2024-10-21

Characterisation of universal qudit gates

Amaro Alcala, David

UNIVERSITY OF CALGARY

Characterisation of universal qudit gates

by

David Amaro Alcalá

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN PHYSICS AND ASTRONOMY

CALGARY, ALBERTA
OCTOBER, 2024

Abstract

To harness the potential of quantum computing, increasing the number of quantum units, a process known as scaling, is critical. Whereas qubits have traditionally been used as the units for quantum computing, the development of multi-level systems (qudits), which offer larger Hilbert spaces and advantages over qubits in cryptography and circuit complexity reduction, requires new methods to characterise the quality of quantum gates and ensure safe scaling. Randomised benchmarking offers a simple and inexpensive method for this characterisation. This thesis reports advances in the characterisation of universal single- and multi-qudit gates.

I introduce the characterisation of universal qutrit gates through the definition of an optimal scheme that requires similar experimental resources as the standard method for non-universal gates. The feasibility of my qutrit scheme is tested numerically using parameters from experimental qutrit implementations. I then generalise my qutrit results and devise a general scheme for a qudit system with arbitrary d. Because using the same construction for qudits with $d > 3$ as in the qutrit case leads to more than two parameters, a different strategy was necessary. I note that my qudit characterisation obtains an estimate of the average error per gate; thus, this characterisation is collective. A more realistic characterisation requires estimating the average gate fidelity of a single non-Clifford gate. In the last part, I generalise my qudit method to individually, in contrast to the previous collective result, characterise non-Clifford gates.

My schemes are relevant to at least two communities: experimental groups with a qudit platform, as my work effectively characterises a complete gate set, and randomised benchmarking theorists, who may be interested both in the gate set I introduce and in the schemes I developed.

Preface

Chapter 4 contains work published in the article by **David Amaro-Alcalá**, Barry C. Sanders, and Hubert de Guise, "Universal qutrit randomised benchmarking", "Phys. Rev. A 109, 012621 (2024)". Chapter 5 consists of work accepted in the "New Journal of Physics": **David Amaro-Alcalá**, Barry C. Sanders, and Hubert de Guise, "Randomised benchmarking for universal qudit gates" "New J. Phys. (2024)". The section on interleaved benchmarking in Chapter 6 was accepted for the conference on multi-valued logic, where the author of this thesis presented the work. The work is published on the preceedings: **David Amaro-Alcalá**, B. C. Sanders and H. de Guise, "Qudit non-Clifford interleaved benchmarking," 2024 IEEE 54th International Symposium on Multiple-Valued Logic (ISMVL), Brno, Czech Republic, 2024. Additionally, the author expects to submit a collaborative paper on the results of cycle benchmarking, which is discussed in §6.3.

Acknowledgements

I express my deepest gratitude to my supervisor, Dr. Barry C Sanders, and my co-supervisor, Dr. Hubert de Guise, for their continuous support, guidance, and education. Dr. Sander's suggestions will be my first-order approach to many problems I encounter.

I am also very grateful to my supervisory committee, Dr. David Feder and Dr. Carlo Maria Scandolo. Their insightful questions challenged me to think critically about key aspects of my research, enhancing both the accessibility of the thesis and my understanding of the material.

Special thanks are due to my internal and external examiners. I thank Dr. Christoph Simon for posing experimental questions that broadened my perspective. I am particularly indebted to Dr. Joseph Emerson, whose precise and thoughtful comments helped me correct misunderstandings and refine my knowledge of randomised benchmarking, improving the content and the presentation of this thesis. I also acknowledge Dr. Jacob Barnet for his question about the motivation for the definition of the Clifford hierarchy, which led to a more precise treatment of this vital gate set.

I would like to thank Dr. Gilad Gour and Dr. Nasser Ahmadi, whose critical questions during my pre-candidacy and field of study examinations deepened my understanding of the thesis material and shaped my academic formation.

I am also thankful to Dr. Michael Wieser and Dr. Timothy Friesen for chairing my examinations and providing valuable assistance during both my field of study and thesis defense.

From the University of Calgary, I would like to thank the instructors of my courses, including Dr. Nasser Ahmadi, Dr. Barry C Sanders, and Dr. Jo-Anne Brown, for their excellent teaching. Especially to Dr. Jo-Anne Brown and her activities in class showed me that it is necessary to make scientific discussion accessible.

I would like to express my gratitude to the administrative staff, particularly Yanmei Fei, Marni Farran, Jacob Hunt, and Jo-Anne Brown, for their help throughout my time at UCalgary.

I am grateful for the hospitality of the research groups I visited, especially Prof. Mario Ziman's group (Dr. Fereshte Shahbeigi, Dr. Michal Sedlák, and Ricardo Rivera Cardoso), and Prof. Karol Życzkowski's group (Albert Rico and Jakub Czartowski).

I would like to thank the journals Physical Review Letters and Nature Communications for granting me permission to use the figures that appear in my thesis. I also want to express my gratitude to Prof. Martin Ringbauer (figures 3.3 and 3.4) and Dr. Alexis Morvan (figure 3.5) for their permission to use the figures of the papers on which they are first authors.

I gratefully acknowledge the financial support from the Government of Alberta and the Natural Sciences and Engineering Research Council of Canada.

To my mom

Table of Contents

Abstract ii

Preface iii

Acknowledgements iv

Dedication vi

Table of Contents vii

List of Figures and Illustrations x

List of Tables xi

List of Symbols, Abbreviations and Nomenclature xii

Epigraph xv

1 Introduction 1
- 1.1 Introduction . . . 1
- 1.2 State-of-the-art . . . 3
- 1.3 Approach . . . 6
- 1.4 Results . . . 8
- 1.5 Organisation . . . 9

2 Characterisation of quantum channels 10
- 2.1 Introduction . . . 10
- 2.2 Representation of quantum channels . . . 10
- 2.3 Representation theory . . . 16
- 2.4 Clifford hierarchy . . . 20
- 2.5 Twirl . . . 21
- 2.6 Context for the introduction of randomised benchmarking . . . 27
- 2.7 Quantum volume . . . 29
- 2.8 Process and gate set tomography . . . 31
- 2.9 Randomised benchmarking and the quantum threshold theorem . . . 34
- 2.10 Interpretation of the parameter estimated . . . 37

2.11 Non-Markovianity . . . 42
2.12 An ion trap qudit platform . . . 43
2.13 Summary . . . 46

3 State-of-the-art methods **47**
3.1 Introduction . . . 47
3.2 Characterisation of Clifford gates . . . 48
3.2.1 Qubit randomised benchmarking . . . 48
3.2.2 Qudit randomised benchmarking for prime-level systems . . . 51
3.3 Characterisation of non-Clifford gates . . . 52
3.3.1 Single-qubit dihedral benchmarking . . . 52
Removing the constant due to SPAM . . . 54
3.3.2 Character randomised benchmarking . . . 56
3.3.3 Multi-qubit dihedral Benchmarking . . . 58
3.3.4 Leakage characterisation . . . 59
3.3.5 Cross-talk characterisation . . . 60
3.4 Individual characterisation . . . 61
3.4.1 Introduction . . . 61
3.4.2 Interleaved benchmarking . . . 62
3.4.3 Clifford interleaved benchmarking . . . 62
3.4.4 Non-Clifford gate characterisation using Clifford gates . . . 63
3.4.5 Interleaved dihedral benchmarking . . . 64
3.4.6 Interleaved benchmarking for qubit gates . . . 66
3.4.7 Cycle benchmarking . . . 67
3.5 Platforms . . . 70
3.5.1 Qudit ion trap . . . 70
3.5.2 Superconductor . . . 73
3.6 Randomised compiling . . . 74
3.7 Summary . . . 75

4 Randomised benchmarking for universal qutrit gates **76**
4.1 Introduction . . . 76
4.2 Construction of the HDG . . . 76
4.3 Representation theory for the HDG . . . 79
4.4 Gate fidelity and sequence fidelity . . . 80
4.5 Scheme description at circuit and SPAM level . . . 81
4.6 Numerical study of my scheme's feasibility . . . 82
4.7 Phase and criteria for universal qutrit randomised benchmarking . . . 84
4.8 Conclusion . . . 88

5 Randomised benchmarking for universal qudit gates **90**
5.1 Introduction . . . 90
5.2 Construction of the qudit rHDG . . . 91
5.3 Representation theory for the rHDG . . . 95
5.4 Expressions for the average gate fidelity and sequence fidelity . . . 100

5.5 Experimental scheme . . . 102
5.6 Multi-qudit case . . . 103
5.7 Conclusion . . . 104

6 Application of the rHDG to extensions of randomised benchmarking 106
6.1 Introduction . . . 106
6.2 Interleaved benchmarking . . . 106
6.3 Shadow estimate for universal gates . . . 112
6.4 Universal cycle benchmarking . . . 118
6.5 Conclusion . . . 119

7 Summary and open problems 121
7.1 Universal qutrit randomised benchmarking . . . 122
7.2 Universal qudit randomised benchmarking . . . 123
7.3 Applications of the real HyperDihedral group . . . 125
7.4 Open problems . . . 126

Bibliography 128

A Proof of an identity for Bell numbers 143

B Miscellaneous computations 149
B.1 Motivation for the Clifford hierarchy . . . 149
B.2 Proof and estimate of the total variation distance . . . 150
B.3 GAP code for qutrit HDG . . . 152
B.4 Proof number of independent parameters in chi-representation . . . 152
B.5 Discussion numerical confidence intervals . . . 154

C Classical randomised benchmarking 159

List of Figures and Illustrations

3.1 Family of circuits used in character randomised benchmarking. 57
3.2 Family of circuits for interleaved dihedral benchmarking. 65
3.3 Sequence fidelity for qudits with $d = 2, 3$, and 5. 71
3.4 Log plot for the sequence fidelity for qutrits. 72
3.5 Average sequence fidelity for a qutrit superconductor platform. 74

4.1 Circuit illustrating the notation for the inversion gate. 82
4.2 Graph of the pairs circuit depth *vs* variance for the HDG. 84
4.3 Contour plot showing the maximum circuit depth m that a gate with average gate fidelity F can be composed of before the phase contribution changes the behavior of the single exponential. 87

5.1 Geometric figure used to illustrate the ${}_d$rHDG. 95

6.1 Circuit corresponding to a run of interleaved benchmarking. 111

List of Tables

5.1 Table with phases in the proof 98

B.1 Table reporting the confidence values for the estimate of the parameter η_0. The noise corresponds to a randomly sampled channel with average gate fidelity 0.89. 154

B.2 Table reporting the confidence values for the estimate of the parameter η_0. The noise model is a composition of a totally depolarising and an amplitude damping channel, which is a comprehensive Markovian and completely positive noise model. The value of fidelity corresponds to the parameters 0.01 for the depolarising and amplitude damping channel. 155

B.3 Table reporting the confidence values for the estimate of the parameter η_0. The noise model is a composition of a totally depolarising and an amplitude damping channel, which is a comprehensive Markovian and completely positive noise model. The value of fidelity corresponds to the parameters 0.03 for the depolarising and amplitude damping channel. 155

B.4 Table reporting the confidence values for the estimate of the parameter η_0. The noise model is a composition of a totally depolarising and an amplitude damping channel, which is a comprehensive Markovian and completely positive noise model. The value of fidelity corresponds to the parameters 0.05 for the depolarising and amplitude damping channel. 156

B.5 Table reporting the confidence values for the estimate of the parameter. In contradistinction with the qutrit case, the noise model corresponded to a randomly sampled channel with an average gate fidelity 0.90. 157

B.6 Table reporting the confidence values for the estimate of the parameter. In contradistinction with the qutrit case, the noise model corresponded to a randomly sampled channel with an average gate fidelity 0.99. 157

B.7 Table reporting the confidence values for the estimate of the parameter. In contradistinction with the qutrit case, the noise model corresponded to a randomly sampled channel with an average gate fidelity 0.95. 158

B.8 Comparison of the strategies. The result is based on the confidence intervals. For the noise we used a depolarising channel composed with a phase-damping; the depolarising and dephasing parameters are set to 0.1. 158

List of Symbols, Abbreviations and Nomenclature

Abbreviation	Definition
AGF	average gate fidelity
CPTP	Completely positive and trace preserving
${}_d$rHDG	rHyperDihedral group for a d-level system (qudit)
HDG	HyperDihedral group
irrep	Irreducible representation
gate set	set of gates
POP	Power of a prime
RB	randomised benchmarking
RBS	randomised benchmarking scheme(s)
rHDG	rHyperDihedral group
SPAM	State preparation and measurement
unirrep	Unitary irreducible representation

Symbol	Definition
Γ	Pauli-Liouville representation
γ	Irreducible representation of the ${}_d$rHDG
ϖ	Label of one irrep
$\mathfrak{h}$	Hilbert space
End($\mathfrak{h}$)	Set of linear mappings from $\mathfrak{h}$ to $\mathfrak{h}$.
ϱ	Density operator
$\mathcal{E}$	CPTP mapping
A	Kraus operator
C_k	Cyclic group with k elements
S_k	Symmetric over a set with k elements
$[k]$	List of integers from 0 to $k-1$
σ	Permutation
α	Ring or cyclic group element
d	Number of levels
T	Matrix representing T gate
ω_k	Symbol to denote exp(2πi/k)
$\langle X \rangle$	Group generated by a list of matrices X
$\otimes$	Kronecker or tensor product

$\oplus$	Direct sum
$\chi(g)$	Character for group element g
$\mathbb{E}_{x\in X}$	Average considering the uniform measure of X
f_σ	Number of fixed points of the permutation σ
$\|\varrho\rangle\rangle$	Vectorisation of the density operator ϱ
$\langle\langle\varrho\|$	Complex conjugate of $\|\varrho\rangle\rangle$)
$\circ$	Composition of gates and mappings.
g	Group element or gate
$\tilde{g}$	Noisy version of the gate g
$\hat{g}$	Ideal version of the gate g
F	Average gate fidelity
G	Finite group
G^m	Cartesian product, m times, of G
$\mathbf{g}$	Ordered sequence of group elements.
P	Sequence fidelity
$\eta(\mathcal{E})$	Eigenvalue of the twirl of $\mathcal{E}$
$\mathrm{MS}^{i,j}$	Mølmer–Sørensen gate acting on qudits i and j
$\mathcal{C}_k$	k-th level of the Clifford hierarchy.
F_d	Qudit Fourier matrix
S_d	Qudit Sylvester matrix
$S^{(0)}$	Alternative sequence of gates used to characterise gates
$S^{(1)}$	Alternative sequence of gates used to characterise gates
Y	Qubit Y gate
Z	Qubit Z gate
$\cong$	To denote two objects are isomorphic
$\bigcirc_i$	Composition of objects indexed by i
$\triangleq$	Associate a symbol with a circuit
$\mathcal{C}_d$	Qudit Clifford group
$\delta_{i,j}$	Kronecker delta. Also used as a matrix
Π_ϖ	Projector onto $\mathfrak{h}_\varpi$
$\mathbb{I}$	Identity matrix
$\mathbb{I}_d$	Identity matrix acting on a d-dimensional space
Id	Identity mapping
N	Number of qudits in a system
$[\mathcal{E}]_G$	Twirl of $\mathcal{E}$ by group G
$[\mathcal{E}]_\gamma$	Twirl of $\mathcal{E}$ by irrep γ
diag	Operator to compute a diagonal matrix from a vector
X	Random variable
$\mathbb{E}\,X$	Mean of random variable X
$\mathbb{V}X$	Variance of random variable X
$\chi(\mathcal{E})$	Chi-representation of a channel $\mathcal{E}$
$\bigtimes_{k\in\boldsymbol{k}} C_k$	Cartesian product of sets indexed by $\boldsymbol{k}$
$\mathsf{D}(\boldsymbol{\alpha})$	Representation for a product of cyclic groups
$\mathsf{X}(\sigma)$	Regular representation of the symmetric group

σ	Permutation or representation
ρ	Representation
A_d	Alternating group
$\mathbf{0}$	Null-operator, mapping any vector to the null-vector
$F \restriction_{\mathfrak{h}_\varpi}$	Restriction of the mapping F to the subspace $\mathfrak{h}_\varpi$

[0]Traditional notation in randomised benchmarking literature. Not to be confused with the Choi state [1].

Epigraph

Only presidents, editors, and people with tapeworms ought to have the right to use we.

- (popularised by) Mark Twain.

I have made this longer than usual because I have not had time to make it shorter.

- Blaise Pascal.

Chapter 1

Introduction

The motivation behind this thesis is to develop schemes capable of characterising qudit experimental implementations, of which there has been a recent increase in platforms and applications. This thesis addresses two challenges: generalising dihedral benchmarking and concocting schemes to characterise universal qudit gates; the term dihedral comes from the fact that the set of unitary matrices used in the scheme forms a representation of the dihedral group [2, 3]. By adopting an algebraic approach, I use tools from representation theory, ring theory, and previous randomised benchmarking (RB) schemes to develop my generalisation of dihedral benchmarking. I also use quantum channels and previous experimental implementations of randomised benchmarking schemes to assert the feasibility of my schemes. Furthermore, the gate set resulting from this research is an approximation of a unitary 2 design for qudits, potentially benefiting the study of t-designs.

1.1 Introduction

In this section, I summarise the content of this chapter, which itself is a summary of this thesis. The underlying topic of this work is quantum computing [4]. Quantum computers surpass, currently only in a theoretical proposal, classical machines in several tasks [5]. Crucial to realising this "advantage" are the operations, or gates, within the computer.

There are two essential characteristics of a set of gates: quality and universality. Quality is the maximum error rate among the gates, and universality is the capacity to approximate, by arbitrary composition of members of the gate set, any unitary gate.

As indicated by the quantum threshold theorem [6], the quality of gates determines the capabilities of a quantum computer. Accordingly, characterising gates is of essential importance. Universality is also fundamental since, according to the Gottesman-Knill theorem [7, 8], a non-universal gate set implies that a quantum computer can be efficiently simulated on a classical machine.

Definition 1.1 ([9])**.** By benchmark, [it is meant] a set of quantum circuits (a test suite) together with instructions for how to run them (an experimental design), an analysis procedure for processing the raw results, and finally an interpretation rule for drawing high-level conclusions.

Definition 1.1 generality is necessary to embrace the wide range of benchmarking schemes currently developed and in use. For specificity, I decrease the generality of Definition 1.1 and introduce the definition of benchmarking that is more appropriate for my thesis. Note that, to avoid awkward sentences, I use characterising as a synonym for benchmarking.

Definition 1.2. A gate or set of gates is considered characterised if either the average gate fidelity or an estimate of the average gate fidelity is assigned to the gate or gates.

I delay the introduction of a definition of a benchmarking scheme to Chapter 3, so as to have several examples of randomised benchmarking schemes, therefore, avoiding a complicated definition detached from the literature.

In this thesis, I study quality and universality within the framework of randomised benchmarking schemes. Specifically, my research is about the characterisation of universal gate sets through generalisations of randomised benchmarking [10]. Next, I discuss the content of the present introductory chapter, which expands on the topics mentioned in this paragraph.

In Sec 1.2, on state-of-the-art, I discuss the current knowledge regarding the characterisation gates acting on qudits, which are d-level systems [11], as well as the context in which randomised benchmarking emerged.

In Sec. 1.3 I detail the schemes I aim to generalise and the gate sets that my schemes characterise. Specifically, I discuss the characterisation of non-Clifford qudit gates using different schemes, highlighting the pros and cons of each scheme. I also outline the constraints I impose on my schemes to ensure they are faithful extensions; faithful, meaning the scheme, reduced to the qubit case, is a known scheme. Furthermore, I comment on the computational tools I employed during my research.

Then, in the section on my results—Sec. 1.4, I describe my schemes, highlighting the gaps they addressed. In brief, my schemes represent a generalisation of dihedral benchmarking, interleaved benchmarking, and cycle benchmarking. I specify the gates my schemes characterise, the trade-offs between different schemes, and the specific challenges I overcome in their design.

1.2 State-of-the-art

In this section, I discuss several concepts: a qudit, the historical context of randomised benchmarking, the quantity that randomised benchmarking schemes estimate, and the vanguard of randomised benchmarking schemes to characterise non-Clifford qudit gates.

Most studies in quantum computation focus on qubits [4]. Recently, however, systems with more than two levels have gained interest, and multiple implementations—meaning an experimentally-controlled manipulation of qudit systems, have been reported [12, 13, 14, 15]. Unadorned examples of d-level systems correspond to spin systems and truncated harmonic oscillators [16].

Present qudit implementations include ion traps (for instance, an array of $^{40}Ca^{+}$ atoms) [14], neutral atoms [12], superconductors [13], semiconductors [15], and photons [17]. The appli-

cations of qudits are varied, including quantum communication [18], quantum teleportation [19, 20], quantum memories [21, 22], Bell-state measurements [23], spin chains [24, 13, 20, 25, 23, 26], in enhancing quantum error correction techniques [27, 28], in encoding qubits [29] and qudits [30], simulation of many-body systems [31], quantum key distribution [32, 33, 34], simulation of high-energy physics [35, 36, 37, 38, 39, 40, 41, 42, 43], and quantum computing [17, 44].

With the introduction of qudit systems, qubit-oriented methods need to be extended. Overall, qudit generalisations of qubit techniques are not straightforward—as an example (outside the scope of this thesis) consider the determination of SIC POVM, where the solution to the problem for six-level systems is an open problem [45]. Among the tasks that need generalisation, the characterisation of quantum gates is crucial.

The importance of gate set characterisation emerges from the damage of noise to quantum computing [46] and, more directly, the quantum threshold theorem [6]. This theorem limits the number of operations that can be implemented given a certain "error rate": estimating the error rate is key. If the noise is completely known, as achieved by quantum process tomography, then the computation of the error rate is known [47].

Back in 2008, the characterisation of gates was carried out using quantum process tomography (QPT) [48]. In QPT, gates are reconstructed [48]; the reconstructed matrix is then used to compute the error rate, the quantity used in the quantum threshold theorem to asses the maximum depth circuits can be safely implemented on a given platform.

QPT, however, is practical only for a few qubits [49]; for systems with a large number of qubits, the reconstruction of the noise requires an exponential amount of resources. Moreover, it assumes state preparation and measurement (SPAM) errors to be insignificant [50]. Therefore, to characterise systems with a large number of qubits, an alternative is needed: that alternative is now available. It is called randomised benchmarking.

With the advent of randomised benchmarking in 2008 [49, 51], the two limitations of QPT were addressed. The characterisation of the Clifford group, done by randomised benchmark-

ing, is unaffected by SPAM errors and requires significantly fewer samples [52]. Whereas in tomography, the noise is reconstructed, effectively estimating $d^4 - d^2$ parameters, in randomised benchmarking, a single parameter is estimated: the average gate fidelity [53]. Intuitively, the average gate fidelity quantifies the similarity between the image of a noisy and ideal gate acting on a randomly sampled pure state.

Existing randomised benchmarking schemes are applied to characterise universal qubit gates and qudit Clifford gates [54, 2]. However, due to issues with the quantity estimated by the original RB [55], RBS have been modified to obtain other figures of merit or to simplify experimental requirements, or both [56, 57].

A recent key contribution to the randomised benchmarking ecosystem further supports the utility of these schemes. The technique known as randomised compiling transforms most forms of noise into a Pauli channel [58]. This result not only increases the usefulness of RBS but also links the average gate fidelity to the worst-case characterisation, given by the error rate [59]. This is especially relevant for the construction and design of fault-tolerant devices [6].

The original randomised benchmarking scheme, along with its assumption that every gate has the same noise, estimates the average quality over a gate set. Most implementations, however, make use of composite primitive gates. In practice, there are multiple reasons why this assumption is invalid, e.g. different gates are implemented using different combinations of primitive gates. To address this limitation, interleaved benchmarking was introduced [60].

Interleaved benchmarking estimates the average gate fidelity of a target gate. By using an auxiliary gate set with known average gate fidelity, interleaved benchmarking estimates the average gate fidelity of the composed noise between the auxiliary gate set and the target gate. From the fidelity of the composition of noise, the average gate fidelity of the target gate. The gate that is wanted to separately characterise, is derived.

There are methods within typical randomised benchmarking schemes to characterise an arbitrary gate set [61, 56]. These schemes, however, are feasible only for a low number of

qubits. The reason is that the statistical methods required for such methods are unfeasible for higher-level systems or for multi-qudit systems or both [56].

Another set of theoretical proposals, which I label as single-parameter models, study randomised benchmarking from a matrix theory perspective [62]. These schemes ignore the group structure of the gate set and prove the conditions for which the fidelity has a single parameter. These schemes are important because they show when the group structure of the group is superfluous. Nevertheless, these schemes fail to provide an unbiased characterisation, as these schemes require high fidelity fidelity gates.

Randomised benchmarking schemes can also be used to estimate other quantities. For instance, the amount of leakage, loss, and coherence of quantum gates [63, 64, 65]. However, these studies are mostly relevant for Clifford gates, since these gates are commonly implemented with respect to other gate sets. Therefore, I do not discuss them in detail.

1.3 Approach

In this section, I discuss the approach of my thesis. I list the tools and methodology I use to attain my goals; I include a brief motivation for their application.

My scheme aims to characterise non-Clifford gates [2, 66]. To achieve this, I first generalise dihedral benchmarking [2], which characterises a gate set featuring a T gate [55]. To obtain a separate and more precise characterisation of a non-Clifford gate, I extend cycle benchmarking [67]. The techniques I require to extend dihedral benchmarking and cycle benchmarking differ, although both characterise a non-Clifford gate.

To ensure that my schemes characterise any diagonal qudit T gate, I investigate the most general form of a diagonal T gate; that is, any diagonal non-Clifford gate that, together with the Clifford gate set, generates a universal gate set. I use the T gates I identify in my construction of the gate set that generalises the dihedral group. Because of this constraint, my schemes characterise any diagonal T gate.

I make my results practical to most forms of noise by using two representations of quantum channels widely discussed in textbooks [4, 68]. Specifically, I use the Pauli-Liouville and Kraus representations to analyse my schemes [69, 70]. Additionally, I use the χ-representation [48] to approximate the fidelity as the composition of channels, which I apply in my generalisation of interleaved benchmarking.

I use group theory and representation theory of semidirect products [71]. Using the semidirect product, I specify a natural generalization of the dihedral group. Furthermore, through my investigation, I provide a novel simplification of a result from the theory of induced representations [71].

I rediscover a mathematical identity between the Bell numbers as the average over partitions of an integer [72]. I use tools from combinatorics in my proof [72, 73]. Specifically, I employ generalised generating functions, requiring the manipulation of series, where each term is, in turn, an infinite product.

I consider an extension, of dihedral benchmarking or cycle benchmarking, successful based on two criteria: fewer gates required than Clifford randomised benchmarking and only two parameters needed to estimate the average gate fidelity. Both constraints are critical for the scalability of the scheme. The restriction on the number of parameters is motivated by the following observation: using an arbitrary gate set, the number of parameters increases as $O(d^{2n})$ [56, 10], where n is the number of qudits.

To produce a feasible scheme, I need to estimate the number of samples required in each experiment. I use tools from statistics to estimate this number in a randomised benchmarking experiment. In particular, I apply Hoeffding's inequality and Chernoff bound [74] to estimate the sampling required to obtain confidence intervals for the average gate fidelity.

Also, to ensure the feasibility of my schemes, I investigate and extract the experimental resources required from reported implementations [13]. I study the experimental capabilities, such as maximum circuit depth, number of repetitions per circuit, the time a gate takes to apply, and so forth, that experimental groups report. I use these results, along

with my numerical simulations, to ensure my scheme is feasible with current experimental implementations.

I employ fundamental notions of ring theory, which I need to study non-prime level systems. In particular, I use a generalization of row-reduction for matrices with entries in a ring. I apply this procedure to construct of the generators of the gate set I propose [75]; this generalization is not a straightforward adaptation for complex matrices: various key steps of the algorithm are not applicable in the ring case [75]. I use this row-reduction scheme to compute the minimal generating set of a group. I also use the language of C^*-algebras and bounded operators for finite-dimensional spaces to present my results in the most general form and do not depend on a specific representation.

1.4 Results

My first result is a scheme to characterise any diagonal qutrit T gate. By using the natural generalisation of the X gate and any diagonal T gate, I provide the resulting randomised benchmarking scheme, which includes: circuit design, number of samples required, and data analysis.

My second achievement is the development of a gate set, akin to the dihedral group for qubits but tailored for qudits. I call this gate set the real hyperdihedral group (rHDG). An outstanding property of the gate set is that its Pauli-Liouville (PL) representation decomposes into three real irreducible representations; this property is unaffected by the dimensional of the system and the number of qudits. This property is fundamental for my next result.

My third result is the generalisation of dihedral benchmarking for single and multi (controlled) qudit gates. This is my most important result. My scheme, using the rHDG, faithfully generalises dihedral benchmarking; 'faithfully' means that it shares all the properties of the original scheme except being d-dimensional instead of two-dimensional.

Next, I designed a scheme to characterise a T gate individually; 'individual' means I assume the noise of the T gate is different from the rest of the gates. My previous three results characterised a gate set that included a T gate; for this result, I now only characterise the T gate, not the entire gate set. The characterisation is more precise.

Then I present some issues in the literature: the misuse of a representation of the ququart Clifford gate set. These issues are important since these schemes are being used [54, 14] and reported. However, as I show here, the characterisation resulting from these schemes is flawed.

