null terms in \( G_T \).
Figure 4-4: (semi-log) Examples of the numerical calculation of \( f^{LE}(\Gamma, \rho_0, t) \) with \( N = 100 \) realizations for 8 qubits, initially all in the \( |0\rangle \) state. Here took \( G_T \) with one-body terms and only first-neighbor two-body terms. The fitting corresponds to formula (4.29) with \( f_{co} = 0.1 \) (see text).

\[
\begin{align*}
\uparrow & \text{: type U, } P \text{ constant, } \alpha = 0.05, \beta = 0.02. \\
\square & \text{: type U, } P \text{ constant, } \alpha = 0.02, \beta = 0.05. \\
\circ & \text{: type U, } P \text{ constant, } \alpha = \beta = 0.05. \\
\ast & \text{: type SC, } P \text{ Gaussian with } \alpha = \beta = 0.05, \sigma_\alpha = \sigma_\beta = 0.005. \\
\ast & \text{: type SC, } P \text{ Gaussian with } \alpha = \beta = 0, \sigma_\alpha = \sigma_\beta = 0.05. \\
\diamond & \text{: type LC, } P \text{ Gaussian with } \alpha = \beta = 0.05, \sigma_\alpha = \sigma_\beta = 0.005.
\end{align*}
\]

Figure 4-5: (semi-log) Examples of the numerical calculation of \( f^{LE}(\Gamma, \rho_0, t) \) with \( N = 100 \) realizations for 8 qubits; initially, qubits 1 and 2 are in the \( I/2 \) state while the rest is in the \( |0\rangle \) state. Here we worked with long-time correlated noise (LC), with a Gaussian distribution centered around 0. The fitting corresponds to formula (4.29) with \( f_{co} = 0.05 \) (see text); the legend shows the obtained values of \( \zeta \) with their relative error.

\[
\begin{align*}
\ast & \text{: one-body and first-neighbor terms in } G_T, \sigma_\alpha = \sigma_\beta = 0.03. \\
\ast & \text{: one-body and first-neighbor terms in } G_T, \sigma_\alpha = \sigma_\beta = 0.05. \\
\circ & \text{: with all the terms in } G_T, \sigma_\alpha = \sigma_\beta = 0.03. \\
\uparrow & \text{: with all the terms in } G_T, \sigma_\alpha = \sigma_\beta = 0.05.
\end{align*}
\]
Fig. 4-6 shows the decay rate $\zeta$ of formula (4.29) as a function of the strength $\eta$, for different noise scenarios. It’s not surprising to find that the decay is faster for higher $\eta$. Notice also the different proportionality between $\zeta$ and $\eta$ for the different cases. On the other hand, this exponential-like decay suggests we can estimate the decay rate $\gamma$ from the initial linear decay. As anticipated in Sec. 4.3, in our data we observe $f^{LE}(\Gamma, p_0, t) \approx f_0(1 - \gamma t)$ for $t$ sufficiently small. Taking the same source for the data $\zeta$ vs. $\eta$ in Fig. 4-6 (that we used with $f_{co} = 0.1$), we obtained $\gamma$ vs. $\eta$, but fitting with a straight line the points with $f > f_{lim} = 0.9$. The resulting fit is shown in Fig. 4-7. Naturally $\zeta$ (Fig. 4-6) and $\gamma$ (Fig. 4-7) don’t agree as the generalized fidelity decay is not a simple shifted exponential. However the observation of the exponential-like behavior and the study of $\zeta$ encouraged further work on the concept of a decay rate. As we will see in the next Section, $\gamma$ is a very useful quantity for noise characterization.

![Graph showing decay rate $\zeta$ vs. generator strength $\eta$.](image)

Figure 4-6: Values of $\zeta$ fitting the numerical data according to formula (4.29). We worked with 8 qubits initially all in the $|0\rangle$ state.

- **C**: type U, $P$ constant with $\alpha = \beta = \eta$; all the terms in $G_\tau$.
- **E**: type LC, $P$ Gaussian with $\alpha = \beta = 0$, $\sigma_\alpha = \sigma_\beta = \eta$; all the terms in $G_T$.
- **$*$**: idem C, but only one-body and first-neighbor terms.
- **D**: idem E, but only one-body and first-neighbor terms.
- **$\Diamond$**: idem C, but only one-body terms.
- **$\Delta$**: idem E, but only one-body terms.

### 4.4 The initial decay rate $\gamma$

Numerical and analytical evidence supports the conjecture that at least up to a certain noise strength, the initial decay is linear in $t$, with an initial decay rate $\gamma$. Moreover, we conjecture that this initial decay rate depends only on the noise strength (the magnitude of the $\eta$) and not on the particular time variation of $E$. This can be seen in Fig. 4-2, where the initial decay is the same for the different types U, LC and SC as long as the general noise strength is the same. We can put this argument more rigorously for noise with only one-qubit terms,
as discussed in the analysis at the end of Sec. 4.3.1. If the noise is weak enough, we can approximate $\left\langle p_j^t\right\rangle_p \approx \left\langle \left(\langle p_j\rangle_p\right)^t\right\rangle_p$, if $\eta << 1$ so $p(\eta)^t \approx 1 - 4\eta^2t/3 \approx \exp(-4\eta^2t/3)$ holds.

Consistent with this, our analytical and numerical calculations indicate that the dependence of $\gamma$ on the noise strength is quadratic. The first clear evidence comes from the initial slope of the analytical expressions we obtained in Sec. 4.3.1; although these results are limited to noise with only one-body terms. For noise including two-body terms, we fitted the initial decays (as shown in Fig. 4-7), obtaining a quadratic dependence on the governing parameter.

We say then that the initial decay rate is a measure of the noise strength independent of the type of noise. We can actually obtain an analytical expression for the initial decay rate $\gamma$ by calculating the fidelity decay after the first step $\Delta f$,

$$\Delta f = \gamma f_0 = f_0 - f^{LE}(\Gamma, \rho_0, t = 1)$$

(4.30)

up to second order in $\eta$, for a separable initial state. We can also prove that the third order correction in $\eta$ is zero. This expansion in powers of $\eta$ assumes that the map under study is weak enough, that is, close to the identity $I$. Notice that for this calculation we go back to a general noise model with multi-body terms in $G_T$ as in eq. (4.4) with $\eta_l \neq 0 \forall l$ in

Figure 4-7: Values of the initial decay rate $\gamma$ obtained with a linear fitting of $f^{LE}(\Gamma, \rho_0, t)$ in the small $t$ regime. In practice, this linear regime was given by all the points $t$ with $f^{LE}(\Gamma, \rho_0, t) > f_{lim}$. Here $f_{lim} = 0.9$ and $n = 8$. Solid lines: the quadratic fitting.

○: type U, $P$ constant with $\alpha = \beta = \eta$; all the terms in $G_T$.
●: type LC, $P$ Gaussian with $\alpha = \beta = 0$, $\sigma_\alpha = \sigma_\beta = \eta$; all the terms in $G_T$.
*: idem ○, but only one-body and first-neighbor terms.
□: idem ●, but only one-body and first-neighbor terms.
◊: idem ○, but only one-body terms.
▲: idem ●, but only one-body terms.
principle. For an initial state where each qubit is in a pure state \((f_0 = 1)\),

\[
\gamma = c_1 \sum_{j=1}^{n} (\eta_j^{col})^2 + c_2 \sum_{k>j=1}^{n} (\eta_{j,k}^{col})^2 + c_3 \sum_{g>k>j=1}^{n} (\chi_{j,k,g}^{col})^2 + \ldots + O(\eta^4)
\] (4.31)

with \(c_w = 1 - 1/3^w\). In (4.31) we have defined the collective coefficients

\[
(\eta_j^{col})^2 = \sum_{p=x,y,z} (\eta_j^p)^2; \quad (\eta_{j,k}^{col})^2 = \sum_{p,q=x,y,z} (\eta_{j,k}^{p,q})^2; \quad \text{etc.}
\] (4.32)

and \(w = 1, 2, 3, \ldots, n\) for a collective coefficient corresponding to terms with Pauli weight \(w\). This calculation, as well as the other ones that follow in this Section, can be reproduced in detail following Appendix C (notice Sec. C.6).

When the coefficients fluctuate over time the relevant quantity is the average

\[
\langle \gamma \rangle_p = \int \gamma(\vec{\eta}) P(\vec{\eta}) d\vec{\eta}
\] (4.33)

since when we average the realizations of the random rotations, we also average over realizations of the fluctuating coefficients. For example for the Gaussian distribution we described before we would have \(\int (\eta_j^{col})^2 P(\vec{\eta}) d\vec{\eta} = 3(\alpha^2 + \sigma_0^2)\), etc. Therefore for any distribution \(P\) the same equations hold, with the collective coefficients properly replaced by the strength parameters characterizing \(P\) (again, we refer the reader to Appendix C for the detailed calculations).

We can compare \(\gamma\) in eq. (4.31) (as obtained from the fidelity decay after only one step), to the one we would obtain by linearly fitting the observed decay in \(t\) for the points with \(f < f_{lim}\). Take for instance the examples already introduced in Figs. 4-6 and 4-7: Table 4.1 shows some examples of the numerical fitting for \(\gamma = c\eta^2 + O(\eta^4)\). We have a good agreement for \(f_{lim} = 0.9\). This shows the importance of establishing the regime where the noise is weak enough so that we have a sufficient number of steps for which the decay is well described as linear.

Eq. (4.31) shows a decay rate \(\gamma\) that is a weighted sum of the collective coefficients of \(G_T\). However, we are more interested in obtaining a characterization that individualizes

<table>
<thead>
<tr>
<th>Type</th>
<th>(f_{lim} = 0.8)</th>
<th>(f_{lim} = 0.9)</th>
<th>(f_{lim} = 0.95)</th>
<th>Theory: eq. (4.31)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U ◯</td>
<td>25 (8%)</td>
<td>31 (30%)</td>
<td>-</td>
<td>30.22</td>
</tr>
<tr>
<td>LC ●</td>
<td>27 (18%)</td>
<td>31 (19%)</td>
<td>-</td>
<td>30.22</td>
</tr>
<tr>
<td>U *</td>
<td>10.7 (1%)</td>
<td>11.7 (1%)</td>
<td>11.3 (8%)</td>
<td>11.55</td>
</tr>
<tr>
<td>LC □</td>
<td>12.1 (2%)</td>
<td>10.3 (4%)</td>
<td>13 (15%)</td>
<td>11.55</td>
</tr>
<tr>
<td>U ◦</td>
<td>5.69 (2%)</td>
<td>5.45 (0.4%)</td>
<td>5.04 (1%)</td>
<td>5.33</td>
</tr>
<tr>
<td>LC ▲</td>
<td>5.0 (2%)</td>
<td>5.62 (1%)</td>
<td>4.85 (1%)</td>
<td>5.33</td>
</tr>
</tbody>
</table>

Table 4.1: Values for the proportionality constant \(c\) between \(\gamma\) and the strength \(\eta^2\) (see text), for the different types of noise studied in Figs. 4-6 and 4-7. We worked with 8 qubits initially in a pure state. The error of the fitting is shown in parenthesis. The missing values are due to the lack of enough points (leading to errors higher than 100%), taking the points with \(f > f_{lim}\) at increments of 0.01 in \(\eta\).
these coefficients. If we engineer initial states other than $|0\rangle^\otimes n$, the weights of the collective coefficients change; moreover, some vanish if some qubits are initially in the maximally mixed state, $I/2$. For example, consider only the presence of one-body and two-body terms. If

$$
\rho_0 = |0\rangle\langle 0|_a \otimes I/2 \otimes \ldots \otimes I/2,
$$

$$
\Delta f = \frac{4}{3D} \left( \eta_a^{(\text{col})} + \sum_{k \neq a} (\eta_{a,k}^{(\text{col})})^2 \right) + O(\eta^4)
$$

where only the coefficients related to the qubit labeled $a$ are present. On the other hand, if

$$
\rho_0 = (I/2)_a \otimes |0\rangle\langle 0| \otimes \ldots \otimes |0\rangle\langle 0|,
$$

$$
\Delta f = \frac{1}{3} \left( \sum_{j \neq a} (\eta_j^{(\text{col})})^2 + \sum_{k \neq a} (\eta_{a,k}^{(\text{col})})^2 + \frac{4}{3} \sum_{k > j \neq a} (\eta_{j,k}^{(\text{col})})^2 \right) + O(\eta^4)
$$

where only the coefficients related to qubit $a$ are missing. Also, we notice that if we involve more than one qubit, the terms with different Pauli weight appear with a different coefficient, just as in eq. (4.31), but also depending on the relative different initial state of each qubit.

For example,

$$
\rho_0 = |0\rangle\langle 0|_a \otimes |0\rangle\langle 0|_b \otimes I/2 \otimes \ldots \otimes I/2,
$$

$$
\Delta f = \frac{8}{3D} \left( \eta_a^{(\text{col})} + \eta_b^{(\text{col})} + \sum_{k \neq a,b} (\eta_{a,k}^{(\text{col})})^2 + (\eta_{b,k}^{(\text{col})})^2 + \frac{2}{3} (\eta_{a,b}^{(\text{col})})^2 \right) + O(\eta^4)
$$

To make use of this feature efficiently, we can calculate the fidelity of the state of just a few qubits. Let’s call $M$ the set of $m$ qubits that is going to be measured ($m \leq n$), and $\overline{M}$ its complementary (Fig. 4-8). We then re-write eqs. (4.1) considering a measurement on the reduced system in $\mathcal{H}_M$ only. Denoting the reduced density matrices by $\rho^{(X)} = \text{Tr}_{\overline{X}}[\rho]$,

$$
\langle \langle \rho^{(M)}(R, E, t)|\rho_0^{(M)} \rangle \rangle = \text{Tr}_M[\rho^{(M)}(R, E, t)|\rho_0^{(M)}] = \text{Tr}_M[\text{Tr}_{\overline{M}}[\rho^{(M)}(R, E, t)|\rho_0^{(M)}] = \text{Tr}_M[\text{Tr}_{\overline{M}}[R_1 \ldots R_i E_i R_t \ldots E_1 R_1 \rho_0 R_1^\dagger E_i^\dagger R_t^\dagger E_1^\dagger R_1 \rho_0 R_1^\dagger E_i^\dagger R_t^\dagger E_1^\dagger R_1 \rho_0 R_1^\dagger E_i^\dagger R_t^\dagger E_1^\dagger R_1 \rho_0 R_1^\dagger E_i^\dagger R_t^\dagger E_1^\dagger R_1 | \text{Tr}_{\overline{M}}[\rho_0]]]
$$

$$
f^{(M)}(\Gamma, \rho_0, t) = \int \int P(\tilde{\eta}) \text{Tr}_M[\rho^{(M)}(R, E(\tilde{\eta}), t)|\rho_0^{(M)}] dR d\tilde{\eta} = \left\langle \text{Tr}_M[\rho^{(M)}(R, E, t)|\rho_0^{(M)}] dR \right\rangle_p
$$

Correspondingly, we denote as $\gamma^{(M)}$ the initial decay rate of $f^{(M)}(\Gamma, \rho_0, t)$.

It can be shown (see Appendix C) that $\gamma^{(M)}$ is independent of the initial state of the qubits not being measured. This is indeed a desirable feature since then we do not have to worry about experimentally initializing them - as long as the separability of the initial state of the $m$ qubits in $\rho_0^{(M)}$ is guaranteed. We show below the results for measuring arbitrary one, two and three qubits, which we have labeled $a$, $b$ and $c$; these qubits are each initially in an arbitrary pure state. More general formulas are given in Appendix C; the following
Figure 4-8: Circuit representation of the algorithm focussing on a subset of $m$ qubits.

will suffice to show the usefulness of $\gamma^{(M)}$ in characterizing the map $\Gamma$.

$$\gamma^{(a)} = \frac{2}{3} \left( (\eta_a^{\text{col}})^2 + \sum_{j \neq a} (\eta_{a,j}^{\text{col}})^2 + \sum_{k > j \atop j,k \neq a} (\eta_{a,j,k}^{\text{col}})^2 + \ldots \right) + O(\eta^4) \quad (4.34)$$

$$\gamma^{(a,b)} = \frac{2}{3} \left( (\eta_a^{\text{col}})^2 + (\eta_b^{\text{col}})^2 + \sum_{j \neq a,b} (\eta_{a,j}^{\text{col}})^2 + \sum_{k > j \atop j,k \neq a,b} (\eta_{a,j,k}^{\text{col}})^2 + \ldots \right) + \frac{8}{9} \left( (\eta_{a,b}^{\text{col}})^2 + \sum_{j \neq a,b} (\eta_{a,b,j}^{\text{col}})^2 + \ldots \right) + O(\eta^4) \quad (4.35)$$

$$\gamma^{(a,b,c)} = \frac{2}{3} \left( (\eta_a^{\text{col}})^2 + (\eta_b^{\text{col}})^2 + (\eta_c^{\text{col}})^2 + \sum_{j \neq a,b,c} (\eta_{a,j}^{\text{col}})^2 + (\eta_{b,j}^{\text{col}})^2 + (\eta_{c,j}^{\text{col}})^2 + \sum_{k > j \atop j,k \neq a,b,c} (\eta_{a,j,k}^{\text{col}})^2 + (\eta_{b,j,k}^{\text{col}})^2 + (\eta_{c,j,k}^{\text{col}})^2 + \ldots \right) + \frac{8}{9} \left( (\eta_{a,b}^{\text{col}})^2 + (\eta_{a,c}^{\text{col}})^2 + \sum_{j \neq a,b,c} (\eta_{a,b,j}^{\text{col}})^2 + (\eta_{a,c,j}^{\text{col}})^2 + (\eta_{b,c,j}^{\text{col}})^2 + \ldots \right) + \frac{26}{27} (\eta_{a,b,c}^{\text{col}})^2 + O(\eta^4) \quad (4.36)$$

It’s expected that for weak noise (or equivalently short times $T$) only terms with low Pauli weight will be present in $G_T$. In this case, by measuring the initial decay rate of a few qubits, the value of selected coefficients of $G_T$ can be extracted. For example, if terms with
Pauli weight $\geq 3$ are negligible, the combination
\[ \gamma^{(a)} + \gamma^{(b)} - \gamma^{(a,b)} = \frac{4}{9} (\eta_{a,b})^2 \]
(4.37)
allows us to establish whether any two-body term between an arbitrary pair of qubits $a$ and
$b$ is present in $G_T$. Notice that the measurements return the value of a given coefficient aver-
gaged over the distribution $P(\eta^P_j, \eta^Q_j)$ (refer to eq. (4.33)), giving thus its strength according
to the parameters of $P$. In this way we can probe any two-qubit collective coefficient we are
interested in, or conduct a fair sampling of some of them.

The systematic protocol to obtain all the collective coefficients is the following:

\( i. \) Measure the $n$ initial decay rates $\gamma^{(j)}$ for all the qubits individually. That is: apply
one step of the algorithm given in Fig. 4-8 measuring only qubit $j$ initially in a pure
state, thus obtaining $f^{(j)}(\Gamma, \rho_0, t)$. From this the initial decay rate can be obtained
by subtracting $f_0^{(j)} = 1$. The accuracy of this procedure can be set by measuring the
fidelity decay at a few points $t = 1, 2, 3, \ldots$ and doing a linear fit. The quality of this
least squares fitting should set the accuracy of the measurement. In such case, the
choice of $f_{\text{fit}}$ is determined by the goodness of the fit.

\( ii. \) Measure the $n(n-1)/2$ initial decay rates for all the possible pairs, $\gamma^{(j,k)}$. This is just
as in \( i \) but now measuring qubits $j$ and $k$, each initially in a pure state.

\( iii. \) With this data, all the two-body coefficients can be determined using (4.37).

\( iv. \) All the one-body coefficients can be extracted by subtracting the two-body coefficients
from the initial decay rates of one qubit, according to (4.34).

This protocol does not distinguish between different product operators $P_j$ with the same
Pauli weight for a given subset of the $n$ qubits, since all the corresponding coefficients add
up to form the collective coefficients as expressed in (4.32).

If terms with higher Pauli weight are present in the generator, we can extend the method,
but of course the number of initial decay rates required to map out the $\eta$ increases, eventually
becoming exponential in $n$. The advantage of this approach is that, when higher order terms
are negligible, it makes good use of this information. In addition, it provides a procedure
to measure selected coefficients instead of having to measure all of them. It is possible, for
example, to probe the importance of three-body terms in $G_T$ (neglecting terms with Pauli
weight $\geq 4$) with the combination
\[ \gamma^{(a)} + \gamma^{(b)} + \gamma^{(c)} - \gamma^{(a,b)} - \gamma^{(a,c)} - \gamma^{(b,c)} + \gamma^{(a,b,c)} = \frac{8}{27} (\eta_{a,b,c})^2 \]
More details on the analytical calculation of the initial decay $\gamma^{(M)}$ is given in Appendix C.

### 4.5 Closing remarks

As described in [58], fault-tolerant QIP (whether quantum computing or quantum com-
munication) requires the magnitude of the noise affecting the implementation of a gate or
transmission channel to be smaller than a certain critical value. The quantity measuring
the noise magnitude and its threshold value depend on the structure of the noise, where
by “structure” we mean which multi-body terms are negligible and how this scales with the
number of qubits. For example, Figs. 4-9 and 4-10 illustrate two possible structures of
non-negligible two-body terms between 6 qubits, represented by lines between them.
In [58] for example, the fault-tolerance threshold presented deals with up to at most two body terms, and for fault-tolerant QIP to be possible, these couplings must obey $|\eta_{i,j}^z| < |\delta|/|j - i|^z$, with $z > D$ and $\delta$ sufficiently small. This means that we require the noise to remain “local enough”, an assumption that needs to be probed experimentally in our QIP setups.

Fault-tolerant thresholds of this type are formulated in terms of a Hamiltonian $\mathcal{H}$ responsible for the errors, acting for the time $\mathcal{T}$ required to implement the gate or channel. The generator $\mathcal{H} \mathcal{T}$ includes the interaction with an external environment (it thus generates both unitary and non-unitary processes), and has support on the system’s Hilbert space $\mathcal{H}_D$ as well as outside of it (the environment’s space, or extra degrees of freedom, as described in Chap. 3). In practice, however, we expect to have access only to the system’s space; the intention of our approach is actually to characterize a generator resulting from the action of $\mathcal{H}$ in the system of $n$ qubits: we have presented here a protocol to analyze the noise structure with at most two-body terms in the generator. Notice the method can be extended if higher order multi-body terms are present, at the price of compromising its scalability. In any case, the method offers a way to probe the importance of these higher order terms.

Our proposal relies on certain assumptions which could limit its reach. The first assumption is, as already mentioned, that terms with high Pauli weight can be neglected. Nevertheless, our method offers a way to probe this assumption. Moreover, this should be generally expected for systems with Hamiltonians that couples the qubits in small groups (low order multi-body terms) and for short enough $\mathcal{T}$.

Another assumption is that we have taken the operators $E$ to be unitary, thus confining the non-unitary processes we considered to the subset of unital processes. This assumption is reasonable enough considering the time scale of typical non-unital processes (the so-called $T_1$-processes or relaxation); unital processes are expected to occur faster thus the corresponding coefficients are expected to be larger. Nevertheless, we will present a more general scenario in the next Chapter. Another limitation of this method is that it deals only with noisy channels, not noisy gates. That is, the map to be characterized is expected to be the identity operator, so the coefficients in $G_T$ are small. We will remove this constraint in the next Chapter. Nevertheless, the identity operator is a primordial gate that gives a fair idea of the noise and errors of the setup. This is particularly true for systems that have an internal Hamiltonian that is always on, and implementing $\mathbb{I}$ means implementing a time-suspension sequence [59, 60] that in principle is composed of several gates modulating the internal Hamiltonian of the system. Also, if we want to analyze the structure of the noise resulting from the implementation of a particular gate $U_g$, we can always take $E = U_gU_g^\dagger$ and perform a gate+gate reversal study.
To conclude this Chapter we revisit the following point: our work indicated that the generalized fidelity decay is initially linear in time, also in agreement with the results already published by Emerson et al. [13]. While this seems to contradict previous well established results reporting a universal quadratic decay for the Loschmidt echo [23], these two statements do not in fact contradict each other, since the random dynamics studied here is not considered by previous work, which presumes the use of a constant evolution operator. In our case, the evolution is given by random rotations which vary at each step and in each realization, and we then study the evolution of an ensemble-averaged state. Even when some relations between our operator $E$ and the perturbation considered in the Loschmidt echo can be drawn, the nature of the calculation is different.
Chapter 5

Experimental implementation of a one-step protocol to characterize spatial correlations

In the previous Chapter we discovered that the decay rate $\gamma$ of the generalized fidelity encoded information about the structure of the map under study. Since the initial decay is linear, we estimated the decay rate by calculating the change in the fidelity after only one step, eq. (4.30).

This actually suggest using a one-step fidelity decay protocol, whose outcome will be exactly the one calculated analytically. In the previous Chapter we targeted a more ambitious task: obtaining the fidelity decay under randomization $f^{LE}(\Gamma, \rho, t)$ for arbitrary $t$ under a one-qubit twirl. But we will see now that, using just one step, the problem becomes analytically tractable while at the same time we can drop many restrictions on what type of processes we can study. Also, in this Chapter we introduce the Clifford twirl to take the place of the Haar twirl. The reasons for doing this, as we will see, are practical: although the theory was developed with a Haar twirl, the experimental implementation is considerably simplified by the use of the Clifford gates.

We present this one-step protocol in a concise form, directly geared towards the measurement of the $\langle |\eta|^{2} \rangle_{\rho}$ coefficients when the interactions involving a large number of qubits simultaneously can be neglected. Even more than in Chapter 4 we leave aside the rather cumbersome mathematical calculations required to prove the formulas we use. Nevertheless, all these results are developed in detail in Appendices B and C.

We implemented this protocol in a liquid-state NMR 4-qubit QIP. This experimental work will be the core of this Chapter. We present an analysis of the implementation, including practical details and an assessment of implementation errors: our basic method, like others proposed [1, 13, 14, 15, 16, 18, 20], assumed error-free implementation stages. This idea is of course unrealistic in practice, and implementation errors complicate the task of reliable quantum process characterization. Thus here we have included an analysis of their effect.
5.1 Theoretical statement of the protocol

We start by describing the action of a general map \( \Gamma \) on the state of an \( n \)-qubit system \((D = 2^n)\), described by an initial state \( \rho_0 \), as in eq. \((2.4)\),

\[
\Gamma(\rho_0) = \int P(\vec{\eta}) \ E(\vec{\eta}) \ \rho_0 \ E(\vec{\eta})^\dagger \ d\vec{\eta}.
\] (5.1)

The vector \( \vec{\eta} \) denotes the \( D^2 \) complex coefficients \( \{ \eta_k, \eta_l; l = 1, \ldots, D^2 - 1 \} \) that parameterize \( E \), an arbitrary (not necessarily unitary) operator in the Hilbert space \( \mathcal{H}_D \), as

\[
E = \eta_0 \mathbb{1} + \sum_{l=1}^{D^2-1} \eta_l \ P_l
\] (5.2)

where the \( P_l \) are the generalized Pauli operators, as already introduced in Sec. 4.2. The difference from the previous Chapter is that the \( \vec{\eta} \) parameterize \( E \) directly, not through a generator \( G_T \). Also, in this case these coefficients are allowed to be complex, and the \( l = 0 \) term is included with a coefficient \( \eta_0 \). When \( P(\vec{\eta}) \) is real, this describes the most general open quantum system dynamics that can occur in \( \mathcal{H}_D \), and it encompasses any arbitrary CP map when \( P \) is a nonnegative distribution. The condition

\[
\int P(\vec{\eta}) |\vec{\eta}|^2 d\vec{\eta} = 1
\] (5.3)

guarantees the preservation of \( \text{Tr}[\rho] \); it can be deduced by taking the trace of eq. \((2.4a)\). (See also Appendix Sec. C.2 a more detailed explanation.)

If the \( \eta_l \) with \( l > 0 \) are small, these coefficients in \((5.1)\) relate to the ones of eq. \((4.4)\). This can be seen from a Taylor expansion of \( \exp(-iG_T) \) (see Appendix Sec. C.6 for more details). Again, when necessary, we shall denote the \( \eta_l \) in more detail as \( \eta^{pq\ldots}_{jk\ldots} \), where \( j > k > \ldots \) label qubits, and \( p, q, \ldots = x, y, z \).

As in the previous Chapter, we measure a subset of \( m \) qubits at a time, which will allow us to extract the magnitude of the coefficients involving that subset. So we now break the system in two: \( m \) qubits belonging to the Hilbert space \( \mathcal{H}_M \), which will be the ones measured, and \( \bar{m} = n - m \) qubits belonging to the complementary space \( \mathcal{H}_{\bar{M}}, \mathcal{H}_D = \mathcal{H}_M \otimes \mathcal{H}_{\bar{M}} \). As before, we require the initial state to be separable in these two spaces, and within \( \mathcal{H}_M \), thus

\[
\rho_0 = \rho^{(M)}_0 \otimes \rho^{(\bar{M})}_0 \quad \text{and} \quad \rho^{(M)}_0 = \bigotimes_{j \in M} |0\rangle_j \langle 0| = |0\rangle_M \langle 0|_M
\] (5.4)

where we have decided to prepare the qubits to be measured in the \( |0\rangle_M \) computational state. We now perform a \( U(2)^{\otimes m} \) twirl of the target map \( \Gamma \), as in Fig. 2-1(b). Unlike Chapter 4, where we worked with a Haar twirl, we state the twirl in terms of Clifford gates:

\[
\rho_1 = \frac{1}{K^m} \sum_{k=1}^{K^m} \mathcal{C}_k \Gamma \left( C_k \rho_0 \ C_k^\dagger \right) C_k
\] (5.5)

Here, the \( \mathcal{C}_k \) are \( m \)-fold tensor products of the one-qubit Clifford operators, for the \( m \) qubits in \( \mathcal{H}_M \). The equivalence of the Haar twirl and the Clifford twirl, already stated in Chapter 2, is explained in more detail in Appendix B. In particular, as shown in Appendix Sec. B.3,
for the task we propose a subset of $K = 6$ operators from the Clifford group for one qubit will suffice. We refer to this minimum set of operators required to apply the twirl (5.5) as the $6^{th}$-Clifford element pool.

The reduced density matrix $\rho^{(M)}_1 = \text{Tr}_M[\rho_1]$ is the state of the $m$ qubits we measure. As seen in Chapter 4, the fidelity decay from $\rho_0$ to $\rho_1$ will encode information about the map $\Gamma$ we are characterizing. In particular, considering only the qubits we will measure and expressing $\rho^{(M)}_0 = (I \otimes \rho_{\text{dev}}^{(j)})/2^{m}$, we obtain

$$\text{Tr}[(\rho^{(M)}_0)^2] - \text{Tr}_M[\rho^{(M)}_0 \rho^{(M)}_1] = 1 - M \langle 0 | \rho^{(M)}_1 | 0 \rangle = \gamma^{(M)} - \Omega^{(M)}_{\text{dev}}$$

(5.6)

with

$$\gamma^{(M)} = \sum_{l \geq 1} (|\eta_l|^2)P \left( \prod_{j \in M} \mathcal{P}_j - \prod_{j \in M} \mathcal{E}_j(l) \right)$$

(5.7)

and $\Omega^{(M)}_{\text{dev}} = 0$ if $\rho^{(M)}_{\text{dev}} = 0$. The quantity defined in the LHS of (5.6) is the fidelity decay rate of the previous Chapter. In eq. (5.7) we denote again $\langle ... \rangle_P = \int P(\vec{\eta}) \cdots d\vec{\eta}$, $\mathcal{P}_j = \text{Tr}[(\rho^{(j)}_0)^2]$ the initial purity of each of the $M$-qubits, and

$$\mathcal{E}_j(l) = \begin{cases} (2/3)(1-\mathcal{P}_j/2) & \text{if } P_{l(j)} = \sigma_x, \sigma_y, \sigma_z \\ \mathcal{P}_j & \text{if } P_{l(j)} = I \end{cases}$$

(5.8)

For the derivation of eqs. (5.6-5.7) we use the equivalence between a Clifford twirl and a Haar twirl (see [15], also Appendix B) and apply the tools developed in [94, 95, 96]. These calculations are developed in full detail in Appendix C: in Sec. C.7 we have calculated explicitly $\Omega^{(M)}_{\text{dev}}$ with $\rho^{(M)}_{\text{dev}} \neq 0$. If $\rho^{(M)}_{\text{dev}}$ is known, then this can be used to bound the value of $\Omega^{(M)}_{\text{dev}}$ with respect to $\gamma^{(M)}$. Note that if we twirl all the $n$ qubits but measure only the $M$-set, we don’t need to worry about $\rho^{(M)}_0$: the $M$-qubits can be in any state. This is proven in Sec. C.4. Thus we can either prepare the maximally mixed state $\rho^{(M)}_0 = (I/2)^{\otimes m}$ and not twirl the $M$-qubits, or twirl all the qubits and not prepare $\rho^{(M)}_0$ in any particular state. Here we work with the former approach, since the preparation of the maximally mixed state in liquid-state NMR is simple.

We have established a closed form for the decay rate $\gamma$ (compare with the previous Chapter, where the formulas (4.34)-(4.36) were up to $O(\eta^4)$). In eq. (5.7), a $l$-term vanishes when $P_l$ is the identity operator for the $m$ qubits being measured. By systematically choosing different sets of $M$-qubits, it is possible to leave out certain $\eta_l$ in a given $\gamma$, as we saw in the previous Chapter. Note however that $\mathcal{E}_j$ does not distinguish the direction of the Pauli matrices, so there is an implicit coarse-graining of all the $P_l$ that have the identity $I$ for the same subset of $m$ qubits. This is the origin of what we called “collective coefficients” $\eta_\text{col}^{(p)}$, which now hold for complex coefficients as

$$|\eta_\text{col}^{(p)}|^2 = \sum_{p=x,y,z} |\eta_p^{(p)}|^2; \quad |\eta_\text{col}^{(j,k)}|^2 = \sum_{p,q=z,y,z} |\eta_{p,q}^{(j,k)}|^2; \quad |\eta_\text{col}^{(j,j',j'')}|^2 = \sum_{p,q,z,x,y} |\eta_{p,q,j,j',j''}^{(j,j',j'')}|^2; \quad \text{etc.}$$

57
By combining the $\gamma$'s from different $M$-sets it is possible to further isolate the collective coefficients. If we prepare a and b in a pure state, so $\mathcal{P}_a = \mathcal{P}_b = 1$, we obtain

$$\frac{9}{4} \left( \gamma^{(a)} + \gamma^{(b)} - \gamma^{(a,b)} \right) = \langle |\eta_{a,b}^{\text{col}}|^2 \rangle_P + \sum_j \langle |\eta_{a,b,j}^{\text{col}}|^2 \rangle_P + \ldots \tag{5.9}$$

The detailed formulas relating to this example can be found in Appendix Sec. C.9.

If 3-body and higher multi-body terms can be neglected, eq. (5.9) gives $\langle |\eta_{a,b}^{\text{col}}|^2 \rangle_P$. Similarly, the combination of the seven $\gamma^{(j)}$, $\gamma^{(j,j')}$, and $\gamma^{(j,j',j'')}$ for a set of three qubits $a, b, c$ would return $\langle |\eta_{a,b,c}^{\text{col}}|^2 \rangle_P$, and so on (again, see Appendix Sec. C.9). As explained in Sec. 4.5, the collective coefficients report on the spatial correlations (interactions involving pairs of qubits) occurring in the quantum process under study.

### 5.2 The protocol

The protocol is depicted in Fig. 5-1. The systematic recipe for measuring the collective coefficients involving the $m$ qubits of a particular subset is the following:

1. Prepare each of the qubits to be measured in the initial state $|0\rangle$. Prepare each of the other $m$ qubits in the maximally mixed state $I/2$.

2. Apply one of the $m$-fold Clifford operators from the $6^m$-Clifford element pool.

3. Implement the target gate or noise $\Gamma$ under study.

4. Invert the Clifford operator applied in ii.

5. Measure the projection of the resulting state on the initial state $|0\rangle$, for each of the qubits being measured.

To implement the twirl, we repeat $i - v$ each time taking a different operator from the $6^m$-Clifford element pool, and average the results of the measurement.

![Circuit representation of the one-step protocol to measure $\gamma^{(M)}$.](image)

Figure 5-1: Circuit representation of the one-step protocol to measure $\gamma^{(M)}$. We use Clifford gates to implement the twirl on the $M$-qubits, and prepare all the qubits accordingly.

In a canonical QIP, the implementation of this protocol to measure the decay rates involving $m$ qubits will require $N$ realizations. This will take care of:

1) preparing the desired initial state as in step $i$, starting from the $|0\rangle^\otimes n$ state, and randomly flipping the $m$ qubits we do not measure (so they are effectively in the $I/2$ state);
2) measuring the fidelity decay through repeated projective measurements of the $m$ qubits as prescribed in step v. Notice that $M \langle 0 | \rho_1^{(M)} | 0 \rangle_M$ in eq. (5.6) is just the probability of finding the $m$ qubits back in the initial state (i.e.: a survival probability);

3) implementing the twirl approximately, by randomly drawing the twirl operators for steps ii and iv from the corresponding $6^m$-Clifford element pool (or alternatively from an infinite pool of one-qubit random rotations).

With this strategy the outcome of the measurement step $v$ are binary strings, and thus the required $N$ can be estimated solely from usual statistics. This makes $N$ independent of the size of the system (the number of qubits $n$) thus allowing the protocol to be scalable, as we already discussed in Sec. 2.3 when introducing the idea of sampling a twirl. Further details on the basic statistics on which this idea is based can be found in Appendix A.

In the case of liquid-state NMR ensemble QIP, the state preparation allows for initialization in the $I/2$ state over the ensemble of molecules (cf. point 1), and ensemble measurements avoid the need of repeated realizations in order to perform step $v$ by using QST (cf. point 2). The number of experiments required to perform QST would be $O(2^{2n})$. If $n$ is small, QST will be practical and we will need $O(2^{2n}6^m)$ experiments to implement the twirl exactly (cf. point 3) and measure $\gamma^{(M)}$. This is the case in our experimental work with a liquid-state NMR processor, where we worked with $n = 4$ and $m = 2$ and implemented the two-fold one-qubit twirl exactly, instead of using $N$ realizations with $N$ according to the minimum number of realizations for a given statistical error threshold.

We notice also that in liquid-state NMR QIP we do not need to perform full QST on the $n$ qubits to measure the projection of the final state $\rho_1^{(M)}$ onto $|0\rangle_M$ for $m$ qubits. Since we only need to measure the $2^m$ coefficients of $\rho_1^{(M)}$ in the generalized Pauli operator (Product Operator) basis, in principle we only need $O(2^m)$ experiments (instead of $O(2^{2n})$).

This and other features related to the NMR implementation will be explained further in Sec. 5.3. Nevertheless, we can see that for liquid-state NMR QIP the scaling is different. This should not be taken as the scalability of the method, which should be evaluated in terms of canonical (non-ensemble) QIP. We explain this further in the next Section.

5.2.1 Scalability of the method

The measurement step $v$ retrieves the information to calculate the decay rates for the $m$ qubits and for any smaller subset of them. For example, twirling qubits $a$ and $b$ only ($m = 2$), we can obtain $\gamma^{(a)}$, $\gamma^{(b)}$ and $\gamma^{(a,b)}$ simultaneously, and calculate $\langle |\eta_{a,b}|^2 \rangle$ as in eq. (5.9) - neglecting the higher multi-body terms. This procedure can be repeated for the $\binom{n}{2} = n(n - 1)/2$ pairs of qubits, and by doing so all the collective coefficients for one-body and two-body terms can be extracted.