1.5 Organisation

The thesis is logically organised as follows. Chapters 2 and 3 contain background material. Chapters 4 through 6 are about my results. Chapter 5 is the most important since it presents the main result and positively answers the original research question: Can dihedral benchmarking be extended to qudit systems?

In a bit more of detail, the content of each chapter discuss the following topics:

- The problem of characterising gates, and current methods (Chap. 3) shows the need to characterise T gates and the need for the
- HDG for qutrits (Chap. 4), which does not generalize so well to qudits so we move to
- rHDG for qudits (Chap. 5), which actually works so well that we move to
- applications of rHDG to other schemes (Chap. 6).

Each step in the sequence comes with a peer-review publication, either already published or accepted for publication.

In the last chapter, I present a summary of my results and a list of open problems. These open problems go beyond randomised benchmarking by shedding light on practical limitations of quantum computing, such as universal gate sets for qudit systems.

Chapter 2

Characterisation of quantum channels

2.1 Introduction

In this first background chapter, I discuss the mathematical methods and figures of merit that randomised benchmarking schemes estimate. I review the twirling procedure, which is critical in randomised benchmarking schemes. Next, I discuss how a randomised benchmarking scheme estimates the quality of quantum gates. I follow with a discussion on the parameters estimated by a randomised benchmarking scheme. I conclude with an exposition of the link between the sequence fidelity and the average gate fidelity and the role of the twirling procedure in this relation. I also discuss the scheme called gate set tomography and cross-entropy benchmarking to compare the pros and cons of using randomised benchmarking for gate characterisation; for historical context, I discuss process tomography. Throughout this section, I use the notation that will be used later in the thesis.

2.2 Representation of quantum channels

In this section, I swiftly review quantum channels [68]. Then I discuss the representations of quantum channels that are used in randomised benchmarking schemes.

As my thesis focuses on finite Hilbert spaces $\mathfrak{h}$, it is known that linear operators acting

on $\mathfrak{h}$ can be expressed as $d \times d$ matrices [76]. In quantum mechanics, this can be done more explicitly. The bra-ket notation $\varrho = |\psi\rangle\langle\psi|$ highlights that the density operators reside in a tensor space formed by the pair state $|\psi\rangle$ and its dual $\langle\psi|$, which represents the set of linear mappings from $\mathfrak{h}$ to $\mathbb{C}$.[1]

Now, I discuss the sets of states and mappings between states that I use in my thesis. Let $\mathfrak{h}'$ denote the Hilbert space with elements given by $d \times d$ complex matrices and inner product $\mathrm{tr}\big(uv^\dagger\big)$ for any $u, v \in \mathfrak{h}'$, where † denotes the Hermitian conjugation operation. The set of density matrices is defined as

$$\mathfrak{h} \coloneqq \{\varrho \in \mathfrak{h}' \colon \ \mathrm{tr}\,\varrho = 1, \varrho \geq 0, \varrho \text{ is Hermitian}\}. \tag{2.1}$$

Next, I define the set of completely positive and trace preserving (CPTP) mappings from $\mathfrak{h}'$ to $\mathfrak{h}'$, which I denote $\mathcal{B}(\mathfrak{h}')$. For any $\mathcal{E} \in \mathcal{B}(\mathfrak{h}')$ there is a finite set of $d \times d$ complex matrices $\mathbf{A} = \{A_0, \ldots, A_r\}$ $(0 \leq r < d^2)$ satisfying $\sum_{A \in \mathbf{A}} A^\dagger A = \mathbb{I}$, such that for any ϱ

$$\mathcal{E}(\varrho) = \sum_{A \in \mathbf{A}} A \varrho A^\dagger. \tag{2.2}$$

These matrices A are known as Kraus matrices, and the mapping $\mathcal{E}$ is used to compute the representations of semigroups in the case of quantum channels, as discussed in the following sections.

In randomised benchmarking schemes, two ubiquitous representations of quantum channels are commonly used: Pauli-Liouville and Kraus-operator representation. In this subsection, I review the Pauli-Liouville representation, which is the representation most commonly used in randomised benchmarking schemes.

Before discussing the representation, I review the set of qudit Pauli matrices, which forms a basis for d-square complex matrices. For qubits, the Pauli group forms the foundation of

[1] Consider the ordered basis for $\mathfrak{h}$ be $\{|i\rangle : i \in \{0, 1, \ldots, d-1\}\}$. The mapping ϕ, which maps the linear operator $|i\rangle\langle j|$ to the matrix $\delta_{i,j}$, where all entries are zero except for the i, j entry, which equals one, is an isomorphism.

the stabiliser formalism [4]. This formalism efficiently represents certain classes of states and their evolution with respect to the normaliser of the Pauli group, known as the Clifford gate set.

Moreover, error-correcting schemes exploit this efficiency to study noise in the form of Pauli channels [4]. These schemes leverage the compact representation provided by the stabiliser formalism to simplify the analysis and correction of errors. Now, I use the Pauli group to define a representation of matrices; the Pauli group also appears in my discussion of the Clifford group.

The first ingredient to define the Pauli-Liouville representation is the pair of matrices X and Z, known as clock and shift matrices, introduced by Sylvester [77] and re-introduced, among others, by Schwinger [78]. These matrices are defined in terms of their action on the computational basis:

$$X\left|i\right\rangle \coloneqq \left|i+1\right\rangle \text{ and } Z\left|i\right\rangle \coloneqq \omega_d^i\left|i\right\rangle, \tag{2.3}$$

where $i \in [d]$ and

$$\omega_d \coloneqq \exp(2\pi \mathrm{i}/d). \tag{2.4}$$

Using X and Z, the d^2 Heisenberg-Weyl matrices are defined as

$$W_{d(i-1)+j} \coloneqq X^i Z^j; \tag{2.5}$$

the matrices W_i are the unitary generalisation of the Pauli matrices [79, 4]; a Hermitian generalisation is the set of generalised Gell-Mann matrices [80].

Definition 2.1 ([70])**.** Let $\mathcal{E}$ be a CPTP mapping. Then the Pauli-Liouville representation of $\mathcal{E}$ is a matrix $\Gamma(\mathcal{E})$ with entries

$$\Gamma(\mathcal{E})_{i,j} \coloneqq d^{-1/2}\operatorname{tr}(W_i^\dagger \mathcal{E}(W_j)). \tag{2.6}$$

Definition 2.1 should be applied for a channel $\mathcal{E}$ in its Kraus form; otherwise, applying

Eq. (2.6) is invalid. This issue can be seen in the computation of the Pauli-Liouville representation of the depolarising channel.

Now, I verify that the mapping Γ is a homomorphism. Consider two channels, $\mathcal{E}_0$ and $\mathcal{E}_1$. The composition of two channels is a channel with action $\mathcal{E}_0 \circ \mathcal{E}_1(\varrho) \coloneqq \mathcal{E}_0(\mathcal{E}_1(\varrho))$, where I use $\circ$ to denote composition. From the definition of the Pauli-Liouville representation [68], it can be verified that for any pair of channels $\mathcal{E}_0$ and $\mathcal{E}_1$, Γ satisfies:

- $\Gamma(\mathcal{E}_0 \circ \mathcal{E}_1) = \Gamma(\mathcal{E}_0)\Gamma(\mathcal{E}_1)$.
- $\Gamma(\mathcal{E}^\dagger) = \Gamma(\mathcal{E})^\dagger$.

Since the set of channels forms a semigroup, calling Γ a representation is justified.

With respect to the Pauli-Liouville representation, I also need to define a representation of the states. Thus, I implicitly define the representation for a density operator ϱ with respect to the Heisenberg-Weyl basis as

$$|\mathcal{E}(\varrho)\rangle\!\rangle = \Gamma(\mathcal{E})|\varrho\rangle\!\rangle; \tag{2.7}$$

the entries

$$|\varrho\rangle\!\rangle_i \coloneqq d^{-1/2}\,\mathrm{tr}\Big(W_i^\dagger \varrho\Big). \tag{2.8}$$

This concludes the description of the Pauli-Liouville representation.

The next representation, which I call the Kraus representation [81], is more useful in the context of representation theory since it appears naturally in the isomorphism between the set of endomorphisms of $\mathfrak{h}$ and the tensor product of $\mathfrak{h}$ and its dual. Computing this representation requires the Kraus matrices in Eq. (2.2) and the vectorisation of density matrices, which I now discuss.

For any ϱ with entries

$$\varrho = \begin{bmatrix} \varrho_{0,0} & \cdots & \varrho_{0,d-1} \\ \vdots & \ddots & \vdots \\ \varrho_{d-1,0} & \cdots & \varrho_{d-1,d-1} \end{bmatrix}, \tag{2.9}$$

the reshaping operation vec yields

$$\mathrm{vec}(\varrho) \coloneqq (\varrho_{0,0}, \ldots, \varrho_{0,d-1}, \ldots, \varrho_{d-1,0}, \ldots, \varrho_{d-1,d-1}). \tag{2.10}$$

Note that from Eq. 2.10 $\mathrm{vec}(|i\rangle\langle j|) = |ij\rangle$. Consider three $d \times d$ matrices A, ϱ, and C. The vectorisation in Eq. 2.10 yields

$$\mathrm{vec}(A\varrho C) = A \otimes C^{\top} \mathrm{vec}(\varrho). \tag{2.11}$$

Now, I prove Eq. (2.11).

$$\begin{aligned} \mathrm{vec}(ABC) &= \mathrm{vec}(\sum_{ij} (ABC)_{ij} \, |i\rangle\langle j|) && (2.12) \\ &= \mathrm{vec}(\sum_{ijk} (AB)_{ik} C_{kj} \, |i\rangle\langle j|) && (2.13) \\ &= \mathrm{vec}(\sum_{ijkl} A_{il} B_{lk} C_{kj} \, |i\rangle\langle j|) && (2.14) \\ &= \sum_{ijkl} A_{il} B_{lk} C_{kj} \, \mathrm{vec}(|i\rangle\langle j|) && (2.15) \\ &= \sum_{ijkl} A_{il} B_{lk} C_{kj} \, |ij\rangle \, . && (2.16) \end{aligned}$$

On the other hand,

$$(A \otimes C^\top)\,\mathrm{vec}(B) = (A \otimes C^\top)\,\mathrm{vec}(\sum_{ij} B_{ij}\,|i\rangle\langle j|) \tag{2.17}$$

$$= \sum_{ij} B_{ij}(A \otimes C^\top)\,|ij\rangle \tag{2.18}$$

$$= \sum_{ij} B_{ij}((\sum_{uv} A_{uv}\,|u\rangle\langle v|) \otimes C^\top)\,|ij\rangle \tag{2.19}$$

$$= \sum_{ij} B_{ij}((\sum_{uv} A_{uv}\,|u\rangle\langle v|) \otimes (\sum_{kl} C_{kl}\,|l\rangle\langle k|))\,|ij\rangle \tag{2.20}$$

$$= \sum_{ij,uv,kl} B_{ij}((A_{uv}\,|u\rangle\langle v|) \otimes (C_{kl}\,|l\rangle\langle k|))\,|ij\rangle \tag{2.21}$$

$$= \sum_{ij,uv,kl} B_{ij}A_{uv}C_{kl}\,|ul\rangle\langle vk|\,|ij\rangle \tag{2.22}$$

$$= \sum_{ij,uv,kl} B_{ij}A_{uv}C_{kl}\delta_{v,i}\delta_{k,j}\,|ul\rangle \tag{2.23}$$

$$= \sum_{ij,u,l} B_{ij}A_{ui}C_{jl}\,|ul\rangle \tag{2.24}$$

by re-labelling the dummy indices:

$$= \sum_{ij,l,k} B_{lk}A_{il}C_{kj}\,|ij\rangle\,, \tag{2.25}$$

which is equal to Eq. (2.16). Therefore, the identity written in Eq. (2.11) is proven.

Now, I illustrate the application of the vectorisation operation in the study of quantum channels. Using the vectorisation of a matrix ϱ in Eq. (2.10), the Kraus representation of $\mathcal{E}$ is implicitly defined by

$$\mathrm{vec}(\mathcal{E}(\varrho)) = \sum_{A\in\mathbf{A}} A\varrho A^\dagger = \sum_{A\in\mathbf{A}} (A \times \bar{A})\,\mathrm{vec}(\varrho) = \Gamma_{\mathrm{K}}(\mathcal{E})|\varrho\rangle\!\rangle_{\mathrm{K}}, \tag{2.26}$$

where $\Gamma_{\mathrm{K}}(\mathcal{E}) \coloneqq \sum_{A\in\mathbf{A}} A \times \bar{A}$, the bar $^-$ denotes complex conjugation, and $\times$ denotes the Kronecker product. The vectorisation consistent with the Kraus representation is computed

as

$$(|\,|\psi\rangle\langle\psi|\rangle\rangle_{\mathrm{K}})_{(i-1)d+j} \coloneqq \langle i|\psi\rangle\,\langle\psi|j\rangle\,. \tag{2.27}$$

Next I discuss how to obtain Γ_{K} from Γ.

The Pauli-Liouville representation is linked to the Kraus representation by changing the basis used in the computation of the representation. Given a channel $\mathcal{E}$, $\Gamma(\mathcal{E})$ can be obtained by replacing the HW matrices in Eq. (2.6) with the matrices $\delta_{i,j}$. Having the representations Γ and Γ_{K}, the next step is to review the twirling procedure.

In the following section, I study how to reduce the number of parameters in a channel. This reduction is fundamental for the practicality of randomised benchmarking schemes.

2.3 Representation theory

In this section, I introduce several notions related to representation theory that are used in my thesis. I start with basic definition of reducible and irreducible irreps, followed by the celebrated Schur's lemma. Then I conclude with the rearrangement theorem and the importance of inequivalent irreps in my thesis.

In this section, every definition starts with the most formal version, and I conclude with a simplified version that reduces the tediousness. I do so to define a group and a representation. A group is a pair $G, \bullet$, where G is a finite set, and $\bullet$ is a binary mapping (takes two arguments) such that:

- for any $a, b, c \in G$, $a \bullet (b \bullet c) = (a \bullet b) \bullet c$.
- there is some $e \in G$ such that for every $g \in G$ e satisfies $g \bullet e = e \bullet g = g$.
- for every element $g \in G$ there is some g^{-1} such that $g \bullet g^{-1} = g^{-1} \bullet g = e$.

In the remainder of this thesis, I do not use the pair to denote a group; instead, I only use G.

Let ρ be a mapping taking as arguments elements of a group G to the set of linear mappings on a vector space $\mathscr{V}$. Then ρ is a representation of G if it satisfies, for any $a, b \in G$,

$$\rho(ab) = \rho(a)\rho(b), \tag{2.28}$$

which makes ρ a homomorphism.

A representation ρ is reducible if there is a non-trivial subspace $\Sigma \subset \mathscr{V}$ such that for all $g \in G$ and for all $v \in \Sigma$, $\rho(v) \in \Sigma$, which is concisely written as

$$\rho(\Sigma) \subset \Sigma. \tag{2.29}$$

A representation is irreducible if the only subspaces Σ that satisfy Eq. (2.29) is the complete and the null subspaces. For example, consider a representation ρ acting on $\mathscr{V}$ such that

$$\mathscr{V} = \Sigma_0 \oplus \Sigma_1, \tag{2.30}$$

where Σ_0 and Σ_1 are two minimal invariant subspaces. Minimal means that any other invariant subspace W is either Σ_0 or Σ_1 or W strictly contains Σ_0 or Σ_1. Similarly to the decomposition of vector spaces, the mapping ρ is decomposed. I denote the decomposition with the same symbol as for vector spaces:

$$\rho = \rho_0 \oplus \rho_1, \tag{2.31}$$

where ρ_i is the restriction of ρ into Σ_i, with $i \in \{0, 1\}$.

In physics, the subspace on which a representation ρ acts is also called a representation. In particular, the minimal invariant subspaces are also known as irreducible representations or irreps, for short. This abuse of language, for the term representation, comes in handy in the study of representation decomposition. Decomposition of a representation refers to identifying the invariant subspaces, and for each invariant subspace, a basis. The basis of

each subspace is then used as a basis for the whole Hilbert space, and the states are labelled according to the invariant subspace they belong. Then, it is said that ρ decomposes into two irreps Σ_0 and Σ_1.

A somehow simple result from group theory is useful in many parts of this thesis. Consider a group G. The rearrangement theorem states that, considering any ordered list of elements of G:

$$(g_0, g_1, \ldots, g_{|G|}), \qquad g_0, g_1, \ldots, g_{|G|} \in G \tag{2.32}$$

multiplying, without loss of generality, by the left each element of the sequence with another fixed element of G, say $h \in G$, only results in a permutation of the group elements

$$h \bullet (g_0, g_1, \ldots, g_{|G|}) = (h \bullet g_0, h \bullet g_1, \ldots, h \bullet g_{|G|}) = (g_{\sigma(0)}, g_{\sigma(1)}, \ldots, g_{\sigma(|G|)}), \tag{2.33}$$

where σ is some permutation.

The next item in the representation theory agenda is Schur's lemma. Consider a linear mapping M acting on a vector space $\mathscr{V}$. Given an irrep ρ of a group G, Schur's lemma states that if for any $g \in G$

$$\rho(g)M = M\rho(g), \tag{2.34}$$

then M is a multiple of the identity: $M = \alpha\mathbb{I}$ [82].

Now, I construct a linear mapping that is used in randomised benchmarking. Again , consider a linear mapping M acting on a vector space $\mathscr{V}$. Now for a representation ρ compute

$$[M]_\rho \coloneqq \sum_{g \in G} \rho(g) M \rho(g)^{-1}. \tag{2.35}$$

Now, I prove that $[M]_\rho$ commutes with any representative $\rho(h)$:

$$\rho(h)[M]_\rho = \sum_{g\in G} \rho(h)\rho(g)M\rho(g)^{-1} = \sum_{g\in G} \rho(h\bullet g)M\rho(g)^{-1}, \tag{2.36a}$$

$$= \sum_{g\in G} \rho(h)\rho(g)M\rho(g)^{-1} = \sum_{g\in G} \rho(h\bullet g)M\rho(h^{-1}\bullet h\bullet g)^{-1} \tag{2.36b}$$

$$= \sum_{g\in G} \rho(h)\rho(g)M\rho(g)^{-1} = \sum_{g\in G} \rho(h\bullet g)M\rho(h\bullet g)^{-1}\rho(h^{-1})^{-1} \tag{2.36c}$$

$$= [M]_\rho\rho(h). \tag{2.36d}$$

This shows that $[M]_\rho$ is a multiple of the identity for any M. Thanks to the cyclic property of the trace, I can compute the proportionality constant.

Compute the trace of $[M]_\rho$:

$$\mathrm{tr}[[M]_\rho] = \sum_{g\in G} \mathrm{tr}\big[\rho(g)M\rho(g)^{-1}\big] = \sum_{g\in G} \mathrm{tr}[M] = |G|\,\mathrm{tr}[M]. \tag{2.37}$$

By Schur's lemma, on the other hand, I know that $\mathrm{tr}[[M]_\rho] = \alpha d$, where d is the dimension of $\mathscr{V}$. Thus $\alpha d = |G|\,\mathrm{tr}[M]$, therefore

$$\alpha = \frac{|G|}{d}\,\mathrm{tr}[M], \tag{2.38}$$

thus now $[M]_\rho$ is completely known.

Now, I discuss the version of the previous result (the form of the operator $[M]_\rho$) that is used in randomised benchmarking. I need to discuss the concept of inequivalent irreps in the decomposition of a representation. For this discussion, a representation refers to the mapping and not to the vector space on which it acts. Two irreps are inequivalent if they are not isomorphic; that is, one cannot be transformed into the other by a change of basis. A representation decomposes into inequivalent irreps if each irrep in the decomposition is non-isomorphic to the rest of irreps.

Let ρ decompose into n irreps ρ_i, with $i \in \{0, \ldots, n-1\}$, such that

$$\mathscr{V} = \mathscr{V}_0 \oplus \cdots \oplus \mathscr{V}_{n-1}, \tag{2.39}$$

where each ρ_i acts invariantly on $\mathscr{V}_i$. Then, for any linear operator M,

$$[M]_\rho = \alpha_0 \mathbb{I} \oplus \cdots \oplus \alpha_{n-1} \mathbb{I}, \tag{2.40}$$

where $\alpha_i \coloneqq \frac{|G|}{d_i} \operatorname{tr}[M]$, i is the dimension of $\mathscr{V}_i$. In following sections, unless explicitly mentioned, the subspaces $\mathscr{V}$ are obviated and only the representation ρ is used. Additionally, I use Σ_i to denote the subspaces instead of $\mathscr{V}_i$ to keep the notation that I used in my papers.

I conclude this section with a brief review of the basic groups that appear later in this thesis. These groups are cyclic, symmetric, and dihedral groups [3]. To avoid dealing with unnecessary abstract definitions for a group, the groups I use in my thesis are introduced with either a geometric representation or an algebraic representation, whichever is simpler. Let $n \in \mathbb{N}$. The elements of the cyclic group C_n are powers of $\exp(2\pi \mathrm{i}/n)$. The symmetric group S_n is the group of all permutations of n elements. The dihedral group D_n is the group of transformations that leave a regular polygon with n sides invariant.

2.4 Clifford hierarchy

In this section I explain the Clifford hierarchy. Clifford gates have as motivation the teleportation scheme. In this scheme, the operations that can be teleported are Clifford operations [83]. The Pauli operations are denoted by $\mathcal{C}_1$. The Clifford gate set is defined as

$$\mathcal{C}_2 \coloneqq \{U \in U(d) \colon UPU^\dagger \in \mathcal{C}_1\}. \tag{2.41}$$

The set $\mathcal{C}_2$ is also known as the normaliser of the Pauli gate set. In the teleportation scheme, the normalisation property is exploited to apply the gate U to the teleported state. In the

same fashion, the next level of the Clifford hierarchy can be defined as

$$\mathcal{C}_3 \coloneqq \{U \in U(d) \colon UPU^\dagger \in \mathcal{C}_2\}. \tag{2.42}$$

Note that any gate G in $\mathcal{C}_3$ but not in $\mathcal{C}_2$ requires the composition of two teleportation circuits to apply G. In this thesis, my primary objective is to introduce a scheme to characterise operations in $\mathcal{C}_3 \setminus \mathcal{C}_2$, that is, operations in the third level of the Clifford hierarchy that are not Clifford.

2.5 Twirl

In this section, I discuss the context in which the averaging process known as twirling emerged. Then I provide the definitions and notations used in the rest of this thesis for mappings between channels.

The procedure known as twirling was introduced in the context of entanglement purification [84]. In entanglement purification, the goal is to use several copies of a noisy bi-qubit state to obtain a high-fidelity Bell state. In this procedure, a state invariant under any bi-qubit unitary transformation is required. Consider a state $\varrho_{SS'}$ on a bi-partite system SS'. Then, $\varrho_{SS'}$ is invariant under any bi-qubit unitary transformation iff

$$\varrho_{SS'} = (U \otimes U)\varrho_{SS'}(U^\dagger \otimes U^\dagger), \forall U \in U(2). \tag{2.43}$$

As consequence of the rearrangement theorem discussed on Eqs. (2.36), the invariance with respect to any bi-qutrit unitary can be achieved by averaging over the conjugates of the state ϱ with respect to each group representative.

$$[\varrho_{SS'}]_{\mathcal{O}} = \mathop{\mathbb{E}}_{U \in \mathcal{O}} U \varrho_{SS'} U^\dagger, \tag{2.44}$$

where the notation $\mathbb{E}_{s\in\boldsymbol{S}}$ refers to the average over the uniform measure over $\boldsymbol{S}$:

$$\mathop{\mathbb{E}}_{s\in\boldsymbol{S}} f(s) \coloneqq \frac{1}{|\boldsymbol{S}|}\sum_{s\in\boldsymbol{S}} f(s). \tag{2.45}$$

This procedure is known as twirling, and $[\varrho_{SS'}]_{\mathcal{O}}$ is the twirl of $\varrho_{SS'}$ with respect to the group representatives of $\mathcal{O}$, where representative means the image of a representation [82]. Now, I discuss another role in which twirling appears in the context of quantum information.

The concept of a superselection rule in quantum mechanics, that emerges to explain the lack of experimental evidence of certain transitions between states in a physical system. While standard quantum mechanics, based on the three "modern" postulates [4, 85], allows transitions between states with different eigenvalues, such as transitions between states of different charges, this has not been observed in practice. For instance, consider a system with a charge operator: although quantum mechanics allows transitions between states of different charges, no experimental evidence supports these transitions. A consequence of this is that states being in a linear superposition of states with different charge are prohibited.

This discrepancy is addressed by the introduction of superselection rules, which are motivated by the need to exclude unphysical phenomena. A superselection rule ensures that only transitions between states with the same eigenvalue (such as charge) are allowed, thereby prohibiting transitions between states of different eigenvalues. Consequently, this prohibition becomes a defining characteristic of superselection rules for a given physical system [86].

In the context of quantum information, the lack of a reference frame (an undefined phase, for example) induces a superselection rule [87]. The superselection rule, in this case, is to forbid transitions between states with different phase. One consequence of a superselection rule is that the states must commute with some symmetry operators. As in the previous case of entanglement purification, a state that commutes with every operator in the symmetry group is constructed by twirling [87].

In this subsection, I discuss the role of twirling in randomised benchmarking schemes. I

highlight the Kraus representation and its relation to the twirl, as well as the importance of non-degenerate representations. In the rest of the thesis, I assume representations with no multiplicity, meaning, only inequivalent irreps, which is a concept introduced in Sec. 2.3, appear in the decomposition of a representation [3]. Designing schemes with multiplicity-free representations simplifies the study and is also a constraint that I impose on my schemes. The twirling procedure requires four ingredients: a Hilbert space $\mathfrak{h}$, a group G, an irrep γ of G acting on $\mathfrak{h}$, and a matrix M acting on $\mathfrak{h}$. Before proceeding further, I introduce a notation I use throughout this thesis. Consider a finite set $\boldsymbol{S} \coloneqq \{s_0, \ldots, s_l\}$ and a mapping $f\colon \boldsymbol{S} \to \mathbb{C}$. If the set $\boldsymbol{S}$ is known from the context, I omit writing it and just write $\mathbb{E}_s$. Then the twirl of M by γ is, as computed in Eq. (2.38),

$$[M]_G \coloneqq \mathop{\mathbb{E}}_{g\in G} \gamma(g) M \gamma(g)^\dagger = \mathbb{I}\,\mathrm{tr}(M); \tag{2.46}$$

the identity $\mathbb{I}$ should be seen as the projector onto $\mathfrak{h}$.

Now, I compute the form of the twirl using the concepts introduced in Sec. 2.3. Suppose instead γ is reducible and decomposes into two inequivalent irreps: $\gamma = \gamma_0 \oplus \gamma_1$, such that γ_0 and γ_1 are inequivalent. If the irreps are equivalent, then, as I discuss in the section on my qutrit results, the sequence fidelity has an oscillatory contribution, which is difficult to statistically recover requiring more sophisticated fitting resources [56]. This decomposition implies γ has two invariant subspaces in $\mathrm{end}(\mathfrak{h})$, namely $\mathfrak{h}_0$ and $\mathfrak{h}_1$:

$$\mathrm{end}(\mathfrak{h}) = \mathfrak{h}_0 \oplus \mathfrak{h}_1. \tag{2.47}$$

I label the projectors from $\mathrm{end}(\mathfrak{h})$ onto $\mathfrak{h}_i$ as Π_i. The twirl of M with respect to γ is

$$[M]_G = \mathop{\mathbb{E}}_{g\in G} \gamma(g) M \gamma(g)^\dagger = \sum_i \Pi_i \frac{\mathrm{tr}(\Pi_i M)}{\dim(\mathfrak{h}_i)} = \sum_i \Pi_i \eta_{\mathfrak{h}_i}(M), \tag{2.48}$$

which for a single irrep decomposition simplifies to Eq. (2.46). The value of $\eta_{\mathfrak{h}_i}$ is ob-

tained similarly to the one in Eq. (2.38). Originating in randomised benchmarking for gate-dependent noise [88, 89], the quantity

$$\eta_{\mathfrak{h}_i}(M) \coloneqq \frac{\mathrm{tr}(\Pi_i M)}{\dim(\mathfrak{h}_i)} \tag{2.49}$$

is known as the eigenvalue associated with irrep $\mathfrak{h}_i$ [88]. These eigenvalues, as I discuss later, are proxy quantities to estimate the average gate fidelity.

Now, I describe the connection of a unitary representation with the Kraus representation; for a given unitary representation γ, the corresponding Kraus representation is

$$\Gamma_{\mathrm{K}}(g) \coloneqq \gamma(g) \otimes \bar{\gamma}(g) \in \mathrm{end}(\mathfrak{h}). \tag{2.50}$$

Assuming Γ_{K} decomposes as a sum of l (an integer between 1 and d^2) inequivalent irreps:

$$\mathrm{end}(\mathfrak{h}) \cong \mathfrak{h}_1 \oplus \mathfrak{h}_2 \oplus \cdots \oplus \mathfrak{h}_l, \tag{2.51}$$

the twirl of channel $\mathcal{E}$ by a group G with one irrep γ is defined as

$$[\mathcal{E}]_G \coloneqq \mathop{\mathbb{E}}_{g\in G} \Gamma_{\mathrm{K}}(g)\Gamma_{\mathrm{K}}(\mathcal{E})\Gamma_{\mathrm{K}}(g)^\dagger = \sum_{i\in[l]} \Pi_i \eta_{\mathfrak{h}_i}(\Gamma_{\mathrm{K}}(\mathcal{E})) \tag{2.52}$$

The decomposition in Eq. (2.52) is key to my investigation.