The scalability of the method goes as follows. If we neglect the multi-body terms above a certain Pauli weight $w$, and $N$ is the number of realizations required to measure the fidelity $\gamma$ for $m \leq w$ qubits, then with $N \binom{w}{2} \leq Nn^w/w!$ experiments we can estimate all the non-negligible coefficients. This should be compared against the overhead in the number of experiments required for QPT: $N2^{4w}$ (see Table 5.1). We emphasize here that our proposal seeks to characterize the correlations among up to $w$ qubits in order to establish the range of the noise and not to characterize the process fully. The reason is the following: to use our protocol, the negligibility of multi-body terms above a certain Pauli weight $w$ must be established a priori. In a canonical QIP we could apply our protocol to measure all the
$n$ qubits, obtain all the decay rates $\gamma^{(j)}$, $\gamma^{(j,k)}$, \ldots, $\gamma^{(1,\ldots,n)}$ and extract all the collective coefficients after only $N$ experiments (with $N$ independent of $n$). In this way we can handle all the Pauli weights, from 0 to $n$. But the error in the decay rates $\sigma_\gamma$, which will fulfill $\sigma_\gamma \leq 1/\sqrt{N}$ independently of $n$, propagates into the $\eta$ for $m$ qubits inefficiently, roughly as $\sigma_\eta^2 \propto \sqrt{\sum_{j=0}^{w} \binom{n}{j}} \sigma_\gamma$. So although all the $\gamma$'s for $n$ qubits can be efficiently estimated, the error in the $\eta$ will scale inefficiently with $w$. Therefore our strategy of looking for the $\eta$ after establishing a cut-off Pauli weight $w$. To our knowledge, neither QPT nor other proposals so far are able to make use of the negligibility of high order correlations in order to gain further insight (as reflected in Table 5.1). Note that we neglect multi-body terms above a certain $w$, but the multi-body terms with lower Pauli weight can still involve any group of $m \leq w$ qubits among the $n$ qubits.

One approach to establishing this cut-off $w$ is to apply the method developed by Emerson et al. [16] which gives the probability of multi-body terms in $\Gamma$ distinguishing them only by Pauli weight (i.e., an average of all the $\binom{n}{w}$ collective coefficients having Pauli weight $w$). This proposal uses the same resources as our protocol. We will return to this point in Chapter 6.

<table>
<thead>
<tr>
<th>Task</th>
<th>QPT</th>
<th>Our method, canonical QIP</th>
<th>Our method, liquid-state NMR QIP with full QST</th>
<th>Our method, liquid-state NMR QIP with partial QST</th>
</tr>
</thead>
<tbody>
<tr>
<td>measuring a</td>
<td>$N \times 2^N$</td>
<td>$N$</td>
<td>$\text{min}(N, 6^w) \times 2^n$</td>
<td>$\text{min}(N, 6^w) \times 2^w$</td>
</tr>
<tr>
<td>$w$-body coefficient</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>measuring all</td>
<td>$N \times 2^N$</td>
<td>$(\binom{n}{w}) \times N$</td>
<td>$(\binom{n}{w}) \times \text{min}(N, 6^w) \times 2^n$</td>
<td>$(\binom{n}{w}) \times \text{min}(N, 6^w) \times 2^w$</td>
</tr>
<tr>
<td>the $m$-body</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>coefficients</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>with $m \leq w$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.1: Scaling of the overhead of the cost of performing a characterization of the $\eta$ coefficients, assuming that the coefficients corresponding to a Pauli weight higher than $w$ can be neglected. Our method includes a factor $\binom{n}{w}$ that arises as the number of possible sets of $w$ qubits given $n$. Notice that $\binom{n}{w} \leq n^w/w! \leq n^w$, so this factor is polynomial in $w$. $\text{min}(N, 6^w)$ represents the choice of implementing the Clifford twirl exactly or by sampling, which is a reasonable alternative in liquid-state NMR QIP where weak measurements avoid the need of repeated realizations to obtain the outcome of a measurement (partial or full determination of the state of the system). Notice that in canonical QIP we already need many realizations to perform the measurement, so we may as well sample the twirl too. The last column represents the option of having partial QST (again, this option is irrelevant in canonical QIP when repeated realizations are required for partial or full state measurement). The feasibility of such process depends on finding a mapping between the $2^m$ coefficients that characterize $\rho^{(M)}_1(0)\rho^{(M)}_1(0)$ (instead of the $2^n$ that characterize $\rho_1$) and the outcome of the measurement; in any case, the scaling should be at most of $O(2^m)$.  

5.3 Experimental QIP in liquid-state NMR

5.3.1 The basics

In liquid-state NMR QIP, the spin-1/2 nuclei of a given molecule constitute the two-level systems that play the role of qubits. The molecules are in a liquid-state sample, so there are not one but many of them (of the order of $O(10^{21})$), and are not individualized: NMR operations affect all of them simultaneously, and in liquid-state NMR QIP the collective spin state of the sample is measured as a whole. When talking about a system of $n$ qubits being in the state $\rho$, we are referring in fact to $O(10^{21})$ indistinguishable $n$-fold two-level systems. This is why it is called ensemble QIP. The idea of using liquid-state NMR systems for QIP in this fashion emerged in [66, 67].

![Figure 5-2: The crotonic acid molecule. Adapted from [61]. The M stands for the 3 hydrogens of a methyl group. $C_1$ (carbon # 1) labels the carbonyl.](image)

In particular, in the experimental work reported here we use the four $^{13}$C-labeled carbons of crotonic acid, depicted in Fig. 5-2. This molecule has been widely used as a 4-qubit quantum information device (see for example [34, 61, 62, 63, 64]). Recent work seeks to include the hydrogen nuclei to increase the size of the processor to up to 7 qubits (see for example [65]).

The sample is placed in a NMR spectrometer, that is, under a static magnetic field $B_0$ taken to be along the $\hat{z}$ direction. In our experiments we used a 400MHz Bruker spectrometer ($B_0 = 9.4$ T). The dynamics of the spins are controlled by applying a radio-frequency (rf) magnetic field, which translates a Hamiltonian

$$\mathcal{H}^{L}_{\text{rf}} = \sum_{j=1}^{4} \omega_j(t) \cos(\Omega(t) t + \phi(t)) \sigma_x^{(j)}$$ (5.10)

where $\omega_j(t) = -\gamma C B_1(t)$ is related to the magnitude $B_1$ of the applied transversal (perpendicular to $\hat{z}$) rf field, with $\gamma C$ being the gyromagnetic ratio for carbon $^{13}$C. $\Omega(t)$ and $\phi(t)$ are the frequency and phase of the carrier rf signal. $B_1(t)$, $\Omega(t)$ and $\phi(t)$ are varied in time, and are the parameters through which we can control the system. The superscript $L$ stands for laboratory frame. It is convenient to move our description to a rotating frame (a form of an interaction picture) that takes into account the intrinsic rotation given by $\Omega$. In this frame, the system is described by a density matrix $\rho(t) = R^d \rho^{L}(t) R$, and

$$R = \exp \left( -i \Omega_0 t \sum_{j=1}^{4} \sigma_x^{(j)} \right)$$ (5.11)

$$\mathcal{H} = R^d \mathcal{H}^{L} R - \Omega_0 \sum_{j=1}^{4} \sigma_x^{(j)}$$ (5.12)
The carbon $^{13}$C Larmor frequency $|\gamma C B_0|$ at 9.4 T is 100.658 MHz. $\Omega_0$ is set near this value in order to operate close to resonance. Notice the four carbons have slightly different Larmor frequencies (due to their different chemical environment given their positions in the molecule), thus we pick $\Omega_0$ to be near those values although not exactly matching any of them. Specifically, in our experiments we worked with $\Omega_0 = -100.623$ MHz. The action of $\mathcal{H}_{r_f}^L$ is usually referred as pulsing/pulses.

The effective internal Hamiltonian $\mathcal{H}_{\text{int}}$ of the molecule in the rotating frame is

$$\mathcal{H}_{\text{int}} = \sum_{j=1}^{4} \frac{\omega_{\delta,j}}{2} \sigma_z^{(j)} + \sum_{k>j=1}^{4} \frac{\pi J_{jk}}{2} \sigma_+^{(j)} \cdot \sigma_-^{(k)} \quad (5.13)$$

where the $\omega_{\delta,j}$, at our 400MHz spectrometer, are: $\omega_{\delta,1}/\pi = 6650.6$ Hz, $\omega_{\delta,2}/\pi = 1695.8$ Hz, $\omega_{\delta,3}/\pi = 4210.0$ Hz, $\omega_{\delta,4}/\pi = -8796.7$ Hz, following the labeling of Fig. 5-2. These frequencies arise from the original Larmor frequencies $\omega_{L,j} = -\gamma C (1-\sigma_j^z) B_0$ after $\Omega_0$ is subtracted. The $\sigma_j^z$ represent a change in the field perceived by each nuclei due to the electromagnetic environment created by the molecule (a chemical shift). The first sum in eq. (5.13) are the Zeeman energies (arising from the presence of an external static magnetic field $B_0$ along $\hat{z}$), while the second sum are the so-called J-coupling (also scalar coupling) terms, which arise from the electronic bonding between nuclei, and are always present in the molecule. The J-couplings are $J_{12} = 72.6$ Hz, $J_{23} = 69.8$ Hz and $J_{14} = 7.1$ Hz, while $J_{24} = 1.6$ Hz, $J_{13} = 1.3$ Hz, and $J_{34} = 41.6$ Hz.$^\dagger$

$\mathcal{H}_{r_f}^L$ must be taken to the rotating frame according to eq. (5.12). Notice this is not simply achieved by subtracting the rotation given by $\Omega_0$. Instead, the $\mathcal{H}_{r_f}^L$ of eq. (5.10) is decomposed it into a counterclockwise component and a clockwise component, and the latter is neglected once the resonance condition $\Omega(t) - \Omega_0 \approx \omega_{\delta,j} << \Omega_0$ (at any instant $t$) is assumed, together with the negligibility of the Bloch-Seigert shift as $\gamma C B_1(t) << \Omega_0$ (at any instant $t$). These conditions are met experimentally, since the magnitude of the transversal magnetic field is much smaller than the static one, $B_1(t) << B_0$ (at any instant $t$), and $\Omega(t)$ can be tuned close enough the Larmor frequencies $\Omega_0 + \omega_{\delta,j}$ of the nuclei with $\Omega_0 \approx -\gamma C B_0$. This is called the rotating wave approximation (RWA).

In such scenario, the effective rf field Hamiltonian in the rotating frame can be taken as

$$\mathcal{H}_{r_f} = \omega(t) \left( \sum_{j=1}^{4} \cos(\Omega'(t)t + \phi(t))\sigma_z^{(j)} + \sin(\Omega'(t)t + \phi(t))\sigma_y^{(j)} \right) \quad (5.14)$$

with $\Omega'(t) = \Omega(t) - \Omega_0$, and $\omega(t) = \gamma C B_1(t)$, neglecting already the effect of the $\sigma_j^z$ in $\mathcal{H}_{r_f}$ but not in the Zeeman terms in $\mathcal{H}_{\text{int}}$ (since $B_1(t) << B_0$ at any instant $t$).

Engineering and control of closed quantum systems deals with the problem of using control Hamiltonians (like $\mathcal{H}_{r_f}$) to overcome the natural evolution of internal Hamiltonians and in general to fully govern the evolution of the system. In liquid-state NMR, $\mathcal{H}_{\text{int}} + \mathcal{H}_{r_f}$ guarantee complete control over the carbons in crotonic acid. We can, for example, turn on and off $\mathcal{H}_{r_f}$ (piece-wise constant $\omega(t)$), with a constant carrier frequency and phase. Such pulses are called hard pulses and constitute basic building blocks for more general pulsing

$^\dagger$The absolute value of the J-couplings are according to a characterization we did of the sample we prepared for the experiments - crotonic acid in deuterated chloroform (CDCl₃).
techniques, such as composite pulses [68]. General control over the system can be achieved and further improved using sophisticated pulsing profiles. These profiles can be basically searched numerically starting from a guess shape with so-called pulse finding/making methods. In particular, we will be using the Strongly Modulating Pulses (SMP) [34] and GRadient Ascent Pulse Engineering (GRAPE) [69] methods. More information about these methods and about closed quantum systems control can be found in [70].

The sample is initially at (room temperature) thermal equilibrium, thus its state is

\[
\rho_{th} = \frac{\exp(-\frac{H_{Lat}}{K_B T})}{\text{Tr}[\exp(-\frac{H_{Lat}}{K_B T})]} \approx \frac{I}{D} + \sum_{j=1}^{4} \Omega_0 + \omega \delta_g \sigma_z^{(j)}
\]

\[
\approx \frac{I}{D} + \frac{\epsilon}{D} \sum_{j=1}^{4} \sigma_z^{(j)} := \frac{I}{D} + \epsilon \rho_{th}^{dev}
\]

(5.15)

(5.16)

where \( D = 2^4 \) for our 4 qubits, and \( K_B \) is the Boltzmann constant and \( T \) the (room) temperature, in Kelvin. The approximation comes from considering the high-temperature limit, which makes the last term in (5.15) relatively small and thus such Taylor approximation is valid (i.e. \( \epsilon = \frac{\Omega_0}{K_B T} << 1 \)). We have defined in (5.16) a deviation density matrix: \( \rho_{th}^{dev} \).

It is evident that something other than simple pulsing needs to be done in order to have the system initialized in a pure state, which is a requirement for the realization of general QIP. Since the purity of the thermal equilibrium state is \( \text{Tr}[\rho_{th}^2] = (1 + \epsilon^2)/D < 1 \), we would need a controlled non-unitary operation to take it to a pure state\(^\dagger\). This was one of the major points addressed in [66, 67], which lead to the concept of pseudo-pure states: we assume that everything that will happen in our quantum processor will be a unital process, then all the dynamics occur in \( \rho_{th}^{dev} \), since unital processes take the identity to itself. This is the case with all the operations in liquid-state NMR, because the only non-unital processes (typically, but not exclusively, \( T_1 \) relaxation) have a time scale much larger than the unital ones (any operation achieved by rf field pulsing, \( T_2 \) relaxation, etc.) and thus can be neglected. This a well-known fact in the NMR community, where actually the state of a system is directly described by a traceless matrix \( \rho^{dev} = \rho - \frac{I}{D} \) rather than by \( \rho \) itself. Thus in preparing the system in a pure state \( |\varphi\rangle \), we are actually taking \( \alpha \rho_{th}^{dev} \) to \( \rho_{\varphi}^{dev} = |\varphi\rangle \langle \varphi| - \frac{I}{D} \), up to a constant factor \( \alpha \). This constant factor can be then arbitrarily set to match the desired state. The key is that later on, when we perform other operations on the system after having prepared it initially in \( \rho_{\varphi}^{dev} \) \( \rightarrow |\varphi\rangle \), we should use this same factor \( \alpha \) to compare back with the initial state. Keeping track of this is critical when calculating the fidelity decay, as we will calculate indeed changes on the initial state.

On the other hand, we have mentioned the desire to prepare some of the qubits in the maximally mixed state \( I/2 \). Again, if the starting point is the equilibrium state (5.16), we need a clever trick as again there is a change in purity. One solution is the use of magnetic field gradients. The spectrometer is able to implement controlled spatial variations in the static field along \( \hat{z} \), which are called gradients. The \( B_0 \) field is then inhomogeneous across the sample in a well defined manner. We can consider that during the implementation of a

\(^\dagger\)Unitary operations do not change the purity of a state. If we transform a state \( \rho \) into \( U \rho U^\dagger \), we have \( \text{Tr}[(U \rho U^\dagger)^2] = \text{Tr}[U \rho U^\dagger U \rho U^\dagger] = \text{Tr}[U \rho^2 U^\dagger] = \text{Tr}[\rho^2] \).
gradient, we add an additional Hamiltonian (in the rotating frame)

$$\mathcal{H}_{\text{grad}} = \frac{g(x,y,z)}{2} \sum_{j=1}^{4} \sigma_z^{(j)}$$

(5.17)

Since eventually we choose not to distinguish the $O(10^{23})$ nuclei that collectively form a qubit, the net state should always be taken as the average over the spatial degrees of freedom of the sample. Take the initial $\rho_{th}^{\text{dev}}$ and rotate the k-th qubit from $\sigma_z^{(k)}$ to $\sigma_x^{(k)}$. This can be done with a unitary operator implemented experimentally by a pulse profile generated with the SMP or GRAPE methods. Now let it evolve under $\mathcal{H}_{\text{grad}}$ with a $g$ that has a $z$-dependence. The other qubits will remain in $\sigma_z^{(j)}$, but the k qubit will evolve from $\sigma_z^{(k)}$ to $\cos(g(z)t)\sigma_x^{(k)} + \sin(g(z)t)\sigma_y^{(k)}$. If $g(z)$ is conveniently chosen so $\int e^{i g(z) t} dz = 0$, the average over the $z$ direction can completely suppress this term\(^\dagger\). This is how we obtain the state $I/2$ on the k-th qubit.

We must be careful and note that the average occurs only at the end of the experiment (at the measurement step), so when using more than one gradient fields, we must guarantee that they all produce the desired average at the end. Typically, when using only two or three gradients, we simply choose the variation of $g$ at each one to be along perpendicular directions. We also note that gradients are unitary operations.

We also point out that the use of gradients will also be required to prepare pseudo-pure states, since deviation density matrices like $|\varphi\rangle\langle\varphi| - I/D$ cannot be obtained from $\rho_{th}^{\text{dev}}$ using solely unitary transformations (even up to a constant $\alpha$). This will be illustrated clearly when we discuss the initial state preparation in our experiment, in Sec. 5.4.2. This strategy of using gradients for pseudo-pure state preparation is called ‘spatial labeling’ (as opposed to temporal or logical labeling), and it was introduced in [66].

### 5.3.2 Measurement

The measurement process in NMR is such that the spectrometer gives a signal proportional to the magnetization of the whole sample in the transverse plane. If the state of the collective system at some final time $t_f$ is $\rho_f$, using quadrature detection on the pick-up coil signal we measure simultaneously the net magnetization along $\hat{x}$ and $\hat{y}$ in the rotating frame ($q = x, y$),

$$M_q(t_{ac}) = N \gamma_C \text{Tr} \left[ e^{-i \mathcal{H}_{\text{int}} t_{ac}} \rho_f e^{i \mathcal{H}_{\text{int}} t_{ac}} \left( \sum_{j=1}^{4} \sigma_z^{(j)} \right) \right]$$

(5.18)

where the acquisition time $t_{ac}$ is relative to $t_f$, and is typically a few seconds (that is why the evolution of the system during that time, given by $\mathcal{H}_{\text{int}}$ must be taken into account). $N$ is the number of nuclei per unit of volume.

We decompose $\rho_f$ in terms of the Pauli operators as $\rho_f = I/D + \sum_{l=1}^{D^2-1} c_l P_l$. The $P_l$ term will evolve to $P_l^{\text{meas}}(t_{ac}) = e^{-i \mathcal{H}_{\text{int}} t_{ac}} P_l e^{i \mathcal{H}_{\text{int}} t_{ac}}$. Let’s approximate the couplings in $\mathcal{H}_{\text{int}}$ as weak couplings.$^\S$ In that case, we see that any $P_l$ that only has $I$ and/or $\sigma_z$ factors

\(^\dagger\)To be more precise, if the sample is inhomogeneous we must also take the spin density into account in the average.

\(^\S\)The weak coupling approximation consists in taking the $J$-coupling terms in $\mathcal{H}_{\text{int}}$ as simply $\sigma_z^{(j)} \sigma_z^{(k)}$
commutes with $\mathcal{H}_{\text{int}}$ and will render $M_q = 0$ after taking the trace. Also, it is trivial to see that a $P_l$ that has two or more factors that are $\sigma_x$ or $\sigma_y$ will evolve into a $P_{\text{meas}}^t(t)$ which is a combination of Pauli operators other than $\sigma_q \otimes (I/2)^{\otimes n-1}$. In that case $\text{Tr}[P_{\text{meas}}^t(t)\sigma_q] = 0$. In both these situations, the corresponding $c_l$ will be inaccessible. Finally, it is easy to prove that a $\mathcal{H}_{\text{int}}$ with only $\sigma_z$ and $\sigma_z$ terms will give out a non-null oscillating function (characterized by frequencies $\omega_{k,j}$ and $2\pi J_{jk}$) for only the $P_l$ that have $\sigma_q$ for one qubit and either $I$ or $\sigma_z$ for the rest.

These $n \times 2 \times 2^{n-1} = nD$ terms are “the observables” (the observable Product Operators [56]). By Fourier-analyzing the measured signal, a typical NMR spectrum of the state of the sample is obtained. If the frequencies given by the combination of the $\omega_{k,j}$ and $2\pi J_{jk}$ are well resolved (in particular, if the $J$-couplings are large enough compared to the typical linewidth of the spectra), the $c_l$ coefficients for these $nD$ terms are obtained from the processing of this spectral data.

To obtain the other $D^2 - 1 - nD$ coefficients in order to describe the full state, another strategy is required. This is using the so-called readout pulses. Instead of measuring the system at the final stage given by $\rho_f$, a readout (ro) pulse is applied before the measurement so the state of the system is $\rho'_f = U_{\text{ro}} \rho_f U_{\text{ro}}^\dagger$. If the ro pulse is chosen cleverly, the “non-observable” $P_l$ can be transformed to observable ones. In this way, all the coefficients can be mapped out to observable terms and be measured. This is how QST is performed in liquid-state NMR. To our knowledge, there is no systematic way to find out the readout pulses. In the best case scenario, if we can map out a different subset of $nD$ coefficients with each different readout pulse, we would need $(D^2 - 1 - nD)/nD = O(D/n)$ pulses. The method is then, as expected, non-scalable. However, this is of course feasible for small systems. For $n = 4$ this formula indicates a minimum of 7 pulses; our best guess of an actual set of readout pulses is at this point 18.

In this analysis we have assumed that the couplings in $\mathcal{H}_{\text{int}}$ are weak. If we take into account strong coupling, some other $P_l$ become observable in principle [56], but based on empirical evidence and numerical simulation, we can rely on the weak coupling assumption to choose a set of readout pulses. After the pulse finding, the most general scenario (see following Secs. 5.3.3, 5.3.4 and 5.3.5) was considered when testing the pulses, and more importantly, when reversing them to extract which are the non-observable $c_l$ mapped out to observable terms by a given ro pulse.

We can increase the number of readout pulses arbitrarily to $N_{\text{ro}}$, to obtain a $nDN_{\text{ro}} \times (D^2 - 1)$ matrix that contains how the $nD$ coefficients measured after each of the $N_{\text{ro}}$ readout pulses map back to the $D^2 - 1 c_l$, after the inversion of the readout. The pulses, acting under any experimental conditions, remain a complete set as long as the rank of this matrix turns out to be $D^2 - 1$. The weak coupling assumption comes into play only to orient us in choosing the readout pulses that will form such a complete set.

We see that performing QST in liquid-state NMR is very different from the standard projective measurement (which is the type of measurement we have discussed so far in the theory of the protocols). This is one of the main differences between ensemble (liquid-state NMR) QIP and canonical QIP. Quantum computation and information theory work on the basis of projective measurements, so protocols and algorithms are developed in that frame instead of $\hat{\sigma}^{(j)} \cdot \hat{\sigma}^{(k)}$ [strong coupling]. This is valid for a pair $j,k$ as long as $|\omega_{L,j} - \omega_{L,k}| >> |J_{jk}|$, where the $\omega_{L}$ are the Larmor frequencies in the Zeeman terms of $\mathcal{H}^{L}_{\text{int}}$. 
(hence our analysis of the scalability of our method in Sec. 5.2.1). Nevertheless ensemble QIP is equally valid and has a long history as test-bed for quantum control and study of system dynamics: although the initial state preparation and measurement are different from the ones of a canonical quantum computer, the system manipulation in between is incredibly rich and developed, and useful as test ground for experimental QIP.

In this regard, one important point is that there is no need to repeat an experiment associated with a readout, unlike with canonical QIP devices where we must have $N$ realizations to conclude the value of any coefficient in which we have expanded the state $p_f$. In principle, to perform QST we need $O(N \times D^2)$ realizations, while in liquid-state NMR we only need $O(D)$. Although they are both non-scalable, there is no need for repeated realizations in the readout. Since we have established already that in liquid-state NMR we also do not need many realizations to prepare maximally mixed states (as we required in the description of the protocol in 5.2), the only reason to do many realizations in our experiment would be to sample the $6^m$-Clifford pool. Again, if $m$ is small enough, we can implement the twirl exactly and then there is no reason to use repeated realizations at all. This is reflected in Table 5.1.

### 5.3.3 Simulation

The main idea behind numerical pulse finding methods is to guess an initial profile, then make an educated guess proposing a variation, and test the new variation to see if it actually performs better that the previous one or not. Successive iterations of this will retrieve a locally optimal profile with respect to parameter space. The typical quantity that evaluates the goodness of a pulse profile is in fact the average gate fidelity we have been discussing so far, as defined in eqs. (2.5), (2.7). This is what is used by the SMP and GRAPE algorithms. However, as we have seen, the calculation of this quantity is non-scalable, which makes pulse searching strategies non-scalable in general. Moreover, even if we sample the average fidelity (as it is one of the main ideas discussed in this thesis), the calculation of each element of the average would require the numerical calculation of the action of the Hamiltonians on a particular state, using $2^n \times 2^n$ complex matrices. This is usually refereed as the (classical) simulation of a quantum process. Naturally, such calculations will be impossible without enough (classical) computer memory.

Nevertheless, for reasons other than the non-scalability of simulations for pulse finding, experimental QIP is still at the level of even fewer qubits. In liquid-state NMR QIP for example, we have the problem of the exponential decrease of the signal-to-noise ratio in the preparation of pseudo-pure states [66]. Moreover, it still remains a challenge to find a molecule with many nuclei with a well resolved spectrum.

In this scenario, where experimental quantum processors are at the level of a few qubits, we find that we can simulate our quantum processor with a classical one (note that simulations are nothing but classical numerical calculations involving large complex matrices). It is with these simulations that pulse profiles are searched with the SMP or GRAPE methods. Consequently, each pulse (or sequence of several of them) will have an associated fidelity $F_g$, which is the outgoing entanglement fidelity (2.8) after a numerical simulation of the performance of the designed pulsing against the ideal (desired) gate. This $F_g$ indicates the expected goodness of our pulse profile. The liquid-state NMR systems used in QIP have been well studied, and the simulations account for the known occurring dynamics. Typically, these simulations will include not only the presence of the full Hamiltonian of the system,
but also other features such as relaxation or the ones we explain below (presence of the hydrogens and rf field inhomogeneity). When presenting our experimental work in Sec. 5.4, we will indicate these conditions.

### 5.3.4 Incoherent sum treatment of the hydrogens

We must acknowledge that the full Hamiltonian of the molecule, in the rotating frame,\(^\dagger\) is

\[
H = H_{\text{int}} + H_H + H_{CH} \tag{5.19}
\]

\[
H_H = \sum_{j=5}^{9} \frac{\omega_{L,j}}{2} \sigma_{z}^{(j)} + \sum_{k>j=5}^{9} \frac{\pi J_{j,k}}{2} \sigma_{z}^{(j)} \cdot \sigma_{z}^{(k)} \tag{5.20}
\]

\[
H_{CH} = \sum_{j=1}^{4} \sum_{k=5}^{9} \frac{\pi J_{j,k}^{CH}}{2} \sigma_{z}^{(j)} \sigma_{z}^{(k)} \tag{5.21}
\]

where we directly take the couplings between the different species carbon and hydrogen to be simply \(\sigma_z \sigma_z\) instead of \(\vec{\sigma} \cdot \vec{\sigma}\) thanks to the fact that \(|\omega_{L,j} - \omega_{L,k}| \gg J_{j,k}\) for different species. (As we mentioned, the Larmor frequencies of carbons are around 100MHz, while the ones for Hydrogens are around 400MHz.) Acknowledging the presence of the hydrogens also alters the expression for the equilibrium thermal state (5.16),

\[
\rho_{th}^{\text{full}} \approx \rho_{th}^{\text{full}} + \epsilon_{th}^{\text{dec}} \otimes \left( \frac{1}{2} \right)^{\otimes 5} + \epsilon_H^{\text{dec}} \left( \frac{3}{2} \right)^{\otimes 4} \otimes \sum_{j=5}^{9} \sigma_{z}^{(j)}. \tag{5.22}
\]

If we approximate the coupling between hydrogens also as weak, we have that the Zeeman states \(|\chi_m\rangle\) of the hydrogens\(^*\) are eigenstates of the full Hamiltonian: \(H |\chi_m\rangle = |\chi_m\rangle (H_{\text{int}} + H_H + H_{CH})\) (notice \(H_{CH}\) is diagonal in the Zeeman basis)\(^\dagger\). Since we are interested only in the evolution of the carbons, we can trace out the hydrogens after the evolution of the system. Consider \(H'\) that contains not only \(H\) but also any other external Hamiltonians acting on the carbons, such as \(H_{rf}\) or \(H_{\text{grad}}\). Still \(|\chi_m\rangle\) is an eigenstate of \(H'\), and moreover

\(^\dagger\)We now use a doubly rotating frame given by \(R_C \otimes R_H\), where \(R_C\) is given by eq. (5.11) and \(R_H\) is just like that, but with an \(\Omega_0\) that will be around the typical Larmor frequency of the hydrogens: 400MHz.

\(^*\)The Zeeman basis for \(n\) qubits is just the \(2^n\) tensor products of the eigenstates of \(\sigma_z\) for each qubit, and each of its elements is typically labeled by a binary string of length \(n\) indicating the corresponding eigenvalues. For the hydrogens, \(n = 5\) and \(m = 0, 1, \ldots, 2^5\).

\(^\dagger\)Given a tensor operator \(A \otimes B\) in Hilbert space \(\mathcal{H}_a \otimes \mathcal{H}_b\), and given two states \(|\chi_a\rangle\) and \(|\chi_b\rangle\) in \(\mathcal{H}_b\), we have that \(\langle \chi_a | A \otimes B | \chi_b \rangle = \langle \chi_a | B | \chi_b \rangle \cdot \langle \chi_a | A | \chi_b \rangle\). An operator in \(\mathcal{H}_a\) only. \(H_{CH}\) and \(H_H\) represent the residual operator corresponding to \(H_{CH}\) and \(H_H\).
\[
[H', \sum_{j=5}^9\sigma_z^{(j)}] = 0. \text{ Then the evolution of the carbon subsystem will be}
\]

\[
\text{Tr}_H[p^{\text{full}}(t)] = \text{Tr}_H[e^{-iH' t}p^{\text{full}} e^{+iH' t}]
\]

\[
= \frac{i}{D} + \epsilon \text{Tr}_H \left[ e^{-iH' t} p_{\text{th}} \otimes \left( \frac{i}{2} \right) \otimes^5 e^{+iH' t} \right] + \frac{\epsilon H}{2^5} \left( \frac{i}{2} \right) \otimes^4 \text{Tr}_H \left[ \sum_{j=5}^9 g^{(j)} \right]
\]

\[
= \frac{i}{D} + \epsilon \sum_{m=1}^{2^5} \langle \chi_m | e^{-iH' t} p_{\text{th}} \otimes \left( \frac{1}{2} \right) \sum_{k=1}^{2^5} |\chi_k\rangle \langle \chi_k | e^{+iH' t} |\chi_m\rangle
\]

\[
= \frac{i}{D} + \frac{\epsilon}{2^5} \sum_{m=1}^{2^5} e^{-iH'_m t} p_{\text{th}} e^{+iH'_m}
\]

(5.23)

where \( H'_m = \langle \chi_m | H' |\chi_m\rangle \) is a \( 2^4 \times 2^4 \) matrix that simply has the Hamiltonians we defined above for the carbons only, shifted by a diagonal matrix \( H_{CH} + H_H \).

We refer to this strategy as incoherent sum treatment of the hydrogens. By using this procedure we take into account the magnetically active hydrogens in the molecule, while still keeping the problem in the space of four qubits. Although the scalability of the problem remains the same of course, this approach proves to improve the pulse finding speed notably, as in this way the memory usage of the computer is more effective: we take into account the presence of an extra \( 2^5 \times 2^5 \) space but still working only the \( 2^4 \times 2^4 \) space of interest. The strategy now is to guess a pulse profile (generated by \( H_f \) in \( H' \)) and calculate its fidelity \( F_y \) as

\[
\frac{\text{Tr}[\hat{U}^{-1}\hat{\Gamma}_\text{pulse}]}{D^2} = \frac{1}{D^2} \text{Tr}[ \left( \hat{U}^\dagger \text{ideal} \otimes \hat{U}^\dagger \text{ideal} \right) \sum_m \hat{A}_m \otimes \hat{A}_m]
\]

\[
= \sum_m \frac{|\text{Tr}[\hat{U}^\dagger \text{ideal} \hat{A}_m]|^2}{D^2} = \sum_m \frac{|\text{Tr}[\hat{U}^\dagger \text{ideal} \text{T} [e^{-i\int_0^t \hat{H}'_{\text{m}} t dt}]\hat{A}_m]|^2}{2^5 D^2}
\]

where we used eq. (2.8) and, following (5.23), took the \( \text{T} [e^{-i\int_0^t \hat{H}'_{\text{m}} t dt}]/\sqrt{2^5} \) as Kraus operators like in (2.6). \( \hat{T} \) represents the Dyson time-ordering operator specifying how the integral should take into account the non-commutativity of \( H'_m \) at different times. Typically all the time-varying parameters in \( H_f \) are piece-wise constant, so the numerical calculation of the net propagator is not a problem. \( t_p \) is the length (time duration) of the pulse, determined ultimately by the pulse finding.

\( H_{CH} \) is diagonal in the Zeeman basis for the hydrogens \( |\chi_m\rangle \). \( H_H \) can be taken with a weak coupling approximation, so then it is very simple to evaluate \( H_{CH} + H_H = \langle \chi_m | (H_{CH} + H_H) |\chi_m\rangle \), which is a diagonal matrix in the carbon subspace. Thus, the effect of the hydrogens is to create a series of \( 2^5 \) effective Hamiltonians for the carbons with the diagonal elements shifted by the \( J \)-couplings with and the chemical shifts of the hydrogens. These shifts can be easily computed. Moreover, due to the magnetic equivalence of the 3 carbons in the methyl group (see Fig. 5-2), this number can be reduced to \( 2^3 = 8 \) (a degenerate incoherent sum treatment).

The assumption of weak coupling between hydrogens is justified on empirical evidence: We searched for various pulses in the carbon subspace using this incoherence sum treatment, and then tested the pulses by simulating them in the full Hilbert space of the 9 spins,
including all strong couplings among spins of same species. The fidelity $F_g$ turned out to be the same up to 6 decimal places at least. We can always search for pulses with this incoherent sum treatment and test the validity of the model (and the goodness of the pulse) by doing a one-time simulation over the full space (as opposed to doing a search in the full space directly).

We could retain the strong coupling in the hydrogens and numerically evaluate $\langle \chi_m | (\mathcal{H}_C + \mathcal{H}_H) | \chi_m \rangle$ for each of the $2^5 |\chi_m\rangle$ states. The method holds in this case also because $[\mathcal{H}'', \sum_{j=5}^9 \sigma_j^{(j)}] = 0$ even with strong coupling among hydrogens. But then it is no longer true that two pulses that were independently found with the incoherent sum method will work the same acting back-to-back. The reason is that the hydrogens can be regarded as an environment to the carbons. The initial state (5.22) is indeed separable into the carbon space and the hydrogen space. However, since the $|\chi_m\rangle$ are no longer the eigenstates of $\mathcal{H}_C + \mathcal{H}_H$, we have that $\langle \chi_m | e^{-i\mathcal{H}' t_1} e^{-i\mathcal{H}_m t_2} | \chi_m \rangle \neq e^{-i\chi_m | \mathcal{H}' t_1} e^{-i\chi_m | \mathcal{H}_m t_2}$.

Nevertheless, as we mentioned before, the weak coupling approximation works well and it does not require any extra computation. The incoherent sum treatment is then a practical effective way to take into account the hydrogens when only pulsing on the carbon subsystem (and as long as the hydrogens are initially in the thermal equilibrium state, which will not evolve in this situation).

### 5.3.5 rf field inhomogeneities

Due to the finite size of the coil, there are spatial inhomogeneities in the $B_1$ field across the sample. This gives rise to spatial inhomogeneities in $\omega(t)$ in (5.14), which turn the actual action of the pulse profile into a superoperator of the form

$$\hat{\Gamma}_\text{pulse} = \sum_{m=1}^{2^5} \int P(x, y, z) T[e^{-i \int_0^{t_p} \mathcal{H}_m'' t \, dt}] \otimes T[e^{-i \int_0^{t_p} \mathcal{H}_m'' t \, dt}] \, dx \, dy \, dz$$

(5.24)

$$\mathcal{H}_m'' = \omega(t, x, y, z) \sum_{j=1}^4 \cos(\Omega_j(t) t + \phi(t)) \sigma_j^{(x)} + \sin(\Omega_j(t) t + \phi(t)) \sigma_j^{(y)}$$

$$+ \mathcal{H}_\text{int} + \langle \chi_m | (\mathcal{H}_C + \mathcal{H}_H) | \chi_m \rangle$$

where $P(x, y, z)$ is the distribution that weights the density of spins which perceive a field $B_1(x, y, z, t)$ at $(x, y, z, t)$. To obtain (5.24) we follow eq. (2.6) and include the incoherent treatment of hydrogens. The inhomogeneity can be approximated by a discrete distribution denoting what portion of the sample perceives a given value of $B_1$. This inhomogeneity profile was measured for the carbon coil set of our spectrometer by N. Boulant [71], p. 44.

By searching for a pulse using $\text{Tr}[\hat{U}_{\text{ideal}}^{-1} \hat{\Gamma}_\text{pulse}] / D^2$ as a goodness function (with $\hat{\Gamma}_\text{pulse}$ as in eq. (5.24)), we find pulses that will compensate for both rf inhomogeneity and the presence of magnetically active hydrogens. See also [62] for a first implementation of the rf inhomogeneity compensation in pulse finding.

### 5.4 The experiment

We now present a detailed description of the protocol implementation. Unless otherwise noted, all the fidelities $F_g$ were obtained by a simulation including rf inhomogeneity and an incoherent sum treatment of the hydrogens. The pulse searching, however, may have not
include all these features. But nevermind how we found the pulse profile; what matters is how well it performs under the most realistic conditions.

5.4.1 The processes under study

Our aim was to measure two-qubit collective coefficients \( \langle \eta_{ab}^{\odot l} \rangle_P \), for a series of gates. We chose the following five:

\( i \) A time suspension sequence \( I_E \). It is important to study our ability to “do nothing” in a system with a natural Hamiltonian that is always on. The sequence was 12.2 msec long with \( F_g = 0.9628 \). This gate has ideally \( \langle |\eta_0|^2 \rangle = 1 \), \( \langle |\eta_l|^2 \rangle = 0 \) \( \forall \ l > 0 \).

\( ii \) An engineered error creating a coupling between qubits 1 and 2, of the form \( C_{12}(\beta) = \exp(-i\beta\sigma_z\sigma_z) \). Ideally, the only non-null coefficients are then \( \langle \eta_{12}^{\odot l} \rangle = \langle \eta_{12}^{\odot l} \rangle = \sin^2(\beta) \), and \( \langle |\eta_0|^2 \rangle = \cos^2(\beta) \).

We had \( \beta = 0.1 \) (4.88 msec long and \( F_g = 0.9906 \)), and \( \beta = 0.4 \) -- a concatenation of the previous one, so it was 19.52 msec long with a resulting \( F_g = 0.8662 \).

\( iii \) A CNOT gate between qubits 1 and 2. \( CNOT = 0.5(I + \sigma_z(1) + \sigma_z(2) - \sigma_z(1)\sigma_z(2)) \), so \( \langle \eta_{12}^{\odot l} \rangle = \langle \eta_{12}^{\odot l} \rangle = 0.25 \); it was 11.88 msec long with \( F_g = 0.9900 \). Also, this same gate applied twice: \( CNOT^2 = I \), simulating with \( F_g = 0.9756 \) and again we have \( \langle |\eta_0|^2 \rangle = 0 \) \( \forall \ l > 0 \).