Twirling by the Clifford and Pauli groups is commonly done in the randomised benchmarking literature [90]. The twirl of a channel with respect to the Clifford group is a totally depolarising channel [51]; the twirl with respect to the Pauli group is a Pauli channel [66]. In representation theory terms, it implies that the Kraus representation of the Clifford group decomposes into three inequivalent irreps; the Kraus representation of the Pauli group decomposes into d^2 inequivalent irreps.

This is a good place to discuss the semidirect product as it appears in the literature [91]

and then introduce the way I use it throughout my research. Consider a group N and a subgroup of the group of automorphisms of N called G. I recall that a member of the group N is a bijective mapping from G to G. Because N is finite, a member of N is a permutation over G. The outer semidirect product between G and N is formed as follows. Let $g, h \in G$ and $n, m \in N$, then the product between (g, n) and (h, m) is

$$(n, g) \bullet (m, h) \coloneqq (n \circ m, m(g) \bullet h), \tag{2.53}$$

where $n \circ m$ is the composition of automorphism and $m(g)$ the action of m on g. The set $G \times N$ together with the product defined in Eq. (2.53) is known as the semidirect group between G and N and is denoted by $G \ltimes N$.

Instead of using the abstract definition of the previous paragraph, I use a practical form with the help of representations; this is the form mostly used by physicists [92]. Consider a group N with a representation γ. Consider another group G with a representation ρ. Further assume representations γ and ρ satisfy the following property: for any $g \in G$ and any $n \in N$ $\rho(g)\gamma(n)\rho(g)^\dagger = \gamma(n')$, for some n'. Then the set of matrices spanned by $\{\gamma(n) \colon n \in N\}$ and $\{\rho(g) \colon g \in G\}$ is a representation of $G \ltimes N$: every element is written as $\{\rho(g)\gamma(n) \colon (g, n) \in G \times N\}$.

Now, I illustrate the definition of a semidirect product based on representatives with the case of the dihedral group D_{16} found in the paper on qubit universal benchmarking [2, 66]. Consider the matrices

$$T = \begin{bmatrix} \omega_{16} & 0 \\ 0 & \omega_{16}^7 \end{bmatrix} \text{ and } X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{2.54}$$

where ω_{16} is $\exp(2\pi \mathrm{i}/16)$ and $\overline{\omega}_{16}$ its complex conjugate. A representation $\gamma_{D_{16}}$ of the dihedral group is

$$\gamma_{D_{16}}(x, t) \coloneqq X^x T^t, \tag{2.55}$$

where $(x, t) \in C_2 \times C_{16}$ denotes an element of $D_{16} \cong C_2 \ltimes C_{16}$; I use both x and t to denote

elements of a cyclic group (C_2 and C_{16}) as integers of a ring ($\mathbb{Z}_2$ and $\mathbb{Z}_{16}$).

I identify the terms of the representation $\gamma_{D_{16}}$ according to the notation I introduced in this section: $G = C_2$, $N = C_{16}$, $\gamma(x) = X^x$, and $\rho(t) = T^t$. The multiplication rule is

$$X^{x_1}T^{t_1}X^{x_2}T^{t_2} = X^{x_1+x_2}T^{(-1)^{x_2}t_1+t_2}, \tag{2.56}$$

which is identified using the following algebraic manipulations

$$X^{x_1}T^{t_1}X^{x_2}T^{t_2} = X^{x_1}X^{x_2}X^{-x_2}T^{t_1}X^{x_2}T^{t_2}, \tag{2.57}$$

$$= X^{x_1+x_2}(X^{-x_2}T^{t_1}X^{x_2})T^{t_2}, \tag{2.58}$$

by comparison of the entries resulting from conjugating with respect to X^{x_2}:

$$= X^{x_1+x_2}T^{(-1)^{x_2}t_1}T^{t_2}, \tag{2.59}$$

$$= X^{x_1+x_2}T^{(-1)^{x_2}t_1+t_2}. \tag{2.60}$$

Using the representations γ and ρ, a representation of D_{16} is constructed.

I conclude this section by defining some notation used throughout my thesis. Firstly, G denotes an abstract group [91], and $g \in \mathrm{G}$ is a group element. Given a representation γ of G, $\gamma(g)$ is the representative of g.

More importantly, in the context of randomised benchmarking schemes, are the concepts of noisy and ideal gates. I discuss first ideal gates. Let G be a finite group with elements $\{g\}$. Let ϱ be a density matrix and γ a representation of G. The ideal gate (denoted by a caret) corresponds to the map

$$\hat{g}(\varrho) \coloneqq \gamma(g)\varrho\gamma(g)^\dagger. \tag{2.61}$$

Noise models for gates are divided into two categories: gate-independent and gate-dependent, which refer to whether the error is the same or varies accross the gate set.

Noisy gates are denoted with a tilde above them, but the explicit form depends on the application. The most common form of noisy gate is

$$\widetilde{g}(\varrho) \coloneqq \mathcal{E}_g \circ \hat{g}(\varrho), \tag{2.62}$$

where the subscript $_g$ denotes dependence on the gate; in the case of gate-independent noise, $\widetilde{g}(\varrho) = \mathcal{E} \circ \hat{g}(\varrho)$ ($\circ$ denotes composition). For a T gate (or, generally, any non-Clifford gate), the noisy gate is of the form

$$\widetilde{T} \coloneqq \hat{T} \circ \mathcal{E}_T. \tag{2.63}$$

The inversion in order of application for the T gate is valid and exclusively done by convenience [93, 60]. The only condition is to be consistent with the ordering throughout the analysis. Other forms of including noise are ad-hoc and are discussed in their context.

2.6 Context for the introduction of randomised benchmarking

In this section, I write down the main actors present during the time of the formulation of randomised benchmarking. Whereas the notion of twirl is fundamental and thus discussed in its own section, two main actors were also present: process tomography and unitary designs.

Process tomography [94, 48, 95] is a well-known technique for the reconstruction of quantum operations [4]. This technique requires total knowledge of the input states of the systems or assumes the input states are ideal. From this assumption, standard state tomography allows for the recovery of the matrix entries of the χ-representation of a CPTP mapping. For single qudit gates, the tomographic procedure requires estimating $d^4 - d^2$ real parameters, see my proof in Appendix B.4. The number of experiments employed in tomographic techniques scales exponentially with the number of qudits, n, and so becomes rapidly impractical in terms of resources (such as computing time and lab time) required.

The motivation for randomised benchmarking becomes clear in light of my summary of process tomography. The exponential scaling justifies the claim that tomography does not scale well. One of the main advantages of randomised benchmarking over tomography is its much better scaling with d and n. Additionally, assuming perfect states is inconsistent and unrealistic. The reason is that to prepare states, gates are applied to a unique state, usually $|0\rangle$. Thus, by using process tomography, the characterization becomes inconsistent as it always relies on either improperly characterised gates or improperly characterised states.

The second important ingredient for randomised benchmarking, as it is currently known, is unitary 2-designs [96]. Based on spherical designs [97], Dankert et al. introduced the concept of unitary 2-designs to estimate second moments with respect to the Haar measure of the unitary group. Furthermore, they realised that the Clifford group, for prime dimensions, is a unitary 2-design. Notably, twirling with respect to the Clifford group was introduced earlier by DiVincenzo et al. [98].

A unitary 2-design is formally defined as a finite set of matrices $\mathbf{U} = \{U_i \colon i \in I\}$, where I is an index set, that satisfies the following condition:

$$\int \mathrm{d}U U^{\otimes 2} \otimes (\bar{U})^{\otimes 2} = \sum_{i \in I} U_i^{\otimes 2} \otimes (\bar{U}_i)^{\otimes 2}, \tag{2.64}$$

where $\mathrm{d}U$ is the Haar measure over the unitary group. The importance of Eq. (2.64) and unitary 2-designs in randomised benchmarking can be seen through the vectorisation transformation. For a channel $\mathcal{E}$, let the twirl with respect to the unitary group and a unitary 2-design $\mathbf{U}$ be denoted $[\mathcal{E}]_{U(d)}$ and $[\mathcal{E}]_{\mathbf{U}}$, respectively. Then:

$$[\mathcal{E}]_{\mathbf{U}} = \mathop{\mathbb{E}}_{i \in I} \Gamma_K(\hat{U}_i) \Gamma_K(\mathcal{E}) \Gamma_K(\hat{U}_i)^\dagger, \tag{2.65}$$

applying the vectorisation operation:

$$\mathrm{vec}([\mathcal{E}]_{\mathbf{U}}) = \mathop{\mathbb{E}}_{i \in I} U_i \otimes \bar{U}_i \mathrm{vec}(\Gamma_K(\mathcal{E})), \tag{2.66}$$

which, by the definition of unitary 2-design:

$$= \int \mathrm{d}U U \otimes \bar{U} \mathrm{vec}(\Gamma_K(\mathcal{E})), \tag{2.67}$$

applying the inverse vectorisation transformation:

$$\int \mathrm{d}U \Gamma_K(\hat{U}) \Gamma_K(\mathcal{E}) \Gamma_K(\hat{U})^\dagger = [\mathcal{E}]_{U(d)}. \tag{2.68}$$

In summary, I have shown that twirling with respect to a unitary 2-design produces the same operator as twirling with respect to the unitary group.

One result from Dankert et al. [96] is that the Clifford group forms a unitary 2-design. This fact was later applied to show that twirling any CPTP mapping reduces it to a totally depolarising channel. Showing that twirling results in a totally depolarising channel is equivalent to the fact that a gate set is a unitary 2-design, as derived by Magesan [99]. This is particularly relevant in my thesis since, as the totally depolarising channel has only one parameter, the smallest number of parameters in the twirl of a gate set that is not a 2-design must be 2. Thus, my goal is to generate a gate set that is close, in the sense discussed above, to a unitary 2-design.

2.7 Quantum volume

Currently, there are two other schemes as alternatives to randomised benchmarking for characterising quantum gates: one is the determination of quantum volume, and the other is gate set tomography. I discuss gate set tomography in the following section. In this section, I focus on how quantum volume is used and the abstract procedure employed to quantify the quality of a set of gates on any platform.

To provide more context to the state-of-the-art in quantum gate characterisation, I discuss quantum volume [100]. Quantum volume corresponds roughly to the number of qubits and

circuit depth that can implement an arbitrary unitary matrix with decent quality. The figure of merit for this task is dictated by the heavy output problem, which I will describe.

The heavy output problem takes as input two circuits, U and U', which are assumed to be unitary matrices. The output is a boolean value: true if U' is heavy, false otherwise. Suppose U is an n-qubit circuit. Then compute the following list of probabilities:

$$\vec{p} \coloneqq \{p_U(x) \colon p_U(x) \coloneqq \mathrm{tr}\Big[|x\rangle\langle x|\, \hat{U}(|0^n\rangle\langle 0^n|)\Big], x \text{ is a string with } n \text{ binary values } \}. \tag{2.69}$$

Consider sorting the values $\vec{p}$ from least to greatest: $p_0 \geq p_1 \geq \cdots \geq p_{2^n-1}$. The median probability is computed as $p_{\text{med}} \coloneqq (p_{2^{(n-1)}} + p_{2^{(n-1)}-1})/2$. From $\vec{p}$, a subset is formed as $\vec{p}_H \coloneqq \{p \in \vec{p} \colon p > p_{\text{med}}\}$. Given U, an output x is heavy if $p_U(X) > p_{\text{med}}$. Now, the circuit U' is heavy if the probability of drawing a heavy output x is greater than $2/3$.

In practice, U is an ideal unitary matrix, and U' is the approximation of U using a restricted gate set with the gates available on the platform. Additionally, in the context of estimating the volume, U' depends on the number of qubits used; thus, I denote it by $U'(n)$. The depth is then increased until a non-heavy label is assigned. The resulting depth is labelled $d(n)$. Iterating over distinct values of n, a set of pairs $\{(m, d(m))\}$ is generated. From this list of values, the quantum volume is implicitly defined as:

$$\log_2 V_Q = \underset{n}{\mathrm{argmax}} \min\left(n, d(n)\right). \tag{2.70}$$

The interpretation of $\log_2 V_Q$ is the largest n from the list of minimums computed from the list of pairs $(n, d(n))$. Current values for quantum volume range from up to 32 for IBM devices and up to 512 for Quantinuum [101]. IonQ, Oxford Quantum Circuits (OQC), and Rigetti perform up to 8 for V_Q.

While V_Q is easy to measure, its appropriateness for determining the quality of quantum gates is mathematically weak, as it has little connection to the error rate [59]. Therefore, V_Q is useful for benchmarking and comparing distinct platforms but not for ensuring fault-

tolerant devices, which requires a formal estimate of the error rate of quantum gates [4]. In the next section, I discuss gate set tomography, which, because it allows the reconstruction of gate sets, can be used to determine whether a device is fault-tolerant or not.

2.8 Process and gate set tomography

In this section, I discuss process and gate set tomography. These two schemes aim to reconstruct a gate or gate set. However, the approaches of these methods are quite different, and recently, process tomography has been considered obsolete in light of gate set tomography. Nevertheless, gate set tomography is a relatively new technique, and to introduce it properly, I will also discuss process tomography as a starting point.

Process tomography is a modification of state tomography. By replacing the state with the image of an imperfect gate $\mathcal{E}$ and assuming perfect states and measurements, the matrix elements of the χ-representation of $\mathcal{E}$ are estimated. This procedure requires many measurements. Additionally, process tomography has two main issues. First, the requirement of perfect state preparation and measurement is unfeasible and inconsistent because preparing a non-native state requires applying one of the gates being characterised. Second, the number of experiments required grows exponentially. Remedies such as compressed sensing are invalid for arbitrary noise [102]. To address the inconsistency issue for process tomography, gate set tomography was introduced.

Gate set tomography[2] estimates the Pauli-Liouville representation of a gate set from experimental data [103]. The Pauli-Liouville representation of the gates to be characterised is parametrised in such a way that the specific type of noise affecting the gates can be studied. The input consists of a set of gates and the preparation and measurement of some state ϱ. In contrast to process tomography, gate set tomography is robust against SPAM (state preparation and measurement) issues, so having an unknown initial state is

[2]This section is not an exhaustive exposition of gate set tomography. This method already includes several optimisations, which could hide a fair comparison with randomised benchmarking. The purpose of this section is to highlight the differences and similarities with randomised benchmarking.

not problematic. In this description, I assume the fitting procedure is given and do not go into detail, but I highlight that eliminating SPAM issues increases the classical computational cost. Even reconstructing two-qubit gates can be challenging for standard computers [104]. The procedure is divided into two parts: short sequences and long sequences of gates.

The minimisation process uses empirical frequencies from experiments for each sequence and compares them with simulated probabilities using the original model. This is repeated, with the previous set of parameters used in the next iteration of the minimisation. The basic unit for constructing a sequence is not usually a single gate but rather a germ. A germ is a sequence of gates that, when repeated, amplifies particular parameters. Since arbitrary germs may not lead to amplification, an optimisation procedure is necessary to select germs that amplify specific parameters. This optimisation process is numerical and based on the computation of singular vectors of a Jacobian matrix obtained from the partial derivatives of each matrix with respect to the parameters.

The first estimate proceeds as follows. Let s' label a particular measurement corresponding to a germ g_s composed $L > 0$ times, with initial and final states $|f_i\rangle$ and $\langle f_j|$. Then $f_{s'} = \mathrm{tr}[|f_j\rangle\langle f_j|\, g_{s'}(|f_i\rangle\langle f_i|)] = \mathrm{tr}\big[|f_j\rangle\langle f_j|\, g_s^L(|f_i\rangle\langle f_i|)\big]$. The χ^2 statistic is given by

$$\chi^2_{s'} = N\frac{(p_{s'} - f_{s'})^2}{p_{s'}} + N\frac{(p_{s'} - f_{s'})^2}{1 - p_{s'}}, \tag{2.71}$$

$$= N(p_{s'} - f_{s'})^2 \left(\frac{1}{p_{s'}} + \frac{1}{1 - p_{s'}}\right), \tag{2.72}$$

$$= \frac{N(p_{s'} - f_{s'})^2}{p_{s'}(1 - p_{s'})}, \tag{2.73}$$

where $p_{s'}$ is computed as $f_{s'}$ but using the model for the germs, and N is the number of repetitions used to estimate each $f_{s'}$. Summing over each possible sequence of germs (for a fixed L), we obtain $\chi_S \coloneqq \sum_{s'} \chi^2_{s'}$. Minimising χ^2_S yields an estimate of the parameters, which is used in a subsequent estimate of χ^2_S, with the model based on the gate set resulting from the previously estimated parameters.

The output of the minimisation over χ_S, repeated over different L values, provides the set

of parameters. These parameters are used in the next stage of the scheme. Having explained the minimisation with respect to χ^2, it is now simpler to explain the second estimate. The only difference is that L is larger, often ranging from hundreds to thousands of repetitions. The expression to maximise is:

$$\log \mathcal{L}_{\mathcal{S}} = \sum_{s' \in \mathcal{S}} \log \mathcal{L}_{s'}, \tag{2.74}$$

where $\log \mathcal{L}_{s'} = N f_{s'} \log(p_{s'}) + N(1 - f_{s'}) \log(1 - p_{s'})$. Similar to the χ^2 minimisation, the maximisation of $\log \mathcal{L}_{\mathcal{S}}$ is iterated to reduce the error in the parameters. From this procedure, the parameters of each germ are estimated, and from the germs, a set of gates is reconstructed.

Due to the large number of measurements required, this method is primarily intended for use with up to two qudit gates. Increasing the number of levels or qudits increases the number of parameters, requiring many more experiments. Another cost is associated with choosing the germs. If the rank of the Jacobian is not equal to the number of parameters, the set of germs needs to be modified, incorporating one germ into another and verifying again that the Jacobian satisfies the rank constraint.

The Pauli-Liouville representation serves as the common ground for both randomised benchmarking and gate set tomography. In randomised benchmarking, the estimate is the trace of the Pauli-Liouville representation, while in gate set tomography, the entire gate representation is estimated. The number of parameters (real and independent quantities) in a Pauli-Liouville representation is $d^{4n} - d^{2n}$, where n is the number of qudits. This can be seen by imposing the trace-preserving condition [93]. In contrast, randomised benchmarking estimates a single parameter. This, once again, shows that randomised benchmarking should be used as a first test for the quality of gates, followed by another characterisation mechanism. I also need to clarify that the two parameters my scheme aims to estimate are used to calculate the average gate fidelity, which is a single parameter of the Pauli-Liouville representation.

A one-to-one comparison of the number of experiments required (classical complexity)

might suggest process tomography is better since it does not require optimisation techniques. Nevertheless, as mentioned earlier, process tomography is inconsistent and should not be used. On the other hand, gate set tomography allows for the incorporation of knowledge about the noise affecting the gates, which could reduce the number of experiments required.

2.9 Randomised benchmarking and the quantum threshold theorem

In this section I discuss the characterisation carried out by a randomised benchmarking scheme. I review concepts from quantum information such as error rate and fidelity; quantities used to characterise quantum gates. I conclude the section with a discussion of the effects of different kinds of noise on the average gate fidelity.

The importance of randomised benchmarking schemes is that—by their use—an experimental group estimates the quality of their gates inexpensively compared to tomography. Why is it relevant to know the quality of quantum gates?—the reason is the quantum threshold theorem (QTT) [59]. The QTT is stated as follows [4, 59]:

Theorem 2.2 ([6])**.** *Consider a quantum circuit $\mathscr{Q}$, with depth t (number of time steps), size s (number of one and two qubit gates), and with n (number of qubits). If the worst local error rate [lower-bounded by the diamond distance] η is below a threshold η_0, then, for all $\varepsilon > 0$, there is a quantum circuit $\mathscr{Q}'$ with a polylog overhead in t, s, n, and $1/\varepsilon$, such that the [diamond distance] between $\mathscr{Q}$ and $\mathscr{Q}'$ is less than ε.*

What is the relation between randomised benchmarking schemes and QTT? The estimate of η. The quality of gates, corresponding to the diamond distance of the noise with respect to the identity map necessary for the application of the QTT, is estimated with the diamond distance between the ideal gate $\mathcal{E}_{\text{id}}$ and the actual gate $\mathcal{E}_{\text{ac}}$. This quantity is known as the error rate [59].

The experimental estimate of the diamond distance requires the reconstruction of the gates via quantum tomography. However, as discussed before, the implementation of tomography does not scale [105]. A convenient method to characterise quantum gates is randomised benchmarking.

Randomised benchmarking arises as a method to estimate an alternative quantity that still provides a sense of quality: the average gate fidelity [106]. I now discuss the figures of merit motioned in this introduction and then discuss the adequacy of the average gate fidelity to characterise gates. The average gate fidelity is defined for a pair of quantum channels, $\mathcal{E}$ and $\mathcal{E}'$. Let $\mathrm{d}\psi$ denote the Fubini-Study measure over pure states in a d-dimensional system [107]. Then the average gate fidelity between $\mathcal{E}$ and $\mathcal{E}'$ is [107]

$$F(\mathcal{E},\mathcal{E}') \coloneqq \int \mathrm{d}\psi \,\mathrm{tr}(\mathcal{E}'(|\psi\rangle\langle\psi|)\mathcal{E}(|\psi\rangle\langle\psi|)). \tag{2.75}$$

Now, I review a simpler formula to compute the average gate fidelity for finite dimensional systems [53]. Let $\mathcal{E}$ be a channel. Then, the average gate fidelity of $\mathcal{E}$ is

$$F(\mathcal{E}) \coloneqq F(\mathcal{E},\mathbb{I}) = \frac{d\,\mathrm{tr}(\Gamma(\mathcal{E})) + d^2}{d^2(d+1)} = \frac{d\,\mathrm{tr}(\Gamma_{\mathrm{K}}(\mathcal{E})) + d^2}{d^2(d+1)}. \tag{2.76}$$

Now, I proceed to prove the equality of Eq. (2.76). First, I note that the integral over the measure of states is equivalent to integrating over the Haar measure of the unitary group. Therefore, the first step in my proof is

$$F(\mathcal{E},\mathcal{E}') = \int_U \mathrm{d}U \,\mathrm{tr}\Big[\mathcal{E}\circ\hat{U}(|0\rangle\langle 0|)\mathcal{E}'\circ\hat{U}(|0\rangle\langle 0|)\Big]. \tag{2.77}$$

The next step is to show that the twirl with respect to the unitary group is present in

Eq. (2.77).

$$\int \mathrm{d}U \langle\langle \mathcal{E} \circ \hat{U}(|0\rangle\langle 0|) || \mathcal{E} \circ \hat{U}(|0\rangle\langle 0|) \rangle\rangle = \int \mathrm{d}U \langle\langle 0|\Gamma(\mathcal{E} \circ \hat{U})^\dagger \Gamma(\mathcal{E}' \circ \hat{U})|0\rangle\rangle \tag{2.78}$$

$$= \int \mathrm{d}U \langle\langle 0|\Gamma(U)^\dagger \Gamma(\mathcal{E})^\dagger \Gamma(\mathcal{E}')\Gamma(U)|0\rangle\rangle \tag{2.79}$$

Here, I identify the twirl as the integral with respect to U:

$$= \langle\langle 0| \left(\int \mathrm{d}U \Gamma(U)^\dagger \Gamma(\mathcal{E})^\dagger \Gamma(\mathcal{E}')\Gamma(U) \right) |0\rangle\rangle \tag{2.80}$$

$$= \langle\langle 0|[\mathcal{E}^\dagger \circ \mathcal{E}']_{U(d)}|0\rangle\rangle. \tag{2.81}$$

Now, the twirl with respect to the unitary group of a channel $\mathcal{E}$ is a depolarising channel with depolarising parameter $p_\mathcal{E}$ [53], which is given by

$$p_\mathcal{E} \coloneqq \frac{d^2 - \mathrm{tr}[\Gamma(\mathcal{E})]}{d^2 - 1}. \tag{2.82}$$

$$F(\mathcal{E}, \mathcal{E}') = \langle\langle |0\rangle\langle 0| \, |[\mathcal{E}^\dagger \circ \mathcal{E}']_{U(d)}| \, |0\rangle\langle 0| \rangle\rangle \tag{2.83}$$

$$= (1 - p_\mathcal{E}) + \frac{p_\mathcal{E}}{d} = \frac{d \, \mathrm{tr}(\Gamma(\mathcal{E}^\dagger \circ \mathcal{E}')) + d^2}{d^2(d+1)}, \tag{2.84}$$

This concludes the proof.

The formula in Eq. (2.76) is widely used in randomised benchmarking schemes; twirling does not change the trace, which is related to the average gate fidelity [53]. Therefore, if twirling "simplifies" the noise channel (reducing the number of independent entries), then the estimate of the average gate fidelity is simpler by using the "twirled" noise. In this thesis, a gate is considered as characterised if it has been assigned an average gate fidelity. I use the word assigned since some schemes, particularly cycle benchmarking, by construction estimate an approximation of the average gate fidelity.

The average gate fidelity can differ, for most forms of noise, from the error rate (the

diamond distance between the noise and the identity) by several orders of magnitude [59]. In general, there is no direct relationship between these quantities: the error rate neither bounds nor is bounded by the average gate fidelity. This discrepancy poses a significant challenge to randomised benchmarking schemes since the ultimate goal of any characterisation scheme should be to approximate or establish a lower bound for the error rate.

However, a solution to this problem is now available. The technique known as randomised compiling [58] converts any CPTP mapping into a unital[3] Pauli channel. Because there is a known relationship between the average gate fidelity and the error rate for Pauli channels [59], this result is crucial. Now, if the randomised compiling scheme can be applied, any form of noise can be characterised and correlated with the error rate.

The benefits of randomised compiling extend beyond linking the error rate with the average gate fidelity. Transforming arbitrary noise, and particularly coherent noise, into a Pauli channel is also advantageous. It has been demonstrated that coherent noise leads to worse scaling, in terms of the decay of fidelity for deeper circuits, compared to incoherent noise [59]. Thus, applying the randomised compiling scheme allows to mitigate the decrease in fidelity characteristic of coherent noise [108, 58].

2.10 Interpretation of the parameter estimated

In this section I discuss the figure of merit estimated by randomised benchmarking schemes. I go over the original example, which assumes gate-independent noise, and then contrast it against the gate-dependent noise scheme [106, 88]. I discuss how the average gate fidelity is estimated from the sequence fidelity as done in the original formulation of randomised benchmarking. I conclude this section and chapter by explaining the link between the twirl, the trace of the Pauli-Liouville representation, and the average gate fidelity.

The first formulation of randomised benchmarking schemes assumes that every member of the gate set to characterise has the same noise [106, 2]. Under this assumption, the quantity

[3] A unital channel is defined as a CPTP map that maps the identity to the identity [68].

estimated by randomised benchmarking is the average gate fidelity of each gate set member. In the earlier randomised benchmarking schemes [51], it was suggested to use a perturbative approach to weaken the assumption of gate-independent noise [60]; I only discuss it briefly in this paragraph. Such a suggestion involved adding a small gate-dependent perturbation to the "average" noise $\mathcal{E}$. Due to the more complicated expression for the sequence fidelity, the perturbation method is no longer discussed in the literature. Moreover, it has been shown to be superfluous [55, 89].

What does the original randomised benchmarking estimate for the gate-dependent case? From the analysis done with gate-dependent noise [88, 89], two conclusions can be drawn. First, the average gate fidelity estimated does not, in general, correspond to the average fidelity over a gate set; second, the average gate fidelity remains gauge-dependent. Gauge-dependent means that depending on a similarity transformation the characterisation could change. Therefore, randomised benchmarking probably should only be used for gate sets with similar noise.

The Fourier method, also known as the convolution method [88], is a theoretical argument that, even in the presence of gate-dependent noise, a single-exponential decay curve should be obtained for the sequence fidelity. In this subsection I discuss the main points of the gate-dependent formulation, enough to later justify that my schemes are compatible with it. The Fourier transform method [88, 89] (a different method than the one discussed at the back of a well-known textbook [4]) justifies the similarity of the average (over the gates in the gate set) of the average gate fidelity and the value estimated from randomised benchmarking experiments.

Now, I discuss the gate-dependent noise scheme following Wallman's [89]. Consider a set of CPTP mappings labelled by elements of the Clifford gate set G; I denote the noisy representatives (of G) by $\widetilde{g}$ and the ideal representatives by $\hat{g}$. Two matrices l and r are

computed to obtain a gate-dependent approximation of $\widetilde{g}$:

$$\widetilde{g} \coloneqq \hat{l}_g \circ \hat{g} \circ \hat{r}_g \approx \hat{l} \circ \hat{g} \circ \hat{r}; \tag{2.85}$$

which means having a gate-independent approximation to the gate-dependent problem; note that $\hat{l}$ and $\hat{r}$ do not depend on the gate and could not even be channels, just mappings, since these mappings could fail to be completely positive or trace preserving.

Now, I discuss an issue affecting randomised benchmarking schemes, which is known as gauge-dependence [55]. Neglecting a gauge transformation in the SPAM leads to confusing a similarity transformation with coherent errors [55]. This misunderstanding results in orders of magnitude differences between the average gate fidelity and the parameter fitted. I discuss two attempts to address this issue.