These gates are more complex than one-qubit operations (which are typically less than 1 msec long) and they all involve refocusing idle times (periods of free evolution under \( H_{int} \) or pulsing at very low \( B_1(t) \) values) in their pulse sequences.

The time-suspension sequence was implemented by composing one-qubit pulses and free-evolution periods (denoted by \( -\tau - \)) as follows:

\[
\tau - \pi \ \sigma_z^{3,4} - \tau - \pi \ \sigma_x^{2,2} - \tau - \pi \ \sigma_z^{3,4} - \tau - \pi \ \sigma_x^{1,4} - \tau - \pi \ \sigma_z^{3,4} - \tau - \pi \ \sigma_x^{1,4} - \tau - \pi
\]

\( \theta \)\( j,k \ldots \) denotes a \( \theta \)-pulse in the \( p \)-axis for each of the qubits \( j, k, \ldots \). This is, for each qubit, a propagator in \( U(2) \) that is a rotation around \( \hat{p} \) in an angle \( \theta \). If the pulses are perfect, this sequence perfectly refocuses the evolution of an internal Hamiltonian with weak couplings [72]. The individual pulses were searched with the SMP method and then combined using a “delay optimization” technique: we simulated the whole sequence varying \( \tau \) around 1 msec, looking to maximize the fidelity \( F_g \), considering the imperfections in the pulses and taking into account the strong coupling**.

The pulse profiles for the \( C_{12}(0.1) \) and CNOT gates where searched as a whole using the GRAPE method. The starting point for the search was a profile resembling \( \pi \) and \( \pi/2 \) pulses with free evolution periods in between. An example of a final pulse profile is given in Fig. 5-3.

**We used MATLAB’s fminsearch routine on \( 1 - F_g \).
5.4.2 The initial state preparation

There are a total of six two-body collective coefficients for the four carbons of crotonic acid. We chose the pairs [1, 2], [2, 3] and [1, 4]: the pair [1, 2] is the one targeted by the C12 and CNOT gates, while the other two are the pairs involving qubits 1 or 2 with the highest J-coupling. We expect the errors for the three chosen pairs to be larger (due to internal evolution that is not perfectly refocused). The pulse sequence

\[
\frac{\pi}{3} x^2, \frac{\pi}{2} x^3, 4 - \text{grad}_x - \frac{\pi}{4} x^1 - T - \frac{\pi}{4} x^2, 4 - \text{grad}_x
\]

prepares the state \( \rho_0 = |0\rangle \langle 0|_1 \otimes |0\rangle \langle 0|_2 \otimes (I/2)^{(3)} \otimes (I/2)^{(4)} \) for measuring \( \langle |\eta_{1,2}^{\text{col}}|_P \rangle \). The periods denoted by \(-T-\) are periods of free evolution for a time \( T = 1/(8J_{12}) \), while \( \text{grad} \) denotes the implementation of a gradient across the given direction. We simulated the overall pulse sequence (the single pulses were searched individually) varying the time \( T = 1/(8J_{12}) \) looking to maximize the correlation\(^5\) between the outcome of the simulation and the theoretical state, \( C(\rho_{\text{sim}}, \rho_{\text{theo}}) = \text{Tr}[\rho_{\text{sim}} \rho_{\text{theo}}] / \sqrt{\text{Tr}[\rho_{\text{sim}}^2] \text{Tr}[\rho_{\text{theo}}^2]} \). Again, we apply the Delays optimization technique, but this time aiming to optimize a state-to-state transformation.

The sequence above is not unique. For example, many of the pulses can actually be taken along \( x \) or \( y \). In practice, we searched for alternative pulses and used the ones that gave the best simulated correlation.

To prepare the initial state for the other pairs of qubits, we only need to cycle through \( 1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1 \). The net pulse sequence for preparing the pair [1, 2] was 12 msec long,

\(^5\)This correlation function is motivated in [34].
the one for [2,3] was 11 msec and the one for [1,4] 73 msec (substantially longer since the $J_{14}$ coupling is very weak).

Notice that the delay optimization must be performed every time we use a different pulse profile, either because we are trying alternative pulses, or because we are working with a different pair of qubits.

The experimental initial state preparation over the 4 qubits reported a correlation with the theoretical one that was on average 0.98 (0.97 the lowest). The correlation for the targeted qubits (a pair of qubits) was 0.99 in each case. The purity of the state of qubits to be measured was always 1.00 (the constant $\alpha$ in the pseudo-pure state preparation was set to guarantee this). Fig. 5-4 shows an example of the typical correlations for the experimental initial state.

![Graph showing correlation over 60 realizations](image)

Figure 5-4: Initial state correlation (theory vs. experiment), for the pair [1, 2], across different realizations of the same characterization experiment performed across several months. ■: correlation of the subsystem of qubits 1 and 2; ▲: correlation of the subsystem of qubits 3 and 4; ○: correlation of the full carbon system.

### 5.4.3 The twirl

We implemented the twirl of pairs of qubits exactly using 36 Clifford operators. We searched for pulses for each of the 12 Clifford gates for each carbon individually (see the first 12 operators in Table B.1 in Appendix B). From the NMR pulsed point of view, we only need to find six of them, since we can change the direction of a pulse in the transverse plane changing the phase of the transmitter: adding a 90° phase shift taxes $x \rightarrow y$, $y \rightarrow -x$, $-x \rightarrow -y$ and $-y \rightarrow x$. Nevertheless, we searched for 12 different pulses, and chose the one that performed the best experimentally, which we then implemented with phase shifts accordingly.

The simulated fidelities $F_g$ of the chosen pulse profiles found with the SMP method were between 0.98 – 0.99 (except for the pulses for Clifford gates $C_9$ and $C_{12}$ on carbon #1
(the carbonyl), where it was 0.97). On the other hand, the pulses found with GRAPE had $F_g = 0.99 - 1.00$. As a quantitative indication of the experimental performance of a pulse we took the correlation between the expected and actual outcome of the pulse acting on the equilibrium state (or occasionally on another prepared initial state). Table 5.2 below shows the correlations obtained when implementing the best 12 pulse profiles on the thermal equilibrium state (5.16), for each of the qubits (each of the carbons).

<table>
<thead>
<tr>
<th>Clifford gate #</th>
<th>carbon # 1</th>
<th>carbon # 2</th>
<th>carbon # 3</th>
<th>carbon # 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.98*</td>
<td>0.98*</td>
<td>0.98</td>
<td>0.96</td>
</tr>
<tr>
<td>2</td>
<td>0.99*</td>
<td>0.98*</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td>3</td>
<td>0.96</td>
<td>0.97</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>4</td>
<td>0.96</td>
<td>0.99</td>
<td>0.98</td>
<td>0.99</td>
</tr>
<tr>
<td>5</td>
<td>0.98*</td>
<td>0.98*</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td>6</td>
<td>0.97</td>
<td>0.98</td>
<td>0.98</td>
<td>0.99</td>
</tr>
<tr>
<td>7</td>
<td>0.98*</td>
<td>0.99*</td>
<td>0.99</td>
<td>0.97</td>
</tr>
<tr>
<td>8</td>
<td>0.97</td>
<td>0.98</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>9</td>
<td>0.96</td>
<td>0.97</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>10</td>
<td>0.98*</td>
<td>0.98*</td>
<td>0.97</td>
<td>0.99</td>
</tr>
<tr>
<td>11</td>
<td>0.97*</td>
<td>0.98*</td>
<td>0.97</td>
<td>0.99</td>
</tr>
<tr>
<td>12</td>
<td>0.98*</td>
<td>0.97</td>
<td>0.98</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Table 5.2: Experimental performance of the Clifford gates acting on each qubit (each carbon), on the thermal equilibrium state (5.16). As measure of quality, we list the correlation $C(\rho_{\text{exp}}, \rho_{\text{theo}}) = \frac{\text{Tr}[\rho_{\text{exp}}\rho_{\text{theo}}]}{\sqrt{\text{Tr}[\rho_{\text{exp}}^2]\text{Tr}[\rho_{\text{theo}}^2]}}$. The * indicates a GRAPE pulse; the rest are SMP pulses.

These are the pulse profiles we used in our experiments. We find that the typical performance of one-qubit gates on the spectrometer is 1-2% below the average simulated fidelity $F_g$. In particular, both GRAPE and SMP perform similarly in the experiments, although the simulated $F_g$ are in general higher for GRAPE pulses. Also, we see that a same pulse profile applied with a transmitter phase shift returns correlations within 1% difference (in the simulations, a transmitter phase shift does not alter $F_g$). Moreover, taking the same pulse profile and applying it on different experiments (that is, on different days across several months) may result in a change in the correlation of at most 0.01.

We must choose one of the four available 6-element subsets of the 12 Clifford operators for each qubit (see Sec. B.3). If we randomize the state of the system (state-twirl $^1\mathcal{T}$) in $U(D)$ (in particular with $D = 2$), we have

$$^1\mathcal{T}(\rho) = \int R\rho R^t dR = I/2$$

(5.26)

This can be easily proved for one qubit using the expression for the random rotation $R$ given in (4.2); alternatively see Appendix C. We used this as a criterion to choose the best set between set $A = \{C_1, C_2, C_5, C_6, C_9, C_{10}\}$ and set $B = \{C_3, C_4, C_7, C_8, C_{11}, C_{12}\}$: for each of the carbons, we chose the set that best takes the equilibrium state to the $I/2$ state for the qubit being twirled, while leaving the others the most unchanged.
As an example, we show the simulated and experimental performance of the Clifford gates on the carbonyl. With $1^{(j)}_{X}$ denoting the one-qubit state-twirl over the $j$-qubit using the set $X$ of Clifford operators, we had

**simulation**: \[ C(\rho_{\text{theo}}, \rho_{\text{sim}}) = 1.00 \]

\[ 1^{(j)}_{A+B}(\alpha \rho_{\text{th}}^{\text{dev}}) = 0.99 \sigma_z^{(2)} + \sigma_z^{(3)} + \sigma_z^{(4)} - 0.03 \sigma_z^{(1)} \sigma_z^{(2)} + O(0.03) \]

**simulation**: \[ C(\rho_{\text{theo}}, \rho_{\text{sim}}) = 1.00 \]

\[ 1^{(j)}_{A}(\alpha \rho_{\text{th}}^{\text{dev}}) = \sigma_z^{(2)} + \sigma_z^{(3)} + 0.99 \sigma_z^{(4)} - 0.03 \sigma_y^{(2)} \sigma_y^{(3)} + 0.03 \sigma_x^{(2)} \sigma_x^{(3)} + O(0.03) \]

**simulation**: \[ C(\rho_{\text{theo}}, \rho_{\text{sim}}) = 1.00 \]

\[ 1^{(j)}_{B}(\alpha \rho_{\text{th}}^{\text{dev}}) = 0.99 \sigma_z^{(2)} + \sigma_z^{(3)} + \sigma_z^{(4)} - 0.04 \sigma_z^{(1)} \sigma_z^{(2)} + O(0.03) \]

**experiment**: \[ C(\rho_{\text{theo}}, \rho_{\text{exp}}) = 0.99 \]

\[ 1^{(j)}_{A+B}(\alpha \rho_{\text{th}}^{\text{dev}}) = \sigma_z^{(2)} + 0.93 \sigma_z^{(3)} + \sigma_z^{(4)} + 0.09 \sigma_z^{(1)} - 0.09 \sigma_x^{(1)} + O(0.08) \]

**experiment**: \[ C(\rho_{\text{theo}}, \rho_{\text{exp}}) = 0.99 \]

\[ 1^{(j)}_{A}(\alpha \rho_{\text{th}}^{\text{dev}}) = \sigma_z^{(2)} + 0.93 \sigma_z^{(3)} + 0.99 \sigma_z^{(4)} - 0.12 \sigma_y^{(1)} + O(0.1) \]

**experiment**: \[ C(\rho_{\text{theo}}, \rho_{\text{exp}}) = 0.98 \]

\[ 1^{(j)}_{B}(\alpha \rho_{\text{th}}^{\text{dev}}) = 0.98 \sigma_z^{(2)} + 0.93 \sigma_z^{(3)} + \sigma_z^{(4)} + 0.15 \sigma_z^{(1)} + O(0.1) \]

where we chose $\alpha$ so $\sigma_z^{(1)}$ has a coefficient equal to 1 in $\alpha \rho_{\text{th}}^{\text{dev}}$.

Finally, we note that imperfections in the gate design result in $(\sigma_k^{\dagger})_{\text{exp}}(\sigma_k)_{\text{exp}} \neq \mathbb{I}$ in general. This appears already in the simulations, indicating that the flaw is already in the pulse design, not strictly in the experimental implementation. For example, if we simulate $(\sigma_k^{\dagger})_{\text{sim}}(\sigma_k)_{\text{sim}} = \mathbb{I}_{k}^{\text{perfect}}$ (considering the hydrogens but, for simplicity, without rf inhomogeneity), and then calculate the entanglement fidelity of this map, we obtain for example for Clifford gates on the carbonyl: $F_e = 0.99$ for $k = 1, 2, 5, 7$, $F_e = 0.98$ for $k = 11$, $F_e = 0.97$ for $k = 9, 10, 12$ and $F_e = 0.94$ for $k = 3, 4, 6, 8$.

The non-unitary behavior of the experimental gates is shown for example in Fig. 5-5, where we plot the correlation between the experimental initial state and the state after we have also applied $(\sigma_k^{\dagger})_{\text{exp}}(\sigma_k)_{\text{exp}}$ for carbons $\neq 2$ & 3. If the experimental gates were unitary, this correlation would be equal to 1, independently of other imperfections in the gates and/or in the initial state preparation. This could also be used as a criterion to choose the best 6-element subset. For comparison, we also plot the typical correlation between the theoretical and the experimental initial state (which ideally should also be equal to 1). This shows that the drop in the correlation due to the application of the gates is even larger, as it is harder to find a good unitary than a good state-to-state transformation. Moreover, the concatenation of two imperfect gates retrieves an even more imperfect net gate.

It is also interesting to observe that the initial state preparation takes in general much longer than the implementation of one or two gates (GRAPE pulses take 1 msec, and SMP pulses take less in general, typically 0.5 msec). However, the drop in the correlation is larger in general due to the implementation of a pulse. It is not so the case in the pair [1,4], since
the initial state preparation takes indeed seven times longer (see Fig. 5-6).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure5-5.png}
\caption{The \(\Delta\) show \(C(\rho_A, \rho_B)\) when \(\rho_A\) is the experimental initial state for the qubit pair [2,3] and \(\rho_B = (C_1^\dagger)_{\text{exp}}(C_k)_{\text{exp}}\rho_A(C_1^\dagger)_{\text{exp}}(C_k)_{\text{exp}}\). Each \(C_k\) represents a product of the elements in the sets \(B\) for carbons \# 2 & 3 (therefore we have 36 Clifford pairs). For comparison, we have also plotted with \(\bullet\) the correlation between the ideal and the experimental state.}
\end{figure}

\subsection{5.4.4 The measurement}

To perform QST we used a set of \(N_{\text{ro}} = 19\) readout pulses. We used the set given in [71], p. 58\(^4\), plus an extra 19-\textit{th} readout pulse which was simply the identity: we directly measured the observable terms of \(\rho_f\). Notice that in this case the implementation of the readout pulse and its reversal in the data processing (both are “do nothing”) are perfect; that is why we added the data coming from this experiment to each QST procedure.

The number of experiments required to measure one collective coefficient \(\langle |\eta_{a,b}|^2 \rangle\) for a given pair for a particular gate under study was then \(19 \times 36 = 684\), plus 19 experiments to characterize the initial state. In practice, however, the preparation and characterization of the initial state was done every 1 or 2 days as a control to account for variations in the spectrometer (in particular, variations in the shimming). We performed QST of the full system in order to have a broader knowledge of the experimental performance, but this is not required by the protocol; only the target qubits must be measured.

It is difficult to quantify the performance of the readout process alone. We must notice though that all the correlations \(C(\rho_{\text{exp}}, \rho_{\text{theo}})\) we have reported already include measurement errors, since we obtain the experimental \(\rho_{\text{exp}}\) after implementing our QST process.

Part of our experimental work was to improve the “tomography code”, that takes the spectral data from the \(N_{\text{ro}}\) experiments and process it to retrieve a valid density matrix given a suitable constant \(\alpha\). The main improvements were in the fitting of the spectral data and manipulation of the readout pulses.

\(^4\)We actually used \(\pi/2)^{\dagger}\) instead of \(\pi/2)^{\dagger}.\)
Figure 5-6: The ▲ show $C(\rho_A, \rho_B)$ when $\rho_A$ is the experimental initial state for the qubit pair [1,4] and $\rho_B = (C_k^\dagger)_{\exp}(C_k)_{\exp}\rho_A(C_k^\dagger)_{\exp}(C_k)_{\exp}$. Each $C_k$ represents a product of the elements in the set $A$ for carbons $\neq 1$ and set $B$ for carbon $\neq 4$ (therefore we have 36 Clifford pairs). For comparison, we have also plotted with ● the correlation between the ideal and the experimental state.

5.4.5 Negligibility of higher order multi-body terms

In order to apply our protocol on pairs of qubits, the negligibility of three-body and four-body terms must be established a priori. In a simple model, these gates consist basically of periods of free evolution of length $\tau$ (the corresponding propagator is $U_\tau = \exp(-i\mathcal{H}_{\text{int}}\tau)$) separated by $\theta$-pulses on some of the qubits (single qubit rotations). Assuming weak coupling between the $n$ spins, any time-suspension sequence, or gate of the form $\exp(-ia\sigma_z)$ (one-qubit) or $\exp(-ib\sigma_x\sigma_z)$ (two-qubit) in an $n$-qubit system can be constructed in this way [72]. Then any arbitrary one-qubit or two-qubit operator can be obtained by combining them with $\theta$-pulses on the transversal plane [55, 56]. Let’s denote $U_\theta = \exp(-i\mathcal{V})$, $\mathcal{V} = \sum_j(\theta_j/2)\sigma_{p_j}^{(j)}$, with $p_j$ in the transverse plane and $(j)$ indicating qubits.

Using the Baker-Campbell-Hausdorff (BCH) formula [73] we can calculate the structure of the building block $U_\tau U_\theta$,

$$U_\tau U_\theta = \exp \left(-i\mathcal{H}\tau - i\mathcal{V} - \frac{\tau}{2}[\mathcal{H}, \mathcal{V}] + \frac{i}{12} \left(\tau^2[\mathcal{H}, [\mathcal{H}, \mathcal{V}]] + \tau[[\mathcal{H}, \mathcal{V}], \mathcal{V}]\right) + \ldots \right) \tag{5.27}$$

The typical delays $\tau$ are 1–2 msec long, while $\theta \sim \omega t \theta$. A typical pulse takes no more than 1 msec, which makes $\omega$ (the frequency related to the transversal $B_1$ field) of the order of KHz for typical $\theta = \pi, \pi/2, \pi/3$, etc. Thus in order of magnitude we can say $\theta \sim 1$. On the other hand, the weight of the one-body terms in $\mathcal{H}$ is approximately $\omega_3 \tau \sim \pi$ 1 KHz $\times$ 1 msec $\sim 1$ at least, while the weight of the two-body terms is $\pi J\tau \sim \pi$ 70 Hz $\times$ 1 msec $\sim 10^{-1}$ at most. (See values given for (5.13).)

Looking at (5.27): $\mathcal{H}$ has one-body and two-body terms, with weights $\omega_3 \tau \sim 1$ and $\pi J\tau \sim 10^{-1}$, respectively. $\mathcal{V}$ has only one-body terms with weight $\theta \sim 1$. Then $[\mathcal{H}, \mathcal{V}]$ has
one-body terms with weight $\omega_\delta \tau \theta \sim 1$ and two-body with weight $\pi J \tau \theta \sim 10^{-1}$. $[[\mathcal{H}, \mathcal{V}], \mathcal{V}]$ has still only one-body terms with $\omega_\delta \tau \theta^2 \sim 1$ and two-body with weight $\pi J \tau \theta^2 \sim 10^{-1}$. Finally, $[[\mathcal{H}, \mathcal{H}, \mathcal{V}]$, $\mathcal{V}]$ has one-body terms with weight $\omega_\delta^2 \tau^2 \theta \sim 1$, two-body terms with $\omega_\delta^3 \tau \theta \sim 1$, and three-body terms with weight $\pi^2 J^2 \tau^2 \theta \sim 10^{-2}$. This means that in the exponent in (5.27), three-body and higher order terms will appear with a factor that makes them at least ten times smaller with respect to possible one-body and two-body terms. We can then repeat this to combine all the building blocks, to still have any three-body or higher multi-body term with a factor of $10^{-1}$ respect to the others. In this situation, where $\tau$ is small enough so we can regard $J \tau \sim 10^{-1}$, the generation of higher order multi-body terms is then unlikely (although not absolutely forbidden).

On the other hand, the simulation of the engineered pulse sequences used in the experiment showed that all the three-body and four-body terms appear with collective coefficients $\left| \langle \eta_{j, k, j'}^{\text{col}} \rangle \right|^2 = 0.005$ for the CNOT and $C_{12}(0.4)$ gates, and $< 0.002$ for the rest. As we will see next, these are much smaller than the differences between measured and predicted values of $\left| \langle \eta_{j, k}^{\text{col}} \rangle \right|^2$. The latter can be explained as implementation errors in the protocol or genuine one-body and two-body error in the gate implementation arising from imperfect refocusing, rather than arising due to the effect of non-null multi-body terms.

### 5.5 Experimental results

The results on the measurement of the collective coefficients for the qubit pairs $[1,2]$, $[2,3]$ and $[1,4]$ are presented in Table 5.3, where also the theoretical values are displayed. A smaller version of this table was published in [22].

In order to analyze the results obtained, we also performed several simulations (accounting for rf inhomogeneity and the presence of the hydrogens). They are represented in Table 5.3 as follows:

- **simG** Simulation of the Gate: calculation of the propagator corresponding to the gate given by the engineered pulse sequence, and also the simulation of the experiment using this propagator with a perfect initial state and Clifford operators;

- **simETI** Simulation of the whole Experiment, using the Theoretical Initial state: simulation of the experiment simulating the pulse sequences corresponding to both the gate and the Clifford operators, but using a perfect initial state;

- **simESI** Simulation of the whole Experiment, using the Simulated Initial state: simulation of the experiment simulating the pulse sequences corresponding to the gate and the Clifford operators, and also simulating the initial state preparation with pulses starting from the thermal equilibrium state;

- **simEEI** Simulation of the whole Experiment, using the Experimental Initial state: simulation of the experiment with the pulse sequences corresponding to both the gate and the Clifford operators, and using the experimental data for the initial state.

Notice that the largest differences between measured and predicted coefficients appear on the pair $(1,2)$ and on the most complex gates: $CNOT^2$, $CNOT$, $C_{12}(0.4)$. This indicates that these deviations are due to the errors expected from spurious processes in our QIP.
particularly an imperfect refocusing during the gate sequence, rather than from an imperfect implementation of the protocol.

It worth mentioning that typically in NMR QIP the main errors occurring in one-qubit gates are correlated one-body errors (i.e., a one-qubit rotation that is slightly off), and they do not introduce two-body errors, which are the main target of a spatial correlation analysis. This can be noticed also in the fact that the simG values are similar to the simEPI-simESI values (a change of 0.00–0.01, except in two cases where we found 0.02 and 0.03). Moreover, there is a contribution arising from $T_2$ relaxation. However, a calculation over the theoretical propagator and the numerics over the simulated one show a change on the order of 0.01. The little change between simEPI and simESI indicates that the error is mainly in the implementation of the Clifford gates and not in the initial state preparation.

Elements outside our system model, which would explain further the gap between theory and experiment, are $B_0$ (static) field inhomogeneities, the presence of transients and residual non-linearities in the spectrometer circuitry, and an imperfect spectral fitting of the measured signal. These are well-know issues in liquid NMR QIP, whose effect falls within the $1–2\%$ error. We believe also there is a contribution from the measurement process beyond the spectral fitting, since we must use readout pulses that also include errors in their design. These pulses carry an error of similar importance to the ones of the Clifford gates, and that is not included in any of the simulations.

An attempt to include the measurement errors is given in simEEI, where we used as input initial state the data corresponding to the experimentally characterized initial state (which includes then errors induced by the QST). In this case now the coefficients are predicted with a much higher error, sometimes even surpassing the measured ones. This can be understood by considering that there are time-correlations between the different steps of the process. The rf inhomogeneity indeed creates a form of incoherence (cf. Sec. 3.5.2) that correlates one step with the following ones. By using the experimental initial state, we are including the QST errors in it, but at a different stage in the process. So in that sense the result does not reflect the actual process. This is also related to the fact that pulses designed to compensate rf inhomogeneity do not necessarily perform consequently when we concatenate one after the other, as we can see from eq. (5.24) that $\hat{\Gamma}(\mathcal{H}_1^{t_1})\hat{\Gamma}(\mathcal{H}_2^{t_2})$ is not necessarily equal to $\hat{\Gamma}(\mathcal{H}_1^{t_1}+\mathcal{H}_2^{t_2})$.

We must differentiate between the implementation errors in the protocol (initial state preparation, one-qubit twirl and readout), and the errors in the gate under study: the former ones affect the accuracy of protocol. As discussed, rf field inhomogeneities, the presence of hydrogens and $T_2$ relaxation already give an error bar $\sigma_{\eta}^2 \approx 0.03$. There are still other sources of error mentioned above that could make $\sigma_{\eta}^2$ larger, but within that order of magnitude.
| Gate    | \(||n_{1,2}|^2\) | \(||n_{2,3}|^2\) | \(||n_{1,4}|^2\) |
|---------|----------------|----------------|----------------|
| \(I_E\) | \(0.02\)     | \(0.02\)     | \(0.01\)     |
| 12.2 msec | \(0.00\) | \(0.00\) | \(0.00\) |
| \(F_g = 0.96\) | \(0.01\)     | \(0.02\)     | \(0.00\)     |
| \(simG\) | \(0.01\)     | \(0.02\)     | \(0.00\)     |
| \(simEPI\) | \(0.01\)     | \(0.02\)     | \(0.00\)     |
| \(simESI\) | \(0.01\)     | \(0.02\)     | \(0.00\)     |
| \(simEEI\) | \(0.01\)     | \(0.02\)     | \(0.00\)     |
| \(C_{12}(0.1)\) | \(0.02\)     | \(0.02\)     | \(0.01\)     |
| 4.88 msec | \(0.01\) | \(0.00\) | \(0.00\) |
| \(F_g = 0.99\) | \(0.02\)     | \(0.00\)     | \(0.00\)     |
| \(simG\) | \(0.02\)     | \(0.00\)     | \(0.00\)     |
| \(simEPI\) | \(0.02\)     | \(0.00\)     | \(0.00\)     |
| \(simESI\) | \(0.02\)     | \(0.00\)     | \(0.00\)     |
| \(simEEI\) | \(0.02\)     | \(0.00\)     | \(0.00\)     |
| \(C_{12}(0.4)\) | \(0.26\)     | \(0.03\)     | \(0.03\)     |
| 19.52 msec | \(0.15\) | \(0.00\) | \(0.00\) |
| \(F_g = 0.87\) | \(0.23\)     | \(0.01\)     | \(0.00\)     |
| \(simG\) | \(0.23\)     | \(0.01\)     | \(0.00\)     |
| \(simEPI\) | \(0.23\)     | \(0.01\)     | \(0.00\)     |
| \(simESI\) | \(0.24\)     | \(0.02\)     | \(0.02\)     |
| \(simEEI\) | \(0.25\)     | \(0.04\)     | \(0.05\)     |
| \(CNOT\) | \(0.32\)     | \(0.01\)     | \(0.03\)     |
| 11.88 msec | \(0.25\) | \(0.00\) | \(0.00\) |
| \(F_g = 0.99\) | \(0.25\)     | \(0.01\)     | \(0.00\)     |
| \(simG\) | \(0.27\)     | \(0.02\)     | \(0.01\)     |
| \(simEPI\) | \(0.28\)     | \(0.02\)     | \(0.02\)     |
| \(simESI\) | \(0.33\)     | \(0.06\)     | \(0.07\)     |
| \(simEEI\) | \(0.33\)     | \(0.06\)     | \(0.07\)     |
| \(CNOT^2\) | \(0.07\)     | \(0.05\)     | \(0.04\)     |
| 23.76 msec | \(0.00\) | \(0.00\) | \(0.00\) |
| \(F_g = 0.97\) | \(0.01\)     | \(0.01\)     | \(0.00\)     |
| \(simG\) | \(0.01\)     | \(0.02\)     | \(0.00\)     |
| \(simEPI\) | \(0.01\)     | \(0.00\)     | \(0.00\)     |
| \(simESI\) | \(0.01\)     | \(0.02\)     | \(0.00\)     |
| \(simEEI\) | \(0.01\)     | \(0.02\)     | \(0.00\)     |

Table 5.3: Measured (meas) collective coefficients for selected pairs of qubits, for the various
gates studied. theo are theoretical coefficients. The sim values correspond to different numerical
simulations (see text).
Moreover, a fiducial initial state preparation is critical to the success of the algorithm. If we call $\varepsilon_0$ the error in the initial state preparation, and similarly we call $\varepsilon_1$ the error in the implementation of the Clifford gates, an error propagation in the formula for the $\gamma$'s gives $\sigma_2^2 \leq \varepsilon_0^2(1 + 4\gamma) + \varepsilon_1^2$. We can propagate this into the formulas for the $|\eta_2^2|$: for example, for a pair of qubits $a, b$ we follow eq. (5.9) and obtain $\sigma_{\eta^2} = \frac{9}{4} \sqrt{\frac{\sigma_a^2}{\varepsilon_a^2} + \frac{\sigma_b^2}{\varepsilon_b^2} + \frac{\sigma_{ab}^2}{\varepsilon_{ab}^2}}$. These $\varepsilon$'s account for non-statistical errors (typically correlated one-body errors) and set the accuracy of the method. Given the low complexity of initial state preparation and one-qubit operators, these $\varepsilon$'s are smaller than the errors in target operations (a fact reflected, for example, in the gate fidelities). This is why even though the theory was developed for error-free initial state preparation and twirl, the actual implementation can still retrieve information about the target operations.

5.6 Closing remarks

We have presented a method to characterize the spatial correlations occurring in a gate or quantum process under study applied in one step, showing its potential through liquid-state NMR QIP.

To our knowledge, neither QPT nor other proposals so far are able to make use of the negligibility of higher order correlations in order to gain further insight. In general, even after it has been established that a subset of parameters is null, it is not trivial to direct the measurement procedure to target the non-null ones exclusively. Our protocol focuses on this. At the same time, the assumption of negligible high order multi-body terms can be probed with the protocol developed by Emerson et al. in [16]. It is notable that the work presented there requires exactly the same resources, which makes it completely compatible with ours. This will be the topic of the next Chapter, where we will present a comparative and unifying view of the available protocols for quantum process characterization that use the concept of sampling a one-step twirl.

Furthermore, we have analyzed the relevance of implementation errors, showing the need for strategies that are not only scalable but also robust. This has been already pointed out by Knill et al. in [17], where they presented a method that, in principle, does not require error-free stages, allowing for certain types of errors to occur during the computation. Unfortunately it is neither evident how to take their scheme beyond one-qubit QIP (for example, as analyzed in [19]), nor is it clear how the protocol works for general errors in the twirl operations.

Nevertheless, our experimental work has shown that it is possible to extract information about complex dynamics using the relatively simpler one-qubit Clifford gates. In such a scenario it is realistic to assume that the implementation part of the protocol could be accurate enough to reveal information about the process under study without interfering with it. At this point, we notice that there is a clear need for experimental feedback in order to assess the several proposed characterization strategies.
Chapter 6

Comparison between different protocols using twirling

At this point we have established that the complete characterization of a map describing an arbitrary quantum process requires resources that scale exponentially with \( \log(D) \), being \( D \) the dimension of the Hilbert space \( \mathcal{H}_D \) of the quantum system [1]. The recent works by the groups at University of Waterloo [15, 13, 16], University of Buenos Aires [20, 74] and ours (Chapters 4 & 5; [14, 22]) have demonstrated that by twirling the map under study it is possible to extract relevant information in an efficient way - that is, at a cost that scales at most polynomially with \( \log(D) \).

In this Chapter we take the results from [16] and [22] and present them in a new compact form as one method that allows us to study the diagonal elements of the \( \chi \)-matrix characterizing any quantum process in the basis formed by the generalized Pauli operators. This is a re-interpretation of the protocol presented and implemented in Chapter 5 that allows it to be combined with [16] to obtain a more powerful algorithm. Also, this new presentation allows us to compare this strategy to the ones presented in [15, 20, 74], which are protocols that operate on the same principle of twirling but retrieve information differently.

6.1 Re-stating the basics

The concept of twirling a map was presented in Chapter 2, in particular see Fig. 2-1. We have a quantum process characterized by a map \( \Gamma \) that acts on a system, as depicted in Fig. 2-1(a). We twirl the map by applying an operator \( U \) before the map, and \( U^\dagger \) after, as in Fig. 2-1(b). Different families of \( U \)'s will return different types of twirls; the sum over the different \( U \) results in a net map \( \Gamma^T \), the twirled map. In general, the number of elements in these families of operators is infinite or grows exponentially with \( \log(D) \) \[83\] (see also Appendix Sec. B.1).

However, as it was initially suggested in [13, 14] and already introduced in Chapter 2, we can approximate the twirl by randomly sampling over the family of \( U \)'s, say \( N \) times, as depicted in Fig. 2-2. The probability of finding the system in a particular state after the twirling can then be measured efficiently (since \( N \) is independent of the size of the system). We will see that combining the outcome probabilities we can obtain substantial information about the quantum process under study.

This time we describe a general quantum process as the action of an arbitrary map \( \Gamma \) on
the state $\rho$ of an $n$-qubit system in $\mathcal{H}_D$ ($D = 2^n$), by the following $\chi$-matrix representation

$$\Gamma(\rho) = \sum_{l,l'=0}^{D^2-1} \chi_{l,l'} P_l \rho P_{l'} \quad (6.1)$$

where as before the $P_l$ are generalized Pauli operators. Again, $P_0 = \mathbb{I}$ is the identity operator in $\mathcal{H}_D$, and for $l > 0$ the $P_l$ are traceless. For this process to be hermiticity-preserving, the $\chi$-matrix must be an hermitian matrix. This description includes any completely positive map, and it is an alternative to the representations given by eqs. (2.3) or (2.4). However, like eq. (2.4) (and unlike eq. (2.3)), it represents an arbitrary linear hermiticity-preserving map, therefore it can describe any possible open quantum dynamics occurring in the system\footnote{After our discussion in Chapter 3, we are in position to state that the most general open quantum dynamics that may occur is given by a map that takes $\rho_S(0) = \text{Tr}_E[\rho_{SE}(0)]$ to $\rho_S(t) = \text{Tr}_E[U \rho_{SE}(0) U^\dagger]$, where $\rho_{SE}(0)$ is any valid density matrix describing the state of the System + Environment, and $U$ stands for an arbitrary unitary evolution. Shabani and Lidar have shown in [28] that this is equivalent to a linear hermiticity-preserving map acting on the Hilbert space of the system.} [27]. In addition, for this process to be trace preserving, we must have

$$\sum_{l,l'=0}^{D^2-1} \chi_{l,l'} P_l P_{l'} = I \quad \text{which implies} \quad \sum_{l=0}^{D^2-1} \chi_{l,l} = 1.$$ 

This condition reduces the $D^4$ real independent parameters in the $\chi$-matrix to $D^4 - D^2$.

We switch to this description since it will allows us to calculate the effect of a one-qubit twirl very straightforwardly. Moreover, this representation (which was first used in this context in [20]) will retrieve a new interpretation of what is the information we are extracting from the system (Sec. 6.4).

We now introduce a way of labeling the $P_l$ in more detail. Each index $l$ will carry the following information: $w$, $\nu_w$, $i_w$. $w$ is the Pauli weight of $P_l$, that is, how many of the factors in $P_l$ are nonidentity. The index $\nu_w \in \{1, \ldots, \binom{n}{w}\}$ counts the number of distinct ways that $w$ Pauli matrices can be distributed over the $n$ factor spaces. The index $i_w$ is a vector of length $w$ of the form $i_w = (i_{1w}, i_{2w}, \ldots, i_{nw})$ with each component being $1 = x$, $2 = y$ or $3 = z$ to denote which Pauli matrix occupies that respective factor position in the tensor product forming $P_l$. There are $3^w$ of these $i_w$ for given $w$ and $\nu_w$.

After a Pauli twirl (PT), the map $\Gamma$ becomes

$$\Gamma^{\text{PT}}(\rho) := \frac{1}{D^2} \sum_{m=0}^{D^2-1} P_m \Gamma(P_m \rho P_m) P_m = \frac{1}{D^2} \sum_{m=0}^{D^2-1} \sum_{l,l'} \chi_{l,l'} P_m P_l \rho P_m P_l \rho P_m \quad (6.2)$$

$$= \sum_{l=0}^{D^2-1} \chi_{l,l} P_l \rho P_l \quad (6.3)$$

This result was proven in [85]. It can be also seen as follows: for $l = l'$, $P_m P_l \rho P_m P_l = P_l \rho P_l$ since each $P_l$ either commutes or anti-commutes with each $P_m$. And if $l \neq l'$, for each $j$-th factor in which they differ, we have $P_m^{(j)} P_l^{(j)} P_l^{(j)} P_m^{(j)} = P_l^{(j)} P_m^{(j)}$, with each sign happening for half of the four possible $P_m^{(j)}$. Thus they cancel out in the sum.
We consider now a Symplectic one-qubit twirl (S1T) of the form
\[
\Gamma^{S1T}(\rho) := \frac{1}{3^n} \sum_{m=1}^{3^n} S_m^\dagger \Gamma (S_m \rho S_m^\dagger) S_m \quad \text{with} \quad S_m = \bigotimes_{j=1}^{n} S_m^{(j)}
\]
(6.4)
where each \( S_m^{(j)} \) is an element of the set of Symplectic operators given by \( \{ \exp(-i(\pi/4)\sigma_p), \ p = x, y, z \} \).

It is straightforward to show that
\[
\frac{1}{3} \sum_{m=1}^{3} S_m^{(j)} \sigma_j S_m^\dagger \rho S_m^\dagger \sigma_j S_m = \frac{\sigma_x \rho \sigma_x + \sigma_y \rho \sigma_y + \sigma_z \rho \sigma_z}{3}
\]
so after a Clifford (Pauli + Symplectic) one-qubit twirl (C1T)\(^\dagger\) we get
\[
\Gamma^{C1T}(\rho) = \frac{1}{3^n} \sum_{m=0}^{3^n} S_m^\dagger \Gamma^{PT} (S_m \rho S_m^\dagger) S_m = \sum_{w=0}^{3^n} \sum_{\nu_w} \chi^{col}_{w,\nu_w} \left( \sum_{i_w} P_{w,\nu_w,i_w} \rho P_{w,\nu_w,i_w} \right)
\]
(6.5)
where the collective coefficients \( \chi^{col}_{w,\nu_w} \) are just the diagonal \( \chi \)-matrix coefficients \( \chi_{i,i} \), relabeled \( \chi_{w,\nu_w,i_w} \), after disregarding (averaging over) the information given by \( i_w \):
\[
\chi^{col}_{w,\nu_w} = \sum_{i_w} \chi_{w,\nu_w,i_w}
\]
(6.6)
This is so far what was presented in [16], which can also be proven as in [22] using a different set of tools [14, 94, 95, 96] after establishing the connection between the Clifford twirl and a Haar twirl [15].

The circuit representation of this procedure, presented in Fig. 6-1, is a modification of Fig. 5-1, which was what we used for the experimental determination of spatial correlations in Chapter 5. The key now is to twirl over the full space, and prepare an initial state where all the qubits are simply in a pure state. The spatial correlations are of course the two-body collective \( \chi \)-coefficients.

\[\rho_0 = |0\rangle \otimes^n\]

Figure 6-1: Circuit representation of the general one-step protocol with one-qubit twirling. The \( C \)
denote one-qubit Clifford operators randomly drawn from a pool of 12 elements (see text).