The quantity known as circuit fidelity [109] was introduced to provide meaning to the parameter estimated in randomised benchmarking schemes. To discuss circuit fidelity, consider a gate set G. For $\mathbf{g} \coloneqq (g_0, \ldots, g_m) \in G^m$ define $\hat{S} \coloneqq \hat{g}_0 \circ \cdots \circ \hat{g}_m$ and $\widetilde{S} \coloneqq \widetilde{g}_0 \circ \cdots \circ \widetilde{g}_m$. Thus, the circuit fidelity for the gate set G at depth m is defined as:

$$F(G; m) \coloneqq \mathop{\mathbb{E}}_{(g_0,\ldots,g_m)\in G^m} F(\hat{S}, \widetilde{S}). \tag{2.86}$$

This quantity is what is now believed to be estimated by randomised benchmarking schemes [109].

Now, I discuss how the sequence fidelity is used to estimate the average gate fidelity; this is a basic task in any randomised benchmarking scheme [56]. Moreover, I discuss how the sequence fidelity can be approximately estimated from an experiment with low design complexity. Given a sequence of gates $\mathbf{g}$, the inversion gate $\hat{\mathbf{g}}_{\text{inv}}$ is the composition and then the inverse of the elements of $\mathbf{g}$. As an observation, for ideal gates, the composition of $\mathbf{g}$ and $\hat{\mathbf{g}}_{\text{inv}}$ results in the identity gate; in the presence of noise, only an approximation of the ideal gate is obtained. This is the most explicit description of a randomised benchmarking scheme in my thesis.

One of the two important quantities in a randomised benchmarking scheme is sequence fidelity. Here I introduce the sequence fidelity for a set of gates with same noise, which is the configuration used in Chapters 2 and 3. Consider a sequence of m noisy gates $\mathbf{g}$ (concluding with an inversion gate) with noise $\mathcal{E}$, an initial state $|\psi\rangle$, and a measurement $\langle\phi|$. The sequence fidelity associated with the previous ingredients is

$$P(m;\mathbf{g},|\psi\rangle,\langle\phi|) \coloneqq \operatorname{tr}\left(|\phi\rangle\langle\phi|\,\mathcal{E}\circ\hat{\mathbf{g}}_{\text{inv}}\circ\mathcal{E}\circ\hat{g}_m\cdots\mathcal{E}\circ\hat{g}_0(|\psi\rangle\langle\psi|)\right). \tag{2.87a}$$

Given a gate set G and considering the average over G for each $\hat{g}_i$ we obtain

$$P(m;G,|\psi\rangle,\langle\phi|) \coloneqq \mathop{\mathbb{E}}_{\mathbf{g}\in G^{\times m}} \operatorname{tr}\left(|\phi\rangle\langle\phi|\,\mathcal{E}\circ\hat{\mathbf{g}}_{\text{inv}}\circ\mathcal{E}\circ\hat{g}_m\cdots\mathcal{E}\circ\hat{g}_0(|\psi\rangle\langle\psi|)\right). \tag{2.87b}$$

In an abuse of language, both Eqs. (2.87) are called sequence fidelity, with the average in Eq. (2.87b) understood from context.

Now, I describe a run of a randomised benchmarking experiment and discuss the quantity that the measurement approximates. The computations presented in this section are an intrinsic part of any randomised benchmarking scheme [106, 2], which I introduce using the notation from this chapter. I present these computations here because in later chapters I discuss modifications to them.

1. Prepare state $|0\rangle$; I call this the input state.
2. Randomly draw m group elements $\mathbf{g} \leftarrow (g_0,\ldots,g_{m-1})$.
3. Compute the inverse of the ideal sequence $\hat{\mathbf{g}}_{\text{inv}} \leftarrow (g_0\cdots g_{m-1})^{-1}$.
4. Apply the sequence (including the inversion gate) of noisy gates onto the initial state: $\varrho \leftarrow \tilde{\mathbf{g}}_{\text{inv}}\circ\widetilde{g}_0\circ\cdots\circ\widetilde{g}_{m-1}(|0\rangle\langle 0|)$.
5. Measure ϱ with respect to the initial state, which is $|0\rangle$.
6. Repeat the above steps K times.

7. Average the results of the repetitions.

Let me denote by G^m the multiset—as there could be repetitions—of sequences $\boldsymbol{g}$ with every sequence of depth m. Note that $\mathbf{g}$ is a tuple and $\mathbf{g}_{\text{inv}}$ a group element. The quantity thus estimated in the previous procedure is

$$P(m;|0\rangle,\langle 0|,\mathcal{E}) \coloneqq \mathop{\mathbb{E}}_{\mathbf{g}\in G^m} \operatorname{tr}\{|0\rangle\langle 0|\,\widetilde{\mathbf{g}}_{\text{inv}}\circ\widetilde{g}_0\circ\cdots\circ\widetilde{g}_{m-1}(|0\rangle\langle 0|)\}, \tag{2.88a}$$

$$= \mathop{\mathbb{E}}_{\mathbf{g}\in G^m} \langle\!\langle 0|\Gamma(\widetilde{\mathbf{g}}_{\text{inv}}\circ\widetilde{g}_0\circ\cdots\circ\widetilde{g}_{m-1})|0\rangle\!\rangle, \tag{2.88b}$$

$$= \mathop{\mathbb{E}}_{\mathbf{g}\in G^m} \langle\!\langle 0|\Gamma(\mathcal{E})\Gamma(\hat{\mathbf{g}}_{\text{inv}})\Gamma(\mathcal{E})\Gamma(\hat{g}_0)\cdots\Gamma(\mathcal{E})\Gamma(\hat{g}_{m-1}))|0\rangle\!\rangle, \tag{2.88c}$$

where $\mathcal{E}$ is the noise accompanying the ideal application of the accompanying gate g_i or $\boldsymbol{g}_{\text{inv}}$.

I now show Eq. (2.88a) averaging over g_i results in the expression for the sequence fidelity. Consider first the average over each g_i

$$P(m;|0\rangle,\langle 0|,\mathcal{E}) = \mathop{\mathbb{E}}_{\mathbf{g}}\langle\!\langle 0|\Gamma(\mathcal{E})\Gamma((\hat{g}_{m-1}\circ\cdots\circ\hat{g}_0)^{-1})\Gamma(\mathcal{E})\Gamma(\hat{g}_0)\cdots\Gamma(\mathcal{E})\Gamma(\hat{g}_{m-1})|0\rangle\!\rangle; \tag{2.89a}$$

note here the fact the inversion gate is a member of the gate set. Therefore, I can assign the same noise $\mathcal{E}$ to $(g_0\cdots g_{m-1})^{-1}$. Now, I expand the inversion operation over the gates

$$P(m;|0\rangle,\langle 0|,\mathcal{E}) = \mathop{\mathbb{E}}_{\mathbf{g}}\langle\!\langle 0|\Gamma(\mathcal{E})\Gamma(\hat{g}_{m-1})^\dagger\cdots\Gamma(\hat{g}_0)^\dagger\Gamma(\mathcal{E})\Gamma(\hat{g}_0)\cdots\Gamma(\mathcal{E})\Gamma(\hat{g}_{m-1})|0\rangle\!\rangle. \tag{2.89b}$$

Averaging over g_0 we obtain

$$P(m;|0\rangle,\langle 0|,\mathcal{E}) = \mathop{\mathbb{E}}_{\mathbf{g}}\langle\!\langle 0|\Gamma(\mathcal{E})\Gamma(\hat{g}_{m-1})^\dagger\cdots\Gamma(\hat{g}_1)^\dagger[\mathcal{E}]_G\Gamma(\mathcal{E})\Gamma(\hat{g}_1)\cdots\Gamma(\mathcal{E})\Gamma(\hat{g}_{m-1})|0\rangle\!\rangle. \tag{2.89c}$$

The operator $[\mathcal{E}]_G$ commutes with every g_i:

$$P(m;|0\rangle,\langle 0|,\mathcal{E}) = \mathop{\mathbb{E}}_{\mathbf{g}}\langle\!\langle 0|\Gamma(\mathcal{E})\Gamma(\hat{g}_{m-1})^\dagger\cdots[\mathcal{E}]_G\Gamma(\hat{g}_1)^\dagger\Gamma(\mathcal{E})\Gamma(\hat{g}_1)\cdots\Gamma(\mathcal{E})\Gamma(\hat{g}_{m-1})|0\rangle\!\rangle. \tag{2.89d}$$

Averaging over g_1:

$$P(m;|0\rangle,\langle 0|,\mathcal{E}) = \mathop{\mathbb{E}}_{\mathbf{g}}\langle\langle 0|\Gamma(\mathcal{E})\Gamma(\hat{g}_{m-1})^{\dagger}\cdots[\mathcal{E}]_G^2\Gamma(\mathcal{E})\Gamma(\hat{g}_2)\cdots\Gamma(\mathcal{E})\Gamma(\hat{g}_{m-1})|0\rangle\rangle. \tag{2.89e}$$

Continuing averaging over each gate in the sequence, I obtain the twirl:

$$P(m;|0\rangle,\langle 0|,\mathcal{E}) = \langle\langle 0|\Gamma(\mathcal{E})[\mathcal{E}]_G^m|0\rangle\rangle. \tag{2.89f}$$

This is the final reduction of the expression $P(m;|0\rangle,\langle 0|,\mathcal{E})$. Depending on the form of the twirl $[\mathcal{E}]_G$ the expression changes; if the gate set is Clifford, then the expression is of the form $\frac{1}{2}+\frac{1}{2}\eta(\mathcal{E})^m$. For a general gate set the expression is more complicated and is a combination of more than one exponential [56].

2.11 Non-Markovianity

Whereas non-Markovianity is usually associated with noise dependent on time or previous gates applied, a formal description is that noise at different steps is correlated and thus cannot be described as a CPTP mapping acting only on the system of interest. This condition means that channels can only be written as acting on a tensor product between the system and an environment.

Non-Markovian gate-independent noise means that the noise associated with a gate interacts with an environment; that is, the action of the mapping is not restricted to the intended system. This form of noise can modify the expressions for sequence fidelity. In the Markovian approximation, as outlined in the original formulation of randomised benchmarking, the sequence fidelity is limited to being a sum of exponentials [56]. On the contrary, non-Markovian noise allows complicated forms, such as the ones reported by Figueroa-Romero *et al* [110].

The process-tensor framework [111, 112] provides a general formalism to separate the

'uncontrollable' interactions and the gates to be implemented. Using the process-tensor framework, it is possible to write formal expressions, similar in construction but not in simplicity, for the sequence fidelity under non-Markovian but gate-independent noise [111].

Contrary to Markovian randomised benchmarking, the sequence fidelity expression is too complicated to extract information related to the average gate fidelity. Among other issues, it is proven that the sequence fidelity expression contains SPAM contributions dependent on the circuit depth. Therefore, non-Markovian randomised benchmarking aspires only to detect signatures of non-Markovianity.

The detection of signatures of non-Markovianity is formally done by computing the minimal distance between the actual implementation and every possible Markovian circuit. In practice, this is impossible, and the comparison is restricted to sensible candidates.

The most tangible signature of non-Markovianity is a non-exponential form of the sequence fidelity [111]. Although, in principle, there could be non-Markovian noise that is capable of an exponential sequence fidelity, in that case, Markovian randomised benchmarking is applicable. This gives rise to a justification for applying randomised benchmarking: if the sequence fidelity follows, approximately, an exponential decay, randomised benchmarking offers a valid characterization. I apply this rule of thumb in future examples.

Finally, I must highlight that randomised compiling has been shown to mitigate non-Markovian noise effects [110, 58, 113]. Therefore, even under such adverse conditions, traditional randomised benchmarking remains useful, although with more overhead on the experimental side.

2.12 An ion trap qudit platform

In this section, I discuss in detail the implementation by Ringbauer et al. [14] of an ion trap qudit platform. My discussion of this platform is to illustrate the state-of-the-art of qudit implementations. Later, in Sec. 3.5, I also use this platform to justify my claims that the

assumptions of randomised benchmarking are currently satisfied by state-of-the-art qudit platforms.

I chose this qudit ion-trap platform for two reasons. First, ion-trap platforms are a familiar implementation, which facilitates describing them and assessing their physical workings [114]. Second, it has available a universal set of gates with state-of-the-art fidelity. Additionally, as is the nature of ion-trap platforms, the long coherence time allows us to implement long circuits, which helps in computing many values for the sequence fidelity. Also, the particular implementation I discuss explicitly reports the statistical details of their randomised benchmarking implementation.

This platform uses a $^{40}Ca^{+}$ ion as the physical qudit. The system is called an ion trap because it confines ions for prolonged periods [115]. Due to Earnshaw's theorem, maintaining the position of ions requires the use of a non-constant (time-dependent) electromagnetic field. Imposing mechanical equilibrium of the ions leads to conditions in the expression required for the electric field potential. However, the conditions on the electric potential from the mechanical condition are incompatible, for static electromagnetic fields, with Laplace's equation. Therefore, a time-dependent electric potential is required [116, 117, 118].

I discuss a classical picture of the system. Computing the force ions feel, from a time-dependent electric potential, leads to Mathieu's equation. Thus, the system can be modelled as a set of spring-coupled of multi-level systems. Each member of the string is used as a qudit and the collective motion, via normal modes, is used for entangling gates.

I discuss the encoding, the way states are labelled, used in Ringbauer et al. [14]. The states are labelled by the hyperfine Zeeman states; the splitting is caused by a magnet close to the atoms. A set of magnets generates a magnetic field splitting $S_{1/2}$ into two levels and $D_{5/2}$ into six levels. The first two levels are encoded in $|0\rangle = S_{1/2,-1/2}$ and $|1\rangle = D_{5/2,-1/2}$. The allowed transitions are $\Delta m = 0, \pm 1, \pm 2$, and these transitions form the basis for the native single-qudit gates.

Four stages of preparation are used. First, Doppler cooling and polarization gradient

cooling are used to put the system in one of the normal modes with lowest energy. The other two preparation procedures are optical pumping and resolved sideband cooling, these methods are used to reach a specific Zeeman sublevel.

The Hamiltonian describing the interaction between the laser and a single atom is [119, 120, 121]

$$\mathscr{H} = \hbar\Omega\sigma_+ e^{-\mathrm{i}[(\omega-\omega_{\mathrm{eg}})t-\varphi]} e^{\mathrm{i}\eta(ae^{-\mathrm{i}\nu t}+a^\dagger e^{\mathrm{i}\nu t})} + \mathrm{H.c.}, \tag{2.90}$$

where ω and ϕ denote the laser frequency and phase, ν is the motional mode frequency, ω_{eg} is the qudit transition frequency being addressed, and "H.c." denotes the Hermitian conjugates. Other parameters are Ω for the Rabi frequency associated with the transition, η is the Lamb-Dicke parameter, $a^\dagger$ is the phonon creation operator, and σ_+ denotes the atomic spin. In the Lamb-Dicke approximation or regime [121], evolution due to $\mathscr{H}$ is a unitary operation for a single qudit $U(\alpha) = e^{-i(\alpha/2)\vec{n}\cdot\vec{\sigma}}$.

For some appropriate values of the values $\vec{n}$ and α, the native single-qudit gates are

$$\mathrm{R}^{i,j} = \exp\bigl(-\theta\sigma_\phi^{i,j}/2\bigr);$$

i and j label different states and $\sigma_\phi^{i,j} = (\cos(\phi)\sigma_x^{i,j} \pm \sin(\phi)\sigma_y^{i,j})$, σ_x and σ_y are two qubit Pauli matrices embedded into a $d \times d$ matrix, where the acting on the rest of the level is the identity operation.

Additionally, the platform is capable of implementing the following entangling gate:

$$\mathrm{MS}^{i,j} = \mathrm{MS}^{i,j}(\theta,\phi) = \exp\left(-\mathrm{i}\frac{\theta}{4}\left(\sigma_\phi^{i,j}\otimes\mathbb{I} + \mathbb{I}\otimes\sigma_\phi^{i,j}\right)^2\right).$$

This gate is realised by coupling the levels of each individual atom with the normal modes of the chain. The realisation uses spin-dependent optical dipole forces [114]. The laser parameters are chosen to ensure the motional state is independent of the spin states (internal states of the atom) at the end of the interaction [121, 122, 123]. With access to any unitary

operation and the entangling gate $\mathrm{MS}^{i,j}$, the platform has access to a universal gate set.

2.13 Summary

In this first background chapter, I introduced the main tools and concepts in randomised benchmarking schemes. However, my presentation is clearly biased towards a representation theory approach. In a reductionist fashion, randomised benchmarking is a method to estimate the trace of the Pauli-Liouville representation of the noise affecting quantum gates. In this chapter, I showed the tools that are used to achieve this purpose: every randomised benchmarking scheme at least uses one of the ingredients I mentioned here. The star of this chapter, this thesis, and randomised benchmarking schemes is the procedure of twirling; the goal of twirling is to reduce the number of parameters required to estimate the trace. In the next chapter, which is also a background chapter, I discuss the many ways in which the tools of the present chapter are used to characterise gates for different gate sets.

Chapter 3

State-of-the-art methods

3.1 Introduction

In this chapter, I describe the state-of-the-art benchmarking schemes. These schemes characterise qudit gates or universal gates, or both. First, I discuss the qudit randomised benchmarking scheme that characterises a Clifford and qudit gate set. Then I discuss four schemes to characterise a $\mathcal{C}_3 \setminus \mathcal{C}_2$ gate.

For brevity's sake, I use the phrase "a randomised benchmarking characterises a gate (or a gate set)" to refer to the following process: an experimental group conducts the randomised benchmarking experiment using the circuit, states, measurements, and sampling specified by the scheme. Next, a data analysis, following the scheme specification, is done; the outcome is the average gate fidelity of either a gate or a gate set.

Before starting my survey, I introduce terminology to standardise the language present in different sources. The terminology refers to either individual or collective characterisation; the distinction is based on the noise associated with the gates used in the randomised benchmarking scheme for characterisation.

The aim of a collective characterisation is to determine the average gate fidelity of every gate in a gate set; this kind of characterisation is based on the assumption that each gate

has the same noise. By contrast, in an individual characterisation, all but one gate have the same noise. The gate with distinct noise is called the target gate. The rest of the gates are called the auxiliary gate set. The aim of an individual scheme is to estimate the average gate fidelity of the target gate assuming the auxiliary gate set has been characterised, i.e., the average gate fidelity of the auxiliary gates is known.

3.2 Characterisation of Clifford gates

3.2.1 Qubit randomised benchmarking

In this subsection, I discuss the first scheme to estimate the average gate fidelity of a gate set. In particular, I discuss the assumptions of the scheme, the circuit design, and the quantities estimated. The notation I follow differs from the original [50].

The assumptions of the scheme include the ability to sample an arbitrary multi-qubit gate U from the Haar measure of the unitary group. Additionally, it assumes the inversion gate is prepared on demand. The noise $\mathcal{E}$ is considered to act between the two gates. Let U be the drawn gate. The circuit is mathematically described as:

$$\hat{U} \circ \mathcal{E} \circ \hat{U}^{\dagger}; \tag{3.1}$$

that is, the noise of both gates is combined into a single channel $\mathcal{E}$.

I now explain the procedure and the figure of merit obtained. The fidelity between the initial state $|0\rangle\langle 0|$ and $\hat{U} \circ \mathcal{E} \circ \hat{U}^{\dagger}(|0\rangle\langle 0|)$ is given by:

$$\mathrm{tr}\Big[\hat{U} \circ \mathcal{E} \circ \hat{U}^{\dagger}(|0\rangle\langle 0|)\Big]. \tag{3.2}$$

Since the result of twirling with respect to the Clifford gate set is a depolarising channel,

the average fidelity over U becomes:

$$\mathop{\mathbb{E}}_{U} \operatorname{tr}[|0\rangle\langle 0|\, S(|0\rangle\langle 0|)] = p + \frac{1-p}{2^N}, \tag{3.3}$$

where N is the number of qubits and

$$p \coloneqq \frac{\operatorname{tr}[\Gamma(\mathcal{E})] - 1}{2^N - 1}. \tag{3.4}$$

Note that the average is computed with respect to the Haar measure of the unitary group.

Similar to current schemes, randomly sampled unitary operations are composed, and the fidelity between the initial state and the final state is obtained. In this case, the gate set is the whole unitary group. In Emerson et al. work, two distinct forms of the sequence of operations are introduced. First, I define them and then discuss how they are used to estimate the fidelity and the assumptions imposed.

Given a set of m unitary matrices $\{U_i \colon i \in \{0, \ldots, m-1\}\}$, the first kind of sequence is

$$S^{(0)}(\{\mathcal{E}_i, U_i\}) \coloneqq \hat{U}_{m-1}^{\dagger} \circ \mathcal{E}_{m-1} \circ \hat{U}_{m-1} \circ \cdots \circ \hat{U}_0^{\dagger} \circ \mathcal{E}_0 \circ \hat{U}_0. \tag{3.5}$$

To avoid confusion with the sequence fidelity in randomised benchmarking schemes, I call the following quantity the $S^{(0)}$-fidelity:

$$P^{(0)}(\{\mathcal{E}_i, U_i\}) = \operatorname{tr}\big[\varrho S^{(0)}(\{U_i\})(\varrho)\big]. \tag{3.6}$$

Assuming each $\mathcal{E}_i$ is the same (for all i, $\mathcal{E}_i = \mathcal{E}$), the resulting average $S^{(0)}$-fidelity is

$$\mathop{\mathbb{E}}_{U_i} P^{(0)}(\{\mathcal{E}, U_i\}) = p^m \operatorname{tr}\big[\varrho^2\big] + \frac{1-p^m}{2^N}. \tag{3.7}$$

The procedure to obtain Eq. (3.7) is the same as in the Clifford case; see Sec. 2.9. Therefore, by using circuits of the form of Eq. (3.5), a characterisation of the whole unitary group can

be done.

The next sequence requires applying all the gates in sequence and then the inverse of each element in reverse order. The sequence has the form

$$S^{(1)'}(\mathcal{E}, \{U_i\}) \coloneqq \mathcal{E} \circ \hat{U}_0^\dagger \circ \mathcal{E} \circ \hat{U}_1^\dagger \circ \cdots \circ \mathcal{E} \circ \hat{U}_{m-1}^\dagger \circ \mathcal{E} \circ \hat{U}_{m-1} \circ \cdots \circ \mathcal{E} \circ \hat{U}_1. \tag{3.8}$$

The authors of this study claim [50], based on numerical evidence, that instead of using $S^{(1)'}$, the following sequence can be used instead with little deviation from the resulting fidelity:

$$S^{(1)}(\mathcal{E}, \{U_i\}) \coloneqq \hat{U}_0^\dagger \circ \hat{U}_1^\dagger \circ \cdots \circ \hat{U}_{m-1}^\dagger \circ \mathcal{E}_{m-1} \circ \hat{U}_{m-1} \circ \cdots \circ \mathcal{E}_0 \circ \hat{U}_0. \tag{3.9}$$

Defining

$$P^{(1)}(\{\mathcal{E}_i, U_i\}) = \mathrm{tr}\big[\varrho_0 S^{(1)}(\{U_i\})(\varrho_0)\big], \tag{3.10}$$

and assuming each $\mathcal{E}_i$ has the same parameter p (but not necessarily the same noise), the averaged $S^{(1)}$-fidelity is

$$\mathop{\mathbb{E}}_{U_i} P^{(1)}(\{\mathcal{E}_i, U_i\}) = p^m + \frac{1 - p^m}{2^N}. \tag{3.11}$$

Thus, a list of pairs $\{(m, P^{(1)}(\{\mathcal{E}_i, U_i\}))\}$ can be experimentally estimated. Fitting an exponential then it is possible to estimate the parameter p, which is a proxy quantity to estimate the average gate fidelity over the unitary group.

Therefore, by preparing $|0\rangle\langle 0|$, randomly sampling unitary matrices U_i, and measuring with respect to $|0\rangle\langle 0|$, one can estimate the average gate fidelity over the unitary group. The error in this estimation is of order $O(1/\sqrt{2^N})$ [50]. Compared to contemporaneous methods, this scheme promotes average gate fidelity as a figure of merit for quantum gates and explicitly relies on twirling. For that reason, this scheme is the first randomised benchmarking scheme.

3.2.2 Qudit randomised benchmarking for prime-level systems

In this section, I discuss Clifford randomised benchmarking for qudits [54]; this family of schemes is the latest method for characterising qudit gates. Only in this section, d is a prime number, whereas in the original work d is taken to be any positive integer; certain issues discussed herein require restricting d.

A representation of the Clifford group is spanned by two d-dimensional matrices [124, 125]:

$$\mathcal{C}_d \coloneqq \langle F_d, S_d \rangle, \tag{3.12}$$

where F_d and S_d have entries

$$(F_d)_{i,j} \coloneqq (\omega_d)^{ij}/\sqrt{d}, \tag{3.13}$$

$$(S_d)_{i,j} \coloneqq \delta_{ij}(\omega_d)^{i(i-1)/2}; \tag{3.14}$$

note that F_2 is the Hadamard gate for qubits. The matrices F_d and S_d normalise the qudit Pauli group constructed from the clock and shift matrices; thus, the group generated by F_d and S_d normalises the Pauli group.

In the qudit Clifford randomised benchmarking scheme, the gate set $\mathcal{C}_2$ of Eq. (3.12) is collectively characterised. The circuit design is the same as that of qubit randomised benchmarking [106]. Thus, I do not discuss it. The twirl of an arbitrary channel with respect to $\mathcal{C}_d$ is a totally depolarising channel [53]. Since the totally depolarising channel depends on a single parameter, the expression for the sequence fidelity is a single exponential [106, 54]. Moreover, under the gate-independent noise assumption, the scheme is robust against SPAM errors; the fidelity of the initial state and measurement is irrelevant to the characterisation.

For a gate set that includes non-Clifford elements, the twirl is no longer a totally depolarising channel [2, 56]. This occurs because the unitary representation of the gates is not a unitary 2-design [2, 56]. Therefore, any scheme aiming to characterise a non-Clifford gate set has to deal with more than one parameter to estimate the average gate fidelity.

3.3 Characterisation of non-Clifford gates

3.3.1 Single-qubit dihedral benchmarking

In this section, I review the dihedral benchmarking scheme, which is a method to characterise a qubit non-Clifford gate. This section serves to illustrate the characteristics that I aim to retain from the qubit case in my qudit generalisation.

The dihedral benchmarking scheme aims to collectively characterise a gate set including a specific kind of qubit $\mathcal{C}_3 \setminus \mathcal{C}_2$ gate, as defined in Sec. 2.4. In dihedral benchmarking, the matrix

$$T' = \begin{bmatrix} \omega_{16} & 0 \\ 0 & \omega_{16}^{15} \end{bmatrix} \tag{3.15}$$

is used. This representation T' is not the standard T presented in textbooks [4]. The standard[1] qubit T is

$$T = \begin{bmatrix} 1 & \\ & \omega_8 \end{bmatrix}. \tag{3.16}$$

In this section, I continue using T'.

The following representation γ, mapping D_{16} elements to elements in $\langle X, T\rangle$, is irreducible [126, 2, 66]:

$$\gamma(x,t) = X^x (T')^t, \tag{3.17}$$

with

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{3.18}$$

To compute the decomposition of the Pauli-Liouville representation of γ, I use the character table for this small group; see [127]. From the character table and the orthogonality between characters of irreps [82], I identify three distinct irreps: $\mathfrak{h}_{\mathbb{I}}$, $\mathfrak{h}_0$, and $\mathfrak{h}_+$. From this

[1] This T gate satisfies the most general definition: it is not Clifford and conjugating a Pauli matrix returns a Clifford matrix.

decomposition, I derive expressions for the sequence fidelity and average gate fidelity.

The first step is to discuss the form of the twirl with respect to D_{16}. From the irreps mentioned in the previous paragraph, for any channel $\mathcal{E}$, the twirl with respect to γ in Eq. (3.17) is

$$[\mathcal{E}]_\gamma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \eta_0(\mathcal{E}) & 0 & 0 \\ 0 & 0 & \eta_+(\mathcal{E}) & 0 \\ 0 & 0 & 0 & \eta_+(\mathcal{E}) \end{bmatrix} = \Pi_{\mathbb{I}} + \Pi_1 \eta_0(\mathcal{E}) + \Pi_+ \eta_+(\mathcal{E}), \tag{3.19}$$

where

$$\Pi_{\mathbb{I}} \coloneqq \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \Pi_0 \coloneqq \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \Pi_+ \coloneqq \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{3.20}$$

$$\eta_0 = \mathrm{tr}[\Gamma(\mathcal{E})\mathbb{I}_0], \tag{3.21}$$

and

$$\eta_+ = \mathrm{tr}[\Gamma(\mathcal{E})\mathbb{I}_+]/2. \tag{3.22}$$

From the expression of the twirl of Eq. (3.19), I compute the sequence fidelity and the average gate fidelity.

To compactly write the sequence fidelity and average gate fidelity, I define the state

$|+\rangle \coloneqq H|0\rangle$, where H is the Hadamard matrix.

$$P(m;\mathcal{E},|\varpi\rangle,\langle\varpi|) = \langle\langle\varpi|\Gamma(\mathcal{E})[\mathcal{E}]_\gamma^m|\varpi\rangle\rangle, \quad \varpi\in\{0,+\}, \tag{3.23}$$

by the form of the twirl in Eq. (3.19) (3.24)

$$= A_\varpi + B_\varpi \eta_\varpi(\mathcal{E})^m, \tag{3.25}$$

where A_ϖ and B_ϖ are constants irrelevant to the characterisation. Also, using the form of $[\mathcal{E}]_\gamma$, the average gate fidelity is

$$F(\mathcal{E}) = \frac{2\,\mathrm{tr}[[\mathcal{E}]_\gamma]+2}{12} = \frac{1}{6}\left(1+\eta_0(\mathcal{E})+2\eta_+(\mathcal{E})\right)+\frac{1}{3}. \tag{3.26}$$

The bi-parametric form of the sequence fidelity is one of the features I seek to generalise.