\(^{\dagger}\)The set of twelve operators \( \{ e^{-i\pi\sigma_q/4}, e^{-i\pi\sigma_q/4}\sigma_p \} \) with \( p, q = x, y, z \) generates only half of the Clifford group for one qubit, up to a global phase. Nevertheless, this set is enough to implement the Clifford twirl we want. See Appendix B for details on the Clifford group of operators.
6.2 Extracting the $\chi$-coefficients

Consider the Zeeman state basis $|\bar{v}_h\rangle$, where $\bar{v}_h$ is a boolean vector of length $n$ and Hamming weight\footnote{The Hamming distance between two bit-strings ("codewords of $n$ bits") like $\bar{v}_h$ and $\bar{v}_h'$ is defined to be the number of places at which $\bar{v}_h$ and $\bar{v}_h'$ differ. The Hamming weight $h$ of a bit-string $\bar{v}_h$ is defined to be the distance from the string of all zeros, that is, the number of places at which $\bar{v}_h$ is non-zero ([2] p. 448).} $h$ that labels the $D = 2^n$ states as binary numbers.

The first result we can obtain is that the fidelity of a state $|\bar{v}_h\rangle$ undergoing this transformation is independent of the actual state,

$$
f(\Gamma^{C1T}, |\bar{v}_h\rangle) = \text{Tr}[|\bar{v}_h\rangle\langle\bar{v}_h|\Gamma^{C1T}(|\bar{v}_h\rangle\langle\bar{v}_h|)] = \sum_{w=0}^{n} \sum_{\nu_w} \frac{\chi_{\nu_w}}{3^w} \left( \sum_{i_w} 3^w \langle 0|P_{w,\nu_w,1_w}|0\rangle^2 \right)
$$

To obtain the first expression of the second line, we only need to realize that any computational state $|\bar{v}_h\rangle$ is a result of applying a Pauli operator $P_X^{\nu_h}$ (that has $\sigma_x$ where $\bar{v}_h$ has ones, and identity factors otherwise) to $|0\rangle$. This $P_X^{\nu_h}$ will either commute or anti-commute with $P_{w,\nu_w,1_w}$ (and the $\pm$ will be absorbed by the modulus squared). The last equality is obtained by realizing that the only nonidentity $P_{w,\nu_w,1_w}$ that takes $|0\rangle$ back to it (up to a global phase) is Pauli operator that has $\sigma_z$ in all the positions indicated by $\nu_w$ (thus only one of all the possible $i_w$ given $\nu_w$ and $w$).

We must notice that $f(\Gamma^{C1T}, |\bar{v}_h\rangle)$ is not the average fidelity of the process under study, namely $F(\Gamma) = (Dx_{0,0} + 1)/(D + 1)$. For that to happen, as studied in [13, 15, 20] and stated in eqs. (2.10)-(2.11), the Clifford twirl should have been over the full space $U(D)$, and not on the one-qubit space as we used, $U(2)^\otimes n$. However, the use of this weaker twirl allows us to gain a more detailed insight into the map structure. The first result we point out, presented in [16], is that we can obtain the diagonal elements of the $\chi$-matrix grouped by Pauli weight

$$
p_w = \sum_{\nu_w} \sum_{i_w} \chi_{w,\nu_w,i_w} = \sum_{\nu_w} \chi_{\nu_w}^{col}
$$

The $p_w$ relate to the probability of obtaining any state $|\bar{v}_h\rangle$ with Hamming weight $h$ when measuring the final state $\Gamma^{C1T}(|0\rangle\langle0|)$. We have

$$
\text{Pr}[\bar{v}_h, h] = \text{Tr}[|\bar{v}_h\rangle\langle\bar{v}_h|\Gamma^{C1T}(|0\rangle\langle0|)] = \sum_{w=0}^{n} \sum_{\nu_w} \frac{\chi_{\nu_w}^{col}}{3^w} \left( \sum_{i_w} 3^w \langle 0|P_{w,\nu_w,1_w}|0\rangle^2 \right)
$$

For $\langle 0|P_{w,\nu_w,1_w}|\bar{v}_h\rangle$ to be nonzero (i.e., $\pm 1$), $\nu_w$ must indicate nonidentity factors at least where there are ones in $\bar{v}_h$ (so it must be $w \geq h$). Also the $i_j$ in $i_w$ must be $1 = x$ or $2 = y$ for the qubits with ones in $\bar{v}_h$, and $3 = z$ for the $w - h$ qubits that have zeros in $\bar{v}_h$ but have a nonidentity factor $P_{w,\nu_w,1_w}$. There will be exactly $2^h$ of these operators for given $w$ and
\[ \Pr [\bar{v}_h, h] = \sum_{w=h}^n \sum_{\nu_w = 1}^{(n-h)_w} \frac{2h}{3w} \chi_{w, \nu_h + \nu_w} \]

where \( \nu_h \) indicates a \( \chi_{w, \nu_w}^\text{col} \) for Pauli operators that have a nonidentity factor for at least all the qubits whose corresponding component in \( \bar{v}_h \) is a one. \( \nu_w^n \) labels the \( \binom{n-h}{w-h} \) coefficients with \( w \geq h \) that fulfill this condition. If we now discard the “which qubit” information given by \( \bar{v}_h \), summing over all the \( \binom{n}{h} \) possibilities,

\[ \Pr [h] = \sum_{\bar{v}_h} \Pr [\bar{v}_h, h] = \sum_{w=h}^n \frac{2h}{3w} \left( \frac{w}{h} \right) \sum_{\nu_w = 1}^{(n)_w} \chi_{w, \nu_w} \]

\[ = \sum_{w=h}^n \frac{2h}{3w} \left( \frac{w}{h} \right) \sum_{\nu_w = 1}^{(n)_w} \chi_{w, \nu_w} = \sum_{w=h}^n \frac{2h}{3w} \left( \frac{w}{h} \right) p_w \]

In this way, all the \( p_w \) are related to the probabilities of measuring an outcome with Hamming weight \( h \) by a \( n \times n \) matrix \( R_{h,w} = \frac{2^n}{3^n} \binom{w}{h} \), as stated in [16].

We can also keep the “which qubit” information and use the probabilities \( \Pr [\bar{v}_h, h] \) constructively to gain even more detail. Let’s replace the descriptors \( w \) and \( \nu_w \) by \( \bar{v}_w \), a boolean vector of length \( n \) and Hamming weight \( w \) characterizing a Pauli operator \( P_l \). \( \bar{v}_w \) has a zero in the \( j \)-th position if and only if \( P_l^{(j)} = 1 \), otherwise it has a one. For example, the operator \( \sigma_z^{(1)} \sigma_x^{(3)} \) for \( n = 4 \) qubits has \( \bar{v}_2 = (1, 0, 1, 0) \). There are of course \( \sum_{w=0}^n \binom{n}{w} = 2^n = D \) of these vectors describing the \( P_l \). This is the same type of vector we used to describe the computational states above.

If we start with the probability of having all the qubits flipped in the outcome, and go backwards towards the survival probability (i.e., none of the qubits flipped), we find

\[ \Pr [n] = \frac{2^n}{3^n} \chi_{\bar{v}_n} \]

\[ \Pr [\bar{v}_{n-1}, n - 1] = \frac{2^{n-1}}{3^{n-1}} \chi_{\bar{v}_{n-1}} + \frac{2^{n-1}}{3^n} \chi_{\bar{v}_n} \]

\[ \Pr [\bar{v}_{n-2}, n - 2] = \frac{2^{n-2}}{3^{n-2}} \chi_{\bar{v}_{n-2}} + \sum_{\bar{v}_{n-1}} \frac{2^{n-2}}{3^{n-1}} \chi_{\bar{v}_{n-1}} + \frac{2^{n-2}}{3^n} \chi_{\bar{v}_n} \]

... etc.

So essentially we could determine \( \chi_{\bar{v}_n}^{\text{col}} \) using (6.11a), then insert it in (6.11b) and obtain the \( n \) possible \( \chi_{\bar{v}_n}^{\text{col}} \) from the different \( \Pr [\bar{v}_{n-1}, n - 1] \). And then insert that in (6.11c) and so on and so forth. These equations define a triangular matrix that relates the probabilities \( \Pr [\bar{v}_h, h] \) to the collective coefficients \( \chi_{\bar{v}_w}^{\text{col}} \). Notice there is no need to perform different experiments to obtain the different probabilities: we only need to implement \( N \) realizations of the twirl and keep the outcome of the measurement for each of the realizations. This outcome should be a \( n \)-bit string indicating whether each \( j \)-th qubit was found in \( |0\rangle_j \) or \( |1\rangle_j \).
The problem arises not in obtaining the experimental information, but in its posterior processing. The matrix given by eqs. (6.11) is of size $D \times D$, therefore the cost of the processing would scale exponentially in $n$. For this strategy to work, it is key to relate it hierarchically to the determination of the $p_w$: the experimental information required is the same and can be obtained efficiently by sampling. The idea goes as follows. If we are analyzing a map $\Gamma$ that is close to the identity (a noise channel) or a quantum gate involving a few qubits (typically one or two), then we would expect that above a certain cut-off Pauli weight $w_{co}$, the $p_w$ will be null. This is a reasonable expectation: since $\sum_{w=0} p_w = 1$ (the trace-preserving condition), the $p_w$ cannot all be arbitrarily large, and thus it will be possible to bound the coefficients above the cut-off by a negligible amount. In this scenario, the matrix relating the $\text{Pr}[\bar{\psi}_h, h]$ with the $\chi^{co}_{w}$ will have a size $M_{co} \times M_{co}$, $M_{co} = \sum_{m=0}^{n} (\binom{n}{m})$, which scales polynomially in $n$.\footnote{For example, for $w_{co} < n/2$, it is trivial to prove that $M_{co} \leq (w_{co} + 1)(ne/w_{co})^{w_{co}}$. In T. Worsch, Tech. Rep. IB 31/34, Universität Karlsruhe (1994), it is shown that $M_{co} \leq (1 + \epsilon)(\sqrt{\pi}/4\pi)(ne/w_{co})^{w_{co}}$, with $\epsilon \in O(1/n)$. For a simple case where we find only up to two-body coefficients, then trivially $M_{co} = 1 + n + n(n-1)/2$.}

There is a second caveat though. As explained in [16, 22] respectively, the errors in determining the $p_w$ or the $\chi_{w,p_w}^{co}$ scale inefficiently with $w$, a consequence of the matrices relating them with the corresponding probabilities (eqs. (6.10) and (6.11) respectively). Although the measured probabilities will have a standard error $\leq 1/\sqrt{N}$, this error will propagate into the $p_w$ or the $\chi_{w,p_w}^{co}$ with a factor that grows polynomially with $n$ but exponentially with $w$. Again, we must resort on neglecting the $p_w$ after a certain cut-off. The system can be arbitrary large (arbitrary $n$), but we must have that the $p_w$ are negligible above a certain $w_{co}$.

In order to put in practice this two-fold strategy, we must be able to prepare the initial state $|0\rangle$, measure the probabilities $\text{Pr}[|\bar{\psi}_h\rangle] = \text{Tr}[|\bar{\psi}_h\rangle\langle\bar{\psi}_h|\Gamma^{C1T}(|0\rangle\langle0|)]$ and implement the Clifford one-qubit twirl together with the map $\Gamma$ under study. For the twirl, we must be able to implement the 12 gates arising from the combination of $\{I, \sigma_p\}$ and $\{\exp(-i(\pi/4)\sigma_p)\}$, with $p = x, y, z$, for each qubit.

### 6.3 Comparison with the other methods: twirling in $U(2^n)$ vs. twirling in $U(D)$

Now we want to compare this strategy with the other methods that also use the idea of approximate twirls. The works in [15, 20, 74] utilize a full twirl over $U(D)$, and they show that if the twirl depicted in Fig. 2-2 was over $U(D)$, the survival probability would be the average fidelity of the original map $\Gamma$. The work in [15] also dealt with approximating the twirl not by sampling but by an approximate 2-design.

The comprehensive work in [20, 74] is actually presented not in terms of sampling and averaging over twirl operators but rather over the states of mutually unbiased bases (MUBs): $\{|\psi_{J,m}\rangle, J = 0, \ldots, D; m = 1, \ldots, D\}$. Both these approaches are known to be equivalent for
dimensions that are powers of prime numbers, as it is the case with $D = 2^n$:

$$
\langle \psi | \Gamma^{CDT}(\psi) \langle \psi | \psi \rangle = \frac{1}{|C|} \sum_{l=1}^{\left|C\right|} \langle \psi | C_l \Gamma(C_l | \psi \rangle \langle C_l | \psi \rangle C_l | \psi \rangle
$$

$$
= \frac{1}{D(D+1)} \sum_{J,m} \langle \psi_J,m | \Gamma(| \psi_J,m \rangle \langle \psi_J,m | \psi_J,m \rangle
$$

(6.12)

where the $C_l$ are now the Clifford operators over $U(D)$ and $| \psi \rangle$ is an arbitrary pure state. Both these strategies imply the same cost, as preparing MUB states starting from the computational basis and implementing the $C_l$ require the same resources: $O(n^2)$ one-qubit and two-qubit gates [85, 74]. And again, the number of Clifford operators $|C|$ scales exponentially with $\log(D)$, as does the number of MUB states.

In [20, 74] it was shown how to selectively measure any diagonal $\chi$ element. There is an equivalent to this using a Clifford twirl in $U(D)$. If we implement an extra gate $P_m$ before completing the twirl, as shown in Fig. 6-2, the survival probability is

$$
\text{Tr}[|0\rangle \langle 0 | \Gamma^{CDT}_m (0 \langle 0 |)] = \frac{D\chi_{m,m} - 1}{D + 1}
$$

(6.13)

Figure 6-2: Circuit representation of the action of a map $\Gamma_m$ with $\Gamma_m(\rho) = P_m \Gamma(\rho) P_m$, twirled by $U$.

There are a number of ways to prove this. For example, using the equivalence (6.12) and the references in [20]. But it can be proven independently using the equivalence between the Clifford twirl and the Haar twirl [15] and then the tools in [94, 95, 96], which can be summarized as

$$
\frac{1}{|C|} \sum_{l=1}^{\left|C\right|} \text{Tr}[A_1 C_l B_1 C_l A_2 C_l B_2 C_l] = \frac{\text{Tr}[A_1 A_2]}{D^2 - 1} \left( \frac{\text{Tr}[B_1]}{D} \text{Tr}[B_2] - \frac{\text{Tr}[B_1 B_2]}{D} \right)
$$

$$
+ \frac{\text{Tr}[A_1]}{D^2 - 1} \left( \frac{\text{Tr}[B_1]}{D} \text{Tr}[B_2] - \frac{\text{Tr}[B_1 B_2]}{D} \right)
$$

(6.14)

for any operators $A_1$, $A_2$, $B_1$, $B_2$ in $\mathcal{H}_D$. With $A_1 = |0\rangle \langle 0|$ and $A_2 = |\tilde{v}_h\rangle \langle \tilde{v}_h|$, this formula shows that with a full twirl nothing can be gained by measuring transition probabilities between states that is not already obtained from survival probabilities.

Since the resources required in this case (specifically a twirl in $U(D)$) are more demanding than the ones required for the previously proposed strategy using a one-qubit twirl, we may ask what are the advantages then. The main feature of this approach is that the outcome of the measurement (the survival probabilities (6.13)) are the $\chi_{\nu,\nu,\nu}$ directly. There is no intermediate matrix relating these and the probabilities, as in eqs. (6.10) and (6.11). The error of the $\chi$ coefficients is the error in the measurement, and the problem of an error that grows exponentially with $w$ disappears.

At this point we are able to measure efficiently one $\chi_{11}$ at a time (selective quantum process tomography [20]). We can nevertheless sample the $\chi_{11}$ as we sample the Clifford operators in the twirl, by randomly choosing the $P_m$. Moreover, in this situation we can skip
the implementation of the $P_m$ altogether if we keep track of the bit-string $\bar{v}_h$ we measure. Since for any state $|\bar{v}_h\rangle = P_{X}^{\bar{v}_h}|0\rangle$, the $k$-th realization of the twist in $U(D)$ as in Fig. 2-2 calling to a Clifford operator $C_k$ can also be interpreted as the $k$-th realization of the twist in $U(D)$ as in Fig. 6-2 calling also to a $P_m = C_k P_{X}^{\bar{v}_h} C_k^\dagger$:

$$\text{Tr}[|\bar{v}_h\rangle \langle \bar{v}_h| \Gamma^{CDT} (|0\rangle \langle 0|)] = \text{Tr}[|0\rangle \langle 0| P_X^{\bar{v}_h} \Gamma^{CDT} (|0\rangle \langle 0|) P_X^{\bar{v}_h}]$$

$$= \frac{1}{|C|} \sum_{t=1}^{|C|} \text{Tr} \left[ |0\rangle \langle 0| C_t P_X^{\bar{v}_h} C_t^\dagger \Gamma(C_t|0\rangle \langle 0| C_t^\dagger)(C_t P_X^{\bar{v}_h} C_t^\dagger)C_t \right]$$

And thanks to the Gottesman-Knill theorem, we can calculate $P_m$ efficiently on a classical computer. We believe that this sampling strategy was similarly suggested in [20, 74], in terms of preparing and measuring different MUB states.

Probably the most decisive point when considering the pros and cons of twirling in $U(2)^\otimes n$ or in $U(D)$ is the degree of control we can already achieve in the system at hand. If we have a functional quantum device that implements the Hadamard, Phase and CNOT gates (which are the gates required to implement the Clifford gates [85] or the MUB states [74]; see also Appendix B) and the Pauli operators with enough accuracy, we will be in a position to study more complex maps, in particular, unknown dynamics, with these protocols in $U(D)$. If however we are still aiming to study gates and sequences whose complexity is comparable to the one of a Clifford gate in $U(2^n)$, then the protocol using one-qubit twirls is more suited for the task, assuming those resources are at hand with enough accuracy. An example of such scenario is what we have shown with the experimental in Chapter 5 (also [22]).

### 6.4 What information are we extracting?

We notice that the $\chi$-matrix itself, eq. (6.1), is not a superoperator matrix, that is, its action on a state is not the action of this $D^2 \times D^2$ matrix on a columnized density matrix $|\rho\rangle$. Rather, when the $\chi$-matrix is diagonal in some basis, it gives then the weights of an operator-sum representation, where the operators in the sum are the corresponding basis where the $\chi$-matrix is diagonal. Of course, this basis won’t necessarily be the Pauli operator basis, but in principle a combination of them. Say $\chi^d$ is the diagonalized $\chi$-matrix in the Pauli operator basis\footnote{The $\chi$-matrix will be diagonalizable with real eigenvalues since we want the process to be hermiticity-preserving (so the $\chi$-matrix in the Pauli operator basis is hermitian).}, and let $\beta$ be the change of basis so $\chi_{l,l'} = \sum_{m} \beta_{l,m} \chi^d_{m,m}(\beta^\dagger)_{m,l'}$. Then

$$\Gamma(\rho) = \sum_{m=0}^{D^2-1} \chi_{m,m}^d B_m \rho B_m^\dagger \quad B_m = \sum_{l=0}^{D^2-1} \beta_{l,m} P_l$$

where the $B_m$, just as in an operator-sum representation, are not necessarily unitary (otherwise any process would be unit), nor hermitian nor orthogonal. Thus in general neither the $\chi_{l,l'}$ nor even the $\chi^d_{m,m}$ have a simple interpretation. When the $\chi^d_{m,m}$ are non-negative, we have a CP map; however, they are allowed to be negative in principle (cf. the description presented in [27]).

However, despite the different ways of describing the process under study $\Gamma$ in [14, 15, 16,
20, 22], in all the cases they only target the diagonal elements of the \(\chi\)-matrix of the map in the generalized Pauli operator basis. We ask then what is the meaning of these quantities. It was assumed in [16] that the \(p_w\) represented the probability of an operator of Pauli weight \(w\) happening in the process described by \(\Gamma\). In [22], the \(\chi^{col}_{lw,\nu_w}\) were regarded as indicators of the locality or range of the process, that is, the probability of an operator involving the qubits in \(\nu_w\) happening. These are both quantities that are relevant to quantum error correction and fault-tolerant quantum computing.

Both these interpretations are fair when the \(\chi\)-matrix in the Pauli operator basis is approximately diagonal, at least block-diagonal in blocks characterized by \(w, \nu_w\). This assumes that the Pauli operator basis has a privileged role in the physics of the quantum processes we want to study. The off-diagonal information is completely lost. If there is any information in the off-diagonals that is not somewhat represented also in the diagonal terms, it will not show up in our characterization and the previous interpretation of \(p_w\) and \(\chi^{col}_{lw,\nu_w}\) is arguable: we could have in principle a process involving a set of qubits given by \(w, \nu_w\) that has \(\chi_{ll'} \neq 0\) in that block but \(\chi^{col}_{lw,\nu_w} = 0\) in the diagonal. Nevertheless, if we consider the relation between the \(\chi\)-coefficients and the \(\eta\)-coefficients, we have

\[
\chi_{ll'} = \int P(\bar{\eta}) \, \eta_{l} \, \eta^{*}_{l'} \, d\bar{\eta} = \langle \eta_{l} \, \eta^{*}_{l'} \rangle_{P}
\]

(6.15)

Thus in general for a process that has \(\eta_{l} \eta^{*}_{l'} \neq 0\), then we expect this to also manifest in the diagonal \(|\eta_{l}|^{2}\). Rigorously, if \(P\) is a non-negative distribution, so \(P = \sqrt{P} \, \sqrt{P}\), we can apply the Cauchy-Schwartz inequality for complex-valued functions to obtain

\[
|\langle \eta_{l} \, \eta^{*}_{l'} \rangle_{P}|^{2} \leq \langle |\eta_{l}|^{2} \rangle_{P} \, \langle |\eta_{l'}|^{2} \rangle_{P}
\]

(6.16)

So for CP maps we can guarantee that if \(\chi^{col}_{lw,\nu_w} = 0\), any off-diagonal \(\chi_{ll'}\) for which \(l\) or \(l'\) are associated with those \(w, \nu_w\) is null. Nevertheless, for the general case of \(P\) real, not necessarily non-negative, the average over \(P(\bar{\eta})\) could in principle conspire and take the main features of the process to the off-diagonal terms.

In any case, accessing the diagonal elements provides further insight into the \(\chi\)-matrix beyond the average fidelity of the map (a concept that is quite useless when analyzing a map that is not a noisy channel, or a gate plus its inverse).

It is in order here to point out though the work in [20, 74] also presents a strategy to measure the off-diagonal elements of the \(\chi\)-matrix. However, an errorless qubit ancilla is required for this task, which puts this method in a different category regarding resources and assumptions when it comes to its implementation.

Summarizing, we have presented in this Chapter a compact method that unifies the protocols in [16] and the previous Chapter (also [22]) to measure the diagonal elements of the \(\chi\)-matrix grouped by how many \((w)\) and which qubits \((\nu_w)\) are being affected. Assuming that the \(\chi_{lw,\nu_w,\nu_w}\) are negligible above some cut-off Pauli weight, this strategy efficiently provides further insight into the characterization of any quantum process (not only noisy channels) beyond its average fidelity, and it does so with a minimal set of resources: the sampling of the Clifford twirl in \(U(2)^{\otimes n}\) (utilizing \(12 \times n\) one-qubit gates), and the preparation and measurement of the computational states. The protocols in [15, 20, 74], which we have shown can be casted in terms of sampling of a Clifford twirl too but this time over \(U(2^{n})\), offer an attractive alternative for extracting the diagonal elements of the \(\chi\)-matrix directly
one by one, but this time requiring the implementation of $O(n^2)$ one-qubit operations and the CNOT gates between all the pairs of qubits.
Chapter 7

Generalized fidelity decays for many steps

7.1 Generalized fidelity decays for many steps under progressive randomization

Before moving to a many-step fidelity decay, let’s summarize important features of the one-step fidelity decay under randomization, that led to our interest in a many-step scheme using Clifford gates.

As we have pointed out before, two key features that enable the scalability of the protocols addressed so far are: the sampling of the twirl of the map under study, and the scalable implementation of each element of the sampling ensemble. Considering the unitary operator-twirl approach [13, 14, 15, 16, 22] (as opposed to the nevertheless equivalent state-twirl approach of [20, 74]), we combine eqs. (2.12) and (2.13) to express the first point as

\[ F(\Gamma) = \int dU \langle \langle \psi_0 | \hat{U}^{-1} \hat{\Gamma} \hat{U} | \psi_0 \rangle \rangle \rightarrow \sum_s \langle \langle \psi_0 | \hat{U}_s^{-1} \hat{\Gamma} \hat{U}_s | \psi_0 \rangle \rangle \]  \hspace{1cm} (7.1)

where \( \sum \) represents a twirl that in theory is perfect (so the sum runs over the entire set of twirl operators - which is what we use for our theoretical calculations), but that in practice will be the result of a sampling approach to approximate the twirl with \( N \) realizations (which is what we use in simulations or experiments). The \( \hat{U}_s \) are either random rotations invariant under the Haar measure, or Clifford gates (or, eventually, any other set of operators that form a 2-design; see Appendix B). Typically, the state \( |\psi_0\rangle \) will be one of the computational states\(^5\), and the fidelity measurement at the end will be a projection onto \( |\psi_0\rangle \).

The second point depends specifically on the set of operations \( \hat{U}_s \). An individual Clifford gate can be efficiently implemented using \( O(n^2) \) one-qubit and two-qubit gates [57], thus granting the scalable implementation of each realization. On the other hand, random rotations cannot be implemented efficiently, so beyond a one-qubit twirl we must resort to the use of pseudo-random rotations [39, 40, 41]. Beyond experimental practicalities that may make one set more convenient than the other, the question of whether one strategy

\(^5\)The computational basis is just the Zeeman basis (the \( 2^n \) tensor products of the eigenstates of \( \sigma_z \) for one qubit), although it is usually labeled continuously as \( \{ |q\rangle, q = 0, 1, \ldots, 2^n - 1 \} \), where \( q \) is the decimal representation of the binary string that typically labels the Zeeman basis.
could lead to a faster or better convergence remains open. However, once we consider the strong measurement at the end to determine the fidelity decay, we know in any case that the Chernoff bound bounds the error probability, and the Central Limit Theorem guarantees the convergence to a Gaussian distribution with the sought mean and a variance of \( \leq 1/N \) (see Appendix Sec. A.4).

In retrospective, we can see that a clear breakthrough in scalable process characterization was to use this approach to estimate the fidelities \( F(\Gamma) \), \( F_c(\Gamma) \), for \( \Gamma \) being an arbitrary map acting in between any twirl operation, after preparing the system to be initially in a computational state \( |\psi_0\rangle \). Consequently, a couple of questions immediately arise:

- what is the effect of a faulty initial state preparation and measurement?
- what is the effect of a flawed twirl, that is, errors when implementing the \( U_s \)?
- what if we want to characterize \( \Gamma \) beyond its fidelity?
- what if we want to characterize maps beyond noisy channels, that is, maps that are not expected to be the identity map?

The last two questions were addressed in Chapters 4, 5 and 6 and in [16, 20, 22], which contain all scalable one-step protocols aiming at obtaining and discriminating between the diagonal elements of the \( \chi \)-matrix of \( \Gamma \) in the Pauli group basis.

The questions on using imperfect initial states, imperfect twirl operators \( U_s \) and imperfect measurement were put forward by Knill et al. [17], and remain open—as further discussed in [19]. It was suggested that by repeating the one step fidelity decay, the resulting decay over time would be robust against errors in the initial state preparation and measurement, while the decay strength would provide information on the errors in the twirl operations.

We introduced the idea of a many-step fidelity decay in Sec. 2.4 and worked with it in Chapter 4. We depict it here in Fig. 7-1 again differentiating two possible schemes. In Fig. 7-1(a) the twirl is repeated \( t \) times producing \( \rho_{RT}(t) \) as output, giving us the RT-fidelity \( f^{RT} = \langle \langle \psi_0 | \rho_{RT}(t) \rangle \rangle \) as the outcome of the measurement, where \( RT \) stands for repeated twirl.

In Fig. 7-1(b) we have a sequence in a Loschmidt echo fashion, where we try to refoCUS (reverse) a string of operations, producing \( \rho_{LE}(t) \). This gives the LE-fidelity \( f^{LE} = \langle \langle \psi_0 | \rho_{LE}(t) \rangle \rangle \), where LE stands for Loschmidt echo. In the LE scheme, the reversal gates can be grouped together to form a single reversal gate \( U^R \) that is given by the previous \( t \) twirling gates in that realization, indicated by \( \bar{s} = s_1, \ldots, s_t \). This compression of the reversal motion was first presented in [17]. In that paper and the work which followed it [19, 9], the LE scheme was presented exclusively using Clifford gates, intercalated between Pauli operators. Since the Pauli operators are also Clifford gates, our impression is that such a step is redundant.

The idea is that in the presence of flawed \( U_s \), we will be analyzing the errors from the target map \( \Gamma \) combined with the errors arising from them. If we want to analyze the errors in the \( U_s \) exclusively, then we just don’t apply any intermediate map \( \Gamma \). We will now concentrate on the study of an error map \( A \) arising exclusively from the faulty implementation of the \( U_s \) (\( \Gamma = I \)). The immediate relevance of this is that we would like our twirling gates \( U_s \) to be nearly perfect before attempting to characterize other gates or processes. Once we have a robust twirl, we can then move to the study of an arbitrary map.
Figure 7-1: Circuit representation of two schemes for a fidelity decay over time for $t$ steps, for one particular realization $s$ under (a) a repeated twirl (RT) scheme and (b) a scheme in a Loschmidt echo fashion (LE). We allow now for possibility having flawed twirling gates $\tilde{U}$, and also faulty $\tilde{U}^\dagger$ may not reverse them perfectly. The errors $\Gamma$ could also vary in each step $\tau$. In (b) we have the option of combining all the inverse gates into a single reversal gate $\tilde{U}_s^R$. $\tilde{s}$ is randomly chosen so to sample the pool of twirl operators.

\[ \Gamma. \]

Notice also that if we are using the Clifford gates as the $U_s$, by characterizing the twirl alone we obtain information about gates that are not only twirl operators but also gates of paramount importance in QIP (see Appendix Sec. B.2).

### 7.2 A tale of two schemes

When we first introduced the two many-step generalizations of the fidelity decay in Sec. 2.4, we stated that they were equivalent. This actually holds if the twirl operators are perfect, or if at worst they carry gate-independent errors, so that these errors can be taken as an independent map that is actually perfectly twirled. We develop now the mathematical framework for working with the two schemes in the most general case.

Under the twirl scheme the state of the system evolves as

\[ \rho_{RT}(t + 1) = \sum_s U_s^\dagger A_{s,t} U_s \rho_{RT}(t) U_s^\dagger U_s \]

(7.2)

Here the gates $U_s$ are perfect gates, and the error map $A_{s,t}$ is the one arising from the actual implementation of $U_s^\dagger U_s$ at a given time $t$. From (7.2), the evolution after $t$ steps is just the consecutive application of a “twirl map” $\hat{O}_T$

\[
|\rho_{RT}(t)\rangle = \hat{O}_T(t)|\rho_{RT}(0)\rangle \\
\hat{O}^T(t) = \hat{O}_T^T \ldots \hat{O}_1^T \\
\hat{O}_s^T = \sum_s \hat{U}_s^\dagger A_{s,t} \hat{U}_s
\]
also\(^4\)

\[
\hat{O}^T(t) = \sum_{s_1, \ldots, s_t} \hat{U}_{s_1} \ddag \hat{A}_{s_1,t} \hat{U}_{s_1} \ddots \hat{U}_{s_1} \ddag \hat{A}_{s_1,1} \hat{U}_{s_1}
\]

(7.3)

as depicted in Fig. 7-1(a). Clearly, if the error arising from the implementation of the \(U^\dagger U\) gates is the same at different times so that \(\hat{A}_{s,\tau} = \hat{A}_s \forall \tau\) (i.e.: although there is a gate-dependence, this dependence does not vary in time), then the evolution of the system at time \(t\) is just the superoperator \(\hat{O}^T_1\) applied \(t\) times:

\[
\hat{O}^T(t) = \left(\hat{O}^T_1\right)^t
\]

(7.4)

Otherwise, we will say that the errors have a time dependence.

Meanwhile, under the echo scheme,

\[
\rho_{LE}(t) = \sum_{s_1, \ldots, s_t} U_{s}^R A_{s,t+1}^R \left( A_{s,t} \left( U_{s_1} \ldots A_{s_1,1} \left( U_{s_1} \rho_{RT}(0) U_{s_1}^\dagger \right) \ldots U_{s_t}^\dagger \right) U_{s_t}^R \right)
\]

where again the gates \(U_{s}\) are perfect gates, but now the error map \(A_{s,t}\) is the one arising from the actual implementation of \(U_{s}\) at a given time \(t\) (not of the gate followed by its inverse). Also, there is an “extra” error step \(A_{s,t+1}^R\) corresponding to the reversal gate \(U_{s}^R\) for a given chain of the \(t\) gates in a realization \(s\).

We can define the evolution in terms of an “echo map” \(O^E(t)\)

\[
\|\rho_{LE}(t)\| = \hat{O}^E(t)\|\rho_{LE}(0)\|
\]

\[
\hat{O}^E(t) = \sum_{s_1, \ldots, s_t} \hat{U}_{s}^R \ddag \hat{A}_{s,t+1}^R \hat{A}_{s,t} \ddag \hat{U}_{s_1} \ddots \hat{A}_{s_1,1} \hat{U}_{s_1}
\]

(7.5)

The first point we would like to highlight is that the two schemes characterize different things. The twirl scheme characterizes the error in \(U^\dagger U\), while the echo scheme characterizes the error in \(U\). Errors that are reversed when applying the inverse gate will not show up in the RT-fidelity decay. On the other hand, if the errors plainly “accumulate”, say the magnitude of the error in \(U^\dagger U\) is twice the magnitude of the error in \(U\), then the RT-fidelity will roughly decay twice as fast than the LE-fidelity—at least initially. This idea is illustrated in Fig. 7-2, where we show the simulated fidelity decays arising from the imperfect design of Clifford gates in a liquid-state NMR QIP device (crotonic acid in a 400MHz spectrometer). We observe a good agreement between the two once we have matched the error magnitude per step, assuming that in the RT scheme is twice as large as in the LE scheme. Notice the matching of the two schemes is actually strictly true when the errors are gate-independent.

\(^4\)Note that for a unitary operator \(U\), \(U^\dagger = U^\dagger \otimes U^\dagger = \hat{U}^\dagger = U^\dagger \otimes U^\dagger\).
Figure 7-2: (color) Simulations of generalized fidelity decays in liquid-state NMR QIP, using the 24 Clifford operators on qubit 1 (the carbonyl) of the 4-qubit crotonic acid molecule, and no intermediate map ($\hat{\Gamma} = \mathbb{I}$). We used $N = 200$ realizations to sample the twirl. We have simulated only the 4 carbons and the errors arising from imperfect pulse design over a 4-qubit space (we have not included rf inhomogeneity or the presence of the hydrogens in the molecule, thus the errors are unitary at each step). The simulations under the LE scheme are denoted with filled symbols •, ■, ▲, ●, ▲, while the RT scheme results are denoted with hollow ones ○, □, ◊, ◌ and △. We have plotted $(t, f^{LE}(2t))$ and $(t, f^{RT}(t))$ in order to match the error magnitude per step. The fidelities shown are:

•, ○: $f(t, \Lambda, \psi_0)$ for qubit 1 (using the reduced density matrix of the outcome) for an initial Zeeman state $|\psi_0\rangle = |0\rangle^\otimes 4$

■, □: $f(t, \Lambda, \rho_0)$ for qubit 1 (using the reduced density matrix of the outcome) for an initial state $\rho_0$ with qubit 1 in $|0\rangle$ and the rest in the maximally mixed state $I/2$

◆, ◊: $8 \times f(t, \Lambda, \rho_0)$ of the four carbons, for an initial state $\rho_0$ with qubit 1 in $|0\rangle$ and the rest in the maximally mixed state $I/2$

◆, ◌: $f(t, \Lambda, \psi_0)$ of the four carbons, for an initial Zeeman state $|\psi_0\rangle = |0\rangle^\otimes 4$

▲, △: the generalized entanglement fidelity, $\text{Tr}[\hat{O}(t)]/D^2$
In Fig. 7-2 we have plotted several calculations obtained from the same simulation (for each scheme). Since we are working on a 4-qubit space, the fidelity between two states will depend in principle on the initial state of all four qubits. We have considered two possibilities: $|0\rangle \otimes \mathbb{I}^3$, and $|0\rangle \otimes (\mathbb{I}/2)^\otimes 3$, where the first qubit corresponds to the carbonyl of the molecule. The latter case seeks to avoid changes arising from the evolution of the other carbons. We found that the two calculations match quite well, as better seen in Fig. 7-3. This indicates that the dynamics are separable into to the subspace of qubit 1 and the rest of the molecule. We have also calculated the generalized fidelities for those two initial states on the whole carbon space. When the extra carbons are in the $\mathbb{I}/2$ state, the fidelity matches well to the fidelity for qubit 1 only, which indicates that the dynamics occurring in the extra carbons are unital. However the fidelity for the state $|\psi_0\rangle = |0\rangle \otimes 4$ presents a drop with respect to these, although still showing a similar behavior. Finally, we have also calculated a form of "generalized entanglement fidelity", namely $\text{Tr}[O(t)]/D^2$. The fact that this quantity shows a notably different behavior from the previous can be understood in terms of two facts: there is indeed some dynamics that leaked to the extra carbon space, and since this process is not being twirled, is highly dependent on the initial state.

![Figure 7-3](image_url)

(a) Matched fidelities  
(b) Original fidelities

Figure 7-3: (color) Inset on the simulations of generalized fidelity decays in liquid-state NMR QIP, using the 24 Clifford operators on qubit 1 (the carbonyl) of the 4-qubit crotonic acid molecule, in the same conditions as Fig. 7-2. (Semi-log plot)

The second point we highlight is that in the echo scheme there is an additional drop in the fidelity due to the errors in the reversal map. If we are interested in the overall time-behavior of the fidelity decay, this additional drop should not be particularly meaningful. In general we would expect the effect of $\Lambda R$ to be just a net shift down of $f^{LE}(\Lambda,\psi_0, t)$, which is indeed true for an error map $\Lambda R$ that is gate-independent and time-independent. We can see the negligible effect of the errors in the last step in Fig. 7-4.

We will therefore continue our theoretical analysis of the echo-fidelity decay assuming
that $\Lambda_{s,t+1}^R = \hat{1}$. In that case,

$$\dot{\hat{O}}^E(t) = \sum_{s_1,\ldots,s_t} \hat{U}_{s_1}^\dagger \Lambda_{s_1,t} \hat{U}_{s_1} \cdots \Lambda_{s_1,1} \hat{U}_{s_1} \quad \text{(7.6)}$$

$$= \sum_{s_1,\ldots,s_t} \hat{U}_{s_1}^\dagger \cdots \hat{U}_{s_1}^\dagger \Lambda_{s_1,t} \hat{U}_{s_1} \cdots \Lambda_{s_1,1} \hat{U}_{s_1} \quad \text{(7.7)}$$

$$= \sum_{s_1,\ldots,s_t} \hat{V}_{s_1}^\dagger \Lambda_{s_1,t} \hat{V}_{s_1} \cdots \hat{V}_{s_1}^\dagger \Lambda_{s_1,1} \hat{V}_{s_1} \quad \text{(7.8)}$$

with $V_{s_1} = U_{s_1} U_{s_{t-1}} \cdots U_{s_1}$. Eq. (7.6) is as depicted in Fig. 7-1(b).

For these ideas on the LE scheme to hold, it is key to combine all the backward gates into a single one, as depicted in Fig. 7-1(b). Otherwise, if we actually implement a number of $t$ inverse gates each with errors, the error in the reversal motion cannot be neglected compared with the cumulated errors in the implementation of the forward gates.

It is well known that the Clifford group operators can be simulated efficiently by a classical computer ([2] Sec. 10.5.4, also Appendix Sec. B.2). So by keeping track of the $t$ gates used in the forward motion, we can calculate their product to retrieve a single reversal gate. A product of Clifford gates is a Clifford gate, so if we are able to implement the
forward gates, then we are equally able to implement the resulting reversal gate. Moreover, it is reasonable to expect the magnitude of the error in the forward gates to be $t$ times stronger than in the reversal gate, so for many steps we can neglect the error in $U^R$.

If, on the other hand, we are using random rotations, while the product of random rotations is another random rotation, it is not known to us how to efficiently compute a single reversal gate. This remains the case for pseudo-random rotations. In this situation we could be compromising the scalability of the computation.

Assuming the matching of the maps $\Lambda$ under study in the two schemes, we have that $O^T(t = 1) = O^E(t = 1)$, and that this is true in general only for the first step. By comparing eqs. (7.3) and (7.8), it is clear that the two fidelity decays are in general not the same, unless the error is gate-independent. The $V$ have the same properties as the $U$ gates, since either working with Clifford operators or random rotations, a product of them is just one of them. This is what gives the unitarily invariance of the twirl, mentioned in Sec. 2.4. Thus it is tempting to think that both equations are equivalent. However, an error map $\Lambda_{s,t}$ corresponds to the $U_{s,t}$ in different operators $V_{s',t}$, so we can see qualitatively that each sum over the $V_{s',t}$ is not actually twirling the target error map $\Lambda_{s,t}$, but some other map characterizing these $V$'s. With gate-dependent errors there is a direct relationship between $\Lambda_{s,t}$ and the $U_{s}$, but not between the $\Lambda_{s,t}$ and the $V_{s}$. For (7.8) to be equivalent to (7.3) the sum should run over $s'$, which would require a consequent transformation of the $\Lambda_{s,t}$. The only case we can assure that both fidelities agree is when the errors are gate-independent. In that case $\Lambda_t$ is the error happening at step $t$, either for the $U$'s or for the mathematically artificial $V$'s.

Thus we can say

$$\hat{O}^T(t) = \hat{O}^E(t) = \sum_{s_1, \ldots, s_t} \hat{O}_{s_1} \hat{\Lambda}_{s_1} \hat{U}_{s_1} \ldots \hat{U}_{s_t} \hat{\Lambda}_{s_t} \hat{U}_{s_t}$$

for gate-independent errors

which includes, of course, the case where the $U_{s}$ are perfect and we are just analyzing an error map that is independent of them.

The results in this case were developed in [13]. Consider an error map of the form $\hat{\Lambda}_{s,t} = \hat{\Phi}_t$, that is gate-independent but time-dependent in principle. Summarizing,

$$\hat{O}^T(t) = \hat{O}^E(t) = \hat{\Phi}^{dep}_t \ldots \hat{\Phi}^{dep}_1 \text{ with } \hat{\Phi}^{dep}_t = \sum_s \hat{U}^*_s \hat{\Phi}_t \hat{U}_s$$


(7.9)

$$\Phi^{dep}_t(\rho) = p_t \rho + (1 - p_t) \frac{I}{D}, \quad p_t = \frac{\text{Tr}[\hat{\Phi}_t] - 1}{D^2 - 1}$$

(7.10)

That $\Phi^{dep}_t$ is like in eq. (7.10) can be proven, for example, using eq. (6.14): express $\Phi_t$ either in the $\chi$-matrix form or as in eqs. (5.1)-(5.2), and identify the expression for $\text{Tr}[\hat{\Phi}_t]$. Then decompose $\rho$ in the basis formed by the generalized Pauli Operators $P_m$ and take $A_1 = P_m, A_2 = P_m'$ in eq. (6.14). On the other hand, if we replace $\rho$ expressed in the Pauli basis in (7.10), we find that both expressions are equivalent.

If $p_t = p \forall t$, then simply $f^{RT}(\Phi, \psi_0, t) = f^{LE}(\Phi, \psi_0, t) = p^t (1 - 1/D) + 1/D$, with $f^{RT}(\Phi, \psi_0, t = 1) = f^{LE}(\Phi, \psi_0, t = 1) = F(\Phi)$ indeed. For $p > 0$ the decay is a shifted exponential with an initial decay linear in $t$ ($p < 1$, but for large enough error strength $\text{Tr}[\hat{\Phi}]$, $p$ can be negative and thus the fidelity decay presents a step-by-step oscillation). Notice
also that the fidelity decay will be the same for any pure initial state \( \psi_0 \), which arises from the fact that the net evolution at each step is a depolarizing channel \( \rightsquigarrow \text{eq. (7.10)}. \)

We can prove the relation between the average fidelity \( F \) and the entanglement fidelity \( F_e \) (eq. (2.9)) from these results. We write out \( \Phi^\text{dep}_\tau \) in the Pauli operator basis \( P_l \). We take an inner product between state operators described by a vector \( \tilde{c} \) like \( \rho = \sum_l c_l P_l \) as

\[
\langle \rho | \rho' \rangle = D \tilde{c} \cdot \tilde{c}', \quad \text{while} \quad \langle P_l | \rho \rangle = D \text{Tr}[P_l P_l'] = D \delta_{l,l'}
\]

In this basis \( \Phi^\text{dep}_\tau \) is diagonal, with \( \langle \mathbf{I} | \Phi^\text{dep}_\tau | \mathbf{I} \rangle = 1 \) (since it is a unitary map) and \( \langle P_l | \Phi^\text{dep}_\tau | P_l \rangle = p_r \forall l > 0 \). So clearly by taking the trace of the superoperator (equal in both schemes), the entanglement fidelity is \( F_e^T(t) = F_e^E(t) = (1 + (D^2 - 1)p^l)/D^2 \), and (2.9) holds indeed.

When the errors occurring in the twirl operators \( U_s \) are gate-dependent, the previous results do not necessarily hold.

To start, the RT-fidelity decay and the LE-fidelity decay do not need to be equal. A very simple example can illustrate this: if the error map has a gate-dependence of the form \( \hat{\Lambda}_{s,t} = \hat{U}_s \Phi \hat{U}^\dagger_s \) (where \( \Phi \) is a time-independent and gate-independent map), then

\[
\hat{O}^T(t) = \hat{\Phi}^t \quad \text{while} \quad \hat{O}^E(t) = (\hat{\Phi}^\text{dep})^{t-1} \hat{\Phi}
\]

Although this may be a rather artificial example, it illustrates the point of the two fidelities being substantially different even if the map \( \Lambda \) arising from errors in \( \hat{U}U^\dagger \) in the RT scheme or of \( U \) in the LE scheme was the same. It also already shows that the RT-fidelity decay will not be necessarily exponential-like nor independent of the choice of initial state. For example, if \( \Phi \) is simply a small rotation of the \( j \)-qubit around \( \hat{x} \) with an angle \( \alpha \), then the decay will go like \( \sim \cos(\alpha t) \) if the initial state has the \( j \)-qubit in an eigenstate of \( \sigma_z \), while it will be a constant with value 1 if the \( j \)-qubit was originally in an eigenstate of \( \sigma_x \).

On the other hand, this example also already shows that the LE-fidelity will not necessarily be a perfect exponential decay. However, if \( \Phi \) is a small perturbation from the identity map (a weak noise channel), then with time, the \( \hat{O}^E(t) \) in this example becomes a depolarizing channel effectively, since the factors \( \hat{\Phi}^\text{dep} \) predominate for large \( t \).

The underlying idea in using the generalized fidelities \( f(t, \Lambda, \psi_0) \) for many steps is that we hope they coincide with a "generalized average fidelity" \( F(\Lambda,t) \):

\[
F_e(\Lambda,t) = \text{Tr}[\hat{O}(t)]/D^2 \quad \rightarrow \quad F(\Lambda,t) = \frac{DF_e(\Lambda,t) + 1}{D + 1} : \quad f(t, \Lambda, \psi_0) \leftrightarrow F(\Lambda,t)
\]

where \( \hat{O} \) is the superoperator for either the RT or the LE schemes. If indeed \( f \) coincides with \( F \), by fitting the generalized fidelity curve obtained from experimental data, we can extrapolate to the point \( t = 1 \), where we recover the average fidelity of the faulty twirl under study. Once the twirling channel is perfected, we can then use again the generalized fidelities with an intermediate map \( \Gamma \) to measure its fidelity, again in principle in a way that is robust against errors in the initial state preparation and measurement.

The analytical results we will show in the next Chapter are for a constant error map, so
\[ \hat{\Lambda}_{s,t} = \hat{\Lambda}_s. \] In this case we have

\[ |\rho_{\text{RT}}(t)\rangle \rangle = \hat{O}^E(t)|\rho_{\text{RT}}(0)\rangle \rangle \] (7.11)

\[ O^E(t + 1) = \sum_s \hat{U}_s^\dagger O^E(t) \hat{\Lambda}_s \hat{U}_s \] (7.12)

\[ O^E(t = 1) = \sum_s \hat{U}_s^\dagger \hat{\Lambda}_R \hat{\Lambda}_s \hat{U}_s \] (7.13)

The “nesting” of the maps at consecutive times eq. (7.12) prevents us in general from calculating the evolution as a simple repetitive application of one map, as we are able to do with the RT-fidelity.

Eq. (7.13) is the initial condition for the nesting. We have considered the error in the reversal gate as an error \( \hat{\Lambda}_R \) whose effect can be approximately taken into account at the beginning of the sequence. We can go further and impose the assumption of an errorless reversal, which implies \( \hat{\Lambda}_R = \mathbf{I} \). Depending on the error model, one situation may be more tractable than the other, and they both yield the same result after a few steps. We emphasize though that the analytical results we obtain using eqs. (7.12)-(7.13) are accurate after a few steps, when the approximation of considering the reversal gate in one of these two ways becomes valid.

As final remarks, we revisit two points. On one hand, we notice that the compression of the reversal motion into a single step, or any other approach approximating the error map in the reversal motion, is not essential for obtaining analytical results. We could still define a nesting of the form

\[ O^E(t + 1) = \sum_s \hat{U}_s^\dagger \hat{\Lambda}_R O^E(t) \hat{\Lambda}_s \hat{U}_s \] (7.14)

The question then is what is the process we are characterizing. Nevertheless, our work indicates that the challenge of obtaining analytical results in one case or the other is the same.

On the other hand, when we compress the backward motion, we emphasize that neglecting the error in the reversal gate \( \hat{U}_R \) should not be a critical assumption. The errors in the initial state preparation and measurement can be taken as error maps occurring at the beginning and at the end of the sequence. Therefore, they can be merged with \( \hat{\Lambda}_R \). If we are looking for a protocol that is robust against errors in the initial state and measurement, then indeed the effect of a single error map at the beginning or at the end of the sequence (with an error magnitude similar to the error magnitude present in the other steps) must be negligible and not disturb the shape and magnitude of the fidelity decay.

In the next Chapter we present examples of gate-dependent errors for the case of one qubit. The relevance of the one-qubit case lies first in the fact that, for some error models \( \hat{\Lambda} \), we can obtain analytical results, thus shedding some light into more general situations. Furthermore, we must notice that the protocols in [16, 22] (one-step one-qubit twirl protocols) require one-qubit twirl operators in order to analyze more complex maps. It is reasonable then, from the experimental point of view, to concentrate on perfecting the one-qubit twirl in order to tackle more complex gates afterwards, as we already suggested in Chapter 6.
Chapter 8

Case studies of the generalized fidelity decays using faulty twirl operators, for one qubit

As explained in Sec. 5.3, in liquid-state NMR, the internal Hamiltonian for four qubits (like the carbons in crotonic acid) has the form

$$\mathcal{H}_{\text{int}}^L = -\frac{\gamma_C B_0}{2} \sum_{j=1}^{4} \sigma_z^{(j)} + \sum_{j=1}^{4} \frac{\omega_{\delta,j}}{2} \sigma_z^{(j)} + \sum_{k>j=1}^{4} \frac{\pi J_{jk}}{2} \bar{\sigma}^{(j)} \cdot \bar{\sigma}^{(k)}$$  \hspace{1cm} (8.1)

where $|\gamma_C B_0| >> |\omega_{\delta,j}|$, $|\gamma_C B_0| >> |J_{jk}|$, and $|\omega_{\delta,j} - \omega_{\delta,k}| >> |J_{jk}|$. On the other hand, the propagators for a one-qubit twirl on, say, qubit 1, have the form $U_s = \exp(-i\theta_s \hat{n}_s \cdot \bar{\sigma}/2) \otimes I \otimes I \otimes I$, so it is clear that the action of a control Hamiltonian $\mathcal{H}_{\text{rf}}^L$ must not only result in the correct $\theta_s$ and $\hat{n}_s$, but it must also refocus the net evolution from the $J$-couplings so that the twirl operator contains no two-body terms (nor higher order multi-body terms). Fortunately the magnitude of the control Hamiltonian $B_1(t)$ is $|\gamma_C B_1| >> J_{jk}$, so by pulse design it is possible to achieve the refocusing. In the rotating frame we are left then with a net propagator $U = T[e^{-i J_{\text{rf}}^L \mathcal{H}^R dt}]$, where the predominant effect of $\mathcal{H}^R$ is given by the rf Hamiltonian and the Zeeman terms in (8.1). Working in this simple scenario, we focus on the implementation of the first 12 Clifford gates for one qubit (as listed in Table B.1 in Appendix B). If they were perfect gates, these set of 12 is enough to implement the twirl of an arbitrary map.

Under the RT scheme, at each realization we implement the progressive randomization of the fidelity by applying the Clifford gates chosen at random followed by their inverses, which fall into the same set of 12. For the LE-fidelity, we will implement the forward gates choosing them from these 12 gates and store the sequence for one realization in a vector $\bar{s}$. For each $\bar{s}$, we calculate the final reversal gate $U^R_s$. This reversal gate may fall out of the set of 12 (because they do not form a group), but nevertheless we will assume that we are prepared to implement any of the possible 24 Clifford gates (which do form a group under matrix multiplication, up to a global phase – see Appendix B).

Fig. 8.1 shows the calculation of the generalized fidelities of the previous Chapter, simulating a liquid-state NMR QIP device (crotonic acid in a 400MHz spectrometer). We include here also the simulation using only the first set of 12 Clifford gates, as explained above. We
can see that there are no striking differences in the time-behavior when reducing the sampling pool. We must be aware that these two scenarios do not need to coincide unless the errors are gate-independent, since each Clifford gate carries in principle a unique error. It is nevertheless reasonable that the nature of the error will not be so different as to radically change the shape of the curve. As we see in Fig. 8-1, this seems to be particular true for the LE scheme.

![Graphs showing fidelity decays](image)

(a) Fidelities under the LE scheme  
(b) Fidelities under the RT scheme

Figure 8-1: (color) Simulations of generalized fidelity decays in liquid-state NMR QIP, using the whole Clifford group (24 operators) and half of it (the first 12 of Table B.1). We are twirling only qubit 1 (the carbonyl) of the 4-qubit crotonic acid molecule, with no intermediate map ($\hat{\Gamma} = \mathbb{I}$). $N = 200$ realizations, with no rf inhomogeneity and disregarding the hydrogens. As in the previous Chapter, we have used hollow symbols ○, □, ◄, ◊ for the RT scheme, and filled symbols ●, ■, ◀, ◊ for the LE scheme, when using the whole Clifford group. The semi-filled symbols represent, respectively, the same quantity but with a simulation sampling over the first 12 Clifford gates only. As before,

- ○, f($t, \Lambda, \psi_0$) for qubit 1 for an initial Zeeman state $|\psi_0\rangle = |0\rangle^\otimes 4$
- ■, □, f($t, \Lambda, \rho_0$) for qubit 1 for an initial state with qubit 1 in $|0\rangle$ and the rest in $I/2$
- ◄, ◄, 8 $\times$ f($t, \Lambda, \rho_0$) for the four carbons, for an initial state with qubit 1 in $|0\rangle$ and the rest in $I/2$
- ◊, ◊: f($t, \Lambda, \psi_0$) for the four carbons, for an initial Zeeman state $|\psi_0\rangle = |0\rangle^\otimes 4$.

We will show in this Chapter numerical and analytical results for the generalized fidelity decays for both schemes, considering the implementation of faulty Clifford gates and no intermediate map (this is equivalent to twirling a perfect quantum channel/identity operation). We will consider Clifford gates only and not random rotations, for several reasons:

(i) we can guarantee that the reversal gate $U_i^R$ in the LE scheme can be determined efficiently, and that its potential error should be of the order of the error per gate in the rest of the sequence;

(ii) a study of the Clifford gates is not only a study on twirl operators/2-designs but also a study on basic gates of general relevance for quantum computation and fault-tolerance [57];

(iii) unlike random rotations, the pool of twirl operators to sample from is fixed.

(i) is only relevant with a system of bigger size of course, but we will stick to the principle even when working with one qubit. (iii) refers to the fact that if we manage to characterize
and perfect the $12 \times n$ one-qubit Clifford gates, we are ready to implement the one-qubit Clifford twirl. Having a finite and scalable number of twirl operators simplifies the task of preparing the twirl implementation.

Notice that, for one qubit, the Clifford gates have a simple decomposition as a rotation with a given angle around a certain axis, i.e.: $C_j = \exp(-i\theta_j \hat{n}_j \cdot \hat{a}/2)$ (cf. Table B.1). On the other hand, random rotations are naturally parameterized by angles $\phi$, $\psi$ and $\xi$ (cf. eq. (4.2)), and the way this parameters relate to rotations described by one angle and one direction is

$$R(\phi, \psi, \xi) = \begin{pmatrix} \cos(\phi) e^{i\psi} & \sin(\phi) e^{i\xi} \\ -\sin(\phi) e^{-i\xi} & \cos(\phi) e^{-i\psi} \end{pmatrix} = \exp(-i(\psi - \xi)\sigma_z/2) \exp(-i\phi\sigma_y) \exp(-i(\psi + \xi)\sigma_z/2)$$

Therefore it is not simple to translate the effect of a flawed Hamiltonian $H$ to errors in $\phi$, $\psi$ and $\xi$. Of course, such a relation exists, since $R(\phi, \psi, \xi)$, like any other unitary operator in $SU(2)$, can be written in terms of an angle and a direction (in turn given by two other angles), which are more easily related to physical parameters in $\int_0^T H \, dt$. However in order to obtain analytical results from an error model with physical parameters like the angle and direction of a net rotation (see next Sec. 8.1), the structure of the Clifford operators is better suited.

This is also a good reason to concentrate on the one-qubit twirl. In the previous Chapters we have shown the usefulness of a twirl in $U(2)^{\otimes n}$ for the characterization of quantum processes, specially since one-qubit operations can be implemented with more success than more complex gates. This is one of the main reasons to study the one-qubit twirl. But also, the study of these simple operations can retrieve information about the system that can be more easily linked to physical models of the quantum device dynamics.

The results we will show were obtained exactly as follows. For the RT scheme, the numerical simulations were done calculating $\hat{O}^T(t)$ as in eq. (7.3), where each step $\hat{U}^T \hat{A} \hat{U}$ arises from implementation of a Clifford gate followed by its inverse: $\hat{U}^T \hat{A}^T \hat{U}$, both being faulty. The analytical calculations were also done in this way: since we will work with an error map $\Lambda$ that is time-independent (the error will be exclusively associated to a gate, independently of the moment at which it occurs) we can simply calculate $\hat{O}_I^T$ and exponentiate it to the $t$, as we already mentioned. It is clear then that the dynamics of the system is encoded in the eigenvalues of $\hat{O}_I^T$, namely $\{\lambda_j^T; j = 1, \ldots, D^2\}$. In particular,

$$F(\Lambda) = \frac{1}{D^2} \sum_k (\lambda_k^T)^t$$

(8.2)

On the other hand, for the LE scheme, the numerical simulations were done by calculating $\hat{O}^E(t)$ as in eq. (7.5), where each step $\hat{A} \hat{U}$ arises from a faulty implementation of a Clifford gate alone –and then including a faulty reversal gate at the end. However the analytical calculations were done using eqs. (7.12)-(7.13), where the error map corresponds to the error in the forward gate only, and the error in the reversal gate is taken into account approximately in the first step as in (7.13), or completely neglected (taking $\Lambda_R = \mathbf{I}$ in (7.13)). Both approaches are valid approximations that work well after a few steps.
8.1 Physical model

We take one 1/2-spin in a static magnetic field $B_0$ along $\hat{z}$ (the quantization axis) subjected to a magnetic rf field $B_1(t) \cos(\Omega t + \phi(t))$ along the $\hat{x}$ direction (with constant rf frequency $\Omega$) as system model. Alternatively, we could have a rotating field $B_1(t) \cos(\Omega t + \phi(t)) \hat{x} + \sin(\Omega t + \phi(t)) \hat{y}$. These are quite universal models for any two-level quantum object to which a near resonant field is applied. Assuming that the static field $B_0$ is such that $B_0 << B_1$, moving to a rotating frame with the transformation $\rho = R^t \rho^R$ with $R = \exp(-i\sigma_2 \Omega t)$, and applying the RWA, the Hamiltonian in the rotating frame can be taken as

$$\mathcal{H} = \frac{\omega_\delta}{2} \sigma_z + \omega(t) \left( \cos(\phi(t)) \sigma_x + \sin(\phi(t)) \sigma_y \right)$$ (8.3)

where $\omega_\delta = \omega_L - \Omega$, $\omega_L = \gamma B_0$ is the Larmor frequency of the spin, and $\omega(t) = \gamma B_1(t)$.

The axes along which the Clifford gates should be defined are $\hat{z}$, $\hat{x}$ and $\hat{y}$ in the qubit frame. We can tune $\Omega$ to set $\omega_\delta$ at will, and also we can control $\omega(t)$ and $\phi(t)$ directly. With constant values $\omega(t) = \omega_1$ and $\phi(t) = \phi_1$, we have

$$U = T[e^{-i \int_0^t \mathcal{H} t \, dt}] = \exp \left( -i \frac{\omega_\delta}{2} t \sigma_z - i \omega_1 t \left( \cos(\phi_1) \sigma_x + \sin(\phi_1) \sigma_y \right) \right)$$ (8.4)

and thus any unitary operation can be implemented. In practice, the system is not isolated (to start, we do expect to have a system with many interacting qubits), thus this is an oversimplification. Nevertheless, it helps to gain insight into the nature of typical errors.

Compare eq. (8.4) with the parameterized Clifford gates:

$$C = \exp \left( -i \theta \left( n_z \sigma_z + n_x \sigma_x + n_y \sigma_y \right) \right)$$ (8.5)

We can see then that errors in the parameters $\Omega$, $\omega_\delta$, $\omega_1(t)$ and $\phi_1(t)$ will directly create errors in the angle $\theta$ and/or the direction $\hat{n}$ of the Clifford operators. These are the kind of errors we will consider in the next Sections.

8.2 Clifford gates with their angle off by a fixed amount $a$

In this model every Clifford gate $C_j$ becomes $\tilde{C}_j = \exp(-i(\theta + a) \hat{n}_j \cdot \vec{\sigma})/2) = e^{-ia\tilde{n}_j \cdot \vec{\sigma}/2}C_j$. We notice right away that if no intermediate map is applied, for the RT scheme this error cancels out for $C_1$, $C_2$, $C_5$, $C_7$, $C_9$ and $C_{12}$ (which are the Clifford gates that are $\pi/2$-rotations) because each corresponding inverse gate carries exactly the inverse error. On the other hand, this error “duplicates” for $C_3$, $C_4$, $C_6$, $C_8$, $C_{10}$ and $C_{11}$ (the Clifford gates that are $\pi$-rotations or, rather, Hadamard-like rotations$^5$).

So for the twirl fidelity, we have that

$$\hat{\Lambda}_j^T = \begin{cases} e^{-ia\tilde{n}_j \cdot \vec{\sigma}} \otimes e^{-ia\tilde{n}_j \cdot \vec{\sigma}} \otimes e^{-ia\tilde{n}_j \cdot \vec{\sigma}} & j = 3, 4, 8, 10, 11 \\ \hat{I} & j = 1, 2, 5, 7, 9, 12 \end{cases}$$

---

$^5$The Hadamard gate $H$ is a $\pi$-rotation around the $(\hat{x} + \hat{z})/\sqrt{2}$ axis, i.e.: $H = (\sigma_x + \sigma_z)/\sqrt{2}$. A Hadamard-like gate would be like $H$ but with the $x$ and $z$ axes changed to another pair, including opposite directions. See Table B.1.

104
while $\hat{\Lambda}_j^E = e^{-i\hbar j \cdot \vec{\sigma}/2} \otimes e^{-i\hbar j \cdot \vec{\sigma}/2} \forall j$ for the LE scheme.

![Graphs showing numerical calculations of the fidelity decays produced by an error model considering Clifford gates with their angles off by an amount $a$. The RT-fidelity is on the left, and the LE-fidelity is on the right. Here $N = 200$ realizations, and we used an initial state $|\psi_0\rangle = |0\rangle$ (the eigenstate of $\sigma_z$ with $+1$ eigenvalue). The solid line represents the analytical calculation (see text). Error strength: $\circ a = 0.02$, $\square a = 0.06$, $\diamond a = 0.1$, $\ast a = 0.2$, $\circ a = 0.4$, $\triangle a = 1.0$.]

Figure 8-2: Numerical calculations of the fidelity decays produced by an error model considering Clifford gates with their angles off by an amount $a$. The RT-fidelity is on the left, and the LE-fidelity is on the right. Here $N = 200$ realizations, and we used an initial state $|\psi_0\rangle = |0\rangle$ (the eigenstate of $\sigma_z$ with $+1$ eigenvalue). The solid line represents the analytical calculation (see text). Error strength: $\circ a = 0.02$, $\square a = 0.06$, $\diamond a = 0.1$, $\ast a = 0.2$, $\circ a = 0.4$, $\triangle a = 1.0$.

The maps $O_T$ and $O^E$ are unital (i.e., they take the identity $I$ to itself), and when analyzing the data from numerical simulations (as the ones that produced Fig. 8-2), it is clear that the matrix representation of the map presents a simple structure when written in the basis formed by the Pauli group. In this basis, a density matrix of the form $\rho = c_0 I + \vec{c} \cdot \vec{\sigma}$ has a state-vector representation of the form

$$|\rho\rangle = \begin{pmatrix} c_0 \\ c_x \\ c_y \\ c_z \end{pmatrix}$$

(8.6)

By straightforward calculation, it can be shown that

$$\dot{O}_T(t = 1) = \dot{O}_T^f = \frac{1}{12} \sum_{j=1}^{12} \hat{C}_j^{-1} \hat{\Lambda}_j^T \hat{C}_j = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & A_1 & -B_1 & B_1 \\ 0 & B_1 & A_1 & -B_1 \\ 0 & -B_1 & B_1 & A_1 \end{pmatrix}$$

with eigenvalues $\{\lambda_1 = 1, \lambda_2 = A_1, \lambda_\pm = A_1 \pm i\sqrt{3}B_1 = re^{\pm i\omega}\}$. The eigenvalue $\lambda_1 = 1$ is a consequence of the unitality of the map. $A_1 = (2 + \cos(\omega)) / 3$ and $B_1 = \sin(2\omega) / 6\sqrt{2}$ with this we can plot the pair of complex eigenvalues of $\dot{O}_T^f$, shown in Fig. 8-3(a).

Since the time evolution in the RT-fidelity decay is just the successive application of $\dot{O}_T^f$, the simple eigenvalue structure allows us to calculate

$$\dot{O}_T(t) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & A_t & B_t & B_t' \\ 0 & B_t & A_t & B_t' \\ 0 & B_t' & B_t & A_t \end{pmatrix}$$

$A_t = \left( A_t^t + 2r^t \cos(\omega t) \right) / 3$

$B_t = \left( A_t^t - r^t (\cos(\omega t) - \sqrt{3}\sin(\omega t)) \right) / 3$

$B_t' = \left( A_t^t - r^t (\cos(\omega t) + \sqrt{3}\sin(\omega t)) \right) / 3$
The time dependence of the matrix elements of $O^T(t)$ already indicates an underlying damped oscillation behavior. Indeed, the generalized entanglement fidelity in this scheme, $\text{Tr}[O^T(t)]/D^2$, is clearly $F^T_e(\Lambda, t) = 1/4 + (A_1^T + 2r^t \cos(\omega t))/4$, which has the overall behavior of a damped oscillation. The initial decay will be quadratic in $t$ for small error strength $a$, and linear for large $a$. This is a consequence of the presence of a complex conjugate pair of eigenvalues, as it can be seen in eq. 8.2.

Comparing with a classical damped oscillator, roughly we can say that there are two regimes, $-\log(r) < < \sqrt{\omega^2 + \log(r)^2}/4$ (quadratic initial decay $\leftrightarrow$ undamped oscillation) and $-\log(r) >> \sqrt{\omega^2 + \log(r)^2}/4$ (linear initial decay). The $\cos(a)$ or $\cos^2(a)$, $\sin^2(a)$ dependence of $A_1^T$ gives us an initial decay rate that is mainly quadratic in $a$. We must be aware though, that there is range of rather large $a$ where $A_1$ is negative, thus the decay will behave as a discrete “step-by-step” oscillation resulting from a global factor $(-1)^t$, with an overall exponential decay envelope.

The fidelity decay for a pure state $\psi_0$ with $c_0 = 1/2$ and $|c|^2 = 1/4$ will be $f^{RT}(\Lambda, \psi_0, t) = 1/2 + (A_1^T(1/2 + 4x) + r^t \cos(\omega t)(1 - 4x))/3$, with $x = c_x c_y + c_y c_z + c_z c_x$. So the fidelity will not be the same for all initial states $\psi_0$. In particular, choosing $\psi_0$ so $x = 0$, we immediately have $f^{RT}(\Lambda, \psi_0|x = 0, t) = (2F^T_e(t) + 1)/3$, like in (2.9). So for these particular initial states, the generalized fidelity under randomization coincides with the average fidelity: we can fit the curve and then extrapolate its value at $t = 1$.

On the other hand, if $c_x = c_y = c_z = 1/2\sqrt{3}$, then $f^{RT}(\Lambda, \psi_0|x = 1/4, t) = 1/2 + A_1^T/2$, and the RT-fidelity turns out to be shifted exponential. This state-dependence is quite critical. It means that we can no longer use a twirl and any arbitrary initial state (fixed non-random) to obtain the generalized average fidelity. Of course with any state we would
still extract information about the system, but now to collect this information, we either have to sample different states or have some knowledge in advance that indicates which one is the most convenient. Moreover, this state-dependence found in the RT scheme tell us that the repetition of a step many times does not guarantee that errors that happen only at the beginning or at the end (i.e., errors in the initial state preparation and in the measurement) can be neglected.

For the LE scheme, we find that the structure of the superoperator $O_E$ corresponds to

$$
\hat{O}_E(t+1) = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & A_{t+1} & -B_{t+1} & B_{t+1} \\
0 & B_{t+1} & A_{t+1} & -B_{t+1} \\
0 & -B_{t+1} & B_{t+1} & A_{t+1}
\end{pmatrix}
$$

where the matrix $M_a$ between $A_{t+1}, B_{t+1}$ and $A_t, B_t$ comes from calculating (7.12) explicitly (using the first set of 12 of Clifford), with $O_E(t+1)$ and $O_E(t)$ characterized by $A_{t+1}, B_{t+1}$ and $A_t, B_t$ respectively.

The fidelity decay for a pure state $\psi_0$ with $c_0 = 1/2$ and $|\bar{c}|^2 = 1/4$ will be now $f^{LE}(\Lambda, t, \psi_0) = 1/2 + A_t/2$, independent of the initial state. This is clearly a consequence of having $\langle \sigma_i | O_E(t) | \sigma_i \rangle = -\langle \sigma_i | O_E(t) | \sigma_i \rangle$ for $i \neq i'$, $l = x, y, z$, and $\langle \sigma_i | O_E(t) | \sigma_i \rangle$ the same for all the diagonal elements. Also we have $F_E^{LE}(\Lambda, t) = 1/4 + 3A_t/4$, and since $f^{LE}(\Lambda, \psi_0, t)$ only depends on the diagonal elements of $O_E$, we can already guess there is a relation between the two. Actually, again $f^{LE} = (2F^E_e + 1)/3$, just like in (2.9). So the fidelity decay for an arbitrary pure state coincides with the average fidelity, at $t = 1$ and generalized for many steps.

The expressions for $A_t, B_t$ as a function of $a$ and $t$ can be obtained by diagonalizing the $2 \times 2$ matrix $M_a$ above,

$$
A_t = \frac{\sqrt{3}(\lambda_+^t - \lambda_-^t)(3 + 3 \cos(a) + \sin(a)) + (\lambda_+^t + \lambda_-^t)}{12 \sqrt{\cos^2(a/2)(1 + 2 \cos(a) + \sin(a))}}
$$

$$
B_t = \frac{(\lambda_+^t - \lambda_-^t) \sin(a)}{2 \sqrt{6 \cos^2(a/2)(1 + 2 \cos(a) + \sin(a))}}
$$

$$
\lambda_\pm = \frac{1 + \sqrt{\cos^2(a/2)(1 + 2 \cos(a) + \sin(a))}}{12} \pm \frac{\sqrt{\cos^2(a/2)(1 + 2 \cos(a) + \sin(a))}}{2 \sqrt{3}}
$$

which are rather cumbersome expressions. However, the $\lambda_\pm$ are real for most values of $a$ (see Fig. 8-3(b)), from what we can predict an exponential-like decay behavior in that range. For $a \geq 2.5$, the eigenvalues become complex conjugates, thus giving room for a potential damped oscillation. However in this range the dominant effect is the small value of $|\lambda_\pm|$ which makes $f^{LE}$ decay extremely fast. Plus, the angles are close to $\pm \pi$ implying that $\lambda_\pm$ is negative, and thus we get a “step-by-step” oscillation that decays very fast, as shown in the two examples in Fig. 8-4.

Finally, it is worth discussing the matching of the error magnitude per step in the two
models. When the error in the twirl operators is gate-independent and given by a constant map \( \Gamma \) for all the gates, we have that \( f^{LE}(t, \Gamma, \psi_0) = f^{RT}(t, \Gamma, \psi_0) \). Thus we expect a RT-fidelity that decays twice as fast as the LE-fidelity, as originally pointed out in [13] and discussed in Chapter 7. In the case of gate-dependent errors, this is not necessarily the case. To start, the time-dependence in both schemes can be radically different, as we already saw in the example of the previous Chapter. Nevertheless, we do observe this kind of behavior numerically, by plotting \((t, f^{RT}(t))\) and \((t, f^{LE}(2t))\), so we compare one step of the RT scheme with two of the LE scheme. This is illustrated in Fig. 8-5. The agreement holds for longer times when \( a \) is larger, since in this case the RT-fidelity behaves more like a shifted exponential.

8.3 Clifford gates with their direction rotated by an angle \( b \) around \( \hat{z} \)

In this model every Clifford gate \( C_j \) becomes \( \tilde{C}_j = \exp \left( -i \theta \, \hat{n}'_j \cdot \vec{\sigma} / 2 \right) \), with \( n'_x = \cos(b)n_x - \sin(b)n_y \), \( n'_y = \sin(b)n_x + \cos(b)n_y \), \( n'_z = n_z \). That is: the original direction of the Cliffords is rotated by an angle \( b \) around the \( z \)-axis. We notice right away that if no intermediate map is applied, the twirl operators should remain constant, since \( \tilde{C}_j^{\dagger} \tilde{C}_j = I \). Note that for the Cliffords that are \( \pi/2 \)-rotations, we have that \( \tilde{C}_j = e^{i\pi \hat{n}'_j \cdot \vec{\sigma}/2} = \tilde{C}_j^{\dagger} \), so \( \tilde{C}_j^{\dagger} \tilde{C}_j = \tilde{C}_j^{\dagger} \tilde{C}_j = I \). On the other hand, for the Hadamard-like Cliffords \( \tilde{C}_j = \pm i \hat{n}'_j \cdot \vec{\sigma} \), so \( \tilde{C}_j^{\dagger} \tilde{C}_j = (\hat{n}'_j \cdot \vec{\sigma})^2 = I \).

Although it may seem that the LE-fidelity should exhibit the effect of an error, it also retrieves a constant value for all \( t \). This can be understood by observing that the Clifford gates for one-qubit have a simple geometrical meaning. They represent rotations around directions that distribute evenly over the \( xyz \)-space (see Table B.1). If we apply a global rotation to this space, the action of the set of Clifford gates remains, as a whole, unchanged.
Our numerical simulations confirmed this. The simulation of this error model simply retrieved a perfect flat line for any value of \( b \), for both schemes. This result is obtained analytically by starting with an initial echo map at \( t = 1 \) like in eq. (7.13). Notice that if we had taken \( \hat{A}_R = \mathbb{I} \), we would have obtained an initial small shift downwards in the flat line.

To show somehow the lack of effect of this type of error, we combined it with the angle-off error model. Figs. 8-6 show the simulations using Clifford gates that have their directions rotated an angle \( b \) and also their angle off by an amount \( a \). If we compare with Figs. 8-2, we can see that the rotation of the direction has no effect at all.

We also point out that in fact an arbitrary rotation of the \( \hat{n}_j \), not necessarily around \( \hat{z} \), will yield the same results, as long as all the gates are affected equally, so that we have a global rotation of the reference frame.

### 8.4 Some Clifford gates with their angle off

We consider now that the error model of Sec. 8.2 affects only a subset of Clifford gates. For example, consider that only Cliffords \( C_1, C_5 \) and \( C_9 \) are affected. Since their inverses are errorless (unlike with the global angle-off model), they will produce decay in the RT-fidelity.

In this case, we can calculate

\[
\hat{O}_1^T = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & A & -B & B \\
0 & B & A & -B \\
0 & -B & B & A
\end{pmatrix}
\]

with \( A = (2 + \cos(a))/3 \) and \( B = \sin(a)/6 \). The eigenvalues of \( \hat{O}_1^T \) are again \( \{\lambda_1 = 1, \lambda_2 = A, \lambda_\pm = A \pm i\sqrt{3}B\} \), from which we can already predict a damped oscillation time-
dependence. The eigenvalues $\lambda_2$ and $\lambda_{\pm}$, as a function of the error strength $a$ are plotted in Fig. 8-7(a). Indeed, as we can see in Fig. 8-8, $f^{RT}$ has the shape of a damped oscillation.

After two consecutive steps of the LE scheme, we find that the structure of the superoperator $O_E$ is

$$\hat{O}_E(t+1) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & A_{t+1} & C_{t+1} & B_{t+1} \\ 0 & B_{t+1} & A_{t+1} & C_{t+1} \\ 0 & C_{t+1} & B_{t+1} & A_{t+1} \end{pmatrix}$$

$$\begin{pmatrix} A_{t+1} \\ B_{t+1} \\ C_{t+1} \end{pmatrix} = M_{a'} \begin{pmatrix} A_t \\ B_t \\ C_t \end{pmatrix}, \quad M_{a'} = \begin{pmatrix} 5+\cos(a) & \sin(a) & \sin(a) \\ \frac{\sin(a)}{12} & -\sin(a) & \frac{1-\cos(a)}{12} \\ -\frac{\sin(a)}{12} & 0 & -\frac{\sin(a)}{12} \end{pmatrix}$$

This matrix possesses a pair of complex eigenvalues $\lambda_{\pm} = re^{\pm iw}$ for the whole range of $a$, giving the possibility of an oscillating time-dependence for the matrix elements of $O_E$. However, as we can observe in Fig. 8-7(b), although $w$ is of significant magnitude, $r$ is very small, so $-\log(r)$ is large and we rather have an overdamped oscillation, which can be approximated by an exponential (in particular, we have a linear initial decay). Compare the magnitude of $r$ and $w$ of the complex pair of eigenvalues, for both schemes.

We learn an important lesson here: even when the echo fidelity has traces of an oscillatory behavior, the actual value of the parameters may place the decay in a regime that for all practical purposes can be taken as exponential.