I have now stated the characteristics of dihedral benchmarking that I aim to extend; I discussed them to avoid repeating them for each scheme. The three characteristics my scheme aims to retain are: the bi-parametric form of the sequence and average gate fidelity as in Eq. (3.26); the parameters are accessible via two pure states that are easy to write, such as $|0\rangle$ and $|+\rangle$; each parameter can be estimated by fitting a single exponential as in Eq. (3.25).

A potential problem arises if, because of noise in the preparation of the initial state, the initial state ϱ is not orthogonal to either $|0\rangle$ or $|+\rangle$. Such an issue can give rise to a multi-exponential form for the sequence fidelity, leading to inaccurate characterisation. In the following subsection I describe a method devised to mitigate this practical issue.

Removing the constant due to SPAM

The constant that appear in Eq. (3.23) makes estimating the parameter p more difficult. Therefore, removing the constant would increase the quality of the fit. Now, I show the method developed by Harper et al [128] to remove this constant. This technique is only valid for qubit randomised benchmarking schemes.

I recall the form of the sequence fidelity and of the constant.

$$P(m;\mathcal{E},|0\rangle,\langle 0|) = A^0\eta_0(\mathcal{E})^m + B^0, \tag{3.27}$$

The form of the constant B^0 is

$$B^0 \coloneqq \langle\langle 0|\Gamma(\mathcal{E})(\Pi_{\mathbb{I}} + \Pi_0)|0\rangle\rangle. \tag{3.28}$$

Repeating the randomised benchmarking experiment but with a measurement with respect to $|1\rangle\langle 1| = \hat{X}(|0\rangle\langle 0|)$ I obtain

$$P(m;\mathcal{E},|0\rangle,\langle 1|) = A^1\eta_0(\mathcal{E})^m + B^1, \tag{3.29}$$

with

$$B^0 \coloneqq \langle\langle 1|\Gamma(\mathcal{E})(\Pi_{\mathbb{I}} + \Pi_0)|0\rangle\rangle. \tag{3.30}$$

Note that

$$B^0 + B^1 = 1. \tag{3.31}$$

Now, computing the sum

$$P(m;\mathcal{E},|0\rangle,\langle 0|) - (1 - P(m;\mathcal{E},|1\rangle,\langle 1|)) = A^0\eta_0(\mathcal{E})^m + B^0 - (1 - A^1\eta_0(\mathcal{E})^m - B^1), \tag{3.32}$$

$$= (A^0 + A^1)\eta_0(\mathcal{E})^m + B^0 + B^1 - 1, \tag{3.33}$$

$$= (A^0 + A^1)\eta_0(\mathcal{E})^m. \tag{3.34}$$

In summary, by summing the sequence fidelity of two similar experiments, one the standard randomised benchmarking and the other appending an X gate, the constant appearing in the sequence fidelity can be removed. This technique simplifies the data analysis part of randomised benchmarking qubit schemes.

3.3.2 Character randomised benchmarking

By using the characters of an irrep in the decomposition of the Pauli-Liouville representation, it is possible to recover a single-exponential sequence fidelity, which is particularly useful when the initial state is difficult to prepare. In this subsection, I discuss character randomised benchmarking, a method that recovers the SPAM error independence, as is the case of Clifford randomised benchmarking, to any gate set.

The aim of character randomised benchmarking is to collectively characterise an arbitrary gate set [129]. Characterising an arbitrary gate set can present an important challenge, as it may require estimating d^2 parameters and each of these parameters could appear in the sequence fidelity curve; isolating these parameters is key [56]. Character randomised benchmarking addresses this difficulty by providing a method to obtain a single exponential for each parameter.

The scheme differs in important ways from standard randomised benchmarking schemes. In addition to gates and SPAM, character randomised benchmarking requires knowledge of the character table and irreps of the group associated with the gate set. It is also one of the first methods that requires mixing analytical data with experimental data, as I describe later in this section. The number of samples required also increases.

The circuit is similar to that of the original randomised benchmarking. The main change in this scheme is to add an extra sampling step on top of the random circuit construction. Here is the list of steps:

1. First draw $g \in G$, in standard randomised benchmarking, G is the Clifford group.

2. Then draw the group elements $g_0, \ldots, g_m$ from G.

3. Modify the last gate in the sequence from g_m to gg_m.

4. Apply the gates $\tilde{g}_0, \ldots, \widetilde{gg}_m$.

5. Repeat the draw with g fixed; that is, for l times use the same g but uniformly random

draw $g_0, \ldots, g_m$.

6. After those l are repeated, then we sample another group element g.

The circuit is illustrated in Fig. 3.1, (note that g is not used to compute the inversion gate).

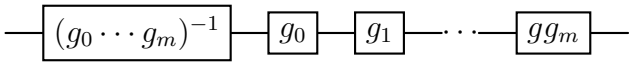


Figure 3.1: Family of circuits used in character randomised benchmarking [129].

From the family of circuits explained in the previous paragraph and illustrated in Fig. 3.1 the sequence fidelity is estimated. The estimate is done by preparing the state $|0\rangle$, applying the drawn gates, and measure with respect to $|0\rangle$; repeating this procedure produces an estimate of

$$P(m; |0\rangle, \langle 0|) = \mathop{\mathbb{E}}_{g} \left(\mathop{\mathbb{E}}_{\mathbf{g}=g_0,\ldots,g_{m-1}} P(\mathbf{g}; g) \right), \tag{3.35}$$

where

$$P(\mathbf{g}; g) \coloneqq \mathrm{tr}\{|0\rangle\langle 0|\, \tilde{g}_0 \circ \cdots \circ \widetilde{gg}_m(|0\rangle\langle 0|)\} \tag{3.36}$$

In a (classical) computer, the approximation of $P(\mathbf{g}; g)$ is multiplied by the character of the irrep and the dimensionality of the irrep. This is the reason behind the name of the scheme [129]. The procedure is repeated for as many circuit depths as necessary; the depths required are not discussed, but they are expected to be the same as required for the standard dihedral benchmarking scheme.

The authors of the original article justify the feasibility by using an argument based on the number of shots required; no formal argument is presented [129]. The experimental implementation of the scheme results in a sequence fidelity with the form of a single exponential. The remaining data analysis is the same as that of the Clifford randomised benchmarking scheme.

3.3.3 Multi-qubit dihedral Benchmarking

This subsection addresses the multi-qubit generalisation of dihedral benchmarking, a technique aimed at characterising controlled-T gates across multiple qubits. This approach is essential for improving the scalability of benchmarking protocols in quantum computing. Notably, the method requires estimating only two parameters regardless of the number of qubits. This efficiency highlights the potential for practical implementation in larger-scale quantum systems.

In the multi-qubit dihedral benchmarking scheme, the standard randomised benchmarking experiment is employed but with a modified gate set. Accordingly, this section will focus on explaining the gate set and the expressions for sequence fidelity and average gate fidelity, which are the scheme's salient features. By understanding these two components, one gains insight into the efficiency and scalability of the multi-qubit scheme. This modification allows for more precise characterisation of multi-qubit systems, a critical step in advancing quantum computing.

The modified gate set corresponds to permutation (or cyclic) gates and T gates on each qubit; for definiteness, I discuss the bi-qubit case. The group is generated by CNOT, $X \times \mathbb{I}$, $\mathbb{I} \times X$, $T \times \mathbb{I}$, and $\mathbb{I} \times T$, where T is defined in Eq. (3.16). The resulting expressions for the average gate fidelity and sequence fidelity are [66]

$$F = \frac{2^N(1 + (2^N - 1)\eta_0 + (2^{2N} - 2^N)\eta_+) + 2^{2N}}{2^{2N}(2^N + 1)} \tag{3.37}$$

and

$$P = a_\varpi \eta(\mathcal{E})_\varpi^m + b_\varpi, \qquad \varpi \in \{0, +\}, \tag{3.38}$$

where $\mathcal{E}$ is the noise affecting the gates. The initial state and the measurement required to have access to the parameters are unspecified [129].

The methodology to prove the scalability result goes as follows: one first decomposes any element of the group into four subgroups. The twirl by the whole group corresponds to a

sequence of twirls. The first twirl is with respect to the Pauli group, which leads to a Pauli channel. The subsequent twirls are conducted by analysing the orbits on the set of Pauli matrices. However, this procedure does not extend to qudits, as it relies on the property of qubit Pauli matrices, where their square equals the identity—a property not shared by qudits.

Equations (3.37) and (3.38) show that the characterisation of universal qubit gates scales. The characterisation is independent of the number of qubits. Also, the standard randomised benchmarking scheme is sufficient to achieve the characterisation. In the next section, I discuss an extension of the schemes that I have presented. The extension considers the target gate to have a different noise than the rest of the gates.

3.3.4 Leakage characterisation

Here, I attempt to give a formal definition of the effect of leakage on a qudit system platform. Assume a system on a space $\mathfrak{h} = \mathfrak{h}_0 \oplus \mathfrak{h}_1$, with projectors Π_0 and Π_1. For a state ϱ_s at step s defined only on $\mathfrak{h}_0$; $\mathrm{tr}[\varrho_s \Pi_0] = 1$. Then a system suffers the effects of leakage if there are two steps in a circuit, s and s' such that $\mathrm{tr}[\varrho_s \Pi_0] \neq \mathrm{tr}[\varrho_{s'} \Pi_0]$ and $\mathrm{tr}[\varrho_s \Pi_1] \neq \mathrm{tr}[\varrho_{s'} \Pi_1]$. In other words, the previous definition indicates that a system suffers from leakage if there is an interchange of probability between two steps in a quantum circuit.

The scheme introduced by Wallman et al. [130] estimates the sum of averages (explained below) of two quantities: interchange from $\mathfrak{h}_0$ to $\mathfrak{h}_1$ and vice versa, averaged over all pure states. Prepared similarly to randomised benchmarking but without an inversion gate and using a totally mixed state $\mathbb{I}/d$, this protocol produces a single exponential sequence fidelity with the decay parameter being the estimated average. Consequently, it enables the quantification of leakage in a given system.

3.3.5 Cross-talk characterisation

In this subsection, I discuss the scheme known as simultaneous randomised benchmarking. This scheme, meant for multi-qubit platforms, aims to detect signatures of cross-talk across different registers on a qubit quantum platform. Although not directly related to the primary objective of this thesis, this scheme demonstrates the versatility of randomised benchmarking techniques in characterising quantum platforms. Cross-talk errors have many experimental origins, making it futile to define them uniformly across platforms [131]. From a mathematical perspective, it is preferable to study these experimental forms of noise in terms of the observed behaviour in the outcome of a quantum circuit experiment.

Two notions are introduced to study cross-talk noise: independence and locality [131]. To avoid additional complexity, the following notations are introduced for a single and bi-qudit systems, though their generalisation to controlled gates is straightforward. A platform is local if every operation that acts on a register k affects only k. A platform is independent if, at step t in the realisation of a quantum circuit, the operation acting on register k is unaffected by the application, at step t, of another gate on register $k' \neq k$.

Whereas the method from [131] efficiently signals cross-talk, it does not quantify its severity. For this purpose, simultaneous randomised benchmarking is used. In simultaneous randomised benchmarking, three randomised benchmarking experiments are used to estimate three parameters: r_k, $r_{k'}$, and $r_{k|k'}$, where k and k' label a qubit register in a given quantum platform. From the parameters r_k, $r_{k'}$, and $r_{k|k'}$, the additional errors induced on subsystem k from controlling k' are computed as

$$\delta r_{k|k'} = |r_k - r_{k|k'}|. \tag{3.39}$$

The parameters are estimated by applying the randomised benchmarking scheme using different gate sets:

- r_k is obtained by using the randomised benchmarking scheme with gate set $\mathcal{C} \otimes \mathbb{I}$, where

$\mathcal{C}$ is the single qubit Clifford gate set and $\mathcal{C} \otimes \mathbb{I}$ refers to the bi-qubit gates, with the first register under the action of a Clifford gate and the second the identity is applied.

- $r_{k'}$ is obtained by using the randomised benchmarking scheme with gate set $\mathbb{I} \otimes \mathcal{C}$.
- $r_{k|k'}$ is obtained by using the randomised benchmarking scheme with gate set $\mathcal{C} \otimes \mathcal{C}$.

Note that the estimate of the last parameter, $r_{k|k'}$, requires the preparation of a non-trivial initial state [132].

To conclude this section, I would like to mention that cross-talk errors can be mitigated by randomised compiling [108, 58, 113]. Therefore, with little overhead in the experimental side, randomised benchmarking can be used to characterise quantum gates under the effects of cross-talking errors.

3.4 Individual characterisation

3.4.1 Introduction

In this section, I discuss the schemes that aim for individual characterisation of a T gate. In a collective characterisation every gate has the same noise [106, 89]. The estimated average gate fidelity is then equal for each gate. In most cases this approximation is unrealistic. Especially for the case of gates acting on encoded qudits; non-Clifford gates are implemented in a different way than Clifford gates. Assuming equal noise is invalid [133].

In an individual characterisation every gate in the gate set, except one, has the same noise [60, 90]: the gate with a different noise is the one being characterised. The scheme's output is the average gate fidelity of the gate with distinct noise [60, 90, 134]. In practical randomised benchmarking schemes the auxiliary gate set is either a Pauli or Clifford gate set; the gate to be characterised is either a Clifford or T gate, respectively [106, 54].

3.4.2 Interleaved benchmarking

There are four different interleaved benchmarking schemes relevant for my discussion. They differ in their auxiliary gates, circuit design and data analysis; one scheme is also defined for multi-qubit (controlled) operations. All of them are different in terms of auxiliary gates, circuit design, and data-analysis; the only point in common is that they require mixing some target gate (to be characterised). The aim of three of the four schemes I describe is explicitly to characterise a T gate.

Notice that in my review, I include one scheme to characterise a Clifford gate; the reason is that it presents what can be considered the standard protocol for an interleaved benchmarking scheme. After that, I review the variety of schemes to estimate the average gate fidelity of a universal non-Clifford gate. For each scheme I only discuss the following points: target gate to characterise, experimental requirements, physical assumptions of the scheme, circuit design, protocol, quantity (or quantities) estimated, and any posterior data analysis that differs from the usual single-exponential case.

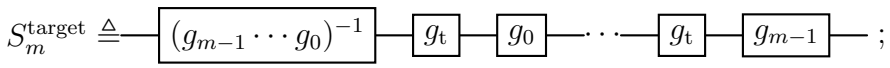
3.4.3 Clifford interleaved benchmarking

The aim of Clifford interleaved benchmarking is to characterise a Clifford gate using as auxiliary gates Clifford gates. Notice the scheme assigns two different noises to the target gate [90], which obviously is an unrealistic assumption. I explain the procedure to estimate the composite noise between the auxiliary and target gates.

I start introducing the necessary notation to describe the scheme; I stick to this notation throughout this thesis. The auxiliary gate set is G and its elements are $g \in G$. Every member of G has noise $\mathcal{E}$, and the target gate has noise $\mathcal{E}_{\mathrm{t}}$; I consider $\mathcal{E} \neq \mathcal{E}_{\mathrm{t}}$. The channel used to estimate the average gate fidelity of the target gate is

$$\widetilde{S}_m^{\mathrm{target}} := \mathcal{E}\circ\hat{g}_m^{-1}\circ\cdots\circ\hat{g}_1^{-1}\bigcirc_{i=1}^{m}\hat{g}_{\mathrm{t}}\circ\mathcal{E}_{\mathrm{t}}\circ\mathcal{E}\circ\hat{g}_i, \tag{3.40}$$

(◯ denotes composition) with circuit representation



where the symbol $\triangleq$ should be read as "corresponds to the circuit". Applying this circuit to $|0\rangle$ and measuring, again, with respect to $|0\rangle$ produces a sequence fidelity that is a single exponential. The decay parameter of the single exponential allows the estimate of the average gate fidelity for the composition of $\mathcal{E}_t \circ \mathcal{E}$. Since $F(\mathcal{E})$ is known, by the approximation of the composition of the fidelity

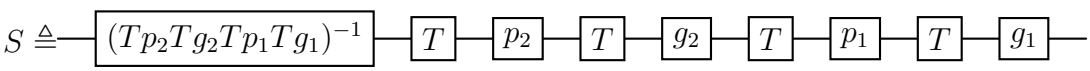
$$F(\mathcal{E}_{\text{t}})\, F(\mathcal{E}) \approx F(\mathcal{E}_{\text{t}} \circ \mathcal{E}), \tag{3.41}$$

the fidelity $F(\mathcal{E}_{\text{t}})$ is estimated. Thus, the average gate fidelity of the target gate is known and thus the target Clifford gate is characterised.

3.4.4 Non-Clifford gate characterisation using Clifford gates

The scheme addressed in this subsection aims to characterise any qubit $\mathcal{C}_3 \setminus \mathcal{C}_2$ gate using, as auxiliary gates, Clifford (including Pauli) gates. This scheme requires estimating two parameters. This scheme has a more general noise configuration (than standard randomised benchmarking); it assigns the same noise to the Clifford and Pauli gates but a different noise to the target gate.

I assume the auxiliary gate set has been characterised and only describes the circuit to estimate the average gate fidelity for the target gate. The noise for the target gate, any gate in $\mathcal{C}_3 \setminus \mathcal{C}_2$, is $\mathcal{E}_T$. Let g_i be a Clifford gate and p_i a Pauli gate. The circuit is



which corresponds to the channel (including noise)

$$\widetilde{S} = \mathcal{E} \circ [p_2, g_2, p_1, g_1]_{\text{inv}} \circ \hat{T} \circ \mathcal{E}_T \circ \mathcal{E} \circ \hat{p}_2 \circ \hat{T} \circ \mathcal{E}_T \circ \mathcal{E} \circ \hat{g}_2 \circ \hat{T} \circ \mathcal{E}_T \circ \mathcal{E} \circ \hat{p}_1 \circ \hat{T} \circ \mathcal{E}_T \circ \mathcal{E} \circ \hat{g}_1. \tag{3.42}$$

The twirl is

$$\mathop{\mathbb{E}}_{p_2,g_2,p_1,g_1} \widetilde{S} = [[\mathcal{E}_T\mathcal{E}]_{\mathcal{P}}\mathcal{E}_T\mathcal{E}]^2_{\mathcal{C}}, \tag{3.43}$$

using the notation of Eq. (2.52). Thus, the fidelity estimated by means of this circuit is $F\left(\mathcal{E}_T^2\mathcal{E}^2\right)$.

Therefore, using the approximation of the fidelity of a composition as the product of the fidelities ($F(\mathcal{E}_T^2\mathcal{E}^2) \approx F(\mathcal{E}_T)\,F(\mathcal{E}_T)\,F(\mathcal{E})\,F(\mathcal{E})$), the data analysis part of this scheme estimates $F(\mathcal{E}_T)$. Notice the product approximation is used twice: first to remove the contribution of the auxiliary gate set and then to get $F(\mathcal{E}_T)$ from $F(\mathcal{E}_T^2)$.

3.4.5 Interleaved dihedral benchmarking

In the original work on dihedral benchmarking [2], the corresponding interleaved benchmarking extension is introduced. Here I discuss, using the notation of dihedral benchmarking, the extension for the individual characterisation of a T gate.

The group $D_{16} \coloneqq \langle X, R(16)\rangle$ has as a subgroup a set of Clifford gates that is invariant under conjugation by the qubit $R(16)$, where

$$R(d) \coloneqq \exp(2\pi \mathrm{i} Z/d). \tag{3.44}$$

Note that $R(8)$ satisfies: $R(8)XR(8)^\dagger = -Y$ and $R(8)ZR(8)^\dagger = Z$. Therefore, $R(8)$ is a Clifford gate. Thus, I consider $T \coloneqq R(16)$ for this discussion on dihedral benchmarking.

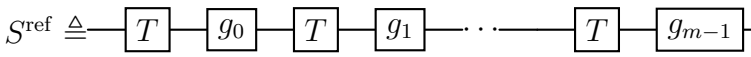
An invariant group, invariant under conjugation by T, is generated by X and $R(8)$; notice $D_8 \cong \langle X, R(8)\rangle$. Importantly, the twirl resulting from the averaging over the group D_8 is equal to twirling with respect to D_{16}.

Assuming the gate set D_8 are characterised, which could be done by assuming every Clifford gate has been characterised, the next step is to estimate the average gate fidelity of the composition of the noise of the T gate and the D_8 gates. I now explain the circuit used to estimate the auxiliary average gate fidelity, which includes the contribution of the

noise of the target gate. An even-length sequence is used. The basic sequence, without the inversion gate, is

$$S^{\mathrm{ref}} = \hat{T} \circ \hat{g}_0 \circ \cdots \circ \hat{T} \circ \hat{g}_{m-1}, \tag{3.45}$$

which corresponds to the circuit



The next step is to include noise and the inversion gate.

Note that D_8 is invariant under T conjugation, that is, for any $g \in D_8$, $TgT^\dagger \in D_8$. I exploit the invariance of D_8 under T conjugation to show that the sequence S^{ref} is equal to $S^{\mathrm{ref}} = \hat{T}^{2m} \bigcirc_i \hat{g}'_i$. Consider the sequence to be Tg_0Tg_1 ($g_0, g_1 \in D_8$), introducing the identity I obtain $T(TT^\dagger)g_0Tg_1 = T^2 g'_0 g_1$, where $g'_0 = T^\dagger g_0 T$ is a Clifford element, because $g_0 \in D_8$ then $T^\dagger g_0 T \in D_8$.

Because T^2 is Clifford, it has the same noise as the elements of D_8. This means that the inversion gate in the sequences is a member of D_8 if the circuit depth is even.

The circuit corresponding to the scheme is:

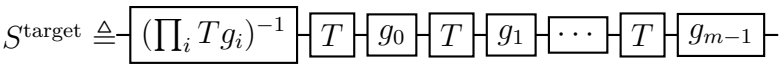


Figure 3.2: Family of circuits for interleaved dihedral benchmarking given in Eq. (3.46).

Including noise, the channel corresponding to the sequence is

$$\widetilde{S}^{\mathrm{ref}} = \mathcal{E} \circ [Tg_0 \cdots Tg_{m-1}]_{\mathrm{inv}} \bigcirc_{i\in[m]} \hat{T} \circ \mathcal{E}_{\mathrm{t}} \circ \mathcal{E} \circ \hat{g}_i = \mathcal{E} \circ \hat{T}^{2m} \bigcirc_{j\in[m]} (\hat{g}_j)^\dagger \bigcirc_{i\in[m]} \hat{T} \circ \mathcal{E}_{\mathrm{t}} \circ \mathcal{E} \circ \hat{g}_i, \tag{3.46}$$

where

$$[Tg_0 \cdots Tg_{m-1}]_{\mathrm{inv}} \coloneqq (Tg_0 \cdots Tg_{m-1})^{-1}. \tag{3.47}$$

By using the fact that the twirl commutes with any group element and the invariance of D_8,

we obtain the twirl

$$\mathop{\mathbb{E}}_{g_0,\ldots,g_m} \widetilde{S}^{\mathrm{ref}} = \Gamma(\mathcal{E})[\Gamma(\mathcal{E}_{\mathrm{t}})\Gamma(\mathcal{E})]^{2m}_{D_8}; \tag{3.48}$$

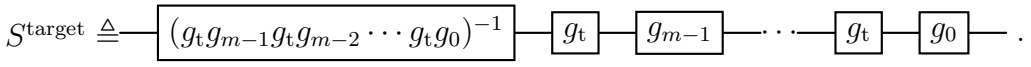
the notation for the twirl is introduced in Eq. (2.46). Therefore, using the randomised benchmarking scheme applied to circuit, the average gate fidelity of $\mathcal{E}_{\mathrm{t}} \circ \mathcal{E}$ is estimated. Dividing by the reference fidelity (of the auxiliary gate set) $\mathcal{E}$, a nice approximation of $\mathcal{E}_{\mathrm{t}}$ is obtained.

3.4.6 Interleaved benchmarking for qubit gates

This technique estimates the fidelity of any single- or controlled-qubit gate [135]. Part of this technique includes constructing a group of symmetries, which is a semidirect product between the following two groups: first, a subgroup of the Clifford group that normalises the target gate; second, a group of permutations that leave invariant the gate acting by conjugation. This scheme has the following requirements: a target gate to benchmark, the SPAM of the state $|0\rangle$, and the single and multiqubit Clifford gate set. This set of requirements is thus in line with standard randomised benchmarking schemes.

The novel assumption of the scheme is that the twirl for the noise of each gate—twirl with respect to the symmetric group—is close to being diagonal as the unitary gate. This assumption is strong; since the twirl commutes with the image of the representation of any group element, it means that every group element should be almost equally diagonal: a block diagonal part with a small non-block-diagonal contribution justified by the symmetries in the Hamiltonian linked to the unitary evolution [135].

The sequence of gates, including the target gate g_{t}, is



The invariance of the auxiliary gate set with respect to the target gate is key to simplifying the twirl of S^{target}; this invariance is common in interleaved benchmarking schemes for non-

Clifford gates.

The noise configuration is also novel: the channel $\mathcal{E}$ is appended to the ideal operation $g_t g$. However, for the inversion gate g_{inv}, the noise used is $\mathcal{E}' \neq \mathcal{E}$. This choice of noise for the inversion gate is irrelevant to the characterisation, since it is absorbed into SPAM instead of the gates.

Therefore, the experimental implementation of the scheme estimates the average gate fidelity of $\mathcal{E}$. The form of the average gate fidelity obtained from the twirl is unspecified. Unlike the standard randomised benchmarking formulation—with a single parameter accessible via $|0\rangle$—this scheme has an unspecified number of parameters. Moreover, it is unclear how these values (the decay parameters of the sequence fidelity) are estimated with the standard randomised benchmarking tool kit.

I conclude this subsection with a summary of the method. By constructing a symmetry group, an individual gate is characterised. The noise corresponds to the composition of the target gate and the symmetry group. The estimate is done using different (non-standard) techniques compared to randomised benchmarking schemes. Note that the scheme does not take into account the many parameters that could appear. This has the impact of requiring distinct initial states and increasing the experimental resources required.

3.4.7 Cycle benchmarking

In this subsection, I discuss cycle benchmarking. The aim of this scheme is to individually characterise multiple simultaneous Clifford gates; these simultaneous gates are called cycles [67]. Using the language of individual characterisation, the auxiliary gate set is the set of Pauli cycles, and the target gate is a Clifford cycle.

A novel technique is required for this scheme. While the quantity obtained is not new, the procedure differs from the other randomised benchmarking schemes I have described so far. I now explain the problem and solution in general terms. Consider Q, a unitary matrix, and ϱ, a density matrix. I now explain how to estimate $\text{tr}[Q\varrho]$ without using tomographic

techniques. The expression is as follows:

$$\mathrm{tr}[Q\varrho] = \sum_z \mathrm{tr}[\lambda_z \left|z\right\rangle\!\left\langle z\right| \varrho] = \sum_z \lambda_z \,\mathrm{tr}[\left|z\right\rangle\!\left\langle z\right| \varrho], \tag{3.49}$$

where $Q = \sum_z \lambda_z \left|z\right\rangle\!\left\langle z\right|$. I follow the notation, which may seem redundant, from the original source [67]. Let B_Q be the linear mapping that diagonalises Q. That is,

$$B_Q(\left|z\right\rangle\!\left\langle z\right|) = U \left|z\right\rangle\!\left\langle z\right| U^\dagger = \left|c\right\rangle\!\left\langle c\right|, \tag{3.50}$$

where U is a unitary matrix, and $\left|c\right\rangle$ is one of the computational basis states. Thus, we obtain

$$\mathrm{tr}[Q\varrho] = \sum_z \lambda_z \,\mathrm{tr}[\left|z\right\rangle\!\left\langle z\right| \varrho] = \sum_c \lambda_z(c) \,\mathrm{tr}\big[\left|c\right\rangle\!\left\langle c\right| U^\dagger \varrho U\big]. \tag{3.51}$$

In an experiment, $\mathrm{tr}\big[\left|c\right\rangle\!\left\langle c\right| U^\dagger \varrho U\big]$ can be estimated[2]. The eigenvalues λ_z are known. Thus $\mathrm{tr}[Q\varrho]$ can be evaluated.