Another feature of the LE-fidelity in this model is that there is in principle a dependence on the initial state. This is due to the fact that we no longer have $\langle \sigma_l | O_E(t) | \sigma_{l'} \rangle = -\langle \sigma_{l'} | O_E(t) | \sigma_l \rangle$ for $l \neq l'$. This feature is stronger if we take an error model where only $C_1$
and $C_5$ have their angles off. In this case, the structure of the superoperator $O_E$ is

$$
\dot{O}^E(t + 1) = \begin{pmatrix}
1 & 0 & 0 \\
0 & A_{t+1} & D_{t+1} & B_{t+1} \\
0 & D_{t+1} & A_{t+1} & -B_{t+1} \\
0 & -G_{t+1} & G_{t+1} & C_{t+1}
\end{pmatrix}
$$

so if we have an initial state of the form (8.6) with $c_0 = 1/2$, the generalized LE-fidelity will be

$$
f^{LE}(t, \Lambda, \bar{c}) = 1/4 + (c_x^2 + c_y^2) A + c_x^2 C + c_x c_y 2D + (c_x - c_y)c_z(B - G) \quad (8.7)
$$

Thus it is clear that, for example, an initial state that is an eigenvalue of $\sigma_x$ (instead of the typical $|0\rangle$) will retrieve a different decay. As discussed above for the RT scheme, this would weaken the usefulness of the LE-fidelity, since the decay will depend on the chosen initial state, and it requires substantial previous knowledge on the error model to choose it conveniently to match the behavior of $F^{E}_e(t)$.

However, when considering the actual values of the decay, we found that the difference between the generalized fidelities in the LE scheme for different initial states is very small and it vanishes with time. That is, the off-diagonal terms of $O^E$ are very small and so is

![Figure 8-7: Plots of the modulus and angle of the eigenvalues characterizing the evolution of the generalized fidelity decay, when the gates $C_1$, $C_5$ and $C_3$ have their angle off. The modulus is denoted by hollow symbols, while the angle is by filled ones. $\lambda_+$ is denoted by $\square$, $\lambda_-$ by $\bigcirc$, $\bullet$, and $\lambda_2$ by $\triangle$.](image)
the difference between the diagonal elements, in the Pauli group basis. As an example, Fig. 8-9 shows this for three different states for the model we just discussed. We found the same behavior in other models.

The difference between the curves in Fig. 8-9 is so small that it cannot be detected by any simulation with \( N \) up to 10,000. We could interpret this as follows. We can regard the different initial states as flawed implementations of |0⟩ (or any other fixed non-random state). Then we find that the decay is indeed robust to faulty preparations of the initial state, as conjectured in [17]. This is an important point, as it is not the case with the RT-fidelity, where we have already found that the state dependence can considerably affect the shape of \( f^{RT} \).

On the other hand, the small values of the off-diagonal elements in \( O^E \) could also potentially suggest that \( O^E \) is actually diagonal in the Pauli group basis, and that the non-null values we have obtained are a consequence of our modeling of the process, specifically eqs. (7.12)-(7.13): as we have already mentioned, the nesting of \( O^E \) at different times is only possible if we neglect the error in the reversal gate, or at most if we approximately account for it at the first step using \( O^E(t = 1) = O^T(t = 1) \) as a start in the nesting. At this point, our simulations (which properly consider a faulty reversal gate at the final step), with \( N < 10,000 \), lack the resolution to settle this issue.

### 8.5 Other models

We also studied numerically other error models. Although the analytical solution for the RT scheme is always easy to obtain (we only need to exponentiate the superoperator for the first step), this is not the case with the LE scheme. Our strategy has been to find a relation, given by a matrix \( \mathbb{M} \), between the matrix elements of \( O^E(t) \) at different times. \( \mathbb{M} \) encodes the action of the nesting, eqs. (7.12)-(7.13), so the time evolution can be obtained by exponentiating it. This strategy is feasible if the number of different matrix elements of \( O^E(t) \) is small, so the size and structure of \( \mathbb{M} \) yields relatively simple expressions for the
eigenvalues. However, when there is not enough symmetry in the error model (or, also, if our parametrization is not the cleverest one), the complexity of \( M \) is such that it is equivalent to just calculate the LE-fidelity numerically. Notice that for one-qubit, the numerical calculation of \( O_1^T = O_E(t = 1) \) (eq. 7.3) and of eqs. (7.12)-(7.13) can be done exactly, as indeed we can run the sums over the 12 Clifford gates, for times up to \( t = 200 \) in less than a minute. Moreover, with the parametrization we have used - which is always in the form of angles, we are guaranteed to explore the full range of error strengths by sampling between \([0, 2\pi]\) (with the zero value corresponding to no error).

We show in Fig. 8-10 the numerical calculation of the generalized fidelities for an error model corresponding to a shift in \( \omega_\delta \) (cf. eq. (8.3)). We also plot the simulation with \( N = 200 \). As the numerical calculation runs over all the Clifford gates, a natural agreement is expected for the RT-fidelity for \( N \) large enough. However, for the LE-fidelity, the comparison with the simulation tests the assumption of neglecting the error in the reversal gate and the validity of eqs. (7.12)-(7.13) as model for the LE scheme. We found indeed a good agreement in this case.

In Fig. 8-11 we show the numerical calculation of the LE-fidelity for an error model where some Clifford gates have their direction rotated, and others do not. This is not a global rotation of frame, thus we expect some decay in the LE scheme. The twirl fidelity remains a flat line, since for an error in the direction, \( \tilde{C}_i \tilde{C}_j = I \) individually for each Clifford. We see that the LE-fidelity again presents a exponential-like decay.

We also ran numerical studies of the rest of the models discussed at the beginning of this
Figure 8-10: Numerical calculations of the fidelity decays produced by an error model considering Clifford gates with $\theta_n z$ disturbed as $(\theta + h)z$. The RT-fidelity is on the left, and the LE-fidelity is on the right. $N = 200$ realizations, and we used an initial state $|\psi_0\rangle = |0\rangle$. The solid line represent the numerical calculations (see text). Error strength: ● $h = 0.02$, □ $h = 0.06$, ◆ $h = 0.1$, * $h = 0.2$, ○ $h = 0.4$, ▲ $h = 1.0$.

Chapter, as well as selective versions (i.e., affecting only some gates and not others, like in Sec. 8.4 and the last example) and also combinations of them. In all the cases we arrived to the same conclusions we have exposed so far. The fact that for a general error (in angle and direction) in only one chosen Clifford gate retrieves an exponential-like decay suggests that the combination of errors in all the gates would return a sum of exponential-like decays. Thus we would still have a linear initial decay and a general exponential shape for $f^{LE}$. However, we must note that we cannot calculate $O^E_j(t)$ arising from a set of faulty Clifford gates as the sum of some partial $O^E_j(t)$ corresponding to having a single faulty Clifford gate $j$. So this is merely an intuitive conjecture.

8.6 Closing remarks

In conclusion, we found that when the errors are gate-dependent, the RT-fidelity and the LE-fidelity are not necessarily the same. This is not only a matter of different amounts of error per step, but also in general they do not have the same behavior.

The RT-fidelity with faulty Clifford gates is highly dependent on the choice of the initial state. Depending also on the error model, $f^{RT}(t, \Lambda, \psi_0)$ can have different behaviors in time. In particular we have found decays that are shifted exponentials of the form $f(t) = (1 - 1/D)e^{-\alpha t} + 1/D$, and damped oscillations like $f(t) = (1 - 1/D) r_T^2 \cos(\beta t) + 1/D$ (with $r_T$, $\alpha > 0$, $\beta$ all real parameters).

The LE-fidelity, on the other hand, presented in all the cases an exponential-like decay of the form $f(t) = (1 - 1/D)r_E^2 + 1/D$ and although analytical results show that we can find damped oscillations, the parameters of it are such that it falls into an overdamped regime that can be seen also as an approximately exponential decay. Our work then supports the conjecture [17, 19] that for weak error strength (so we can guarantee that $r_E > 0$), the LE-fidelity behaves like an exponential.
Figure 8-11: The LE-fidelity $f^{LE}$ for an error model where the Clifford gates $C_1$, $C_2$, $C_3$ and $C_4$ have their direction rotated by an angle $b$ around $\hat{z}$. Here $N = 200$ realizations, and we used an initial state $|\psi_0\rangle = |0\rangle$. The solid line represents the numerical exact calculation (see text). Error strength: • $b = 0.02$, □ $b = 0.06$, ◇ $b = 0.1$, * $b = 0.2$, ◆ $b = 0.4$, ▲ $b = 1.0$.

The LE-fidelity also seems to be either independent of the initial state of the system, or able to absorb the difference between initial states as a small shift that vanishes with time.

These last two features make the LE-fidelity very attractive as a general tool to obtain the fidelities $F(\Lambda)$ and $F_e(\Lambda)$ of the error map $\Lambda$ characterizing the errors in the Clifford twirl: we can fit the decay with a shifted exponential, and then extrapolate the value at $t = 0$, which retrieves the desired fidelities. Knowing the shape of the decay is not as critical as its independence from the choice of initial state. However it is a valuable feature if we are only using a few points of the initial decay for fitting, in which case it is not possible to discriminate clearly between linear or quadratic solely from the data to be fitted. The independence from choice of the initial state (or its negligible effect) is key though. It is only in those cases that we can assure that a curve $f^{LE}(t)$ we measure for a particular state coincides with $(\text{Tr}[O^E(t)]/D + 1)/(D + 1)$, so then by eqs. (2.8) and (2.9), $f^{LE}(t = 1)$ is indeed the average fidelity of the twirl.

Clearly, the open problem at this point is to assure the independence of $f^{LE}$ from the initial state, or at least that the effect of choosing different states is negligible. This is equivalent to showing that $\hat{O}^E$ is diagonal, or approximately diagonal, or with the off-diagonal elements having opposite signs, and all the diagonal elements equal or approximately equal. The second problem would be the exact shape of the curve. This seems to be a more complicated problem, since the curve does not seem to be always an exponential (like it was the case with gate-independent errors), but perhaps an overdamped oscillation that can be nevertheless fitted by an exponential.

We have considered here only unitary errors. It is not difficult to extend this to unital processes, as long as we consider weak errors. Consider that the dynamics in any scheme is given by the eigenvalues $\lambda_k$, which are either the eigenvalues of $\hat{O}_1^T$ for the RT scheme,
or the eigenvalues of the matrix $\mathcal{M}$ characterizing $\hat{O}^E$ (assuming that after a number of steps we can find a recurrence in the structure of $\hat{O}^E$). These eigenvalues will depend on some error strength parameter $c$ (like $a$ in the angle-off model, $b$ in the rotated direction model, etc.). Since the matrix elements of $\hat{O}$ are real (the superoperator is written in the Pauli group basis), and then also $\mathcal{M}$ is real, the eigenvalues will be either real or complex conjugates. Thus in general the decaying part of the trace of $\hat{O}(t)$, for both schemes, would be something like $(\pm 1)^t e^{-g(c)t}$ or $e^{-g(c)t} \cos(h(c)t)$, where $g(c)$ and $h(c)$ are real arbitrary functions of $c$. For general non-unitary but unital processes, we can take that the errors are still represented by unitary operations characterized by an error strength $c$, but with $c$ fluctuating according to some distribution $P(c)$ (cf. the model for a unital map $\Gamma$ in Chapter 4). If the errors are weak, so $g(c)$ and $h(c)$ are small for small $c$, then we can say that for long-time correlated errors,

$$\left\langle e^{-g(c)t} \right\rangle_P = \int P(c)e^{-g(c)t} \, dc \approx 1 - \int P(c)g(c) \, dc \, t + O(g^2) = 1 - \langle g \rangle_P \, t \approx e^{-\langle g \rangle_P t}$$

(8.8)

while for short-time correlated errors

$$\left\langle e^{-g(c)} \right\rangle_P^t = \left( \int P(c)e^{-g(c)} \, dc \right)^t \approx \left( 1 - \int P(c)g(c) \, dc + O(g^2) \right)^t = (1 - \langle g \rangle_P)^t \approx e^{-\langle g \rangle_P t}$$

(8.9)

in the exponential-like regime (where we can take $\cos(h(c)t) \approx 1$). Notice that the approximation in (8.8) is more demanding that the one in (8.9), since the latter requires the net error to be weak considering only one step, while in the former the net error should remain weak when combined after $t$ steps. In general, we can expect the same behavior but governed by the mean values of the characteristic parameters.

Numerical calculations of the previous error models but with their parameters drawn from a Gaussian or a Rectangular distribution confirmed this conjecture. An example is shown in Fig. 8-12, were we simulated having the Clifford gates $C_1$, $C_5$ and $C_9$ with their angles off by an amount $x$ that fluctuates according to a Gaussian distribution $P_G(x, a, \sigma_a) = \exp(-\frac{(x-a)^2}{2\sigma_a^2}) / \sqrt{2\pi}\sigma_a$. These type of simulations were done by sampling the sum over the complete set of 12 Clifford gates. In the case of short-time correlated errors, $x$ is randomly drawn according to $P_G$ for every realization at every step (and kept fixed for the sum over the 12 Clifford gates). In Fig. 8-12 we used $N = 100$ in this case. In the case of long-time correlated errors, the error strength is the same along a realization of $t$ steps, and $x$ fluctuates from one realization to another. To match the sampling of the short-time correlation case, in Fig. 8-12 we took $N = 1200$.

We observe that the fluctuation of the error parameter does not alter the general behavior of the generalized fidelities, although it does change the effective error strength (compared with the decay where $x = a$). If we oversimplify the $x$-dependence of $r$ as $r \approx 1 - \frac{x^2}{2}$ (for small $x$, following Fig. 8-7), we have that the new effective error strength would be $\sqrt{a^2 + \sigma_a^2}$ instead of $x = a$. Thus we expect the decays to be faster, and dominated by the largest of the two parameters. The results of Fig. 8-12 follow this in the LE scheme, when indeed the decay is exponential-like and we can apply the approximation (8.8).
Figure 8.12: (color) Numerical calculations of the LE-fidelity for an error model where the Clifford gates $C_1$, $C_5$ and $C_9$ have their angles offset by an amount $x$ drawn from a Gaussian distribution $P_G(x, a, \sigma_a)$. Initial state $|\psi_0\rangle = |0\rangle$. Top: $\sigma_a = 0.1$; bottom: $\sigma_a = 0.4$. Error strength by symbol shape and color: $\bullet a = 0.04$, $\blacksquare a = 0.07$, $\blacklozenge a = 0.1$, $\blacktriangle a = 0.2$, $\blacklozenge a = 0.4$.

The hollow symbols correspond to a model with short-time correlations ($N = 100$), while the filled ones are with long-time correlations ($N = 1200$). The semi-filled ones shown in the legend are for the unitary transformation (fixed $x = a$).

If we consider $a \sim \sqrt{a^2 + \sigma_a^2}$, then for $\sigma_a = 0.1$, $a = 0.04, 0.07, 0.1, 0.2, 0.4 \sim 0.1, 0.1, 0.1, 0.2, 0.4$; while $\sigma_a = 0.4$ makes $a \sim 0.4, 0.4, 0.4, 0.4, 0.4, 0.6$. The plots show a good agreement with this intuitive idea when the decay is exponential (so particularly for the LE scheme).
Finally, we note that in all the cases we had $f^{RT}, f^{LE} \to 1/D$ when $N \to \infty$. This is a consequence of $|\lambda_k| < 1$ for $k \neq 1$, where $\lambda_1 = 1$ is a result of the unitality of the map (the maximally mixed state is an eigenstate of any unital process). We anticipate that the presence of relaxation or other forms of decoherence in the twirl, that tend to take the state $I/2$ to some $\rho_\infty$, could potentially alter this long-time limit.
Chapter 9

Conclusions

In this thesis we have developed and further studied the characterization of open quantum system dynamics. In particular, we have put an emphasis on the scalability issue that arises when characterizing systems of large size—as desired in quantum information processing.

Our work has concentrated in developing scalable characterization protocols by twirling the map corresponding to the dynamics under study. The key to the scalability has been then the sampling of the twirl, and the scalable implementation of each realization of the sampling ensemble.

Taking the twirling for granted, we concentrated on developing a method that would reveal details of the process beyond its average fidelity. In Chapter 4 found that the fidelity decay under a one-qubit twirl is exponential-like (as with a full twirl in $U(D)$), at least for weak error strength. Moreover, we developed a protocol that was able to extract the magnitude of the multi-body coupling terms in the generator of the process, in the case of a unital noisy channel.

We were then able to extended the protocol to measure the multi-body correlations occurring in any arbitrary quantum process, as presented at the beginning of Chapter 5. This is an important step since we are now able to see further into the dynamics of maps that are not noisy channels or gate+gate reversal operations, even non-CP maps. We implemented our protocol in a liquid-state NMR QIP device, proving its potential and feasibility. More significantly, the experiment also highlighted key points for the design of characterization protocols: that they must account for errors in the implementation (initial state preparation and measurement, and errors in the twirling), but that there is an intrinsic hierarchy of these errors that we can take advantage from. Typically, the degree of imperfection in a gate increases with the number of qubits involved and the length of it. Moreover, characterization protocols require and indeed can assume a certain degree of control (for example, the unlikelihood of high order multi-body terms), but nevertheless they need not (and must not) rely on having a fault-tolerant device.

In the course of the experimental implementation we also deepened our knowledge of liquid-state NMR QIP. In particular, we observed that the measurement process (QST with readout pulses) is a major source of errors. Moreover, although the errors happening in a single experiment are mainly unitary, we observed that the combination of experiments (either to perform QST or in order to sample the twirling) give rise to non-unitary dynamics.
that spoil important features that are relevant to twirls (pure states become mixed, the unitarily invariance is lost). Finally, incoherent processes (due to the ensemble average of coherent dynamics) also contributes significantly to the rise of this non-unitarity.

In parallel to our work, there were also important developments in the QIP community regarding quantum process characterization, using also twirling techniques or similar ideas. We therefore conducted an analysis of the different protocols, which we presented in Chapter 6. We believe that the combination and hierarchization we developed sets the path for a comprehensive and realistic approach to extract the salient features of an arbitrary quantum process.

9.1 Towards efficient experimental quantum process characterization

Before continuing with the summary of the developments presented in this thesis, let’s step back and consider a bigger picture for the characterization of quantum dynamics in QIP. Assume we are interested in implementing a particular operation described by a map \( \Gamma \). Now imagine we can measure the \( \chi \)-matrix of a given process \( \Gamma \) in a QIP device in a scalable and robust way. Then we can use this information in a feedback loop to optimize what may be a faulty map: measure the coefficients, then introduce a change in the parameters, then run the characterization again, and then check whether we have made some improvement in taking the coefficients to their ideal value, so to decide whether to keep the change or reject it. This is the way we search for pulses in liquid-state NMR QIP (a basic concept of control theory, applied in many other settings too) but we would be simply using a quantum computer to “simulate” the operation \( \Gamma \), instead of using a classical one.

At this point, two obstacles are on the way to achieve such a strategy. One is the errors in the characterization protocol itself. The other is the impossibility to measure all the \( \chi \) coefficients. Let’s consider the latter.

If we are looking to implement a typical QIP gate, such operations involve only up to two qubits, giving a corresponding ideal \( \chi \)-matrix with most of its coefficients equal to zero.

If we were guaranteed that the experimental implementation retrieves a flawed but nevertheless completely positive map, we have seen that all the coefficients of the \( \chi \)-matrix can be bound by its diagonal ones, and that in turn we can measure these grouped by subsets (the collective coefficients \( \chi^{col}_{w,v} \) and the \( p_w \)), that we can hierarchically measure in a scalable way. Thus we can use our quantum device on a feedback loop to take all the ideally null coefficients to zero, and the non-null collective ones to their ideal value. To finally perfect the gate, the non-null \( \chi \) can be measured with QPT on only two-qubits: if we are guaranteed that no other qubits participate of the dynamics, we can indeed perform partial QPT.

The remaining issue in this area is the analysis of the information stored in the off-diagonal terms of the \( \chi \)-matrix for a non-CP map, and whether these can be neglected or bounded by some quantity we can determine in an efficient and feasible way. The advantage of the one-qubit twirling strategies over others is that it does provide further insight into the process while requiring a minimum of perfect quantum tasks. Once we have fault-tolerant quantum gates, we could potentially characterize any unknown dynamics with methods that
require twirls in $U(D)$.

Now let’s consider the other obstacle. In Chapters 7 & 8 we turn back to the experimental lessons and launched a study on the characterization of errors acknowledging the use of a faulty device all the way. For this we abandoned the quest for further details and concentrated again on determining the average fidelity of a process, in particular, of the twirling channel we require to use other protocols that give us those details.

Our results indicate that the echo-fidelity could be a robust tool to measure the average fidelity $F$ of the twirling operations. If we compress the reversal gate into a single gate, a value of $F = 1$ signals no error on average in the twirling gates (not in the gates plus their inverse). The usefulness of such tool would be as follows: if we can measure $F$ for the one-qubit twirl of a particular qubit in a $n$-qubit system, using the faulty device we are looking to improve, we could iterate this again in a feedback loop looking to optimize the twirling. Once we have achieved a certain precision, we can go ahead and use this tool to optimize other gates to achieve fault-tolerance as we just described above.

We must acknowledge here that it has been suggested that the LE-fidelity could characterize the twirling in $U(D)$ directly, and also that it can be used to get further insight beyond the average fidelity. This is a possibility, although proving this conjecture is of course a much more ambitious task.

Our work in this area has been preliminary. Our analytical and numerical results back up the potentiality of the LE-fidelity over the RT-fidelity. We have worked though with a limited number of models (in particular, a set of unital errors). There is plenty of room for further research in this area. It is necessary to include a study of non-unital processes. Moreover, it is key to move the analysis of a one-qubit twirl to an $n$-qubit setting, as optimizing the performance of a gate over a subspace does not guarantee that will perform equally well in a larger space, unless we can strongly argue the isolation of the subspace from its complementary.

Also we must add that in our analysis of the LE-fidelity we have left behind any time dependence on the error map $\Lambda$ by assuming $\hat{\Lambda}_{t,s} = \hat{\Lambda}_s$. Such time-dependence implies the presence of memory effects: the error map knows its place in the sequence. This is a more delicate issue that should also be addressed; notice that the one-step protocols are insensitive to this feature.
Appendix A

Measuring probabilities by repeating experiments: Useful statistical properties

Consider a system of dimension $D$. We prepare it in an initial state $|\psi_0\rangle$ that, after implementing a certain process or algorithm, evolves to $\rho_1$. We are interested in measuring the survival probability $\langle \psi_0 | \rho_1 | \psi_0 \rangle$ through a strong measurement in a basis containing the state $|\psi_0\rangle$. For this, we repeat the experiment $N$ times, each time obtaining a random outcome $x_i = 1$ or $0$ ($i = 1, \ldots, N$) with probability $\langle \psi_0 | \rho_1 | \psi_0 \rangle$, corresponding to whether the system was found in $|\psi_0\rangle$ or not. Then, the average

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \quad \xrightarrow{N \to \infty} \quad \mathbb{E}[\bar{x}] = \langle \psi_0 | \rho_1 | \psi_0 \rangle$$

will retrieve the result we want. In this Appendix we expose some useful statistical properties regarding this process.

A.1 The Chernoff bound

The Chernoff inequality states the following. Take a set of independent random variables \{x_i\}, $x_i = 0, 1$ (binary/boolean variables) with probabilities $\Pr[x_i = 1] = \mu_i$, $\Pr[x_i = 0] = 1 - \mu_i$. We calculate the average of $N$ realizations of these variables,

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$$

Notice $\bar{x}$ is another random variable. We want to know how close this average is to its mean value $\mathbb{E}[\bar{x}] = \mu$:

$$\mu := \mathbb{E}[\bar{x}] = \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^{N} x_i \right] = \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}[x_i] = \frac{1}{N} \sum_{i=1}^{N} \mu_i$$

The Chernoff inequality gives a bound for the probability of $\bar{x}$ being bigger than $\mu$ by
an amount $\epsilon$ ($0 < \epsilon < 1 - \mu$):

$$\Pr[\bar{x} - \mu \geq \epsilon] \leq \exp(-2N\epsilon^2) \quad (A.4)$$

This is called the Chernoff bound [75]. The same statement, but now for continuous independent random variables $\{x_i\}$, $0 \leq x_i \leq 1$, is the Hoeffding’s inequality [76], although it is sometimes also referred to as the Chernoff inequality. Here we will work on Hoeffding’s inequality - which of course includes the case when the $\{x_i\}$ are binary.

We present the main proof (Sec. A.1.1), which uses theorems and lemmas presented in Appendix Secs. A.1.3 to A.1.5. All this relies on work developed in [76, 77, 78, 79]. It is then simple to prove (Sec. A.1.2) that

$$\Pr[|\bar{x} - \mu| \geq \epsilon] \leq 2 \exp(-2N\epsilon^2) \quad (A.5)$$

which gives the probability of $\bar{x}$ differing from $\mu$ by an amount $\epsilon$ ($0 < \epsilon < 1$), and it is the most practical form for the bound.

Finally, in Sec. A.2, we present some notes on the practical application of the bound, illustrating the usefulness and limitations of this tool.

### A.1.1 Main proof

\[
\Pr[\bar{x} - \mu \geq \epsilon] = \Pr[\exp(h(\bar{x} - \mu)) \geq \exp(h\epsilon)] \leq E[\exp(h(\bar{x} - \mu))]/\exp(h\epsilon) = E[\exp(h\bar{x})]/\exp(h(\epsilon + \mu)) \quad (A.7)
\]

The first line is true since given a random variable $u$ with probability $P(u)$, then any function of $u$ is also a random variable with the same probability distribution. And $\bar{x} - \mu \geq \epsilon \leftrightarrow \exp(h(\bar{x} - \mu)) \geq \exp(h\epsilon)$ with $h > 0$, since we have just applied a monotonously increasing function to the inequality.

The second line is obtained applying the Markov Inequality (Sec. A.1.3). Now,

\[
E[\exp(h\bar{x})] = E\left[\exp\left(\frac{h}{N}\sum_{i=1}^{N} x_i\right)\right] = E\left[\prod_{i=1}^{N} \exp\left(\frac{h}{N} x_i\right)\right] = \prod_{i=1}^{N} E\left[\exp\left(\frac{h}{N} x_i\right)\right] \leq \prod_{i=1}^{N} \left[1 - x_i + x_i e^{h/N}\right] = \prod_{i=1}^{N} (1 - \mu_i + \mu_i e^{h/N}) \leq \left(\frac{1}{N} \sum_{i=1}^{N} (1 - \mu_i + \mu_i e^{h/N})\right)^N = \left(1 - \mu + \mu e^{h/N}\right)^N \quad (A.11)
\]

The second line is a consequence of the independence of the random variables: if $x_1$ and $x_2$ are random independent variables, then $P(x_1, x_2) = P(x_1)P(x_2)$, which translates to $E[x_1x_2] = E[x_1]E[x_2]$. The third line is obtained using Lemma I (Sec. A.1.4) with $a = 0$ and $b = 1$. The fourth line is obtained using the Arithmetic-Geometric Mean Inequality.
Now we use (A.11) in (A.7) and we obtain
\[
\Pr[\bar{x} - \mu \geq \epsilon] \leq \frac{(1 - \mu + \mu e^{h/N})^N}{\exp(h(\epsilon + \mu))} = F(h)
\] (A.12)

Notice $h > 0$ is a free parameter we introduced, so we can pick it to establish the tightest bound by finding the minimum of $F(h)$,
\[
\Pr[\bar{x} - \mu \geq \epsilon] \leq F(h_{\text{min}}) \leq F(h)
\] (A.13)
where $h_{\text{min}}$ is the value that minimizes $F$. So we calculate $F'$ to find $F'(h_{\text{min}}) = 0$ and $F''(h_{\text{min}}) > 0$, and obtain
\[
h_{\text{min}} = N \log \left( \frac{(1 - \mu)(\epsilon + \mu)}{(1 - \mu - \epsilon)\mu} \right) > 0
\] (A.14)
\[
F(h_{\text{min}}) = \left( \frac{1 - \mu - \epsilon}{1 - \mu} \right)^{N(\epsilon + \mu - 1)} \left( \frac{\mu}{\epsilon + \mu} \right)^{N(\epsilon + \mu)}
\] (A.15)

This is a good bound for $\Pr[\bar{x} - \mu \geq \epsilon]$, but we actually want a bound that is not dependent on $\mu$, which probably will be what we are trying to estimate by calculating $\bar{x}$. In order to do so we re-express $F(h_{\text{min}})$ as $\exp(-N\epsilon^2G(\epsilon, \mu))$. A function $\exp(-v)$ has a maximum where $v$ is minimum, so we look for the minimum of $\epsilon^2G(\epsilon, \mu)$.

First, we minimize it as a function of $\epsilon$, and we find that it should be $\epsilon = 1 - 2\mu$.
\[
G(\epsilon, \mu)|_{\epsilon=1-2\mu} = \log \left( \frac{1 - \mu}{\mu} \right) \frac{1}{1 - 2\mu} = g(\mu)
\] (A.16)

Now we look for the minimum of $g(\mu)$, which happens at $\mu = 1/2$. We evaluate the limit to obtain $g(\mu \to 1/2) = 2$. Thus combining this and (A.12) to (A.16) we have:
\[
\Pr[\bar{x} - \mu \geq \epsilon] \leq F(h_{\text{min}}) = \exp(-N\epsilon^2G(\epsilon, \mu)) \leq \exp(-N\epsilon^2g(\mu)) \leq \exp(-2N\epsilon^2)
\] (A.17)
which proves the bound (A.4).

### A.1.2 Bound including both the lower and upper tails

We define $y_i = 1 - x_i$, $0 \leq y_i \leq 1$. The $\{y_i\}$ are also independent random variables. We have
\[
\bar{y} = \frac{1}{N} \sum_{i=1}^{N} y_i = 1 - \bar{x}
\] (A.18)
and
\[
E[\bar{y}] = 1 - E[\bar{x}] = 1 - \mu
\] (A.19)
We now apply the bound (A.4) to the variable $\bar{y}$, so $\Pr[\bar{y} - (1 - \mu) \geq \epsilon] \leq \exp(-2N\epsilon^2)$. But $\Pr[\bar{y} - (1 - \mu) \geq \epsilon] = \Pr[1 - \bar{x} - (1 - \mu) \geq \epsilon] = \Pr[\mu - \bar{x} \geq \epsilon]$. Therefore
\[
\Pr[-(\bar{x} - \mu) \geq \epsilon] \leq \exp(-2N\epsilon^2)
\] (A.20)
Then we have

\[ \Pr [\lvert \bar{x} - \mu \rvert \geq \epsilon] = \Pr [\bar{x} - \mu \geq \epsilon] + \Pr [-(\bar{x} - \mu) \geq \epsilon] \leq 2 \exp(-2N\epsilon^2) \]  

(A.21)

where each probability is bounded by (A.4) and (A.20) respectively. This gives (A.5), completing the proof.

There are indeed better (tighter) bounds for \( \Pr [\lvert \bar{x} - \mu \rvert \geq \epsilon] \), particularly working out each tail separately. But this is the most practical expression, given its simple dependence on \( N \) and \( \epsilon \), since we probably would like to estimate the number of realizations \( N \) for a certain \( \epsilon \) and a low error probability.

### A.1.3 Markov Inequality

Let \( P(x) \) be the probability distribution for a variable \( 0 \leq x \leq \infty \). Then the probability that \( x \) is larger than certain \( \epsilon > 0 \) obeys the following inequality

\[ \Pr [x \geq \epsilon] \leq E[x] / \epsilon \]  

(A.22)

To prove this, consider the mean value of \( x \),

\[ E[x] = \int_0^\infty xP(x)dx = \int_0^\epsilon xP(x)dx + \int_\epsilon^\infty xP(x)dx \]  

(A.23)

The first integral on the RHS is always positive since \( P \geq 0 \) (it is a probability distribution) and \( x \geq 0 \) by assumption. So it is true that

\[ E[x] = \int_0^\infty xP(x)dx \geq \int_\epsilon^\infty xP(x)dx \]  

(A.24)

But now the remaining integral in the RHS is in the domain \( x \geq \epsilon \), so

\[ \int_\epsilon^\infty xP(x)dx \geq \epsilon \int_\epsilon^\infty P(x)dx = \epsilon \Pr [x \geq \epsilon] \]  

(A.25)

So finally

\[ E[x] = \int_0^\infty xP(x)dx \geq \epsilon \Pr [x \geq \epsilon] \]  

(A.26)

which completes the proof.

### A.1.4 Lemma I

Let \( f(x) \) be a continuous function with \( f''(x) > 0 \) in the interval \( x \in [a,b] \). Then \( f(x) \) is convex in that interval. We can then bound the function in \( [a,b] \) by a straight line that coincides with \( f(x) \) at the ends of the interval:

\[ f(x) \leq \frac{f(b) - f(a)}{b - a}(x - a) + f(a) \]  

(A.27)

To apply this in the main proof, consider the convex function \( f(x) = e^{hx} \), so (A.27)
becomes
\[ e^{hx} \leq \frac{b - x}{b - a} e^{ha} + \frac{x - a}{b - a} e^{hb} \]  
(A.28)

### A.1.5 Arithmetic-Geometric Mean Inequality

The Arithmetic-Geometric Mean Inequality states that for any set of \( N \) nonnegative real numbers \( \{a_n, n = 1, 2, \ldots, N\} \), the following inequality holds:

\[ \frac{1}{N} \sum_{i=1}^{N} a_n \geq \left( \prod_{i=1}^{N} a_n \right)^{1/N} \]  
(A.29)

The quantity on the LHS is the arithmetic mean of the set, and the quantity on the RHS is called the geometric mean.

This theorem is proved first for \( N = 2 \) numbers, then for any power of \( 2, \ N = 2^k \), and finally by induction it is shown that if it is valid for a value \( N \), then it is valid for \( N - 1 \). Therefore it is valid for any \( N \). The complete explicit proof is simple but long, so we refer the reader to [79].

### A.2 Practical use of Chernoff Bound

It is very important to notice that the Chernoff bound is a bound on the error probability and not on the error itself. Essentially, it tells us that if we perform an average of \( N \) realizations, the probability that \( \bar{x} \) will differ from its mean value \( E[\bar{x}] \) by \( \epsilon \) or more is less or equal than \( \delta(\epsilon, N) = 2 \exp(-2N\epsilon^2) \). The bound is independent of details of the variables \( x_i \) - as long as they are random independent variables with \( 0 \leq x_i \leq 1 \). If the \( \{x_i\} \) represent different realizations of the same phenomenon, so to have \( E[x_i] = \mu_i = \mu \ \forall \ i \), then \( \bar{x} \) represents our attempt to measure their mean value \( \mu \).

It is also of practical interest that although the form of the bound \( \delta(\epsilon, N) \) is very simple, it may predict a very large number of realizations, which not only depends on \( \epsilon \) but on the bound we want to impose. It is reasonable to require at least \( \delta = 0.01 \) (a 1\% chance of failure), in which case

\[ N = \frac{\log(2/\delta)}{2\epsilon^2} = \frac{\log(200)}{2\epsilon^2} \]  
(A.30)

For \( \epsilon = 0.05 \) we would have \( N \approx 1,060 \), while for \( \epsilon = 0.01 \) it would be \( N \approx 26,500 \). If we require \( \delta = 0.001 \), then for \( \epsilon = 0.05 \) we would have \( N \approx 1,520 \), while for \( \epsilon = 0.01 \) it would be \( N \approx 38,000 \). Whether the resulting \( N \) is a feasible number or not will depend exclusively on the details of the experimental setup.

To see how the bound operates, let’s take the ubiquitous example of the tossing of a fair coin. With probability 1/2 we will obtain heads (\( x = 1 \)), and with probability 1/2 we will obtain tails (\( x = 0 \)). Clearly \( E[x_i] = \mu_i = \mu = 1/2 \ \forall \ i \).

Now we simulate the tossing numerically, using pseudo-random number generation in MATLAB. We take \( N = 150 \) and \( N = 1,000 \). To illustrate the statistics, we plot the results for many trials of this process.

The following figures show the histograms built with the results from many trials of \( \bar{x} \) with a given \( N \). Fig. A-1 presents the results for \( \bar{x} \) calculated with \( N = 150 \) realizations,
and Fig. A-2 has $N = 1,000$. In both cases we show the statistics first with 100 trials of $N$, and then with 10,000 trials of $N$, to confirm the exhibited behavior.

**Figure A-1:** Coin tossing statistics, using $N = 150$ realizations and 100 (left) or 10,000 (right) trials. The solid bars represent the results from the numerical simulation: a histogram that shows the cumulating $\Pr[|\bar{x} - \mu| < \epsilon]$. The histogram is normalized to 20 non-empty bins. The solid line with ■ represents the Chernoff bound: $2 \exp(-2N\epsilon^2)$.

**Figure A-2:** Coin tossing statistics, using $N = 1000$ realizations and 100 (left) or 10,000 (right) trials. The solid bars represent the results from the numerical simulation: a histogram that shows the cumulating $\Pr[|\bar{x} - \mu| < \epsilon]$. The histogram is normalized to 20 non-empty bins. The solid line with ■ represents the Chernoff bound: $2 \exp(-2N\epsilon^2)$.

The plots show the (cumulating) probability $\Pr[|\bar{x} - 1/2| > \epsilon]$ vs. $\epsilon$, that is, the fraction of events (tossing of $N$ coins) that gave an $\bar{x}$ differing from $1/2$ in more than $\epsilon$. At $\epsilon = 0$, this tends to 1 of course. When moving from low to high $\epsilon$, this probability decreases: it loses an amount of probability equivalent to the fraction of events registered in the given bin at that particular $\epsilon$ (the size of the bins is reported in the caption). The Chernoff bound is fulfilled when for a particular $\epsilon_B$ of our interest, $\Pr[|\bar{x} - 1/2| > \epsilon_B]$ drops below $\delta(\epsilon_B, N)$.

Of course the Chernoff bound $\delta(\epsilon_B, N)$ is lower as $N$ is larger, which in turn means that the probability of getting a value of $\bar{x}$ differing from $\mu = 1/2$ in more than a certain $\epsilon$ becomes lower. On the other hand, at a fixed $N$ the bound becomes higher as $\epsilon$ becomes
lower. This is illustrated with the red solid lines, that show \( \delta(\epsilon_B, N) \) for different \( \epsilon_B \). Notice that for small \( \epsilon_B \), \( 2e^{-2N\epsilon_B^2} > 1 \). This is a consequence of overestimating the bound to include both tails (see Sec. A.1.2). Also, in general for small \( N\epsilon_B^2 \), the Chernoff bound is indeed an overestimation (which we could see already in the mathematical deduction), in particular because of the way we made the inclusion of both tails.