I discuss how the process fidelity is obtained. Let m be the order of the Clifford cycle G. Define

$$m_1 \coloneqq mk_1 \quad \text{and} \quad m_2 \coloneqq mk_2, \tag{3.52}$$

where m is such that

$$G^m = \mathbb{I}, \tag{3.53}$$

k_1 and k_2 are two positive integers; these integers are the only two circuit depths required by the scheme. Draw a gate P from the set of Pauli cycles (Pauli gates acting simultaneously on different qubits). Draw m_i gates from the auxiliary gate set (formed by Pauli cycles): $\mathbf{g}_{m_i} \coloneqq g_1, \ldots, g_{m_i}$. The next part involves the ideal gate $\hat{g}_i$ and the physical (noisy)

[2]Note that the sequence $U^\dagger \varrho U$ may appear erroneous. The reason for this is that, originally, the computation ends with $\mathrm{tr}\big[U \left|c\right\rangle\!\left\langle c\right| U^\dagger \varrho\big]$. However, to make it more appealing to experimental groups that only have access to restricted measurements, it is customary to use the cyclic property of the trace, resulting in $\mathrm{tr}\big[\left|c\right\rangle\!\left\langle c\right| U^\dagger \varrho U\big]$.

implementation $\tilde{g}_i$. The following circuit is constructed:

$$\hat{S}^{\text{cycle}} \coloneqq \hat{g}_1 \circ \hat{G} \circ \hat{g}_2 \circ \cdots \circ \hat{G} \circ \hat{g}_m; \tag{3.54}$$

for any U unitary matrix, $\hat{U}$ acts on a density matrix ϱ by conjugation: $\hat{U}(\varrho) \coloneqq U\varrho U^\dagger$. The noisy version of the previous matrix, is

$$\widetilde{S}^{\text{cycle}} \coloneqq \widetilde{g}_1 \circ \widetilde{G} \circ \cdots \circ \widetilde{G} \circ \widetilde{g}_m, \tag{3.55}$$

where ϱ is a $+1$ eigenstate of the first Pauli matrix drawn ($P\varrho = \varrho$); this condition also reveals an implicit assumption of the scheme: the gates considered are taken to be phase-less.

Using the procedure explained above, the experimental data is of the form

$$\operatorname{tr}\Big[\hat{S}^{\text{cycle}}(P)\widetilde{S}^{\text{cycle}}(\varrho)\Big]. \tag{3.56}$$

Then the process fidelity $\hat{F}$ [67] is estimated as

$$\hat{F} \approx \mathop{\mathbb{E}}_{P\in\mathcal{P}} \frac{\mathbb{E}_{\mathcal{P}^{m_2}} \operatorname{tr}\Big[\hat{S}^{\text{cycle}}(P)\widetilde{S}^{\text{cycle}}(\varrho)\Big]}{\mathbb{E}_{\mathcal{P}^{m_1}} \operatorname{tr}\Big[\hat{S}^{\text{cycle}}(P)\widetilde{S}^{\text{cycle}}(\varrho)\Big]}, \tag{3.57}$$

where the average over $P \in \mathcal{P}$ is the average over all Pauli gates. The details of the approximation are discussed both in the generalisation for universal gates and in the supplementary material of the original article [67]. The process fidelity is equivalent to the average gate fidelity. Whereas the average gate fidelity is related to the trace of the Pauli-Liouville representation of a channel $\mathcal{E}$ as

$$F(\mathcal{E}) = \frac{d\operatorname{tr}(\Gamma(\mathcal{E})) + d^2}{d^2(d+1)}, \tag{3.58}$$

the process fidelity is $\hat{F} \coloneqq d^{-2}\operatorname{tr}(\Gamma(\mathcal{E}))$; thus, determining one determines the other.

To summarise this subsection, cycle benchmarking is a method to estimate the quality of a (single) Clifford qubit cycle. The method use a different circuit and data analysis than

standard randomised benchmarking; notably, "unlike randomised benchmarking protocols, the above protocol does not have an inversion gate" [67]. There is an interesting additional assumption implicit in the lack of an inversion gate[3]. Consider B_Q in Eq. (3.50). The noise of this gate, which is associated with SPAM, is not considered in the analytical expression for the estimated process fidelity.

3.5 Platforms

In this section, I review two qudit platforms. This review serves to give a glimpse of the variety of situations in which randomised benchmarking is used. For the ion trap and superconductor qudits, I reproduce figures corresponding to the sequence fidelity; I use these figures to support the application of Markovian techniques.

3.5.1 Qudit ion trap

In Sec. 2.12, I discussed in more detail the implementation of a qudit on an ion trap. The ion trap platform has access to single-qudit and entangling qudit gates. Now I discuss the way randomised benchmarking is used to characterise this platform [14]. The scheme implemented is Clifford randomised benchmarking; thus Ringbauer et al. assume all the consideration necessary for Clifford randomised benchmarking. They apply this scheme and obtain fidelities of 6×10^{-4}, 2×10^{-3}, and 1.0×10^{-2} for qubits, qutrits and ququints.

There is a marked decrease in the quality of gates by increasing the number of levels. The reason is the way randomised benchmarking was implemented. For qubit gates fewer basic operations, laser pulses, are needed, whereas for ququints, exponentially more are required; the increment on number of two by two unitary matrices follows the triangular number sequence. The errors accumulate, leading to the two order of magnitude discrepancy of the

[3]I note that the requirement in Eq. (3.53) could be considered as the inversion gate for cycle benchmarking. However, I prefer to only use the term "inversion gate" for the last gate in the sequence of a randomised benchmarking experiment

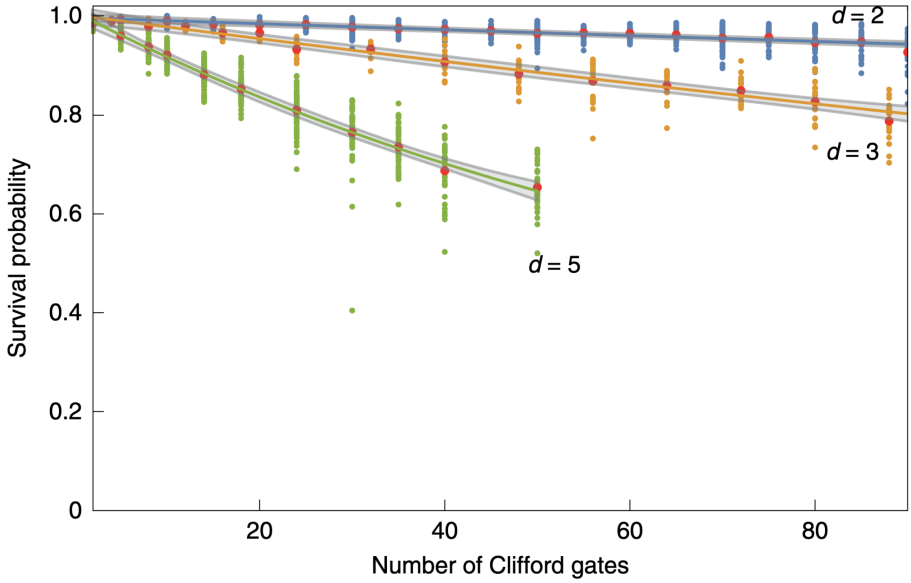


Figure 3.3: **Sequence fidelity for qudits with $d = 2, 3,$ and 5.** Courtesy of Ringbauer et al. [14]. This plot shows the sequence fidelity obtain by implementing qudit randomised benchmarking for $d = 2, 3,$ and 5. Number of Clifford gates refers to the circuit depth and survival probability to the sequence fidelity.

quality between qubits and ququints.

The maximum circuit depth, in this case the number of Clifford gates, is 90, 85, and 50 for qubit, qutrit, and ququint. For each circuit depth, 20 random sequences of gates are used. The number of shots per circuits is 100, assuming the same number of shots as for the entangling gates [14]. As argued in Appendix B.5, the best strategy (suggested by numerical evidence) requires using sequences with depths of up to 100, coupled with 20 shots and random circuits.

In Fig. 3.3, I reproduce the sequence fidelity reported by Ringbauer et al [14] and in Fig. 3.4 I show the log plot for the qutrit data. I discuss these plots as follows. There are two important things to notice in Fig. 3.3. First, the exponential decay justifies the restriction to Markovian techniques [112]. Second, the figure reveals the experimental capabilities of the platform. The circuit depth available and the repetitions support the application of randomised benchmarking in a sensible setting. In particular, asking for circuit depths up to 50 gates and repetitions up to 20 shots should be enough for a proper characterisation.

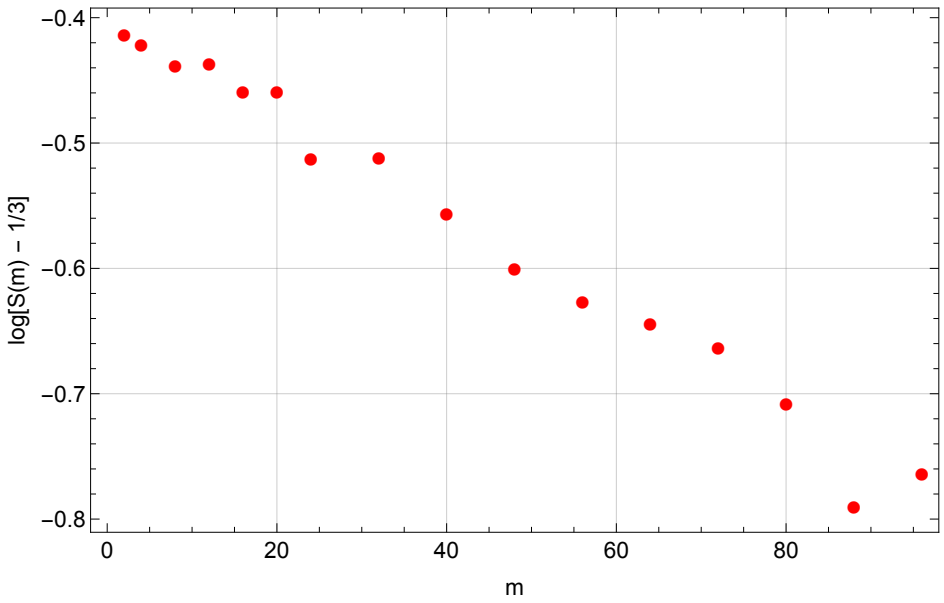


Figure 3.4: **Log plot for the sequence fidelity for qutrits.** Using the supplementary data from Ringbauer et al. [14], I produce a log plot for their sequence fidelity using their qutrit data.

As mentioned in Chapter 2 Sec. 2.11, Markovian noise leads to a single exponential decay. Whereas, in principle, nothing impedes non-Markovian noise to produce a single exponential decay, having a single exponential sequence fidelity allows to estimate the average gate fidelity, which is the only goal of randomised benchmarking and not to determine if the noise is non-Markovian or not. Whereas the study of non-Markovian noise using randomised benchmarking techniques is relevant, such goal is distinct to the goal of randomised benchmarking schemes.

Why is randomised benchmarking used to characterise the gates of this platform? The authors mention that to avoid SPAM interference in the characterisation they opted for randomised benchmarking. Nevertheless, they also implemented process tomography to verify the randomised benchmarking characterisation. This shows that for single-qudit gates, randomised benchmarking and gate set tomography are complementary methods.

3.5.2 Superconductor

The system is a transmon qutrit, which is modelled as an anharmonic oscillator. A transmon qutrit is an artificial system that has a multilevel structure with unevenly spaced energy levels. In the literature, it is sometimes referred to as an 'artificial atom' due to the similarity of its spectrum with that observed in atoms. Briefly, a harmonic oscillator can be implemented in a LC (or resonant) circuit [136]. The anharmonicity is obtained by replacing a linear inductor with a non-linear inductor. Experimentally, this is achieved using a superconducting inductor as a replacement for the linear one. This anharmonicity produces an uneven spectrum. The operations (the quantum gates) are implemented by applying an oscillatory voltage at different frequencies. Some frequencies excite the artificial 'atom'; while others serve to probe or measure its state.

In this platform, any single-qudit unitary gate can be implemented natively. This is done by explicitly constructing the non-diagonal exponentials of the Gell-Mann matrices. The diagonal gates are implemented virtually; that is, the phases are implemented in the next gate. Combining the complex exponentials of Gell-Mann matrices and virtual gates the full unitary group can be implemented. As in the ion trap case, the gates are decomposed into operations that act on only two levels at a time.

As in the ion trap case, Fig. 3.5 is reproduced with a twofold purpose: first, to illustrate the circuit depth that can be achieved with this platform; second, to show that the average sequence fidelity is approximately a single exponential. This supports the assumption that the CPTP and Markovian approximations for the noise are valid, which in turn allows us to apply randomised benchmarking rigorously.

Similar behaviour of the average sequence fidelity is reported in other qudit superconductor platforms [137]. For qutrits [138], Lupascu's group has a qutrit superconductor platform that shows a single exponential decay function [138], and a group in Chicago has a ququart showing the same Markovian behaviour [139].

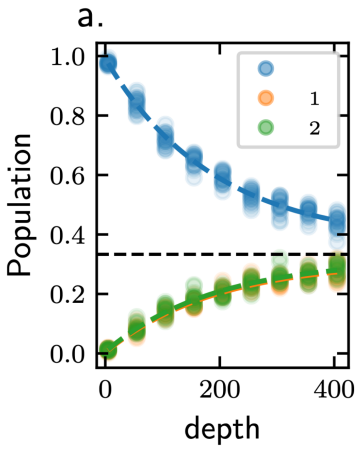


Figure 3.5: Courtesy of A. Morvan et al. [13]. Population decay corresponding to the fidelity between the initial state $Q(|0\rangle)$, with Q a circuit generated from random Clifford gates, and one of the states $|i\rangle$, with $i \in [3]$.

3.6 Randomised compiling

Randomised compiling is initially designed to mitigate coherent errors and increase fault-tolerant thresholds by a procedure that leaves the same theoretical circuit but randomly changes its experimental implementation [58]. The resulting average circuit transforms any Markovian CPTP noise into stochastic (Pauli) noise [58].

The method's ingredients are a physical gate set and an additional "virtual" gate set, implemented by keeping track of the phases until a non-diagonal gate is applied. The virtual gate set twirls gates before the gates are applied. The physical gate set is partitioned into easy and hard gates. Easy gates correspond to gates with low noise and hard gates with high noise.

The scheme goes as follows: a circuit is analytically split into easy and hard gates. Before and after each easy gate, a pair of twirling gates are inserted. In the ideal setting, one of the gates cancels the previous hard gate, then cancels the previous twirling gate, and then restores the previous hard gate. Thus, in the ideal case, the circuit remains the same. The resulting gates, from pre- and post-multiplying, are implemented on the platforms. Thus, these gates should be pre-computed on demand. The action of the virtual gates introduces

a small overhead in the experimental cost.

The result of this procedure is a circuit with only stochastic noise. Moreover, the scheme is beneficial even if the easy gates have gate-dependent noise. Twirling a circuit with gate-dependent noise equals a twirled circuit with gate-independent noise plus a minor correction [58]. This procedure is essential in gate characterisation since now randomised benchmarking can be implemented on top of randomised compiling with the basic assumptions and still yield the desired characterisation.

3.7 Summary

In this chapter, I presented a survey of the state-of-the-art for randomised benchmarking schemes. My survey concentrated on schemes to characterise non-Clifford gates. This survey serves the purpose of illustrating the techniques and results that I aim to generalise for qudit systems.

I discussed schemes for individual and collective characterisation of non-Clifford gates. I presented the dihedral benchmarking scheme as the latest method to obtain the average gate fidelity of a gate set, including a qubit T gate. I also discussed schemes to characterise Clifford gates, such as qudit Clifford randomised benchmarking and cycle benchmarking. In the next chapter, I present my generalisation to qutrit systems of dihedral benchmarking. The features of my scheme are imposed to be equal to those of dihedral benchmarking: only two parameters are necessary to characterise a qutrit T gate.

Chapter 4

Randomised benchmarking for universal qutrit gates

4.1 Introduction

In this chapter, I introduce the hyperdihedral group (HDG) for qutrits. From the associated gate set, I extend the original qubit randomised benchmarking scheme to characterise a qutrit gate set that includes a non-Clifford gate [2]. I start my discussion by enumerating the features a gate set ought to satisfy to be included in a feasible randomised benchmarking scheme. I then verify the HDG satisfies these features. Finally, I show several numerical results illustrating the feasibility of my scheme.

4.2 Construction of the HDG

In this section, I construct the HDG for the qutrit case; the HDG is the group used in my generalisation of the dihedral benchmarking scheme for qutrits. Then I specify features of the HDG, including its generators and its multiplication rule. Both are important in practical implementations.

I discuss the feasibility of my scheme. My feasibility claim is grounded in two points: (i)

the number of parameters required for the characterisation is two and (ii) the variance of the sequence fidelity (using the HDG) is similar to the variance for the Clifford gate set. These two points are discussed later.

Now I describe the gate set construction in detail, starting with the generators. The elements of the group have the form of a product of a cyclic permutation matrix and a diagonal matrix. Therefore, to find the generators of the group, I need to find the generators for the group generated by the matrices obtained by computing every permutation of the diagonal entries of T, which I denote as D.

Now, I argue that finding the set of generators of D is equivalent to finding a basis for the module (or vector space, but over a ring rather than a field) generated by the diagonals of each element of D. To see this, consider d_0 and d_1 in D. Then, since d_0 and d_1 are diagonals with entries of the form $\exp(2\pi i a/b)$ (where b is fixed), matrix multiplication is equivalent to addition of vectors modulo b. Therefore, whereas finding an element of D is equivalent to finding the powers of the generators of D needed to represent the element, the corresponding diagonal should be a linear combination of the basis of the vector space corresponding to the diagonals of the elements of D.

As in the case of finding a basis given a list of vectors, row-echelon reduction is necessary. A small tweak is needed since the matrices have entries in some ring rather than $\mathbb{C}$, but fortunately, this has been developed in Howell's work [75]. Applying such a generalized row-reduction procedure allows me to compute the generators of the HDG.

The generators of the HDG are

$$X = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \omega_9^8 & 0 \\ 0 & 0 & \omega_9 \end{bmatrix}, \text{ and } T' = \begin{bmatrix} \omega_9^2 & 0 & 0 \\ 0 & \omega_9^6 & 0 \\ 0 & 0 & \omega_9 \end{bmatrix}; \tag{4.1}$$

every HDG element is of the form

$$X^x T^{\alpha_0} (T')^{\alpha_1}, \tag{4.2}$$

where $x \in \mathbb{Z}_3$ and $\alpha_0, \alpha_1 \in \mathbb{Z}_9$; these values (x, α_0, α_1) are used to label HDG elements.

Now the multiplication rule between HDG elements. Let V be a diagonal matrix. Conjugating V by X results in a cyclic permutation of the diagonal entries of V. Let $V = \mathrm{diag}[a, b, c]$ $(a, b, c \in \mathbb{C})$; conjugating V by X results in the matrix $XVX^\dagger = \mathrm{diag}[c, a, b]$. Moreover, $\langle T, T' \rangle$ (the set of matrices generated by T and T') is invariant with respect to conjugation by X: for any $V \in \langle T, T' \rangle$, $XVX^\dagger \in \langle T, T' \rangle$.

The multiplication between two HDG elements can be obtained from the observation made in the previous paragraph. Multiply two elements $X^a T^{b_0} (T')^{b_1}$ and $X^{a'} T^{b'_0} (T')^{b'_1}$ of the HDG, introduce the identity in the form $\mathbb{I} = X^{a'} (X^{a'})^\dagger$:

$$X^a T^{b_0} (T')^{b_1} X^{a'} T^{b'_0} (T')^{b'_1} = X^a X^{a'} [(X^{a'})^\dagger T^{b_0} (T')^{b_1} X^{a'}] T^{b'_0} (T')^{b'_1}. \tag{4.3}$$

As observed in the previous paragraph, and noting that a cyclic permutation of the diagonal entries of T' produces a member of $\langle T, T' \rangle$, the matrix $(X^{a'})^\dagger T^{b_0} (T')^{b_1} X^{a'}$ is a member of $\langle T, T' \rangle$.

A trivial algebraic manipulation reveals that the multiplication rule for HDG is

$$X^{x_0} T^{a_0} (T')^{b_0} X^{x_1} T^{a_1} (T')^{b_1} = X^{x_0 + x_1} T^{(a_1 + a'_0)} (T')^{(b_1 + b'_0)}, \tag{4.4a}$$

where

$$\begin{bmatrix} a'_0 \\ b'_0 \end{bmatrix} = \begin{bmatrix} 5 & 8 \\ 4 & 3 \end{bmatrix}^{x_1} \begin{bmatrix} a_0 \\ b_0 \end{bmatrix}. \tag{4.4b}$$

The simplicity of the multiplication rule of Eq. (4.4) is important: using an arbitrary gate set, the computation of the product and the inverse of an arbitrary element is in general expensive. This has led to the development of techniques for cleverly sampling gates [140]. However, as the HDG multiplication rule is explicitly given and is efficient to compute, my scheme does not require such sampling methods, which decreases the difficulty in implementing my scheme.

I denote by HDG the group that has an irrep with representatives $\langle X, T\rangle$. The gate set corresponding to the unitary matrices $\langle X, T\rangle$ is the smallest, verified using `GAP`[1], that produces a bi-parametric fidelity, where

$$T = \mathrm{diag}[1, \omega_9^8, \omega_9]. \tag{4.5}$$

Notice there are only two primitive gates needed to generate the gate set: T' is used only to simplify the computation of composition and inverse.

4.3 Representation theory for the HDG

The objective of this section is threefold. First, it aims to define the irreducible representations (irreps) of the homogeneous differential group (HDG). Second, it seeks to illustrate the decomposition of the Hilbert space in terms of the irreps of the HDG.

Now I start the exposition. The minimal generating set of the HDG is given by the matrices in Eq. (4.1). Then, I define the representation γ as

$$\gamma(x, \alpha_0, \alpha_1) \coloneqq X^x T^{\alpha_0} (T')^{\alpha_1}. \tag{4.6}$$

I now study the decomposition of the Hilbert space according to the invariant subspaces corresponding to the action of $\Gamma = \gamma \otimes \bar{\gamma}$. The decomposition of the set of endomorphisms (from $\mathfrak{h}$ to $\mathfrak{h}$) is fundamental in the computation of the expressions for the gate fidelity and sequence fidelity. The Pauli-Liouville representation of γ (defined in Eq. (2.6)) has five nonequivalent irreps; these five irreps decompose the Hilbert space as

$$\mathrm{end}(\mathfrak{h}) = \mathfrak{h}_{\mathbb{I}} \oplus \mathfrak{h}_0 \oplus \mathfrak{h}_0^* \oplus \mathfrak{h}_+ \oplus \mathfrak{h}_+^*. \tag{4.7}$$

[1] `GAP`is a command line utility that allows to symbolically manipulate groups. For my purposes, it is a repository with access to finite groups including their character tables. From the character tables I verify the results in this chapter and in my paper [10].

As I show in §5.4, the number of irreps is equal to the number of parameters; conjugate pairs of irreps imply conjugate pairs of parameters.

In the next section, I explain how to compute the decomposition in Eq. (4.7). Later in §5.3, I formally compute the decomposition into irreps of the group labelling gates; I use `GAP` to compute the decomposition and the associated projectors. See Appendix B.3 for a list of functions and snippets I wrote for this purpose.

4.4 Gate fidelity and sequence fidelity

I now have all the ingredients (the decomposition of end($\mathfrak{h}$) and the multiplication rule for group elements) to compute the expressions for the gate fidelity and sequence fidelity. I start by computing the expression for the gate fidelity and then for the sequence fidelity. Then I conclude this section by illustrating how the parameters in both fidelities are linked; I show how to estimate them in a randomised benchmarking experiment.

The channel representation decomposes the Hilbert space into five irreps: $\mathfrak{h}_{\mathbb{I}}$, $\mathfrak{h}_0$, $\mathfrak{h}_0^*$, $\mathfrak{h}_+$, and $\mathfrak{h}_+^*$. Let $\varpi \in \{\mathbb{I}, 0, +, 0^*, +^*\}$ be an index corresponding to an irrep. Let Π_ϖ denote the projector onto the irrep ϖ. The twirling map (first defined in Eq. (2.46)) transforms a matrix M to

$$[M]_{\mathrm{HDG}} = \sum_{\varpi \in \{\mathbb{I},0,+,0^*,+^*\}} \eta_\varpi(M)\Pi_\varpi, \tag{4.8}$$

where

$$\Pi_{\mathbb{I}} = \mathrm{diag}(1,0,0,0,0,0,0,0,0), \tag{4.9a}$$

$$\Pi_0 = \mathrm{diag}(0,1,0,0,0,0,0,0,0), \tag{4.9b}$$

$$\Pi_{0^*} = \mathrm{diag}(0,0,1,0,0,0,0,0,0), \tag{4.9c}$$

$$\Pi_+ = \mathrm{diag}(0,0,0,1,1,1,0,0,0), \tag{4.9d}$$

$$\Pi_{+^*} = \mathrm{diag}(0,0,0,0,0,0,1,1,1), \tag{4.9e}$$

and $\eta_\varpi(M) \coloneqq \frac{\mathrm{tr}\left(\Pi_\varpi M^\dagger\right)}{\dim(\Pi_\varpi)}$ is the quantity to be experimentally estimated.

Applied to the Pauli-Liouville representation of a channel $\mathcal{E}$, the twirling map is equal to

$$[\Gamma(\mathcal{E})]_{\mathrm{HDG}} = \sum_{\varpi\in\{\mathbb{I},0,+,0^*,+^*\}} \eta_\varpi(\Gamma(\mathcal{E}))\Pi_\varpi. \tag{4.10}$$

Thus, the average gate fidelity for any channel twirled by the HDG is

$$F(\mathcal{E}) = \frac{3(1 + 2\,\mathrm{Re}(\eta_0(\Gamma(\mathcal{E}))) + 6\,\mathrm{Re}(\eta_+(\Gamma(\mathcal{E})))) + 9}{36}, \tag{4.11}$$

where I used Eq. (2.76).

I discuss the appearance of the real part in the expression for the gate fidelity. Two irreps appearing in the decomposition of $\mathfrak{h}$ in Eq. (4.7) are conjugated: $\mathfrak{h}_{0^*}$ is conjugated to $\mathfrak{h}_0$. A simple computation reveals the eigenvalue η of conjugated irreps are themselves conjugated. Therefore, by adding the two parameters, I get (twice) the real part of the eigenvalue η_0.

4.5 Scheme description at circuit and SPAM level

In this section, I describe my randomised benchmarking scheme. By description, I mean a description of an experimental 'run' of the scheme to obtain a number. Then I illustrate the iterative steps; I conclude with an estimate of the sequence fidelity.

I start the description of the experimental run by describing the ingredients and statistical requirements: I make use of gates, states, and measurements. The gates are labelled by HDG members, the states are $|0\rangle$ and $|+\rangle \coloneqq F_2\,|0\rangle$, and the measurements are with respect to those same states. The parameters of a run are the circuit depth m, the initial state $|\varpi\rangle$, and the measurement $\langle\varpi|$. The circuit depth is the number of gates sampled from the gate set, excluding the inversion gate.

The inversion gate is a gate that, in a noiseless setting, cancels the action of the original circuit, leaving only the identity operation. Assume two gates are drawn (more appropriately

draws and then applies) and I label them by $g_0, g_1 \in \mathrm{HDG}$. Then the inversion gate is labelled by the group element $(g_0 g_1)^{-1} \in \mathrm{HDG}$. In general, for an ordered multiset of gates labelled by group elements $g_0, \ldots, g_{m-1} \in \mathrm{HDG}$, the inversion gate is labelled by $(g_0 \cdots g_{m-1})^{-1}$. A diagram of a run of the experiment is presented in Fig. 4.1). This concludes the description of the experimental setup required.

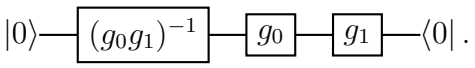


Figure 4.1: Circuit for a run using the HDG gate set. $\langle 0|$ denotes the measurement with respect to $|0\rangle$.

4.6 Numerical study of my scheme's feasibility

Assessing the feasibility of my scheme refers to assessing whether my scheme can be implemented with current experimental resources reported in the literature. In my case, since there are already randomised benchmarking experiments being done, it is enough to compare the resources required by my scheme against known schemes. In particular, I compare my scheme against Clifford randomised benchmarking [13].

My criteria for feasibility have three components [10]: fewer or an approximately equal number of primitive gates required than Clifford schemes, practical random sampling and composition of group elements, and the number of samples required. I explain the last two points. By practical random sampling, I refer to the fact that obtaining an HDG element requires sampling one of the following permutations: $\{e, (123), (132)\}$, and two powers for the diagonal gates. My claim (of practical composition) is justified by the fact that composing two elements of the HDG amounts to the following tasks: 1) compose two permutations, 2) cyclically rotate a vector, and 3) add two vectors. By contrast, an arbitrary finite group given only by its generators requires approximate methods (in terms of a group element), of which a computation complexity is unknown for a general group [141, 142].

The number of samples refers to the number of required runs at a given depth, with

each run using a different randomly drawn circuit. The computation of the number of runs required (for any gate set, including Clifford) is non-trivial and no standard method exists for performing this computation [143]. Here, I explain the method I use to estimate the number of samples required compared to the Clifford case. This estimation is based on the variance of the sequence fidelity, defined as follows.

I assume a uniform probability for drawing one random gate. Therefore, the probability of sampling the ordered multiset $(g_0, \ldots, g_{m-1}) \in G^m$ is $|G|^{-m}$. For a given gate set configuration $(\mathcal{E}, \rho, E)$, I then define the random variable X as the map

$$\mathrm{X}(\mathrm{tr}\{E\mathcal{E}g_0 \cdots \mathcal{E}g_{m-1}\rho\}) = |G|^{-m}; \tag{4.12}$$

This means that the probability of getting the value $\mathrm{tr}\{E\mathcal{E}g_0 \cdots \mathcal{E}g_{m-1}\rho\}$ is $|G|^{-m}$. I use X to define the mean and variance.