### A.3 Mean and variance of \( N \) realizations of a Bernoulli variable

We go back to the particular study of binary variables. Following [80] Sec. 5.1, we will consider the simplest case where \( \Pr[x_i = 1] = p \), \( \Pr[x_i = 0] = 1 - p \) \( \forall i \) with \( 0 \leq p \leq 1 \). These are called Bernoulli variables, and they represent the general case of tossing a biased coin (when \( p = 1/2 \) it reduces to the fair coin case introduced above). We have that 

\[
E[x_i] = p.
\]

Also, \( E[x_i^2] = p \) too, since \( x_i^2 = x_i = 0, 1 \). Then the variance 

\[
\text{Var}[x_i] = E[x_i^2] - (E[x_i])^2 = p - p^2 = p(1 - p).
\]

We are interested in the statistics of the random variable \( \bar{x} \) given by eq. (A.2). If we now consider what is the probability of \( \bar{x} \) being \( n_1/N \), with \( n_1 \) the number of times the \( x_i \) turned up to be 1, we have

\[
\Pr[\bar{x} = n_1/N] = \binom{N}{n_1} p^{n_1}(1 - p)^{N-n_1}
\] (A.31)

where the binomial coefficient \( \binom{N}{n_1} \) gives us the number of ways we can have \( n_1 \) ones in a series of \( N \) realizations. Such event will have a probability \( p^{n_1}(1 - p)^{N-n_1} \) which comes from multiplying the probabilities of all the times the throws gave \( x_i = 1 \) and of the remaining \( N-n_1 \) throws that resulted in \( x_i = 0 \). We see then that the variable \( \bar{x} \) arising from averaging Bernoulli variables obeys a binomial distribution. This distribution has well-known mean and variance:

\[
\begin{align*}
E[\bar{x}] &= \frac{1}{N} \sum_{i=1}^N E[x_i] = \frac{1}{N} \sum_{i=1}^N p = p \\
E[\bar{x}^2] &= \frac{1}{N^2} \sum_{i,j=1}^N E[x_i x_j] = \frac{1}{N^2} \sum_{i\neq j} E[x_i] E[x_j] + \frac{1}{N^2} \sum_{i=1}^N E[x_i^2] = \frac{(N^2 - N)p^2}{N^2} + \frac{p}{N} \\
\text{Var}[\bar{x}] &= \frac{(N^2 - N)p^2}{N^2} + \frac{p}{N} - p^2 = \frac{p(1 - p)}{N}
\end{align*}
\] (A.32)

So the error arising from statistics is the dispersion 

\[
\sigma_{\bar{x}} = \sqrt{\text{Var}[\bar{x}]} = \sqrt{p(1 - p)/N} \leq 1/\sqrt{N}.
\]

Moreover, as \( N \) becomes large, the binomial distribution (A.31) approximates a Gaussian distribution with mean \( p \) and variance \( p(1 - p)/N \), a result known as the Central Limit Theorem ([81] Chap. VII). This is illustrated in Fig. A.3, where we have plotted the histograms corresponding to the data of Figs. A-1 & A-2. We have normalized the histograms so they indicate the fraction of the \( N \) events that fell into the given bin corresponding to that particular \( \bar{x} \) (the size of the bins is reported in the caption). The sum of the height of the 20 bins gives of course 1. We have also plotted the corresponding approximating Gaussian distribution, with mean \( \mu = 1/2 \) and dispersion \( 1/\sqrt{2N} \), showing the agreement. Notice
how the distribution becomes narrower with larger $N$ (and independently of the number of trials of course).

![Histograms](image)

Figure A-3: Coin tossing statistics, using $N$ realizations as indicated (same data source as Figs. A-1 & A-2. From left to right, we increase the number of trials. The solid bars represent the results from the numerical simulation: a histogram that shows $P(\bar{x})$. The histogram is normalized to 20 non-empty bins (size of the bins for $N = 150$: 0.0102 (left), 0.0161 (right); for $N = 1000$: 0.0043 (left), 0.0023 (right)). The ⋄ represent the corresponding Gaussian approximation, which clearly becomes better for large $N$.

The Central Limit Theorem tells us that in general that whatever the probability distribution governing the $x_i$ (as long as it is the same for all), for $N \rightarrow \infty$, $\bar{x}$ will tend to a Gaussian distribution with mean $E[x_i]$ and dispersion $\sqrt{\text{Var}[x_i]}$ [82]. In our particular case the mean value and variance of $\bar{x}$ can be calculated directly without resorting on $N \rightarrow \infty$, eqs. (A.32) and (A.33). However the Central Limit Theorem grants us the statistical properties of a Gaussian distribution, for example its asymptotic confidence intervals and its simple integrability.

### A.4 Application to the measurement of probabilities in twirling protocols

The statistics of tossing a biased coin are completely equivalent to the strong measurement of the state of a two-level system. The probability $\text{Pr}[x_i = 1] = p$ of getting heads is equivalent to the probability $\text{Pr}[\rho_1 = \ketbra{\psi_0}{\psi_0}] = \langle \psi_0 | \rho_1 | \psi_0 \rangle = \mu$. However we must be
careful and acknowledge the two samplings we are doing here. We are sampling not only the measurement of a state, but also the twirl.

Each $x_i$ is the outcome of the strong measurement of $\rho_i = C_i^T \Gamma (C_i |\psi_0 \rangle \langle \psi_0 | C_i^T) C_i$ onto a basis containing the state $|\psi_0\rangle$: if we get $|\psi_0\rangle$ then $x_i = 1$, otherwise $x_i = 0$. $C_i$ is a Clifford gate, $\Gamma$ is the (gate-independent) error map we are characterizing, and the system is initially prepared in the $|\psi_0\rangle$ state. So in this case, the $\Pr[x_i = 1] = \mu_i = \langle \psi_0 | \rho_i | \psi_0 \rangle$ are different at each realization in principle. If the $C_i$ are randomly chosen from a Clifford-element pool of size $|C|$, we will encounter $|C|$ different values of $\mu_i$. In any case, we are facing the sampling of independent random binary variables $x_i$ that obey different distributions. Each of them has a mean value $E[x_i] = \mu_i$ and a variance $\text{Var}[x_i] = \mu_i(1 - \mu_i)/\sqrt{N} \lesssim 1/\sqrt{N}$. We use then a stronger version of the Central Limit Theorem ([81] Sec. X.5), which ensures us that even when the $x_i$ are governed by different probability distributions, the probability distribution for $\bar{x}$ with $N \to \infty$ will tend to a Gaussian distribution with mean $\sum_i E[x_i]/N$ and variance $\sum_i \text{Var}[x_i]/N^2$, as long as the $x_i$ are independent, uniformly bounded\(^5\) and $\sum_i \text{Var}[x_i] \to \infty$.

We notice that $\sum_i\langle \psi_0 | \rho_i | \psi_0 \rangle = \langle \psi_0 | \sum_i \rho_i | \psi_0 \rangle$, so $\sum_i E[x_i]/N$ will tend to $\mu = \langle \psi_0 | \rho_1 | \psi_0 \rangle$, with

$$\rho_1 = \frac{1}{|C|} \sum_{i=1}^{|C|} \rho_i = \frac{1}{|C|} \sum_{i=1}^{|C|} C_i^T \Gamma (C_i |\psi_0 \rangle \langle \psi_0 | C_i^T) C_i$$

(A.34)

So the mean will indeed tend to the fidelity decay with a perfect twirl. On the other hand, the different variances $\text{Var}[x_i]$ are bounded by $1/N$, so we can assure that the variance in the estimation of the mean will also be bounded by $1/N$. Moreover, we can expect that for large $N$, the variance will go like $\mu(1 - \mu)/N$ (if all the $x_i$ obeyed the same distribution, this would be exact).

We illustrate this in Figs. A-4 and A-5. In Figs. A-4 we have simulated the sampling of a twirl for one qubit and of the strong measurement of the projection onto the initial state, for each element of the ensemble.

We took a simple error model, $\Gamma (\rho) = e^{-i\alpha x} \rho e^{i\alpha x}$, and prepared the initial state to be $|0\rangle$ in the Zeeman basis. Thus according to eq. (5.7), we have

$$\langle 0 | \rho_1 | 0 \rangle = \mu = 1 - \gamma = 1 - 2|\eta_{x}|^2/3 = 1 - 2\sin^2(\alpha)/3,$$

and the variance should be then $\sigma^2 = (1 - 2\sin^2(\alpha)/3)(2\sin^2(\alpha)/3)/N$. We see in Figs. A-4 that the plots show the expected Gaussian behavior. In both cases ($\alpha = \pi/6$ and $\alpha = 0.1$) we tried a sampling over the subsets with $|C| = 12$ and 24 (c.f. Appendix Sec. B.3), obtaining the same agreement.

Naturally, when twirling only one qubit, there is not much point in taking a “large $N$” since the twirl can be implemented exactly with $|C| = 12$ realizations, although the sampling works anyways. We thus simulated the measurement of the $\gamma$’s ($\gamma = 1 - \mu$) taking $\Gamma$ to be a perfect CNOT gate acting on qubits $1 \& 2$ within three qubits. Figs. A-5 report the measurement of $\gamma^{(1)}$, $\gamma^{(2)}$, $\gamma^{(1,2)}$ and $\gamma^{(1,2,3)}$ using strong measurements and sampling over the 3-fold tensor products of the 12 Clifford gates for one qubit (thus a pool of $12^3 = 1,728$ Clifford gates). We see that $N = 1,000$ already gives us enough statistics to show the limit statistical behavior.

\(^5\)The $x_i$ are uniformly bounded if there exist a constant $A$ such that $|x_i| < A$ for all $i$. This is indeed true for binary variables.
The Chernoff Bound, on the other hand, can be applied without further ado, since our proof already considered the case where the $x_i$ had different expectation values $\mathbb{E}[x_i] = \mu_i$.

Summarizing, by sampling the twirl together with the strong measurement, we will find that the mean value of the sample $\bar{x}$ converges towards the survival probability we look for: $\mu = \langle \psi_0 | \rho_1 | \psi_0 \rangle$, with $\rho_1$ given by eq. (A.34). Moreover, the distribution will tend to behave like a Gaussian distribution as $N \to \infty$. The dispersion $\sigma$ is expected to go like $\sigma^2 \approx \mu(1-\mu)/N$, and can be rigorously bounded as $\sigma \leq 1/\sqrt{N}$. Thus for a desired precision $\epsilon$, we must require that $1/\sqrt{N} \leq \epsilon$ so $N \geq \epsilon^{-2}$.

On the other hand, if we have a trial of $N$ realizations and we calculate the mean value $\bar{x}$ of it, the probability of it differing from $\mu$ in more than $\epsilon$ is bounded by the Chernoff Bound: $\Pr[|\bar{x} - \mu| > \epsilon] \leq \delta = 2e^{-2N\epsilon^2}$. Since we also want $\delta << 1$, this gives another constraint for $N$, which is $N \geq \log(2/\delta)/(2\epsilon^2)$.

These two conditions do not compete with each other, since one is a bound to the variance of the distribution (in the $N \to \infty$ limit), while the other is a bound to the error probability, (which is a closed result valid for arbitrary $N$). Since we want to fulfill both conditions, we must observe that the Chernoff bound imposes a stronger bound on $N$ when the desired error probability $\delta$ is $\delta < 2e^2 \approx 0.27$. Thus by fulfilling the Chernoff bound with $\delta < 0.2$ we are already guaranteed that the dispersion of the mean value will be $\sigma \leq 1/\sqrt{N}$ in the large $N$ limit.

This concludes our presentation of the statistics of the strong measurement of probabilities, not only taking into account the randomness in the quantum measurement but also acknowledging the approximate implementation of a twirl by sampling.
Figure A-5: Statistics of the sampling of the twirl plus strong measurement, for a one-qubit twirl of three qubits. \( N = 1,000 < |\mathcal{C}| = 12^3 \). Working with the same data in all the cases and taking partial trace over the outgoing \( \rho_i \) in the simulations, we measure the coefficients of a CNOT gate on qubits 1 & 2. According to eq. (5.7), we have
\[
\gamma^{(1)} = 2(|\eta_1|^2 + |\eta_{1,x}|^2)/3 = 0.333 \\
\gamma^{(2)} = 2(|\eta_2|^2 + |\eta_{2,x}|^2)/3 = 0.333 \\
\gamma^{(1,2)} = 2(|\eta_1|^2 + |\eta_{2,x}|^2)/3 + 8|\eta_{1,x}|^2/9 = 0.556 = \gamma^{(1,2,3)}
\]
The solid bars represent the results from the numerical simulation: a histogram that shows \( P(\bar{x}) \), normalized to 20 non-empty bins. The \( \bullet \) represent the Gaussian distribution given by the expected \( \mu = 1 - \gamma \) and \( \sigma^2 \approx \mu(1 - \mu)/N \).
Appendix B

Generalities on the Clifford twirl

B.1 Equivalence between the Clifford twirl and the Haar twirl

The work by Dankert et al. [15] introduced the concept of unitary t-designs. Following [83, 84], we define a unitary t-design as a set \( \{ U_k, k = 1, \ldots, K \} \subset U(D) \) of unitary operators such that

\[
\int_{U(D)} P_t(U) dU = \frac{1}{K} \sum_{k=1}^{K} P_t(U_k)
\]

(B.1)

where \( P_t \) is a polynomial function of degree \( t \) in both the matrix elements of \( U \) and of \( U^\dagger \). In the above \( dU \) denotes the Haar measure on the unitary group \( U(D) \), so \( \int dU = 1 \). Finding the set of \( U_k \) that fulfill this is not a trivial task, in particular for arbitrary \( t \) and for any dimension \( D \) (see discussion in [83]).

We are particularly interested in 2-designs, as they include the operation we have called twirling all through this thesis. In [15, 85] it was proven that the Clifford group of gates \( C_k \) constitutes a 2-design as follows

\[
2T(A, X, B) = \int_{U(D)} dU U^\dagger A U X U^\dagger B U
\]

\[
2T_C(A, X, B) = \frac{1}{|C|} \sum_{k=1}^{|C|} C_k^\dagger A C_k X C_k^\dagger B C_k
\]

\[
2T(A, X, B) = 2T_C(A, X, B)
\]

(B.2)

for any matrices \( A, X \) and \( B \) in a Hilbert space with dimension \( D = 2^n \) (an arbitrary power of 2, or alternatively, a \( n \)-qubit space).

The Clifford group on \( U(D) \) is defined as the set of unitary operators \( C_k \) such that

\[
\{ C_k \in U(D) \mid C_kPC_k^\dagger = P' \}
\]

(B.3)

where \( P, P' \) are elements of the (generalized) Pauli operators excluding the Identity operator.
that is \( \{ P_l, l = 1, \ldots, D^2 - 1 \} \), in \( U(D) \):

\[
P_l = \bigotimes_{j=1}^{n} P_l^{(j)}
\]  

(B.4)

where in turn each \( P_l^{(j)} \) is an element of the Pauli group \( \{ I, \sigma_x, \sigma_y, \sigma_z \} \) for the \( j \)-th qubit, but at least one of them is a Pauli matrix \( \{ \sigma_x, \sigma_y, \sigma_z \} \). We have removed the Identity operator because otherwise any unitary operator would then fulfill the defining condition.

The set of operators that fulfill eq. (B.3) form a group under matrix multiplication, i.e. the product of two of them is also a Clifford operator. This also applies to the generalized Pauli operators \( \{ P_l, l = 0, 1, \ldots, D^2 - 1 \} \). Given eq. (B.3), the Clifford group is said to be the normalizer of the Pauli group, and the Pauli group is a normal subgroup of the Clifford group.

Implementing a Haar twirl as defined on the LHS of eq. (B.2) would require sampling the integral over an infinite pool of random rotations \( U \). The equivalence with the Clifford twirl allows us to implement the twirling exactly with a finite set of \( |C| \) gates. Nevertheless, although not infinite, \( |C| \) is of order \( O \left( 2n^2 \right) \) (see [86, 87] for a explicit formula of \( |C| \)). It is possible to find a subset of the Clifford group that will constitute an exact 2-design in \( U(D) \), or even construct a different set of operations that will do so. However, as demonstrated in [83], the size of such a set is at least \( O(D^4) \).

In conclusion, if we want to implement the twirling efficiently, we have to apply a sampling strategy to do it approximately in any case.

### B.2 Details on the Clifford gates

The Clifford operators will be in general the unitary operators \( C_k \) satisfying eq. (B.3). Up to a global phase factor, which is irrelevant for the purposes of twirling (any arbitrary phases in the \( C_k \) evidently cancel out in eq. (B.2)), any Clifford gate can be decomposed into a sequence of \( O(n^2) \) one- and two-qubit gates. In particular, they can be decomposed into the combination of \( O(n^2) \) steps of solely 3 gates, \( H \), \( P \) and \( CNOT \):

\[
H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \text{Hadamard gate}, \quad P = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} = (\text{single - qubit}) \text{ Phase gate}
\]

\[
CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \text{controlled - NOT gate}
\]

This was demonstrated in [57] Sec. 5.8, where the normalizer of the Pauli group (this is, the Clifford group) plays a major role in quantum error correction codes - in particular the stabilizer circuits are comprised of Clifford operations. It is also concisely stated in [88]
Quantum computations using Clifford gates only can be simulated efficiently in a classical computer. This result is known as the Gottesman-Knill theorem ([2] Sec. 10.5.4). We must notice that this is not enough for universal quantum computation. In particular, the gates $H$, $P$ and CNOT do not form a set of universal quantum gates: for that we would require these gates plus the $\pi/8$-gates [55], or alternatively the use of ancillary qubits and magic state distillation [89, 90].

The fact that computations with Clifford gates can be simulated efficiently has great potential when working with many-step fidelity decay under randomization: the reversal gates for a sequence of Clifford operators can be compressed into one gate resulting from an efficient classical simulation. For further details on the classical simulation algorithms and discussion on the Gottesman-Knill theorem, see [91, 92].

### B.2.1 Clifford gates for one qubit

Up to a global phase, we can write the Clifford operators $C$ for one qubit as products $PS$ or products $SP$ where $P$ is an element of the Pauli group $P = \{I, \sigma_x, \sigma_y, \sigma_z\}$, and $S$ is an element of $S = \{\exp(-i\nu(\pi/3)(\sigma_x + \sigma_y + \sigma_z)/\sqrt{3}), \nu = 0, 1, 2; \exp(-i\pi/4)p, p = x, y, z\}$ (these operators form the so-called Symplectic group, although they are not a group with respect to matrix multiplication). This set of 24 operations forms a group, so the product of any of them is another element of the set (up to a global phase)$^\dagger$.

Let’s parameterize the Clifford operators as $C_k = \exp(-i\theta k/2)\hat{n}_k \cdot \hat{\sigma}$. The following Table B.1 lists all the Clifford gates for one-qubit arising from the products $SP$, up to a global phase. We also list the corresponding reversal gate, and the $PS$ and $SP$ decompositions. With this parametrization it’s easy to visualize the Clifford operators in a sphere of radius 1, where each point corresponds to a unit vector $\hat{n}_k$ in that direction. Then we can see that rotations around these directions, in angles $\theta_k$, take the axes $x$, $y$ and $z$ to one of themselves (the axes would represent the Pauli matrices).

In the literature we can find the $PS$-product expression for the set of the first 12 Clifford gates (which do not form a group) in [16, 17, 19, 22, 93] for example. Alternative expressions to the ones in Table B.1 can be found in [93].

We observe the following transformations

\[
\sigma_p \sigma_j \sigma_p = \sum_q -(-1)^{\delta_j,p} \sigma_q \tag{B.5}
\]

\[
\exp\left(\frac{i\pi}{4}\sigma_p\right) \sigma_j \exp\left(-\frac{i\pi}{4}\sigma_p\right) = \sum_q \sigma_q \epsilon_{j,p,q} + \delta_{j,p} \sigma_j \tag{B.6}
\]

\[
\exp\left(\frac{i\pi}{3}\sigma_x + \frac{\sigma_y + \sigma_z}{\sqrt{3}}\right) \sigma_j \exp\left(-\frac{i\pi}{3}\sigma_x + \frac{\sigma_y + \sigma_z}{\sqrt{3}}\right) = \sum_q \sigma_q \epsilon_{j,q,j+2} \tag{B.7}
\]

\[
\exp\left(\frac{i\pi}{3}\frac{2\sigma_x + \sigma_y + \sigma_z}{\sqrt{3}}\right) \sigma_j \exp\left(-\frac{i\pi}{3}\frac{2\sigma_x + \sigma_y + \sigma_z}{\sqrt{3}}\right) = \sum_q \sigma_q \epsilon_{j,j+1,q} \tag{B.8}
\]

$^\dagger$The size of the Clifford group for $U(2)$ is actually $8 \times 24$ [86, 87]. The factor 8 accounts for the 8 possible different phases of the 24 operators we have introduced. We will disregard these phases, since the twirl involves $C_3$ and $C_3^\dagger$ at the same time (so phases cancel out). So the 24 operators form a group with respect to matrix multiplication up to a global phase.
For any $p, j = x, y, z$. $\epsilon_{j,p,q}$ is the Levi-Civita symbol $(1 = x, 2 = y, 3 = z)$.

As mentioned in [93] and consistently with the above presentation of the Clifford operators, the Clifford gates for one qubit can be obtained as a product of the basic computational gates $H$ and $P$. We show an example of such decomposition in Table B.2.

### B.3 Specifics on the Clifford twirl in $U(2)$

Consider the Clifford twirl as defined in eq. (B.2) but in $U(2)$. Without losing generality, we will take $A = P_a$ and $B = P_b$ two operators from the Pauli group in $U(2)$ chosen arbitrarily

$$
\int_{U(2)} dUU^\dagger P_aUXU^\dagger P_bU = \frac{1}{24} \sum_{k=1}^{24} C_k^\dagger P_aC_kXC_k^\dagger P_bC_k
$$

(B.9)
<table>
<thead>
<tr>
<th>Clifford</th>
<th>Decomposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>$(-i)PHPHP^2H$</td>
</tr>
<tr>
<td>$C_2$</td>
<td>$PHP$</td>
</tr>
<tr>
<td>$C_3$</td>
<td>$(-1)P^2HP^2HPH$</td>
</tr>
<tr>
<td>$C_4$</td>
<td>$HP^2HPHP^3$</td>
</tr>
<tr>
<td>$C_5$</td>
<td>$HP^2$</td>
</tr>
<tr>
<td>$C_6$</td>
<td>$(-i)P^2HP$</td>
</tr>
<tr>
<td>$C_7$</td>
<td>$P^2H$</td>
</tr>
<tr>
<td>$C_8$</td>
<td>$(-i)H$</td>
</tr>
<tr>
<td>$C_9$</td>
<td>$(-i)PHPHPHP$</td>
</tr>
<tr>
<td>$C_{10}$</td>
<td>$(-i)HP^3HPH$</td>
</tr>
<tr>
<td>$C_{11}$</td>
<td>$P^2HP^3HPH$</td>
</tr>
<tr>
<td>$C_{12}$</td>
<td>$P^2HPHPHP^2$</td>
</tr>
</tbody>
</table>

Table B.2: Decomposition of the first 12 Clifford gates for one qubit into the elementary Hadamard and Phase gates (see text). Note the decomposition presented here is just an example; it's not unique nor proven to be minimal in the number of gates used.

Since the Pauli group forms a basis, any arbitrary $A$ and $B$ can be decomposed in terms of operators $P_i$, therefore eq. (B.9) is the basis for the twirl of arbitrary operators.

We now split $C_k$ as products $P_mS_j$, where $m = \{0, 1 = x, 2 = y, 3 = z\}$ spans the Pauli group, and the $S_j$ are the six Symplectic operators in $S$ (described above; take $S_3 = S_x$, $S_4 = S_y$, $S_5 = S_z$):

$$
\frac{1}{24} \sum_{j=0}^{5} \sum_{m=0}^{3} S_j^\dagger P_m P_a P_m S_j X S_j^\dagger P_m P_b P_m S_j := 2T_C(P_a, X, P_b) \quad (B.10)
$$

Each term gives $S_j^\dagger P_m P_a P_m S_j X S_j^\dagger P_m P_b P_m S_j = \pm S_j^\dagger P_a S_j X S_j^\dagger P_b S_j$, with $\pm 1$ depending on the relation between $m$, $a$ and $b$. If $a = b$, then we definitely get $a = +1$. If now $a \neq b$, then $P_m = I$ will give $+1$. If $P_m = P_a$ or $P_m = P_b$, then we get $-1$. Finally, if $P_m \neq P_a$ and $P_m \neq P_b$, then we get $a = +1$ again. These are all the possible cases for one qubit. Thus when $a \neq b$, the 4 different options cancel out. Therefore

$$
2T_C(P_a, X, P_b) = \delta_{a,b} \frac{1}{6} \sum_{j=0}^{5} S_j^\dagger P_a S_j X S_j^\dagger P_a S_j \quad (B.11)
$$

$$
= \delta_{a=0} X + \delta_{a \neq 0} \frac{1}{6} \sum_{j=0}^{5} S_j^\dagger P_a S_j X S_j^\dagger P_a S_j \quad (B.12)
$$

We now split the $S_j$ in $S$ as follows: they can be taken from either $S_1 = \{S_p = \exp(-i(\pi/4)\sigma_p), \ p = x, y, z\}$
or $S_2 = \{ S_{\nu} = \exp (-i\nu(\pi/3)\frac{\sigma_x + \sigma_y + \sigma_z}{\sqrt{3}}),\, \nu = 0, 1, 2 \}$:

$$2T_C(P_a, X, P_b) = \delta_{a,b}^0 X + \delta_{a,b}^{q \ne 0} \left( \frac{P_a X P_a}{6} + \frac{1}{6} \sum_{\nu=1,2} S_{\nu}^j P_a S_{\nu} X S_{\nu}^j P_a S_{\nu} + \frac{1}{6} \sum_{p=x,y,z} S_{\nu}^j P_a S_{\nu} X S_{\nu}^j P_a S_{\nu} \right)$$

where the first term in the sum corresponds to the operator in $S_2$ with $\nu = 0$.

Following eq. (B.6), the piece in $S_1$ will yield $\sum_{\nu=1,2} S_{\nu}^j P_a S_{\nu} X S_{\nu}^j P_a S_{\nu} = \sum_q P_q X P_q (\epsilon_{a,p,q})^2$ if $a \neq p$, and $P_a X P_a$ when $p = a$. On the other hand, following (B.7-B.8), the piece in $S_2$ gives $\sum_{\nu=1,2} S_{\nu}^j P_a S_{\nu} X S_{\nu}^j P_a S_{\nu} = \sum_q P_q X P_q (\epsilon_{a,q,a+2} + \epsilon_{a,a+1,q})$. This means that each piece returns $\sum_{q=x,y,z} P_q X P_q / 6$. We notice then that the Symplectic twirl is redundant: we only need to take $S$ from $S_1$ or $S_2$, as both give the same result. In any case,

$$2T_C(P_a, X, P_b) = \frac{1}{12} \sum_{\delta_{a,b}^0 \neq 0} \sum_{m=0}^3 S_{\nu}^j P_m P_a P_m S_j X S_{\nu}^j P_m P_m S_j$$

$$= \delta_{a,b}^0 X + \delta_{a,b}^{q \ne 0} \frac{1}{3} \sum_{q=x,y,z} P_q X P_q$$

(B.13)

We can check now that eq. (B.13) is indeed equivalent to performing a Haar twirl. For that, we take the results from [94, 95, 96]

$$\int \text{Tr}[U^\dagger P_a U X U^\dagger P_b U Y]dU = \frac{\text{Tr}[P_a] \text{Tr}[P_b]}{3} \left( \frac{\text{Tr}[X \text{Tr}[Y]]}{2} - \frac{\text{Tr}[X Y]}{2} \right)$$

$$+ \frac{\text{Tr}[P_a P_b]}{3} \left( \text{Tr}[X] \text{Tr}[Y] - \frac{\text{Tr}[X Y]}{2} - \frac{\text{Tr}[X Y]}{2} \right)$$

$$= \delta_{a,b}^0 \text{Tr}[X Y] + \delta_{a,b}^{q \ne 0} \frac{2}{3} \left( \text{Tr}[X \text{Tr}[Y]] - \frac{\text{Tr}[X Y]}{2} \right)$$

(B.14)

and the same result is obtained from eq. (B.13):

$$\text{Tr}[^2T_C(P_a, X, P_b) Y] = \delta_{a,b}^0 \text{Tr}[X Y] + \delta_{a,b}^{q \ne 0} \frac{1}{3} \sum_{q=x,y,z} \text{Tr}[P_q X P_q Y]$$

(B.15)

Considering the two cases $X = I$ and $X \neq I$ (so $\text{Tr}[X] = 2$ or $\text{Tr}[X] = 0$, respectively), it is easy to see that eqs. (B.14) and (B.15) are the same.

At this point we notice we only need $K = 12$ Clifford operators to perform the twirl. Going back to Table B.1, the Clifford gates arising from using $S_1$ are $C_k$ with $k = 1$ to 12, while the ones arising from $S_2$ are with $k = 13$ to 24. Each of these two subsets can be used independently. Notice however that these two subsets do not constitute a group on their own.

The size of the required subset can be further decreased to $K = 6$ if we now consider $X = Y$ to be a density matrix $\rho_0$ that is along one of the 3 axes given by the Pauli matrices, as $\rho_0 = (I + \rho \sigma_r)/2$. For these we re-write eq. (B.10) using the $SP$ decomposition rather
than the $PS$ one. Moreover, we already consider we are working with one Symplectic subset:

$$2T_C(P_a, X, P_b) = \frac{1}{12} \sum_{S_1} \sum_{S_2}^{3} P_m S_j P_a S_j^\dagger P_m X P_m S_j^\dagger P_b S_j P_m$$ (B.16)

For the purposes of measuring the fidelity decay, we consider

$$\text{Tr}[2T_C(P_a, \rho_0, P_b)\rho_0] = \frac{1}{12} \sum_{S_1} \sum_{S_2}^{3} \text{Tr}[P_m S_j P_a S_j^\dagger P_m \rho_0 P_m S_j^\dagger P_b S_j P_m \rho_0]$$ (B.17)

$$= \frac{1}{12} \sum_{S_1} \sum_{S_2}^{3} \text{Tr}[P_a S_j^\dagger (P_m \rho_0 P_m) S_j^\dagger P_b S_j (P_m \rho_0 P_m) S_j]$$ (B.18)

Clearly, $P_m \rho_0 P_m = \rho_0$ for $m = 0$ and $m = r$, while $P_m \rho_0 P_m = \rho_0' = (I - \rho \sigma_r)/2$ otherwise. Thus we will have

$$\text{Tr}[2T_C(P_a, \rho_0, P_b)\rho_0] = \frac{1}{12} \sum_{j=1}^{3} \sum_{m=0}^{3} \text{Tr}[P_m S_j P_a S_j^\dagger P_m \rho_0 P_m S_j^\dagger P_b S_j P_m \rho_0]$$

$$= \frac{1}{6} \sum_{j=1}^{3} \text{Tr}[P_a S_j^\dagger \rho_0 S_j^\dagger P_b S_j \rho_0 S_j] + \frac{1}{6} \sum_{j=1}^{3} \text{Tr}[P_a S_j^\dagger \rho_0' S_j^\dagger P_b S_j \rho_0' S_j]$$

The same result can be obtained if instead of taking $P_m$ from the whole Pauli group, we only consider either $P_1 = \{I, \sigma_r\}$ or $P_2 = \{\sigma_r, \sigma_{r'}\}$, with $[\sigma_r, \sigma_{r'}] = \pm \sigma_{r''} 2 \epsilon_{r,r',r''}$.

We emphasize that this last reduction from $K = 12$ to $K = 6$ is only valid for the purposes of calculating (B.18). It does not hold, in general, for calculating arbitrary $\text{Tr}[2T_C(P_a, X, P_b)Y]$ (as does the reduction from $K = 24$ to $K = 12$).

This concludes our remarks on the Clifford twirl in $U(2)$. A Clifford twirl in $U(2)^{\otimes m}$ is what we used in Chapters 5 and 6, which is simply the successive application of a twirl in $U(2)$ for the different qubits. Notice we have proved that to implement a twirl like in (B.2) we only need 12 Clifford operators from either of the two subsets in Table B.1. We have then proven eq. (B.2) only for $D = 2$, not in general. Moreover, our proof relies on using the operators described in Table B.1, rather than using the more general defining property of Clifford operators - eq. (B.3).

The first 12 Clifford gates are the ones we worked with in our experiments. We studied the implementation of the 12 gates, and determined which subset of 6 had the better experimental performance. Notice that there are two possible ways of constructing each of the pools $P_1$ and $P_2$, thus there are 8 possible pools of 6 Clifford operators from the original 24 (considering also the freedom of choosing $S_1$ or $S_2$).

We mentioned at the beginning of this Appendix that either with Clifford operators or random rotations we must resort on sampling the twirl. We can see now that if our aim is to twirl a few qubits, choosing the Clifford twirl may be advantageous, as the pool to be sampled has $6^n$ gates. For $m$ small enough, the twirl could be implemented exactly. If twirling all the qubits, then the decision still relies on the detail of the experimental setup and whether one pool of operators is more robust and/or with an easier implementation.
Appendix C

Analytical calculation of the fidelity decay after one step of a multiple one-qubit twirl

C.1 The system

Consider the space $\mathcal{H}_D = \mathcal{H}_M \otimes \mathcal{H}_{\overline{M}}$ with dimension $D = 2^n$. The index $j$ will label each qubit. $\mathcal{H}_M$ is the space corresponding to the $m$ qubits to be measured. $\mathcal{H}_{\overline{M}}$ is its complement. Consider the initial state:

\begin{align}
\rho_0 &= \rho_0^{(M)} \otimes \rho_0^{(\overline{M})} \\
\rho_0^{(M)} &= \bigotimes_{j \in M} \rho_0^{(j)} \\
\rho_0^{(\overline{M})} &= \bigotimes_{j \in \overline{M}} I^{(j)} / 2 + \sum_u \xi_u Q_u = \frac{I^{(\overline{M})} + \rho_0^{(\overline{M})}}{2n-m} \\
Q_u &= \bigotimes_{j \in \overline{M}} Q_u^{(j)}
\end{align}

where each $Q_u^{(j)}$ is either a Pauli matrix ($\sigma_x, \sigma_y, \sigma_z$) or the identity $I$, but at least one factor in each $Q_u$ is a Pauli matrix (thus the $Q_u$ are traceless). The $\xi_u$ are just valid (real) coefficients. Then $\rho_0^{(M)}$ must be separable; $\rho_0^{(\overline{M})}$ not necessarily. So $\rho_0$ is any density matrix separable in the spaces $\mathcal{H}_M \otimes \mathcal{H}_{\overline{M}}$, and separable within $\mathcal{H}_M$.

C.2 The model for the process $\Gamma$ under study

Consider an operator $E$ as follows:

\begin{equation}
E = \eta_0 \mathbb{1} + \sum_{l=1}^{D^2-1} \eta_l P_l
\end{equation}

143
where the $P_l$ are the generalized Pauli operators (see eq. (4.5)) with $l \geq 1$. $P_0 = \mathbb{I}$ completes the Pauli operator basis. The $P_l$ are similar to the $Q_u$ in (C.3); the difference is that the $Q_u$ have support only on $\mathcal{H}^{(M)}$, while the $P_l$ have support on the full space $\mathcal{H}_D$. Both form orthogonal sets: $\text{Tr}[P_lP_{l'}] = D\delta_{l,l'}$ and $\text{Tr}[ar{Q}_uQ_u] = 2^{n-m}\delta_{u,u}$.

The coefficients $\eta_l$ are arbitrary complex coefficients.

We will consider a model that gives raise to the following map:

$$\Gamma(\rho(t)) = \rho(t + 1) = \int P(\eta) \ E(\eta) \ \rho(t) \ E^\dagger(\eta) \ d\eta$$

(C.6)

For this map to be trace-preserving, we must satisfy the condition $\text{Tr}[\rho(t + 1)] = \text{Tr}[\rho(t)]$ which implies

$$\int P(\eta) \ E^\dagger(\eta) \ E(\eta) \ d\eta = \mathbb{I}$$

(C.7)

By tracing both sides of (C.7) and using the expansion (C.5), we get that

$$\int P(\eta) \ (|\eta_0|^2 + \sum_{l \geq 1} |\eta_l|^2) \ d\eta = 1 \Rightarrow \int P(\eta) |\eta_0|^2 d\eta = 1 - \sum_{l \geq 1} \int P(\eta) |\eta_l|^2 d\eta$$

(C.8)

Notice that this is a necessary but not sufficient condition for (C.7). We must add also $D^2 - 1$ scalar equations that arise from multiplying eq. (C.7) by an arbitrary $P_m$ with $m \geq 1$ and taking the trace. This gives

$$\sum_{l,l' \geq 1} \int P(\eta) \ \eta^*_l \eta_{l'} \ d\eta \ Tr[P_mP_lP_{l'}] = -\sum_{l \geq 1} \int P(\eta) \ (\eta^*_l \eta_0 + \eta^*_0 \eta_l) \ d\eta \ Tr[P_mP_l]$$

$$= -2 \int P(\eta) \text{Re}[\eta^*_m \eta_0] d\eta, \quad m \geq 1$$

(C.9)

for $m = 1, \ldots, D^2 - 1$.

For this map to be hermiticity-preserving (i.e. $\Gamma(\rho) = \Gamma(\rho)$ $\Leftrightarrow \rho^\dagger = \rho$), $P$ must be a real function of the $\eta$ (i.e.: $P = P^*$), as

$$\Gamma(\rho(t))^\dagger = \int P^*(\eta) \ E^\dagger(\eta) \ \rho(t) \ E^\dagger(\eta) \ d\eta$$

(C.11)

And the map is evidently linear. Thus at this point we have a linear, trace-preserving, hermiticity-preserving map. This is actually the most general map that can act on the system in the Hilbert space $\mathcal{H}_D$, including any form of open quantum system dynamics – as recently shown in [27, 28].

It is also simple though to restrict our description to completely positive (CP) maps. For this map to be CP, $P$ must be a nonnegative function of the $\eta$ (i.e.: $P \geq 0$), as in this way we have

$$\langle \Psi | \left( \int P(\eta) \ E(\eta) \ \rho \ E^\dagger(\eta) \ d\eta \right) |\Psi\rangle = \int P(\eta) \langle \eta | \rho | \eta \rangle d\eta \geq 0$$

(C.12)
where $\rho$ is an arbitrary (positive) density matrix and $|\Psi\rangle$ is an arbitrary ket, both in an extended space $\mathcal{H}_D \otimes \mathcal{H}_{extra}$, and $|\psi_\eta\rangle = E^\dagger(\eta)|\Psi\rangle$ is also a ket in the extended space.

Therefore, eq. (C.6), with a real $P$ and together with eq. (C.7) —which includes eq. (C.8)— describes the most general quantum process that may occur. In particular, if $P$ is a non-negative function of the $\eta$, then we have a CP map. Notice that $E$ is not necessarily a unitary operator, since (C.5) describes any operator in $\mathcal{H}_D$. In the case that $E$ is unitary, we would be restricting the model to unital processes. Unital processes would require

$$\int P(\eta) \ E(\eta) \ E^\dagger(\eta) \ d\eta = \mathbb{1} \quad \text{(C.13)}$$

Note that (C.8) is a necessary but not sufficient condition for (C.13).

### C.3 Randomizing the fidelity by twirling

We calculate the fidelity after one step of twirling the map $\Gamma$, measuring only the $M$-qubits:

$$\langle f(M)(t = 1) \rangle = \text{Tr}_M \left[ \rho_0^{(M)} \left\langle \rho^{(M)} \right\rangle \right] \equiv f \quad \text{(C.14)}$$

$$\langle \rho^{(M)} \rangle = \text{Tr}_{\bar{M}}[\rho] = \langle \text{Tr}_{\bar{M}}[\rho] \rangle \quad \text{(C.15)}$$

$$\rho = \left( \bigotimes_{j=1}^{n} R^{(j)} \right) \left[ \int P(\eta) \ E(\eta) \left( \bigotimes_{j=1}^{n} R^{(j)} \right) \rho_0 \left( \bigotimes_{j=1}^{n} R^{(j)} \right) E^\dagger(\eta) \ d\eta \right] \left( \bigotimes_{j=1}^{n} R^{(j)} \right) \quad \text{(C.16)}$$

$$= \int P(\eta) \left[ \left( \bigotimes_{j=1}^{n} R^{(j)} \right) E(\eta) \left( \bigotimes_{j=1}^{n} R^{(j)} \right) \rho_0 \left( \bigotimes_{j=1}^{n} R^{(j)} \right) E^\dagger(\eta) \left( \bigotimes_{j=1}^{n} R^{(j)} \right) \right] d\eta \quad \text{(C.17)}$$

where $R^{(j)}$ is a random rotation on the $j$-qubit. $E$ is not necessarily separable and has support on both $M$ and $\bar{M}$. The average brackets denote the average over the normalized Haar measure as follows:

$$\langle f \left( R^{(1)}, \ldots, R^{(n)} \right) \rangle = \int dR^{(1)} \ldots dR^{(n)} f \left( R^{(1)}, \ldots, R^{(n)} \right) \quad \text{(C.18)}$$

Right away we can take care of the outer set of rotations in (C.16). The ones for qubits belonging to $\bar{M}$ go away because of the cycling property of the trace, in (C.15). At the same time the rotations in $M$ come out of the trace in $\bar{M}$. Thus we have a new expression for $f$:

$$f = \int P(\eta) \left\langle \text{Tr}_M \left[ \left( \begin{array}{c} R^{(M)} \\ \rho_0^{(M)} \end{array} \right) \left( \begin{array}{c} R^{(M)} \end{array} \right)^\dagger \right] \right\rangle_M \ d\eta \quad \text{(C.19)}$$

$$\tilde{\rho}^{(M)} = \text{Tr}_{\bar{M}}[\tilde{\rho}] \quad \text{(C.20)}$$

$$\tilde{\rho} = E(\eta) \left( \bigotimes_{j=1}^{n} R^{(j)} \right) \rho_0 \left( \bigotimes_{j=1}^{n} R^{(j)} \right) E^\dagger(\eta) \quad \text{(C.21)}$$
where we have made explicit the average over the random rotations in $M$ and the one over the random rotations in $\overline{M}$. As shorthand notation, we have used:

$$R^{(M)} = \bigotimes_{j \in M} R^{(j)} \quad (C.22)$$

We will use also the equivalent for $\overline{M}$.