I now illustrate the computation of the mean and variance. The mean is

$$\mathbb{E}(\mathrm{X}) = |G^m|^{-1} \sum_{\boldsymbol{g} \in G^m} \mathrm{tr}\{E\mathcal{E}g_0 \cdots \mathcal{E}g_{m-1}\rho\}; \tag{4.13}$$

Note that $\mathbb{E}(\mathrm{X})$ is equal to the sequence fidelity. The variance is

$$\mathbb{V}(\mathrm{X}) = |G^m|^{-1} \sum_{\boldsymbol{g} \in G^m} \mathrm{tr}\{E\mathcal{E}g_0 \cdots \mathcal{E}g_{m-1}\rho\}^2 - \mathbb{E}(\mathrm{X})^2 = \mathbb{E}(\mathrm{X}^2) - \mathbb{E}(\mathrm{X})^2. \tag{4.14}$$

Figure 4.2 illustrates the qualitative behaviour of the variance for unital noise or noise with high fidelity $F > 0.999$. The figure is typical of the variance behaviour for any kind of noise. High-fidelity noise makes the channel approximately unital. As the variance curves are qualitatively similar, the number of samples for the HDG should be similar to the number of samples for Clifford.

Confidence intervals and the number of samples are closely related. The variance of the data directly affects the confidence interval [52, 143], which in turn determines the required

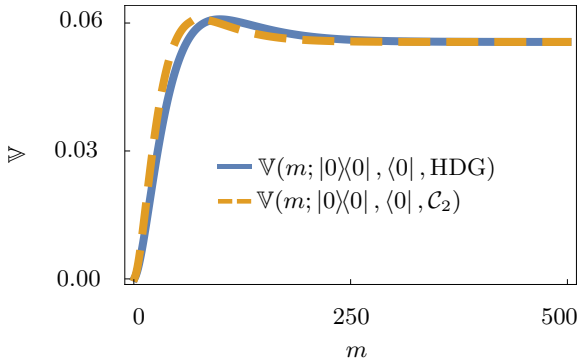


Figure 4.2: Plot comparing the variance of the HDG sequence fidelity with the variance of the Clifford sequence fidelity; I used the analytical expression (using Eq. (4.14)) to compute the plot. The fidelity of the noise considered is 0.99.

number of samples. Given a certain number of samples, the error and confidence level for the average result, such as the average sequence fidelity, can be established. This relationship helps to ensure that sufficient samples are taken to achieve the desired accuracy.

Here is how the number of samples required is related to the variance. In general, it is possible to set the confidence and error and, from the expression of the confidence interval, estimate the minimum number of samples required. Tight bounds are commonly obtained from expressions containing the variance of the sampled statistic. Therefore, similar variances imply similar numbers of samples.

4.7 Phase and criteria for universal qutrit randomised benchmarking

In this section, I discuss my results for the qutrit case, which became redundant in light of my qudit results. These include the analysis of whether the phase can be neglected and the criteria for selecting the HDG over other groups. I will also explain the reasons for the obsolescence of these results. By comparing the two cases, it became clear that the qudit approach offers more comprehensive solutions.

As discussed in Sec. 4.3, the HDG decomposes into five irreps, two of which appear as

conjugate irreps. This decomposition leads to an oscillatory contribution to the sequence fidelity. This contribution becomes negligible when dealing with a high-fidelity gate. Therefore, the oscillatory effects do not significantly impact the accuracy of the results.

To prove this claim, I first express the phases as a function of the chi-representation. The chi-representation helps relate the phase to the average gate fidelity. I then calculate the contributions from both the real and imaginary parts of the sequence fidelity. The result shows that the imaginary part is several orders of magnitude smaller than the real part.

With this form, I can find the minimum and, thus, the value at which the deviation from an exponential becomes noticeable. The non-exponential contribution is given by

$$\cos\varphi_0 = \frac{1}{\sqrt{1+(v/u)^2}}, \tag{4.15}$$

where

$$v \coloneqq \frac{2}{\sqrt{3}}\left(\chi_{33}+\chi_{44}+\chi_{55}-\chi_{66}-\chi_{77}-\chi_{88}\right), \tag{4.16}$$

and

$$u \coloneqq \chi_{00}+\chi_{11}+\chi_{22}-\frac{1}{2}\left(\chi_{33}+\chi_{44}+\chi_{55}+\chi_{66}+\chi_{77}+\chi_{88}\right), \tag{4.17}$$

where the chi-representation is used. The average gate fidelity is a function of χ_{00}. This leads to the definition of four variables that are useful for analyzing the range $\cos\varphi_0$ can take:

$$x_0 = \chi_{00}, \tag{4.18}$$

$$x_1 = \chi_{22}+\chi_{11}, \tag{4.19}$$

$$x_2 = \chi_{33}+\chi_{44}+\chi_{55}, \tag{4.20}$$

and

$$x_3 = \chi_{66}+\chi_{77}+\chi_{88}. \tag{4.21}$$

Using these variables, the new form of $\cos\varphi_0$ is:

$$\frac{v}{u} = \frac{\frac{2}{\sqrt{3}}(x_2 - x_3)}{x_0 + x_1 - \frac{1}{2}(x_2 + x_3)}. \tag{4.22}$$

Whereas the maximum value that $\cos\varphi_0$ can take is 1, the minimum is not clear. Therefore, to identify the range the values can take, it is necessary to compute the minimum.

Observing the form of $\cos\varphi_0$, the minimum is reached when $x_1 = x_2 = 0$ and $x_3 = 1 - x_0$. Thus, the minimum value it can take is

$$\min(\cos\varphi_0) = \frac{1}{\sqrt{\frac{4m^2(\chi_{00}-1)^2}{3\left(\frac{\chi_{00}-1}{2}+\chi_{00}\right)^2} + 1}}. \tag{4.23}$$

Since χ_{00} is a function of F, defined as

$$F = \frac{1}{4}(1 + 3\chi_{00}), \tag{4.24}$$

I substitute the value of χ_{00} as a function of F into $\min(\cos\varrho_0)$:

$$\min(\cos\varphi_0) = \frac{1}{\sqrt{\frac{4m^2\left(\frac{4F}{3}+\frac{1}{3}-1\right)^2}{3\left(\frac{\frac{4F}{3}+\frac{1}{3}-1}{2}+\frac{4F}{3}+\frac{1}{3}\right)^2} + 1}} = \frac{1}{\sqrt{\frac{64(F-1)^2m^2}{27(1-2F)^2} + 1}}. \tag{4.25}$$

To gain intuition, I explore two limits:

$$\lim_{m\to\infty} \min(\cos\varphi_0) = 0, \tag{4.26a}$$

$$\lim_{F\to 1} \min(\cos\varphi_0) = 1. \tag{4.26b}$$

The limits in Eqs. (4.26) indicate that as the gate fidelity increases, the deviation from a single exponential is negligible. However, the limit for m shows that if the fidelity of the gates is expected to be low, then shorter circuits should be used in the randomised benchmarking

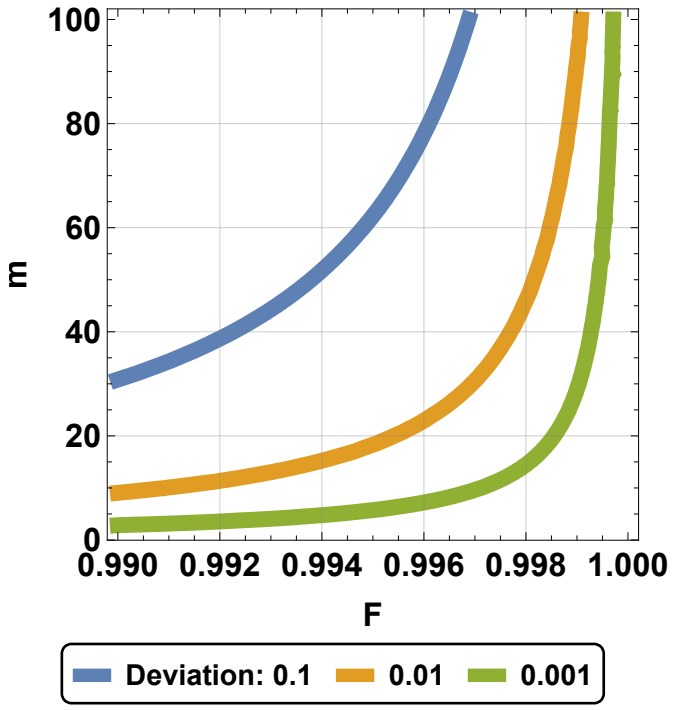


Figure 4.3: Contour plot showing the maximum circuit depth m that a gate with average gate fidelity F can be composed of before the phase contribution changes the behavior of the single exponential.

experiment.

In Fig. 4.3, I show the relation between fidelity, circuit depth, and the deviation result of the phase. It shows that if the target deviation is 0.01, then for a gate with a fidelity of 0.998, the maximum circuit depth to use is 40. Beyond that circuit depth, the deviation surpasses the target deviation, and the sequence fidelity curve deviates significantly from the single exponential. This suggests the following strategy to estimate the fidelity. For not-so-high fidelity gates, it is better to aim for short circuits but sample different gates.

I am able to justify HDG for qutrits based on three criteria. This criterion was used to uniquely identify the group. The three criteria are now listed:

- The T gate is a member of a unitary irreducible representation (irrep) of the group.
- The twirl of an arbitrary matrix should be diagonal in the Heisenberg-Weyl basis.
- The group is the smallest that satisfies the previous two criteria.

- The number of parameters is the smallest.

In summary, our criteria state that the smallest possible group that produces a diagonal twirl with the fewest parameters should be selected. Thus, if two groups have the same number of parameters, the smaller group should be chosen. These points are valid for the qutrit case since the order of the groups that satisfy the first two criteria is small. Therefore, using `GAP`, it made sense to search over small-order groups to identify the best fit.

These criteria are motivated by experimental constraints. The small order is implemented to avoid requiring the preparation of many gates. Additionally, the requirement for a minimal number of parameters is driven by the need to avoid preparing many distinct states. This would otherwise complicate isolating the parameters during experiments, increasing the overall complexity of the setup.

4.8 Conclusion

The results of this chapter are: the HDG for qutrits, the expressions for the fidelities (gate and sequence), and a numerical argument supporting the feasibility of my scheme. These results correspond to a generalisation for qutrits of the dihedral benchmarking scheme.

I now further discuss these results. The group HDG is constructed such that it reduces to the gate set found in dihedral benchmarking. I verified the mathematical properties relevant for my scheme: the number of irreps in the decomposition of the Pauli-Liouville representation and the multiplication rule.

My second result is the computation of the expressions for the sequence and gate fidelities for the HDG. Both expressions share two parameters: these two parameters, in turn, can be computed from experimental data obtained using the standard randomised benchmarking experimental scheme. By fitting an exponential function to the data, the parameters in the sequence fidelity are computed. From these two parameters, the average gate fidelity is computed.

The last discussion of this chapter is the verification of the feasibility of my scheme. I used a numerical approach to compare the resources required to obtain a statistical significance similar to Clifford randomised benchmarking. My numerical study showed that my scheme (which requires fewer resources than Clifford randomised benchmarking in terms of primitive gates) operates under the same statistical parameters (number of samples and number of repetitions) and has a qualitatively similar statistical reliability.

In the following chapter, I introduce my generalisation to qudits of the result presented here. There, I not only generalise the qutrit scheme but also obtain a simpler scheme with no oscillatory contribution to the sequence fidelity. I also present a multi-qudit generalisation.

Chapter 5

Randomised benchmarking for universal qudit gates

5.1 Introduction

This is the main chapter of my thesis. Here I introduce my benchmarking scheme for non-Clifford qudit gates. I start by defining the ${}_d$rHDG, which is constructed for each T gate and forms the gate set used in my scheme. Then comes a representation theory analysis. In this analysis I discuss the decomposition of the Pauli-Liouville representation of the ${}_d$rHDG. I conclude the chapter by obtaining the expressions required in any benchmarking scheme: the average gate fidelity and sequence fidelity for the ${}_d$rHDG.

As I discuss in this chapter, my scheme is useful to characterise any diagonal gate with order higher than 2. However, for non-power-of-prime (POP) dimensions, only the trivial generating set of a universal gate set is known. Therefore, whereas my scheme is useful to characterise diagonal gates in any dimension, I restrict my discussion to POP systems for the important task of characterising the generators of a universal gate set.

5.2 Construction of the qudit rHDG

In this section, I construct a unitary representation of ${}_d$rHDG that I use to compute the Pauli-Liouville representation. I start the construction by defining two auxiliary representations that are needed for the representation of the qudit ${}_d$rHDG. The auxiliary representations are representations of the symmetric and cyclic groups. I start with the representation of the symmetric group, which I denote X. Let S_d denote the symmetric group of d elements [3]. For a permutation $\sigma \in S_d$, its representative by X is

$$\mathsf{X}(\sigma) = \sum_i \delta_{i,\sigma(i)}; \tag{5.1}$$

δ is the Kronecker delta, considered as a matrix with a single entry non-zero and that non-zero entry is equal to 1.

As an example of the representation X, consider the representatives of S_3. For completeness, I include all the permutations, not only the generators.

$$\mathsf{X}(\mathbb{I}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathsf{X}((12)) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathsf{X}((13)) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \tag{5.2}$$

$$\mathsf{X}((23)) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathsf{X}((123)) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad \mathsf{X}((132)) = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

The representation X is called the standard representation of S_d. This representation is reducible and decomposes into two irreps: one is the trivial irrep and the other is called the standard irrep of S_d.

The second auxiliary representation is for a direct product of cyclic groups. I start the definition of this representation by introducing notation for cyclic groups. Elements of the cyclic group are denoted by α. Let p denote the order of α: $\alpha^p = \mathbb{I}$. I denote the cyclic group of order k as C_k; it has elements of the form α^p, with $p \in [k]$. As an example consider, for $k = 3$, C_3 has elements $\{\alpha, \alpha^2, \alpha^3 = \mathbb{I}\}$.

A product of cyclic groups is an important subgroup of the ${}_d$rHDG. Consider a multiset with l elements: $\boldsymbol{k} = (k_0, \ldots, k_{l-1})$, each component of $\mathbf{k}$ is a positive integer. Then I denote the direct product of the cyclic groups with order given by each entry of $\boldsymbol{k}$ as

$$\bigtimes_{k \in \boldsymbol{k}} C_k \coloneqq C_{k_0} \times \cdots \times C_{k_{l-1}}. \tag{5.3}$$

The elements of $\bigtimes_{k\in\boldsymbol{k}} C_k$ are denoted by $(\alpha_0^{p_0}, \ldots, \alpha_l^{p_{l-1}})$, where each $p_i \in [k_i]$ denote the power of the element α_i of the group C_{k_i}. Cyclic groups appear in ${}_d$rHDG in the form of powers of diagonal non-Clifford gates. The representation for $\bigtimes_{k\in\boldsymbol{k}} C_k$ is computed as follows. Given $\boldsymbol{\alpha} = (\alpha_0^{p_0}, \ldots, \alpha_{l-1}^{p_{l-1}}) \in \bigtimes_{k\in\boldsymbol{k}} C_k$, the mapping D is defined as

$$\mathsf{D}(\boldsymbol{\alpha}) = \mathsf{D}(\alpha_0^{p_0}, \ldots, \alpha_{l-1}^{p_{l-1}}) \coloneqq \operatorname{diag}[\omega_{\#\,d}^{p_0}, \ldots, \omega_{\#\,d}^{p_{l-1}}]. \tag{5.4}$$

Now I can define a representation of the ${}_d$rHDG.

Using the representations of Eqs. (5.1) and (5.4), I define the representation for the ${}_d$rHDG. For a pair $(\sigma, \boldsymbol{\alpha}) \in S_d \times \bigtimes_{k\in\boldsymbol{k}} C_k$ the mapping γ is

$$\gamma(\sigma, \boldsymbol{\alpha}) \coloneqq \mathsf{X}(\sigma)\,\mathsf{D}(\boldsymbol{\alpha}). \tag{5.5}$$

The mapping γ is a $d \times d$ unitary irreducible representation (unirrep). In the following

section, I study the properties of γ relevant for my randomised benchmarking scheme.

The rest of this section is devoted to the construction of concrete representations of ${}_d\mathrm{rHDG}$. First, I discuss the form of known T gates for qudits. The matrices that I present are useful for obtaining instances of the ${}_d\mathrm{rHDG}$. I use qudit diagonal T gates [144]. The T gates defined for $d \neq 2, 3$ are

$$T = \sum_{j \in \mathbb{Z}_d} \omega_d^{j^3} \delta_{j,j}. \tag{5.6a}$$

For $d = 3$ the diagonal T gate I use is

$$T = \sum_{j \in \mathbb{Z}_d} \omega_{3d}^{j^3} \delta_{j,j}. \tag{5.6b}$$

For future convenience, I denote the order of the qudit T gate as $\#(d)$; thus

$$\#(d) \coloneqq \begin{cases} 9, & d = 3 \\ d, & \text{otherwise} \end{cases}. \tag{5.7}$$

For power-of-prime dimensions, the corresponding T gate should be constructed as if a multi-qudit gate. For instance, for ququart systems, the T gate is $T \otimes \mathbb{I}$, where T is the qubit T gate. Next step is to compute the group generated by the set of matrices computed from the cyclic permutations of the diagonal entries of T.

Consider $T = \mathrm{diag}[\omega_{\#(d)}^{p_0}, \ldots, \omega_{\#(d)}^{p_{l-1}}]$. Extract the diagonal from T; $D \coloneqq (\omega_{\#(d)}^{p_0}, \ldots, \omega_{\#(d)}^{p_{l-1}})$. Compute every permutations of D: $\boldsymbol{D} = \{(\omega_{\#(d)}^{p_0}, \ldots, \omega_{\#(d)}^{p_{l-1}}), \ldots, (\omega_{\#(d)}^{p_{\sigma(0)}}, \ldots, \omega_{\#(d)}^{p_{\sigma(l-1)}})\}$. Using Howell's algorithm [75], extract from $\boldsymbol{D}$ the generators of the group generated by $\boldsymbol{D}$ along with their orders. The output of Howell's algorithm is the set of permutations $\boldsymbol{\sigma} = (\sigma_0, \ldots, \sigma_{l'})$ and a set of positive integers $\{o_0, \ldots, o_{l'}\}$, one integer per permutation. The minimal generating set is

$$\boldsymbol{T'} \coloneqq \{\mathrm{diag}[\omega_{\#(d)}^{o_i p_{\sigma_i(0)}}, \ldots, \omega_{\#(d)}^{o_i p_{\sigma_i(l-1)}}] \colon i \in [l']\}; \tag{5.8}$$

from $\boldsymbol{T}'$ the form of the group generated by the permutations of the diagonal entries of T can be computed:

$$\langle T_\sigma \colon \sigma \in S_d \rangle \cong \bigtimes_{k \in \boldsymbol{k}} C_k, \tag{5.9}$$

where $\#(d)$ is the order of the T gate, $\sigma \in \boldsymbol{\sigma}$, and $\boldsymbol{k} = \{k_0, k_1, \ldots, k_l'\} = \{\# d/o_0, \ldots, \# d/o_{l'}\}$.

I conclude this section with a qutrit example of the ${}_d$rHDG. This construction serves to illustrate the construction for other POP systems. For qutrits, $\#(3) = 9$ as in Eq. (5.7). Thus, the basic diagonal gates are:

$$\mathsf{D}\left(\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2)\right) = \mathrm{diag}[{\omega_9}^{\alpha_0}, {\omega_9}^{\alpha_1}, {\omega_9}^{\alpha_2}], \tag{5.10}$$

where $\alpha_i \in [3]$. The T gate for this case is $T = \mathrm{diag}[\omega_9, 1, 1]$.

Knowing the representatives for diagonal and permutations, I can write the general representative for a ${}_d$rHDG element. Consider for instance

$$\gamma((12), \boldsymbol{\alpha}) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathrm{diag}[{\omega_9}^{\alpha_0}, {\omega_9}^{\alpha_1}, {\omega_9}^{\alpha_2}] = \begin{bmatrix} 0 & {\omega_9}^{\alpha_1} & 0 \\ {\omega_9}^{\alpha_0} & 0 & 0 \\ 0 & 0 & {\omega_9}^{\alpha_2} \end{bmatrix}. \tag{5.11}$$

In §5.3 I prove that representations such as γ are useful to characterise a T gate for POP-level systems. I never directly exploit the fact that γ is an irrep; the methods that I use to prove the decomposition of the Pauli-Liouville representation can be used to prove that γ is indeed an irrep of ${}_d$rHDG.

Illustrate the group to make it less abstract from a geometric perspective. In the case of a T gate $\mathrm{diag}[\omega_3, 1, 1]$, the resulting group is $S_3 \ltimes C_3 \times C_3 \times C_3$. This group is also the wreath product between S_3 and C_3. This allows us to picture the ${}_d$rHDG for T as the symmetry group of the figure in Fig. 5.1. The symmetry group of the vertices of the triangle in Fig. 5.1 is $S_3 \wr C_3$, which is the group for $T = \mathrm{diag}[\omega_3, 1, 1]$.

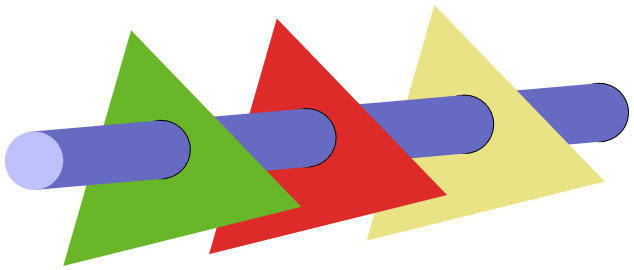


Figure 5.1: Geometric figure used to illustrate the ${}_d$rHDG. The symmetry group of the vertices of the triangles form the ${}_d$rHDG for the qutrit T gate given by diag$[\omega_3, 1, 1]$.

5.3 Representation theory for the rHDG

In this section, I discuss some notation from representation theory applied to the ${}_d$rHDG. The main motivation is to use representation theory to prove that the Pauli-Liouville representation of the representation γ decomposes, indeed, into three inequivalent irreps. This section is basically the proof of this decomposition, including examples of some steps in the proof.

For each representative computed using γ from Eq. (5.5), I define:

$$\Gamma(\sigma, \boldsymbol{\alpha}) = \gamma(\sigma, \boldsymbol{\alpha}) \otimes \overline{\gamma}(\sigma, \boldsymbol{\alpha}) \in \text{end}(\mathfrak{h}), \tag{5.12}$$

where the bar denotes complex conjugate. Now I discuss the key property of Γ: its decomposition into three distinct irreps. It will be shown later that this property is valid for any d, not only POP systems. However, with the current knowledge of universal gate sets, only for POP systems our results are applicable.

The proof of the tripartite decomposition of end$(\mathfrak{h})$ has the following strategy: First, I state a result on characters to count the frequency by which an irrep appears in a reducible representation. Then I compute the character of the irrep γ of Eq. (5.5). Next, using the average of the modulus squared of the character of the irrep, I sum over all the elements of the rHDG. Using Theorem 5.1, below I conclude the proof of the tripartite decomposition.

Theorem 5.1 (Adapted from Serre's textbook [82])**.** *Let χ be the character of a represen-*

tation ρ of a group G acting on the Hilbert space $\mathfrak{h}$. The representation ρ decomposes $\mathfrak{h}$ into the following sum:

$$\mathfrak{h} = \underbrace{\mathfrak{h}_{i_0} \oplus \cdots \oplus \mathfrak{h}_{i_0}}_{\varphi_0 \text{ summands}} \oplus \cdots \oplus \underbrace{\mathfrak{h}_{i_k} \oplus \cdots \oplus \mathfrak{h}_{i_k}}_{\varphi_k \text{ summands}}; \tag{5.13}$$

where φ_i is called the frequency in which the irrep $\mathfrak{h}_i$ appears in $\mathfrak{h}$. Then

$$\mathop{\mathbb{E}}_{g\in G} |\chi_\rho(g)|^2 = \sum_{i\in[k]} \varphi_i^2. \tag{5.14}$$

This result is commonly used to compute the dimensions of the irreps appearing in some representation.

I start the proof the decomposition of the representation of Eq. (5.12) by an observation useful to compute the character of an ${}_d$rHDG element. The elements on the diagonal of any $\mathsf{X}(\sigma)$ correspond to the entries fixed by σ. For instance consider the permutation (13) for an ordered list of three elements $(\diamondsuit, \spadesuit, \heartsuit)$; $(13)(\diamondsuit, \spadesuit, \heartsuit) = (\heartsuit, \spadesuit, \diamondsuit)$: $\spadesuit$ is a fixed element. Then $\langle\spadesuit|\,\mathsf{X}((13))\,|\spadesuit\rangle \neq 0$. Note also that multiplying (left or right multiplication) any $\mathsf{X}(\sigma)$ by a diagonal matrix still keeps the fixed diagonal non-zero entry and the non-zero entries are equal to 1. Another explicit example is given in Eq. (5.11).

The diagonal elements of the rHDG correspond to matrices of the form

$$\mathsf{X}(\mathbb{I})\,\mathsf{D}(\boldsymbol{\alpha}) = \begin{bmatrix} \omega_{\#(d)}^{\alpha_0} & & & \\ & \omega_{\#(d)}^{\alpha_1} & & \\ & & \ddots & \\ & & & \omega_{\#(d)}^{\alpha_{d-1}} \end{bmatrix}. \tag{5.15}$$

More notation is needed for the next part. Let $J(\sigma)$ denote the set of indices of the diagonal entries fixed by σ. Then I know $\mathsf{X}(\sigma)_{i,i} = 1$ if and only if $i \in J(\sigma)$; on the contrary, $\mathsf{X}(\sigma)_{i,i} = 0$ if and only if $i \notin J(\sigma)$. Thus, the diagonal entries of $\mathsf{X}(\sigma)\,\mathsf{D}(\alpha)$ are $\omega_{\#(d)}^{\alpha_i}$ with $i \in J(\sigma)$. I

define the quantity

$$f_\sigma \coloneqq |J(\sigma)|. \tag{5.16}$$

I am ready to compute the characters of γ and Γ,

$$\chi_\gamma(\sigma, \boldsymbol{\alpha}) = \sum_{i\in[J(\sigma)]} \omega_{\#(d)}^{\alpha_i}. \tag{5.17}$$

From Eq. (5.17) I compute the character of the Pauli-Liouville representation: $\chi_\Gamma(\sigma, \boldsymbol{\alpha}) = |\chi_\gamma(\sigma, \boldsymbol{\alpha})|^2$.

From χ_γ the character of the Pauli-Liouville representation:

$$\chi_\Gamma(\sigma, \alpha) = \sum_{i,j\in[J(\sigma)]} \omega_{\#(d)}^{\alpha_i-\alpha_j} = f_\sigma + \sum_{i\neq j} \omega_{\#(d)}^{\alpha_i-\alpha_j} \tag{5.18a}$$

note that $\chi_\Gamma(\sigma, \boldsymbol{\alpha})$ is a real number, thus I omit the modulus appearing in Theorem 5.1. The next step is averaging ${\chi_\Gamma}^2$ (the square of $\chi_\Gamma(\sigma, \boldsymbol{\alpha})$) over every ${}_d$rHDG element and verifying that $\mathbb{E}_{\sigma,\boldsymbol{\alpha}} \chi_\Gamma(\sigma, \boldsymbol{\alpha})^2 = 3$. Then, using Theorem 5.1, show that there are only three inequivalent irreps in the ${}_d$rHDG.

First, I compute $\chi_\Gamma^2(\sigma, \alpha)$:

$$\chi_\Gamma^2(\sigma, \alpha) = (f_\sigma + \sum_{i\neq j} \omega_{\#(d)}^{\alpha_i-\alpha_j})(f_\sigma + \sum_{u\neq v} \omega_{\#(d)}^{\alpha_u-\alpha_v}) \tag{5.18b}$$

$$= f_\sigma^2 + f_\sigma(\sum_{i\neq j} \omega_{\#(d)}^{\alpha_i-\alpha_j}) + f_\sigma(\sum_{u\neq v} \omega_{\#(d)}^{\alpha_u-\alpha_v}) + (\sum_{i\neq j,u\neq v} \omega_{\#(d)}^{\alpha_i-\alpha_j+\alpha_u-\alpha_v}) \tag{5.18c}$$

I note that, for $i \neq j$

$$\mathop{\mathbb{E}}_{\boldsymbol{\alpha}} \omega_{\#(d)}^{\alpha_i-\alpha_j} = 0. \tag{5.18d}$$

Thus

$$\mathop{\mathbb{E}}_{\boldsymbol{\alpha}} \sum \chi_\Gamma^2(\sigma, \boldsymbol{\alpha}) = f_\sigma^2 + \mathop{\mathbb{E}}_{\boldsymbol{\alpha}} \sum_{i\neq j,u\neq v} \omega_{\#(d)}^{\alpha_i-\alpha_j+\alpha_u-\alpha_v} \tag{5.18e}$$

To simplify Eq. (5.18e) I need to isolate the phases α_i appearing in the exponent of ω. Only

Table 5.1: Table with the clasification of the phases.

Configuration				$\alpha_i - \alpha_j + \alpha_u - \alpha_v$
$i \neq v$	$i = u$	$j \neq u$	$j = v$	$2\alpha_i - 2\alpha_j$
$i \neq v$	$i = u$	$j \neq u$	$j \neq v$	$2\alpha_i - \alpha_j - \alpha_v$
$i = v$	$i \neq u$	$j = u$	$j \neq v$	0
$i \neq v$	$i \neq u$	$j = u$	$j \neq v$	$\alpha_i - \alpha_v$
$i \neq v$	$i \neq u$	$j \neq u$	$j = v$	$\alpha_i + \alpha_u - 2\alpha_v$
$i = v$	$i \neq u$	$j \neq u$	$j \neq v$	$\alpha_u - \alpha_j$
$i \neq v$	$i \neq u$	$j \neq u$	$j \neq v$	$\alpha_i - \alpha_j + \alpha_u - \alpha_v$

then I can compute the average over the phases. I decompose the cases $i \neq j$ and $u \neq v$ in Table 5.18.