We introduce here the equations to handle the average over the normalized Haar measure on $U(2)$ of a polynomial function of a random rotation $R$ in $U(2)$:

$$\langle \text{Tr}[AB] \rangle = \frac{1}{2} \text{Tr}[A]\text{Tr}[B] \quad (C.23)$$

$$\langle \text{Tr}[\rho R^\dagger AR\rho R^\dagger BR] \rangle = \frac{1}{3} \text{Tr}[AB] \left( 1 - \frac{\text{Tr}[\rho^2]}{2} \right) + \frac{1}{3} \text{Tr}[A]\text{Tr}[B] \left( \text{Tr}[\rho^2] - \frac{1}{2} \right) \quad (C.24)$$

These equations are a result of the work in [94, 95, 96]. In particular we used the diagrammatic technique developed in [96].

For any operator $A$

$$EAE^\dagger = A|\eta_0|^2 + \sum_{l=1}^{D^2-1} \eta_l^* \eta_0 A P_l + \sum_{l=1}^{D^2-1} \eta_0 \eta_l^* P_l A + \sum_{l,l'=1}^{D^2-1} \eta_l \eta_{l'}^* P_l A P_{l'} \quad (C.25)$$

In general, we will denote any reduced operator as $A^{(X)} = \text{Tr}_X[A]$.

### C.4 Cumbersome calculations

Now identify $\tilde{\rho}$ in eq. (C.21) with $EAE^\dagger$ in eq. (C.25). For each of the 4 terms in (C.25) we have a corresponding term for $\tilde{\rho}$. Let’s see each one separately, following the order of eq. (C.25):

$$\tilde{\rho} = \tilde{\rho}_1 + \tilde{\rho}_2 + \tilde{\rho}_3 + \tilde{\rho}_4 \quad (C.26)$$

$$\tilde{\rho}_i^{(M)} = \text{Tr}_M[\tilde{\rho}_i] \quad (C.27)$$

$$f = \int P(\vec{\eta}) \left( \tilde{f}_1 + \tilde{f}_2 + \tilde{f}_3 + \tilde{f}_4 \right) d\vec{\eta} \quad (C.28)$$

$$\tilde{f}_i \equiv \left\langle \text{Tr}_M \left[ R^{(M)} \tilde{\rho}_0^{(M)} R^\dagger (M) \left\langle \tilde{\rho}_i^{(M)} \right\rangle_{\overline{M}} \right] \right\rangle_M \quad (C.29)$$

where for practical reasons we introduced new quantities $\tilde{f}_i$. Let’s calculate them:
• First term:

\[ \rho_1 = \left( \bigotimes_{j=1}^{n} R^{(j)} \right) \rho_0 \left( \bigotimes_{j=1}^{n} R^{(j)} \right) |\eta_0|^2 \]  

(C.30)

\[ \rho_1^{(M)} = R^{(M)} \rho_0^{(M)} R^{(M)} |\eta_0|^2 = \langle \rho_1^{(M)} \rangle_M \]  

(C.31)

\[ \tilde{f}_1 = \text{Tr}_M \left[ \left( \rho_0^{(M)} \right)^2 |\eta_0|^2 \right] = f_0^{(M)} |\eta_0|^2 \]  

(C.32)

Now we start using the expansion (C.5):

• Second term:

\[ \text{Tr}_\overline{M}[\tilde{\rho}_2] = \sum_l \eta_l^* \eta_0 R^{(M)} \rho_0^{(M)} R^{(M)} P_l^{(M)} \text{Tr}_\overline{M} \left[ \rho_0^{(M)} P_l^{(M)} \right] \]  

(C.33)

with

\[ \text{Tr}_\overline{M} \left[ \rho_0^{(M)} P_l^{(M)} \right] = \text{Tr}_\overline{M} \left[ \left( \bigotimes_{j \in \overline{M}} I^{(j)} \right) R^{(M)} P_l^{(M)} \right] + \sum_u \xi_u \text{Tr}_\overline{M} \left[ R^{(M)} Q_u R^{(M)} P_l^{(M)} \right] \]  

(C.34)

where we have used (C.3) and (C.4). Taking the average over \( \overline{M} \), using (C.23):

\[ \left\langle \text{Tr}_\overline{M} \left[ \rho_0^{(M)} P_l^{(M)} \right] \right\rangle_{\overline{M}} = \frac{1}{2^n - m} \text{Tr}_\overline{M} \left[ P_l^{(M)} \right] + \sum_u \xi_u \prod_{j \in \overline{M}} \text{Tr}_j \left[ Q_u^{(j)} R^{(j)} P_l^{(j)} \right] \]  

(C.35)

where the second term in (C.35) vanishes since each \( Q_u \) is traceless as a whole. Then:

\[ \tilde{f}_2 = \sum_l \eta_l^* \eta_0 \left\langle \text{Tr}_M \left[ R^{(M)} \left( \rho_0^{(M)} \right)^2 R^{(M)} P_l^{(M)} \right] \right\rangle_M \frac{1}{2^n - m} \text{Tr}_\overline{M} \left[ P_l^{(M)} \right] \]  

(C.36)

and using (C.23) with (C.2) for the average over \( M \):

\[ \left\langle \text{Tr}_M \left[ R^{(M)} \left( \rho_0^{(M)} \right)^2 R^{(M)} P_l^{(M)} \right] \right\rangle_M = \prod_{j \in M} \text{Tr}_j \left[ \left( \rho_0^{(j)} \right)^2 \right] \text{Tr}_j \left[ P_l^{(j)} \right] \]  

(C.37)

Finally

\[ \tilde{f}_2 = \sum_l \eta_l^* \eta_0 \frac{1}{2^n} \text{Tr}_\overline{M} \left[ P_l^{(M)} \right] \text{Tr}_M \left[ P_l^{(M)} \right] \text{Tr}_M \left[ \left( \rho_0^{(M)} \right)^2 \right] = 0 \]  

(C.38)

which vanishes since the each \( P_l \) is traceless.
• Third term:

It’s easy to see that $\tilde{f}_3 = 0$, since it would take the same calculations that for $\tilde{f}_2$. This is because $\tilde{p}_2 = \tilde{p}_3$, so $\tilde{f}_3 = \tilde{f}_2$.

• Fourth term:

$$\text{Tr}_{\tilde{M}}[\tilde{p}_4] = \sum_{l,l'}(\eta_1\bar{\eta}_2) P_l^{(M)} R^{(M)} p_0^{(M)} R^{(M)} P_{l'}^{(M)} \text{Tr}_{\tilde{M}} \left[ P_l^{(\tilde{M})} R^{(\tilde{M})} p_0^{(\tilde{M})} R^{(\tilde{M})} P_{l'}^{(\tilde{M})} \right] (C.39)$$

Taking the average over $\tilde{M}$ and using (C.23):

$$\left< \text{Tr}_{\tilde{M}} \left[ P_l^{(M)} R^{(M)} p_0^{(M)} R^{(M)} P_{l'}^{(M)} \right] \right>_{\tilde{M}} = \frac{1}{2^{n-m}} \text{Tr}_{\tilde{M}} \left[ P_l^{(\tilde{M})} P_{l'}^{(\tilde{M})} \right] + \sum_{j \in \tilde{M}} \xi_u \text{Tr}_j \left[ Q_u^{(j)} \right] \frac{\text{Tr}_j \left[ P_l^{(j)} P_{l'}^{(j)} \right]}{2} (C.40)$$

where again the second term vanishes since the $Q_u$ are traceless. Then:

$$\tilde{f}_4 = \sum_{l,l'}(\eta_1\bar{\eta}_2) \left< \text{Tr}_M \left[ R^{(M)} p_0^{(M)} R^{(M)} P_l^{(M)} R^{(M)} p_0^{(M)} R^{(M)} P_{l'}^{(M)} \right] \right> \frac{1}{2^{n-m}} \text{Tr}_{\tilde{M}} \left[ P_l^{(\tilde{M})} P_{l'}^{(\tilde{M})} \right] (C.41)$$

For the average in $M$ we must consider

$$\left< \text{Tr}_M \left[ R^{(M)} p_0^{(M)} R^{(M)} P_l^{(M)} R^{(M)} p_0^{(M)} R^{(M)} P_{l'}^{(M)} \right] \right>_M = \prod_{j \in M} \left< \text{Tr}_j \left[ R^{(j)} p_0^{(j)} R^{(j)} P_l^{(j)} R^{(j)} p_0^{(j)} R^{(j)} P_{l'}^{(j)} \right] \right> (C.42)$$

For this to be non-zero, all the factors in the product must be non-zero. Let’s take an arbitrary factor, say, for qubit $a$, and use (C.24):

$$\left< \text{Tr}_a \left[ R^{(a)} p_0^{(a)} R^{(a)} P_l^{(a)} R^{(a)} p_0^{(a)} R^{(a)} P_{l'}^{(a)} \right] \right> = \frac{1}{3} \left( \text{Tr}_a \left[ P_l^{(a)} P_{l'}^{(a)} \right] \left( 1 - \frac{\mathcal{P}_a}{2} \right) \right. + \text{Tr}_a \left[ P_l^{(a)} \right] \text{Tr}_a \left[ P_{l'}^{(a)} \right] \left( \mathcal{P}_a - \frac{1}{2} \right) \right) (C.43)$$

where $\mathcal{P}_a$ is the purity of the initial state of qubit $a$: $\text{Tr} \left[ \left( p_0^{(a)} \right)^2 \right]$. Let’s say that for qubit $a$ we have $P_l^{(a)} = a$ Pauli matrix. Then this factor of the product will be non-zero only if $P_{l'}^{(a)} = the\ same\ Pauli\ matrix$, becoming $(2/3)(1 - \mathcal{P}_a/2)$. If, on the other hand, $P_{l'}^{(j)} = I$, then for the factor to be non-zero we must have $P_{l'}^{(j)} = I$ too. But in this case it becomes $\mathcal{P}_a$. Notice that, overall, (C.42) is proportional to a delta function $\delta_{l,l'}$. After these considerations, (C.41) gives

$$\tilde{f}_4 = \sum_{l,l'}(\eta_1\bar{\eta}_2) \prod_{j \in M} \epsilon_j(l) \frac{1}{2^{n-m}} \text{Tr}_{\tilde{M}} \left[ P_l^{(\tilde{M})} P_{l'}^{(\tilde{M})} \right] = \sum_l |\eta|^2 \prod_{j \in M} \epsilon_j(l) (C.44)$$

148
where
\[
\mathcal{C}_j(l) = \begin{cases} 
(2/3)(1 - P_j/2) & \text{if } P_l^{(j)} = \sigma_x, \sigma_y, \sigma_z \\
\rho_j & \text{if } P_l^{(j)} = I
\end{cases}
\quad (C.45)
\]

C.5 The fidelity decay after one step

Putting all the results together in eq. (C.26),
\[
f = \int P(\bar{\eta}) \left( \tilde{f}_1 + \tilde{f}_2 \right) d\bar{\eta} = \int P(\bar{\eta})|\eta_0|^2 d\bar{\eta} f_0^{(M)} + \sum_j \int P(\bar{\eta})|\eta_j|^2 d\bar{\eta} \prod_{j \in M} \mathcal{C}_j(l) 
\]
\[
= \langle |\eta_0|^2 \rangle_P \prod_{j \in M} \rho_j + \sum_j \langle |\eta_j|^2 \rangle_P \prod_{j \in M} \mathcal{C}_j(l) 
\quad (C.46)
\]
\[
= f_0^{(M)} + \sum_j \langle |\eta_j|^2 \rangle_P \left( \prod_{j \in M} \rho_j - \prod_{j \in M} \mathcal{C}_j(l) \right) 
\quad (C.47)
\]
\[
\Rightarrow f = f_0^{(M)} + \sum_j \langle |\eta_j|^2 \rangle_P \left( \prod_{j \in M} \rho_j - \prod_{j \in M} \mathcal{C}_j(l) \right) 
\quad (C.48)
\]

where \( \langle \ldots \rangle_P = \int P(\bar{\eta}) \ldots d\bar{\eta} \) and eq. (C.48) is obtained using (C.8).

Thus the fidelity decay defined as \( \Delta f^{(M)} = f_0^{(M)} - f \) is
\[
\Delta f^{(M)} = \sum_j \langle |\eta_j|^2 \rangle_P \left( \prod_{j \in M} \rho_j - \prod_{j \in M} \mathcal{C}_j(l) \right) 
\quad (C.49)
\]

Notice that \( \Delta f^{(M)} \) is not the decay rate \( \gamma^{(M)} \), but \( \langle f^{(M)}(t = 1) \rangle = f_0^{(M)} (1 - \gamma^{(M)}) \). Thus
\[
\gamma^{(M)} = 1 - \frac{\langle f^{(M)}(t = 1) \rangle}{f_0^{(M)}} = \frac{\Delta f^{(M)}}{f_0^{(M)}} 
\quad (C.50)
\]

These two quantities are equal only if \( f_0^{(M)} = 1 \), that is, for an initially pure state \( \rho_0^{(M)} \) in \( \mathcal{H}_M \). In any case, both quantities can be computed from the same quantities that would be determined experimentally: \( f_0^{(M)} \) and \( f = \langle f^{(M)}(t = 1) \rangle \).

Clearly, \( \gamma^{(M)} \) has a physical meaning, since the overall exponential-like behavior of \( f^{(M)}(t) \) is governed by this parameter. However, if our objective is to measure the coefficients \( \langle |\eta_j|^2 \rangle_P \), \( \Delta f^{(M)} \) is what we should compute from the experimental data.
C.6 Characterizing a general CP map $\Gamma$ vs. characterizing a generator $G_T$

We recall the model for a unital map described in Chapter 4, which we state here once again but using coefficients $\tilde{\eta}_l$ (cf. eq. (4.4)):

$$E = \exp(-iG_T)$$  \hspace{1cm} (C.51)

$$G_T = \text{“the generator”} = \sum_{l \geq 1} \tilde{\eta}_l P_l$$  \hspace{1cm} (C.52)

$$E_G = I - i \sum_{l \geq 1} \tilde{\eta}_l P_l - \frac{1}{2} \sum_{l,l' \geq 1} \tilde{\eta}_l \tilde{\eta}_{l'} P_l P_{l'} + O(\tilde{\eta}^3)$$  \hspace{1cm} (C.53)

Comparing (C.53) with (C.5), we have that

$$\eta_m = -i\tilde{\eta}_m - \frac{1}{2} \sum_{l,l' \geq 1} \tilde{\eta}_l \tilde{\eta}_{l'} \frac{\text{Tr}[P_l P_{l'} P_m]}{D} + O(\tilde{\eta}^3)$$  \hspace{1cm} (C.54)

$$|\eta_m|^2 = |\tilde{\eta}_m|^2 - i \sum_{l,l' \geq 1} (\tilde{\eta}_l \tilde{\eta}_{l'} \tilde{\eta}_m^* - \tilde{\eta}_l \tilde{\eta}_{l'\nu} \tilde{\eta}_m^*) \frac{\text{Tr}[P_l P_{l'} P_m]}{D} + O(\tilde{\eta}^4)$$  \hspace{1cm} (C.55)

If the coefficients $\tilde{\eta}_l$ in eq. (C.52) are real (as it is the case for $EE^\dagger = I$ as we required in Chapter 4), then $|\eta_m|^2 = |\tilde{\eta}_m|^2 + O(\tilde{\eta}^4)$. So the formula (C.49) equally applies to the coefficients $\tilde{\eta}_l$ of a generator $G_T$ up to $O(\tilde{\eta}^3)$ inclusive. From this, all the formulas in Chapter 4 can be deduced.

These formulas, nevertheless, were originally deduced independently, with a strategy just like the one developed here. The difference was in eq. (C.25). The approach was to consider that for weak noise, we could expand $E_G$ as

$$E_G = I - iG_T - \frac{1}{2} G_T^2 + \frac{i}{6} G_T^3 + O(G^4)$$  \hspace{1cm} (C.56)

so then

$$EAE^\dagger = A + i(AG_T^\dagger - G_TA) - \frac{1}{2}(A(G_T^\dagger)^2 + G_T^2 A) - i \frac{1}{6} (A(G_T^\dagger)^3 - G_T^3 A)$$  \hspace{1cm} (C.57)

$$+ \frac{i}{2} \left(G_T A G_T^\dagger - G_T (G_T A G_T^\dagger) + G_T A G_T^\dagger - G_T (G_T A G_T^\dagger) \right) + G_T A G_T^\dagger + O(G^4)$$  \hspace{1cm} (C.58)

This would replace eq. (C.25), giving rise to 6 terms in an equivalent eq. (C.26). Each of these terms was then calculated in a similar fashion to what we show here in Sec. C.4. We don’t reproduce these calculations here as they would be redundant.

C.7 Using a partial twirl in $\mathcal{H}_M$

We go back to Sec. C.3 and we choose to do a partial twirl only over the Hilbert space $\mathcal{H}_M$ of the qubits we measure - a twirl in $U(2)^{\otimes m}$. Again we calculate the fidelity after one step,
for the $M$-qubits being measured, but this time

$$\langle f^{(M)}(t = 1) \rangle_M = \text{Tr}_M \left[ \rho_0^{(M)} \left\langle \rho^{(M)} \right\rangle_M \right] \equiv f$$  \hspace{1cm} (C.59)

$$\langle \rho^{(M)} \rangle_M = \text{Tr}_{\Pi} \langle \rho \rangle_M = \langle \text{Tr}_{\Pi} \rho \rangle_M$$  \hspace{1cm} (C.60)

$$\rho = \left( \bigotimes_{j=1}^{m} R^{(j)} \right) \left[ \int P(\vec{\eta}) E(\vec{\eta}) \left( \bigotimes_{j=1}^{m} R^{(j)} \right) \rho_0 \left( \bigotimes_{j=1}^{m} R^{(j)} \right) E^{\dagger}(\vec{\eta}) \, d\vec{\eta} \right] \left( \bigotimes_{j=1}^{m} R^{(j)} \right)$$  \hspace{1cm} (C.61)

$$= \int P(\vec{\eta}) \left[ \left( \bigotimes_{j=1}^{m} R^{(j)} \right) E(\vec{\eta}) \left( \bigotimes_{j=1}^{m} R^{(j)} \right) \rho_0 \left( \bigotimes_{j=1}^{m} R^{(j)} \right) E^{\dagger}(\vec{\eta}) \left( \bigotimes_{j=1}^{m} R^{(j)} \right) \right] d\vec{\eta}$$  \hspace{1cm} (C.62)

The average now is only over $\mathcal{H}_M$:

$$\langle f \left( R^{(1)}, \ldots, R^{(m)} \right) \rangle_M = \int dR^{(1)} \ldots dR^{(m)} \, f \left( R^{(1)}, \ldots, R^{(m)} \right)$$  \hspace{1cm} (C.63)

The rotations in $M$ come out of the trace in $\overline{M}$. Thus we have a new expression for $f$:

$$f = \int P(\vec{\eta}) \left\langle \text{Tr}_M \left[ \left( R^{(M)} \rho_0^{(M)} R^{(M)^\dagger} \right) \tilde{\rho}^{(M)} \right] \right\rangle_M \, d\vec{\eta}$$  \hspace{1cm} (C.64)

$$\tilde{\rho}^{(M)} = \text{Tr}_{\Pi} \tilde{\rho}$$  \hspace{1cm} (C.65)

$$\tilde{\rho} = E(\vec{\eta}) \left( \bigotimes_{j=1}^{m} R^{(j)} \right) \rho_0 \left( \bigotimes_{j=1}^{m} R^{(j)} \right) E^{\dagger}(\vec{\eta})$$  \hspace{1cm} (C.66)

Again, we identify $\tilde{\rho}$ in eq. (C.66) with $EAE^{\dagger}$ in eq. (C.25), and again we define new quantities $\tilde{f}_i$, but unlike before they only have an Haar average over $\mathcal{H}_M$:

$$\tilde{f}_i \equiv \text{Tr}_M \left\langle \left[ R^{(M)} \rho_0^{(M)} R^{(M)^\dagger} \tilde{\rho}_i^{(M)} \right] \right\rangle_M$$  \hspace{1cm} (C.67)

Let’s re-calculate them:

- **First term**:

  $$\tilde{\rho}_1 = \left( \bigotimes_{j=1}^{m} R^{(j)} \right) \rho_0 \left( \bigotimes_{j=1}^{m} R^{(j)} \right) |\eta_0|^2$$  \hspace{1cm} (C.68)

  $$\tilde{\rho}_1^{(M)} = R^{(M)} \rho_0^{(M)} R^{(M)^\dagger} |\eta_0|^2$$  \hspace{1cm} (C.69)

  $$\tilde{f}_1 = \text{Tr}_M \left[ (\rho_0^{(M)})^2 |\eta_0|^2 \right] \equiv f_0^{(M)} |\eta_0|^2$$  \hspace{1cm} (C.70)
• Second term:

\[ \text{Tr}_{\Xi\Phi}[\tilde{\rho}_2] = \sum_l \eta_l^* \eta_0 R^{(M)} P_0^{(M)} R_l^{(M)} P_l^{(M)} \text{Tr}_{\Xi\Phi} \left[ \rho_0^{(M)} P_l^{(M)} \right] \]  

(C.71)

with

\[ \text{Tr}_{\Xi\Phi} \left[ \rho_0^{(M)} P_l^{(M)} \right] = \frac{1}{2n-m} \text{Tr}_{\Xi\Phi} \left[ P_l^{(M)} \right] + \frac{1}{2n-m} \text{Tr}_{\Xi\Phi} \left[ \rho_{\text{dev}}^{(M)} P_l^{(M)} \right] \]  

(C.72)

where we have used (C.3). Then:

\[ \tilde{f}_2 = \sum_l \eta_l^* \eta_0 \left( \text{Tr}_{\Xi\Phi} \left[ R^{(M)} \left( \rho_0^{(M)} \right)^2 R_l^{(M)} P_l^{(M)} \right] \right) \frac{\text{Tr}_{\Xi\Phi} \left[ P_l^{(M)} \right] + \text{Tr}_{\Xi\Phi} \left[ \rho_{\text{dev}}^{(M)} P_l^{(M)} \right]}{2n-m} \]  

(C.73)

and using (C.23) with (C.2) for the average over \( M \), we obtain eq. (C.37) again. Finally

\[ \tilde{f}_2 = \sum_l \eta_l^* \eta_0 \frac{1}{2n} \text{Tr}_{\Xi\Phi} \left[ P_l^{(M)} \right] \text{Tr}_{\Phi} \left[ \left( \rho_0^{(M)} \right)^2 \right] \]

+ \[ \sum_l \eta_l^* \eta_0 \frac{1}{2n} \text{Tr}_{\Xi\Phi} \left[ P_l^{(M)} \rho_{\text{dev}}^{(M)} \right] \text{Tr}_{\Phi} \left[ P_l^{(M)} \right] \text{Tr}_{\Phi} \left[ \left( \rho_0^{(M)} \right)^2 \right] \]  

(C.74)

The first line vanishes, since each \( P_l \) is traceless as a whole. The second line gives us

\[ \tilde{f}_2 = \sum_l \eta_l^* \eta_0 \frac{1}{2n} \text{Tr}_{\Xi\Phi} \left[ \rho_{\text{dev}}^{(M)} P_l \right] \text{Tr}_{\Phi} \left[ \left( \rho_0^{(M)} \right)^2 \right] \]  

(C.75)

• Third term: Again we have \( \tilde{\rho}_2 = \tilde{\rho}_3^\dagger \), so \( \tilde{f}_3 = \tilde{f}_2^\dagger \).

• Fourth term:

\[ \text{Tr}_{\Xi\Phi}[\tilde{\rho}_4] = \sum_{l,l'} (\eta_l \eta_{l'}^*) P_l^{(M)} R^{(M)} P_0^{(M)} R_l^{(M)} P_l^{(M)} \text{Tr}_{\Xi\Phi} \left[ P_l^{(M)} P_{l'}^{(M)} \right] \]  

(C.76)

Using (C.3) we get:

\[ \text{Tr}_{\Xi\Phi} \left[ P_l^{(M)} P_{l'}^{(M)} \right] = \frac{1}{2n-m} \text{Tr}_{\Xi\Phi} \left[ P_l^{(M)} P_{l'}^{(M)} \right] + \frac{1}{2n-m} \text{Tr}_{\Xi\Phi} \left[ P_l^{(M)} \rho_{\text{dev}}^{(M)} P_{l'}^{(M)} \right] \]  

(C.77)

This is going to take the place of \( \frac{1}{2n-m} \text{Tr}_{\Xi\Phi} \left[ P_l^{(M)} P_{l'}^{(M)} \right] \) in eq. (C.41). So we will have

\[ \tilde{f}_4 = \sum_{l,l'} (\eta_l \eta_{l'}^*) \prod_{j \in M} \mathcal{O}_j(l) \delta_{l,l'}^{(M)} \frac{\text{Tr}_{\Xi\Phi} \left[ P_l^{(M)} P_{l'}^{(M)} \right] + \text{Tr}_{\Xi\Phi} \left[ P_l^{(M)} \rho_{\text{dev}}^{(M)} P_{l'}^{(M)} \right]}{2n-m} \]  

(C.78)

We can combine \( \delta_{l,l'}^{(M)} \) and \( \text{Tr}_{\Xi\Phi} \left[ P_l^{(M)} P_{l'}^{(M)} \right] /2n-m \) to obtain a delta over \( \mathcal{H}_P \) for the first term in the fraction. And we can write \( \delta_{l,l'}^{(M)} = \text{Tr}_M \left[ P_l^{(M)} P_{l'}^{(M)} \right] /2^n \) to combine it
with the trace of the second term. Therefore we get

\[
\hat{f}_4 = \sum_l |\eta_l|^2 \prod_{j \in M} \mathcal{C}_j(l) + \sum_{l, l'} (\eta_l \eta_{l'}^*) \prod_{j \in M} \mathcal{C}_j(l) \frac{\text{Tr}[P_l (\mathbb{M}) P_{l'}]}{D}
\]  

(C.79)

**C.8 The fidelity decay after one step with a partial twirl**

Putting the new \( \hat{f}_4 \) together we obtain

\[
f = \int P(\tilde{\eta}) \left( \hat{f}_1 + \hat{f}_2 + \hat{f}_2 + \hat{f}_4 \right) d\tilde{\eta} = \langle |\eta_0\rangle |_P \hat{f}_0^{(M)} + \hat{f}_0^{(M)} \sum_l \langle \eta_l^* \eta_0 + \eta_0^* \eta_l \rangle_P \frac{\text{Tr}[P_l (\mathbb{M}) P_{l'}]}{D} \\
+ \sum_l \langle |\eta_l|^2 \rangle_P \prod_{j \in M} \mathcal{C}_j(l) + \sum_{l, l'} (\eta_l \eta_{l'}^*) \prod_{j \in M} \mathcal{C}_j(l) \frac{\text{Tr}[P_l (\mathbb{M}) P_{l'}]}{D} \tag{C.80}
\]

So now

\[
\Delta f^{(M)} = f_0^{(M)} - f = \sum_l \langle |\eta_l|^2 \rangle_P \left( \prod_{j \in M} \mathcal{P}_j - \prod_{j \in M} \mathcal{C}_j(l) \right) - \Omega^{(M)}_{\text{dev}} \tag{C.81}
\]

\[
\Omega^{(M)}_{\text{dev}} = \sum_l \langle \text{Re}[\eta_l^* \eta_0] \rangle_P \frac{\text{Tr}[\rho^{(\mathbb{M})}_\text{dev} P_l]}{D} + \sum_{l, l'} (\eta_l \eta_{l'}^*) \prod_{j \in M} \mathcal{C}_j(l) \frac{\text{Tr}[P_l (\mathbb{M}) P_{l'}]}{D}
\]

Clearly, (C.81) becomes (C.49) if \( \rho^{(\mathbb{M})}_\text{dev} = 0 \). This means: if we prepare the qubits we do not measure in the maximally mixed state \((I/2)^{\otimes(n-m)}\), there is no need to twirl them. We have then a compromise of resources: we either twirl the \( n \) qubits and do not worry about the initial state of the ones we do not measure, or we prepare them carefully in the \( I/2 \) state, and we do not worry about twirling them.

We can work a bit more on \( \Omega^{(M)}_{\text{dev}} \). We can use eq. (C.3) to expand \( \rho^{(\mathbb{M})}_\text{dev} \). Since eq. (C.9) is valid for any arbitrary \( P_m, m \geq 1 \), in \( \mathcal{H}_D \), it is also valid for the \( Q_u \) that have non-identity factors only in \( \mathcal{H}_\mathbb{M} \). Thus we have

\[
\Omega^{(M)}_{\text{dev}} = \sum_u \frac{\xi_u}{2^m} \left( f_0^{(M)} \sum_l \langle \text{Re}[\eta_l^* \eta_0] \rangle_P \text{Tr}[Q_u P_l] + \sum_{l, l'} (\eta_l \eta_{l'}^*) \prod_{j \in M} \mathcal{C}_j(l) \text{Tr}[P_l Q_u P_{l'}] \right) \\
= \sum_u \frac{\xi_u}{2^m} \sum_{l, l'} \langle \eta_l \eta_{l'}^* \rangle_P \left( \prod_{j \in M} \mathcal{C}_j(l) - f_0^{(M)} \right) \text{Tr}[P_l Q_u P_{l'}]
\]

where the last line comes from using eq. (C.9).

For a \( \rho_0^{(M)} \) that is pure, we can bound \( \left| \prod_{j \in M} \mathcal{C}_j(l) - f_0^{(M)} \right| \leq 1 \). So

\[
|\Omega^{(M)}_{\text{dev}}| \leq \sum_u \frac{\xi_u}{2^m} \sum_{l, l'} \langle \eta_l \eta_{l'}^* \rangle_P \left| \text{Tr}[P_l Q_u P_{l'}] \right| \tag{C.82}
\]
If we manage to prepare a $\rho_0^{(M)}$ that is close enough to the maximally mixed state, we can bound the weight of $|\Omega_{\text{dev}}^{(M)}|$ vs. $|\gamma^{(M)}|$ by the order of the $\xi_a$ coefficients - since both $|\Omega_{\text{dev}}^{(M)}|$ and $|\gamma^{(M)}|$ are of $O(\eta^2)$.

C.9 Expressions of the fidelity decay $\Delta f^{(M)}$ for a few qubits

We now focus on the measurement of $m = 1, 2, 3$ qubits. We will be either twirling the other $n - m$ (thus we will not worry about preparing them in a particular state) or we will prepare them in the $I/2$ state (and so we will not worry about twirling them). As we mentioned above, the $\mathcal{C}_j(l)$ does not distinguish the directions of the Pauli matrices that compose $P_l$, so we will introduce collective coefficients $\eta^{\text{col}}$ that take this into account.

We denote now the $\eta_l$ in more detail as $\eta_j^{p,q,\ldots}$, where $j > k > \ldots$ label qubits, and $p, q, \ldots = x, y, z$. Therefore $\eta_j^{p,q,\ldots}$ labels a term in eq. (C.5) that is a product of $\sigma_p$ for qubit $j$, $\sigma_q$ for qubit $k$, etc., and $I$ for the qubits absent in the subscript. Notice that the number of qubits in the subscript gives the Pauli weight (the number of non-identity operators) of that particular $P_l$. Therefore, the one-body terms (Pauli weight 1) go with coefficients $\eta_j^x, \eta_j^y, \eta_j^z$, two-body terms (Pauli weight 2) are $\eta^{p,q}_j$, etc. To avoid double counting of multi-body terms, the labeling of the qubits must obey $j < k < \ldots$ and so on. For example, the Pauli operator $\sigma_x^{(1)} \sigma_z^{(3)} \sigma_z^{(4)}$ has Pauli weight 3, and the corresponding coefficient is $\eta_j^{x,z,z}$.

Under this notation, we define

$$|\eta_j^p|^2 = \sum_{p=x,y,z} |\eta_j^p|^2; \quad |\eta_j^{p,q}|^2 = \sum_{p,q=x,y,z} |\eta_j^{p,q}|^2; \quad |\eta_j^{p,q,r}|^2 = \sum_{p,q,r=x,y,z} |\eta_j^{p,q,r}|^2; \quad \text{etc.}$$

considering in general complex coefficients. We can now calculate more explicitly the fidelity decay for one, two and three qubits. We just apply eq. (C.49) carefully.

With $\mathcal{A}_j = (2 - \mathcal{P}_j)/3$:

$$\Delta f^{(a)} = (\mathcal{P}_a - \mathcal{A}_a) \left( \langle |\eta_a^{\text{col}}|^2 \rangle_P + \sum_{j \neq a} \langle |\eta_{a,j}^{\text{col}}|^2 \rangle_P + \sum_{k > j \neq a} \langle |\eta_{a,j,k}^{\text{col}}|^2 \rangle_P + \ldots \right) \quad \text{(C.83)}$$

$$\Delta f^{(a,b)} = (\mathcal{P}_a \mathcal{P}_b - \mathcal{A}_a \mathcal{A}_b) \times \left( \langle |\eta_{a,b}^{\text{col}}|^2 \rangle_P + \sum_{j \neq a,b} \langle |\eta_{a,b,j}^{\text{col}}|^2 \rangle_P + \ldots \right)$$

$$+ (\mathcal{P}_a - \mathcal{A}_a) \mathcal{P}_b \left( \langle |\eta_{a,b}^{\text{col}}|^2 \rangle_P + \sum_{j \neq a,b} \langle |\eta_{a,b,j}^{\text{col}}|^2 \rangle_P + \sum_{k > j \neq a,b} \langle |\eta_{a,b,j,k}^{\text{col}}|^2 \rangle_P + \ldots \right)$$

$$+ \mathcal{P}_a (\mathcal{P}_b - \mathcal{A}_b) \left( \langle |\eta_{b,j}^{\text{col}}|^2 \rangle_P + \sum_{j \neq a,b} \langle |\eta_{b,j}^{\text{col}}|^2 \rangle_P + \sum_{k > j \neq a,b} \langle |\eta_{b,j,k}^{\text{col}}|^2 \rangle_P + \ldots \right) \quad \text{(C.84)}$$
\[ \Delta f^{(a,b,c)} = (\mathcal{P}_a \mathcal{P}_b \mathcal{P}_c - \mathcal{A}_a \mathcal{A}_b \mathcal{A}_c) \times \left( \langle |\eta_{a,b,c}^{\text{col}}|^2 \rangle_P + \sum_{j \neq a,b,c} \langle |\eta_{a,b,c,j}^{\text{col}}|^2 \rangle_P + \ldots \right) \]

\[ + (\mathcal{P}_a \mathcal{P}_b - \mathcal{A}_a \mathcal{A}_b) \mathcal{P}_c \left( \langle |\eta_{a,b}^{\text{col}}|^2 \rangle_P + \sum_{j \neq a,b,c} \langle |\eta_{a,b,j}^{\text{col}}|^2 \rangle_P + \sum_{k > j \neq a,b,c} \langle |\eta_{a,b,j,k}^{\text{col}}|^2 \rangle_P + \ldots \right) \]

\[ + \mathcal{P}_a (\mathcal{P}_b \mathcal{P}_c - \mathcal{A}_b \mathcal{A}_c) \left( \langle |\eta_{b,c}^{\text{col}}|^2 \rangle_P + \sum_{j \neq a,b,c} \langle |\eta_{b,c,j}^{\text{col}}|^2 \rangle_P + \sum_{k > j \neq a,b,c} \langle |\eta_{b,c,j,k}^{\text{col}}|^2 \rangle_P + \ldots \right) \]

\[ + (\mathcal{P}_a \mathcal{P}_c - \mathcal{A}_a \mathcal{A}_c) \mathcal{P}_b \left( \langle |\eta_{a,c}^{\text{col}}|^2 \rangle_P + \sum_{j \neq a,b,c} \langle |\eta_{a,c,j}^{\text{col}}|^2 \rangle_P + \sum_{k > j \neq a,b,c} \langle |\eta_{a,c,j,k}^{\text{col}}|^2 \rangle_P + \ldots \right) \]

\[ + (\mathcal{P}_a - \mathcal{A}_a) \mathcal{P}_b \mathcal{P}_c \left( \langle |\eta_{a}^{\text{col}}|^2 \rangle_P + \sum_{j \neq a,b,c} \langle |\eta_{a,j}^{\text{col}}|^2 \rangle_P + \sum_{k > j > j' \neq a,b,c} \langle |\eta_{a,j,j',j''}^{\text{col}}|^2 \rangle_P + \ldots \right) \]

\[ + \mathcal{P}_a (\mathcal{P}_b - \mathcal{A}_b) \mathcal{P}_c \left( \langle |\eta_{b}^{\text{col}}|^2 \rangle_P + \sum_{j \neq a,b,c} \langle |\eta_{b,j}^{\text{col}}|^2 \rangle_P + \sum_{k > j > j' \neq a,b,c} \langle |\eta_{b,j,j',j''}^{\text{col}}|^2 \rangle_P + \ldots \right) \]

\[ + \mathcal{P}_a \mathcal{P}_b (\mathcal{P}_c - \mathcal{A}_c) \left( \langle |\eta_{c}^{\text{col}}|^2 \rangle_P + \sum_{j \neq a,b,c} \langle |\eta_{c,j}^{\text{col}}|^2 \rangle_P + \sum_{k > j > j' \neq a,b,c} \langle |\eta_{c,j,j',j''}^{\text{col}}|^2 \rangle_P + \ldots \right) \]

(C.85)

where \ldots denote the corresponding higher order multi-body terms.

If we prepare the qubits in a pure state, so \( \mathcal{P}_a = \mathcal{P}_b = \mathcal{P}_c = 1 \), we have \( \Delta f^{(M)} = \gamma^{(M)} \) and the combination

\[ \gamma^{(a)} + \gamma^{(b)} - \gamma^{(a,b)} = \frac{4}{9} \left( \langle |\eta_{a,b}^{\text{col}}|^2 \rangle_P + \sum_{j \neq a,b} \langle |\eta_{a,b,j}^{\text{col}}|^2 \rangle_P + \ldots \right) \]

(C.86)

leaves only the collective coefficients involving both the qubits \( a \) and \( b \), while

\[ \gamma^{(a)} + \gamma^{(b)} + \gamma^{(c)} - \gamma^{(a,b)} - \gamma^{(b,c)} - \gamma^{(a,c)} + \gamma^{(a,b,c)} = \frac{8}{27} \left( \langle |\eta_{a,b,c}^{\text{col}}|^2 \rangle_P + \sum_{j \neq a,b,c} \langle |\eta_{a,b,c,j}^{\text{col}}|^2 \rangle_P + \ldots \right) \]

(C.87)

leaves only the collective coefficients involving the three qubits \( a, b \) and \( c \).

Finally, we emphasize, once again, that the formulas (C.49) and (C.81) are valid for coefficients \( \eta \) describing any CP map, with arbitrary large coefficients. \( \Gamma \) does not need to be a noisy channel neither a weak one. In particular, it could be a gate or any process far from the identity operation.
Bibliography


[82] S.M. Ross, op. cit., Sec. 6.3; W. Feller, op. cit., Chap. X.