In Table 5.18 I describe the different combinations of phases that appear in the sum Eq. (5.18e). I decompose the sum-average in the second summand in the right-hand side of Eq. (5.18e):

$$\mathop{\mathbb{E}}_{\boldsymbol{\alpha}} \sum_{i\neq j, u\neq v} \omega_{\#(d)}^{\alpha_i-\alpha_j+\alpha_u-\alpha_v} = \mathop{\mathbb{E}}_{\boldsymbol{\alpha}} \sum_{\substack{i\neq v, i=u,\\ j\neq u, j\neq v}} \omega_{\#\,d}^{2\alpha_i-2\alpha_j} + \mathop{\mathbb{E}}_{\boldsymbol{\alpha}} \sum_{\substack{i\neq v, i=u,\\ j\neq u, j\neq v}} \omega_{\#\,d}^{2\alpha_i-\alpha_j-\alpha_v} \tag{5.18f}$$

$$+ \mathop{\mathbb{E}}_{\boldsymbol{\alpha}} \sum_{\substack{i=v, i\neq u,\\ j=u, j\neq v}} \omega_{\#\,d}^{0} \tag{5.18g}$$

$$+ \mathop{\mathbb{E}}_{\boldsymbol{\alpha}} \sum_{\substack{i\neq v, i\neq u,\\ j=u, j\neq v}} \omega_{\#\,d}^{\alpha_i-\alpha_v} + \mathop{\mathbb{E}}_{\boldsymbol{\alpha}} \sum_{\substack{i\neq v, i\neq u,\\ j\neq u, j=v}} \omega_{\#\,d}^{\alpha_i+\alpha_u-2\alpha_v} \tag{5.18h}$$

$$+ \mathop{\mathbb{E}}_{\boldsymbol{\alpha}} \sum_{\substack{i=v, i\neq u,\\ j\neq u, j\neq v}} \omega_{\#\,d}^{\alpha_u-\alpha_j} + \mathop{\mathbb{E}}_{\boldsymbol{\alpha}} \sum_{\substack{i\neq v, i\neq u,\\ j\neq u, j\neq v}} \omega_{\#\,d}^{\alpha_i-\alpha_j+\alpha_u-\alpha_v}. \tag{5.18i}$$

Note that each sum, except (5.18g), has in the exponent a sum of phases, with each phase different. Thus, averaging over the phases I obtain zero. Only the sum (5.18g) is non-zero:

$$\mathop{\mathbb{E}}_{\boldsymbol{\alpha}} \sum_{\substack{i=v, i\neq u,\\ j=u, j\neq v}} \omega_{\#\,d}^{0} = f_\sigma(f_\sigma - 1). \tag{5.18j}$$

Thus,

$$\mathop{\mathbb{E}}_{\boldsymbol{\alpha}} \sum \chi_\Gamma^2(\sigma, \alpha) = f_\sigma^2 + f_\sigma(f_\sigma - 1) = 2f_\sigma^2 - f_\sigma. \tag{5.18k}$$

Note that the computations needed to arrived at the last equation (Eq. (5.18k)) are valid for any phase order greater than two. This shows my scheme is valid for any diagonal gate with order greater than two.

I am close to the end of my proof of the decomposition. The next step is to average with respect to the permutations $\sigma \in S_d$. I rely on the following Lemma. Let B_k be the k-th Bell number. I need, in a future proof, $B_1 = 1$ and $B_2 = 2$.

Lemma 5.2 ([73])**.** *Let n and k be two positive integers such that $k \leq n$. Then*

$$B_k = \mathop{\mathbb{E}}_{\sigma \in S_n} f_\sigma^k. \tag{5.19}$$

Alternatively to Rota's proof [73], I offer a proof for Lemma 5.2 in Appendix A, which I produced independently.

Proposition 5.3. *For the T gates defined in Eq.* (5.6) *let γ be the corresponding unirrep defined in Eq.* (5.5)*. Then Γ splits* $\mathrm{end}(\mathfrak{h})$ *into three distinct real irreps:*

$$\mathrm{end}(\mathfrak{h}) = \mathfrak{h}_{\mathbb{I}} \oplus \mathfrak{h}_0 \oplus \mathfrak{h}_+. \tag{5.20}$$

Proof. The proof of this proposition amounts to showing that

$$\mathop{\mathbb{E}}_{\sigma, \boldsymbol{\alpha}} \chi_\Gamma(\sigma, \boldsymbol{\alpha})^2 = 3. \tag{5.21}$$

From Eq. (5.18k) and using Lemma 5.2:

$$\mathop{\mathbb{E}}_{\sigma, \boldsymbol{\alpha}} \chi_\Gamma(\sigma, \boldsymbol{\alpha})^2 = \mathop{\mathbb{E}}_{\sigma}(2f_\sigma^2 - f_\sigma) = 2 \cdot 2 - 1 = 3. \tag{5.22a}$$

Since the solution of Eq. (5.14) for the system in Eq. (5.22) is having three different $\varphi_i = 1$

and the rest equal to zero, using Theorem 5.1, I conclude there are only three irreps in the Pauli-Liouville representation of the ${}_d$rHDG: the trivial irrep always appears in the product of one representation and its complex conjugate [10, 82]. The other two irreps are discussed in §5.4. □

5.4 Expressions for the average gate fidelity and sequence fidelity

Knowing there are only three irreps in the Pauli-Liouville representation of ${}_d$rHDG is good because it means that only two parameters are needed to characterise CPTP noise. However, to have a practical scheme, it is important also to know how to access those parameters experimentally. The next step is to relate the eigenvalues η_ϖ of the twirl by the ${}_d$rHDG to the sequence and gate fidelities. To do so, I need to prove that the state $|+\rangle\!\rangle$ is mapped to the null vector by Π_0. First, an auxiliary lemma.

Lemma 5.4. *The average of $f_\sigma - 1$ over all permutations is equal to zero:*

$$\mathop{\mathbb{E}}_{\sigma\in S_d}(f_\sigma - 1) = 0. \tag{5.23}$$

Proof. The quantity $f_\sigma - 1$ is the character of a non-trivial irrep of the symmetric group; it is called standard irrep. The equality follows from the orthogonality of the character with the trivial irrep. □

Lemma 5.5. *The projector Π_0 from $\mathrm{end}(\mathfrak{h})$ to $\mathfrak{h}_0$, maps $|+\rangle\!\rangle$ to the null-vector.*

Proof. For this proof I use the Kraus representation. In this representation $|+\rangle\!\rangle$ has entries equal to

$$(|+\rangle\!\rangle_{\mathrm{K}})_{d(i-1)+j} \coloneqq (|+\rangle\langle+|)_{i,j}. \tag{5.24}$$

Thus, $|+\rangle\!\rangle_{\mathrm{K}}$ has all its entries equal. Therefore, for all σ

$$\left(\mathsf{X}(\sigma)\otimes\overline{\mathsf{X}}(\sigma)\right)|+\rangle\!\rangle_{\mathrm{K}}=|+\rangle\!\rangle_{\mathrm{K}}. \tag{5.25}$$

Thus,

$$\mathop{\mathbb{E}}_{\sigma}(f_\sigma-1)\,\mathsf{X}(\sigma)\otimes\overline{\mathsf{X}}(\sigma)|+\rangle\!\rangle_{\mathrm{K}}=|+\rangle\!\rangle_{\mathrm{K}}\left(\mathop{\mathbb{E}}_{\sigma}(f_\sigma-1)\right)=\mathbf{0}, \tag{5.26}$$

where $\mathbf{0}$ is the null-vector. For the rightmost equality in Eq. (5.26), I use Lemma 5.4. □

Also, the vectorisation of $|0\rangle\!\langle 0|$ is orthogonal to $|+\rangle\!\rangle$. Then I have $\Pi_0|+\rangle\!\rangle=\Pi_+|0\rangle\!\rangle$ and $\Pi_0|0\rangle\!\rangle\neq 0\neq\Pi_+|+\rangle\!\rangle$. Therefore, due to the decomposition in Eq. (5.20):

$$\langle\!\langle\varpi|[\mathcal{E}]_{\mathrm{T}}^{m}|\varpi\rangle\!\rangle=\frac{d-1}{d}\eta_\varpi(\mathcal{E})^m+\frac{1}{d}; \tag{5.27}$$

ϖ labels the irrep and thus take values 0 and +. Thus, from the sequence fidelity the eigenvalues η_ϖ can be estimated.

Now I obtain the form of the average gate fidelity. Because,

$$\mathrm{tr}(\mathcal{E})=\mathrm{tr}([\mathcal{E}]_\gamma)=1+(d-1)\eta_0+(d^2-d)\eta_+, \tag{5.28}$$

the expression for the average gate fidelity is

$$F(\mathcal{E})=\frac{d(1+(d-1)\eta_0+(d^2-d)\eta_+)+d^2}{d^2(d+1)}. \tag{5.29}$$

I now relate the eigenvalues η_ϖ with the sequence fidelity, the last task necessary to define my scheme.

The sequence fidelity is used to estimate η_0 and η_+, which appear in the gate fidelity expression Eq. (5.27). The formal expression of the sequence fidelity comes from the com-

position of the twirl m times $[\mathcal{E}]_\gamma^m$. The sequence fidelity is

$$P(m; |\varpi\rangle, |\varpi\rangle\langle\varpi|) = a + b\eta_\varpi(\mathcal{E})^m; \tag{5.30}$$

a, b depend on SPAM errors, with values close to d^{-1} and $(d-1)d^{-1}$, respectively. Note a and b are quantities that do not appear in the average gate fidelity but are necessary to consider in the fitting procedure. By using the standard randomised benchmarking scheme, the eigenvalues η_0 and η_+ are estimated. Then, using the expression Eq. (5.29), the average gate fidelity is estimated.

5.5 Experimental scheme

In this section, I present the implementation of my scheme using the required ingredients. The requisites of the scheme include gates from ${}_d$rHDG, the Fourier operation F, preparation of $|0\rangle$, and measurement of $\langle 0|$. My explanation overlaps with the one previously given in §2.10.

The basic step of the experiment proceeds as follows:

1. Prepare $|0\rangle$.
2. Randomly draw m ${}_d$rHDG elements: $g_0, \ldots, g_{m-1}$.
3. On a classical computer, identify the gate corresponding to the group element $(g_0 \cdots g_{m-1})^{-1}$.
4. Apply the gates $g_0, \ldots, g_{m-1}, (g_0 \cdots g_{m-1})^{-1}$ to $|0\rangle$, and then measure $\langle 0|$.

I refer to the sequence of gates mentioned in this paragraph as the circuit.

There are two iterative steps in the scheme:

1. The first iteration is on the number of different circuits that need to be realised to estimate Eq. (5.27) by averaging the output of each run. Notice that in an experimental

setting, for each circuit, the number of repetitions, called 'shots', is also required to estimate $P(m; |0\rangle, \langle 0|, \mathcal{E})$ as in Eq. (2.88a).

2. The second iterative step involves changing the values of m, the circuit depth. Thus, after repeating the steps 2-4 for various m values, a graph of circuit depth *vs* averaged sequence fidelity is constructed.

To implement our scheme, besides using ${}_d$rHDG instead of the Clifford group, two randomised benchmarking experiments are needed to conduct, as is the case for Clifford randomised benchmarking. One experiment requires preparation and measurement of $|0\rangle$, and the other requires preparation and measurement of the state $|+\rangle = F|0\rangle$; F is defined in Eq. (3.13). The steps in the experiment using $|+\rangle$ are identical (besides adding at the start of the circuit the gate F and at the end the gate $F^\dagger$) to those using $|0\rangle$.

5.6 Multi-qudit case

As is implicit in the proof of the decomposition in Proposition 5.3, and in the gates used in the multi-qubit dihedral benchmarking scheme [66], our scheme is also useful to characterise a controlled non-Clifford gate. In this section, I describe the required gate set, initial state, and measurement for the characterisation.

For a system of two qudits, the gate set needs to include a representation of S_{d^2}. This group has two generators: a CSUM gate and a single-qudit X gate, where

$$\text{CSUM}(|a, b\rangle) = |a, a \oplus b\rangle\,; \tag{5.31}$$

the $\oplus$ symbol denotes addition modulo d. The single-qudit X gate is simply the tensor product of identities except on the first gate, where X should be put in.

As to the initial states needed to estimate the eigenvalues η_0 and η_+, I note that the argument for projectors and the state $F|0\rangle$ still holds, as shown in Lemma 5.5. Therefore,

the initial states necessary (for an n-qudit system) are $|0\rangle\!\rangle^n$ and $|+\rangle\!\rangle \otimes |0\rangle\!\rangle^{n-1}$.

5.7 Conclusion

In this chapter, I introduced my main contribution: a scheme, based on randomised benchmarking, to characterise universal qudit gates. The first point to note is that, while my construction applies to any diagonal gate with an order equal to or greater than $\#(d)$, I restrict it to POP systems. This restriction is imposed because there is no known universal gate set generator for non-POP systems. By focusing on POP systems, the scheme remains applicable to a more manageable class of quantum systems.

Next, I addressed the mathematical proofs related to the scheme, such as the decomposition of the Pauli-Liouville representation and the states necessary to access a single parameter. These mathematical elements are crucial for understanding how the scheme operates. The key insight is that all the hard work is done by the mathematical identity linking the number of partitions of a finite set with the fixed points of permutations. This connection simplifies many aspects of the proof and shows how combinatorial properties underpin the performance of the scheme.

I finished this section by demonstrating that two initial states, $|0\rangle$ and $|+\rangle$, are members of only one of the two non-trivial irreducible representations (irreps) appearing in the Pauli-Liouville representation. Specifically, $|0\rangle$ belongs to $\mathfrak{h}_0$ and $|+\rangle$ to $\mathfrak{h}_+$. These states thus provide a way to individually estimate the parameters η_ϖ required to characterise a gate set. This feature is particularly useful in practical applications where accurate parameter estimation is key.

I then discussed the individual steps required for an experimental implementation of the scheme. This discussion serves the purpose of illustrating, in detail, the experimental resources needed, ranging from the preparation of gates to the expressions that are fitted to estimate parameters. These parameters ultimately lead to the estimation of the average gate

fidelity, which is a key figure of merit in quantum information theory. By outlining these steps, I provide a clearer understanding of what is required for a successful implementation.

Finally, I concluded the chapter with a discussion of the scheme for multi-qudit gates. I showed how to construct the gate set for multi-qudit systems and identified the initial states that can exploit the bi-parametric nature of the scheme. This extension of the scheme to multi-qudit gates broadens its applicability, allowing for more complex quantum operations to be characterised.

In the next chapter, equipped with the ${}_d$rHDG, I extend several methods that aim to achieve a more useful characterisation than a collective one. These methods take advantage of several properties of the ${}_d$rHDG, which I did not exploit in the collective characterisation addressed in this chapter. This approach promises to yield more nuanced insights into the behaviour of quantum gates and systems.

Chapter 6

Application of the rHDG to extensions of randomised benchmarking

6.1 Introduction

This chapter is about application of the rHDG in alternatives to randomised benchmarking. I concentrate on two scheme alternatives. These alternatives aim is to overcome particular limitations of the standard randomised benchmarking. In this chapter I start by discussing a limitation—in the original randomised benchmarking— and then address the alternative scheme that overcomes it. In both alternatives, the ${}_d$rHDG is substituted for the Clifford gate set.

6.2 Interleaved benchmarking

The original randomised benchmarking scheme estimates the average gate fidelity over a gate set. However, it is easy to imagine the following situation: some Clifford gates—such as Pauli gates—are easy to implement and it is expected to have low noise; this is not the case

for the Hadamard gate, which most of the time require multiple native gates [12]. Therefore in this situation it is reasonable to be unsure about the quantity randomised benchmarking determines. Interleaved benchmarking was developed to overcome this limitation [2, 145, 90, 60].

Interleaved benchmarking is a scheme that allows to estimate the average gate fidelity of an individual gate by using an pre-characterised gate set. The characterisation is achieved by benchmarking the composition of the gate of interest with pre-characterised gate set members. By means of an approximation of the composition of two channels, the average gate fidelity of the gate of interest is computed. In this section I present my generalisation for universal qudit gates of this method; my work generalises qubit results [2].

Three theoretical assumptions are made for the noise of the gates: the noise for the T gate acts before the gate; the noise for members of the Clifford like gate set act after the gate; the third assumption is that the noise resulting from applying T is the same as if applying T multiple times. The requirement to apply the T gate more than once should not be surprising. Newer randomised benchmarking schemes—cycle benchmarking, for instance—have a similar circuit design arrangement. As these schemes are being used today, this assumption is reasonable.

The interleaved benchmarking scheme estimates the noise of the composition of two gates—one from the already characterised gate set and the other from the gate to be characterised. The interleaved benchmarking circuit is designed to achieve this composition. Then an approximation of the composition's average gate fidelity as the product of the average gate fidelity of each noise is used. From the approximation—and the fact one gate set has already been characterised—the average gate fidelity of the non-characterised gate is computed.

In what follows, I design the circuit required to obtain the noise composition, which allows the estimate of the average gate fidelity of the non-characterised gate. Note that this scheme is based on the original work on dihedral benchmarking [2], which also uses the technique of considering T, up to some power, to have the same noise as T.

In this subsection I go over the implementation of my scheme. I assume there is a platform with access to Clifford gates and a T gate. The requirements of my extension of interleaved benchmarking are: a T gate, Clifford gates—specially H, the ${}_d$rHDG, and state and measurement with respect to the state $|0\rangle$.

I now construct the subset of gates that are assumed to be already characterised. Conjugate T using any permutation matrix and form the set

$$\boldsymbol{T} \coloneqq \{X_\sigma T X_\sigma^\dagger \colon \sigma \in S_d\}. \tag{6.1}$$

Compute the group generated by $\boldsymbol{T}$ in Eq. (6.1) and Z (the qudit clock matrix)—call it $\mathcal{C}' \coloneqq \langle \boldsymbol{T}, Z\rangle$. Then compute the normaliser—with respect to the Heisenberg-Weyl group—of $\mathcal{C}'$; I denote it $\mathcal{N}$. Then the group

$$\mathcal{C} \coloneqq \langle \mathcal{N}, X, X_{01}\rangle. \tag{6.2}$$

For convenience, elements of $\mathcal{C}$ are denoted as g. It can be shown that twirling with respect to the $\mathcal{C}$ produces the same twirl as ${}_d$rHDG.

I ensure the circuit has only elements in $\mathcal{C}$. Denote by p the power of the T gate that is a member of $\mathcal{C}$. Notice p depends on the dimension: for qutrits and six-level systems, $\mathrm{p} = 3$; for POP systems with dimension $d = p^k$; for some k, $\mathrm{p} = p$.

The noise for the T gate acts before T, and the noise for the $\mathcal{C}$ gates acts after the ideal gate. The inversion gate is, therefore, also a member of $\mathcal{C}$. The circuit is of the form

$$\mathcal{E}_{\mathcal{C}}(T^{\mathrm{p}} g_0 \cdots T^{\mathrm{p}} g_{m-1})_{\mathrm{inv}} T^{\mathrm{p}} \mathcal{E}_T \mathcal{E}_{\mathcal{C}} g_0 \cdots T^{\mathrm{p}} \mathcal{E}_T \mathcal{E}_{\mathcal{C}} g_{m-1}. \tag{6.3}$$

Notice each ideal gate of the circuit is a member of $\mathcal{C}$.

By the standard analysis of randomised benchmarking, after averaging over the elements

of $\mathcal{C}$, I obtain the following expression for the sequence fidelity

$$P(m;\mathcal{E}_T,\mathcal{E}_\mathcal{C},|\varpi\rangle) = \langle\langle\varpi|\mathcal{E}_\mathcal{C}[\mathcal{E}_T\mathcal{E}_\mathcal{C}]^m_{\mathrm{T}}|\varpi\rangle\rangle, \tag{6.4}$$

where $|\varpi\rangle = |0\rangle, |+\rangle$. Therefore using the qudit universal expressions, of Eq. (5.30), the channel $\mathcal{E}_T\mathcal{E}_\mathcal{C}$ can be characterised. The procedure is explained below.

With the expression for the composite noise I am ready to compute the expression for the average gate fidelity of the T gate. The following approximation is used in interleaved benchmarking schemes [134]. Consider two channels, $\mathcal{E}_0$ and $\mathcal{E}_1$. Then

$$F(\mathcal{E}_0 \circ \mathcal{E}_1) = F(\mathcal{E}_0)\, F(\mathcal{E}_1) + \text{error}; \tag{6.5}$$

since the average gate fidelity is related to the trace of the Pauli-Liouville representation, the approximation (and the associated error) is related to the approximation of the trace of a product as the product of the trace.

The known bounds for the difference between the fidelity of the composition and the product of the composition, smaller the more different $F(\mathcal{E}_0)$ and $F(\mathcal{E}_1)$ are [93]. If the fidelities are expected to be similar, standard randomised benchmarking should be used, since the approximation for similar noise applies. I illustrate my results for a ququart gate set. I link an ideal experimental setting with the expressions required for the characterisation of a set of gates. The justification is that the ququart system requires relatively small circuit depths.

A related interesting result for this example is to show the "correct" generalisation of the Pauli group for ququarts. For ququarts, the normaliser of the Heisenberg-Weyl group is not a unitary 2-design. However, the normaliser of the following set of matrices is indeed a unitary 2-design

$$\mathcal{P} = \langle \sigma_x \otimes \mathbb{I}, \mathbb{I} \otimes \sigma_x, \sigma_z \otimes \mathbb{I}, \mathbb{I} \otimes \sigma_z \rangle. \tag{6.6}$$

From this, then the T gate I use is

$$ {}_4T \coloneqq \mathrm{diag}[1, \omega_8] \otimes \mathbb{I}; \tag{6.7} $$

The rHDG for this particular case of a ququart system is

$$ {}_4\mathrm{rHDG} = \langle X, X_{01}, {}_4T^2 \rangle; \tag{6.8} $$

note that $\langle X, X_{01} \rangle$ normalises $\mathcal{P}$ in Eq. (6.6).

The representatives of the auxiliary group is generated by the following matrices (note this set is not minimal but convenient for the computation of the composition and inverse):

$$ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \omega_4 & & & \\ & 1 & & \\ & & \omega_4 & \\ & & & -1 \end{bmatrix}, \begin{bmatrix} \omega_4 & & & \\ & \omega_4 & & \\ & & 1 & \\ & & & -1 \end{bmatrix}, \omega_4 \mathbb{I}. \tag{6.9} $$

The group generated by these matrices is $A_4 \ltimes (C_4 \times C_4 \times C_4)$. I call the gates associated with the representatives of this group the auxiliary gate set.

Now, I discuss some properties of ${}_4T$. By direct computation, $A_4 \ltimes C_4 \times C_4 \times C_4$ is invariant under conjugation by ${}_4T^2$. Assuming ${}_4T^2$ has the same noise as ${}_4T$, this assumption is sensible because applying a gate twice should result in similar errors as applying the gate once: the reason is that similar techniques are applied. Thus, characterising ${}_4T^2$ implies characterising ${}_4T$. Another useful observation is that ${}_4T^4$ is a member of $A_4 \ltimes C_4 \times C_4 \times C_4$. This ensures the inversion gate, which requires even length gates including ${}_4T^2$, is always in the auxiliary gate set.

With the previous observations, I start discussing the experimental implementation. Consider the following circuit:

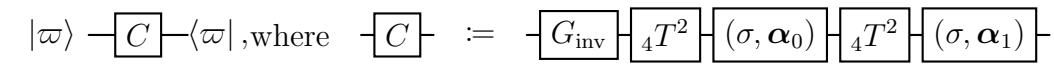


Figure 6.1: Circuit corresponding to a run with circuit depth $m = 2$ of interleaved benchmarking.

I use the circuit in Fig. 6.1 to compute the averaged sequence fidelity that can be obtained by randomly drawing gates from the gate set $A_4 \ltimes C_4 \times C_4 \times C_4$.

From Fig. 6.1, the corresponding channel (appearing between the state $|0\rangle$ and measurement $\langle 0|$) is

$$\widetilde{C} \coloneqq \widetilde{G}_{\text{inv}} \cdot {}_4\widetilde{T}^2 \widetilde{\gamma}(\sigma_0, \boldsymbol{\alpha}_0) \cdot {}_4\widetilde{T}^2 \cdot \widetilde{\gamma}(\sigma_1, \boldsymbol{\alpha}_1), \tag{6.10}$$

where

$$\widetilde{G}_{\text{inv}} = \hat{\gamma}(\sigma_1, \boldsymbol{\alpha}_1)({}_4\hat{T}^2)^\dagger \hat{\gamma}(\sigma_0, \boldsymbol{\alpha}_0)({}_4\hat{T}^2)^\dagger \mathcal{E}. \tag{6.11}$$

Thus, the complete channel is

$$\widetilde{C} = \hat{\gamma}(\sigma_1, \boldsymbol{\alpha}_1)({}_4\hat{T}^2)^\dagger \hat{\gamma}(\sigma_0, \boldsymbol{\alpha}_0)({}_4\hat{T}^2)^\dagger \mathcal{E}\mathcal{E}_T {}_4\hat{T}^2 \hat{\gamma}(\sigma_0, \boldsymbol{\alpha}_0)\mathcal{E}\mathcal{E}_T {}_4\hat{T}^2 \hat{\gamma}(\sigma_1, \boldsymbol{\alpha}_1)\mathcal{E}, \tag{6.12}$$

where I assigned the noise $\mathcal{E}$ to auxiliary gates and $\mathcal{E}_T$ to the target gate. The next step, in the process to characterise the target gate ${}_4T$, is finding the expression for the twirl.

Averaging Eq. (6.12) over $\boldsymbol{\alpha}$ and then σ

$$\begin{aligned}
\mathop{\mathbb{E}}_{\sigma_0,\sigma_1,\boldsymbol{\alpha}_0,\boldsymbol{\alpha}_1} \widetilde{C} = & && \text{(6.13a)}\\
= & \mathop{\mathbb{E}}_{\sigma_0,\sigma_1,\boldsymbol{\alpha}_0,\boldsymbol{\alpha}_1} \hat{\gamma}(\sigma_1, \boldsymbol{\alpha}_1)({}_4\hat{T}^2)^\dagger \hat{\gamma}(\sigma_0, \boldsymbol{\alpha}_0)({}_4\hat{T}^2)^\dagger \mathcal{E}\mathcal{E}_T({}_4\hat{T}^2)\hat{\gamma}(\sigma_0, \boldsymbol{\alpha}_0) && \text{(6.13b)}\\
& \times \mathcal{E}\mathcal{E}_T({}_4\hat{T}^2)\hat{\gamma}(\sigma_1, \boldsymbol{\alpha}_1)\mathcal{E}, && \text{(6.13c)}\\
= & \mathop{\mathbb{E}}_{\sigma_1,\boldsymbol{\alpha}_1} \hat{\gamma}(\sigma_1, \boldsymbol{\alpha}_1)({}_4\hat{T}^2)^\dagger [\mathcal{E}\mathcal{E}_T]\mathcal{E}\mathcal{E}_T({}_4\hat{T}^2)\hat{\gamma}(\sigma_1, \boldsymbol{\alpha}_1)\mathcal{E}, && \text{(6.13d)}\\
= & \mathop{\mathbb{E}}_{\sigma_1,\boldsymbol{\alpha}_1} [\mathcal{E}\mathcal{E}_T]\hat{\gamma}(\sigma_1, \boldsymbol{\alpha}_1)({}_4\hat{T}^2)^\dagger \mathcal{E}\mathcal{E}_T({}_4\hat{T}^2)\hat{\gamma}(\sigma_1, \boldsymbol{\alpha}_1)\mathcal{E}, && \text{(6.13e)}\\
= & [\mathcal{E}\mathcal{E}_T]^2\mathcal{E}, && \text{(6.13f)}
\end{aligned}$$

where I obviate the composition symbol, $\circ$. Therefore, considering the last channel $\mathcal{E}$ to be a SPAM contribution, using the randomised benchmarking scheme with two distinct initial states ($|0\rangle$ and $|+\rangle$), the average gate fidelity of $\mathcal{E} \circ \mathcal{E}_T$ is obtained. Then the approximation $F(\mathcal{E} \circ \mathcal{E}_T) \approx F(\mathcal{E})\, F(\mathcal{E}_T)$ is used to obtain $F(\mathcal{E}_T)$. This concludes the exposition of my generalisation of interleaved benchmarking for qudit systems.

6.3 Shadow estimate for universal gates

The first step is constructing an auxiliary matrix, denoted by A, that determines the quantity estimated by the scheme. There is no unique way of constructing A. Therefore, I present one without claiming its uniqueness, as it is the simplest and does not require further knowledge of the irreps of ${}_d\mathrm{rHDG}$.

The following three lemmas are necessary for the scheme. They appear in the supplementary material of [146] but are not proven there.

Lemma 6.1. *Let G be a finite group and ρ an irrep of G, distinct from the trivial irrep. Then*

$$\mathop{\mathbb{E}}_{g} \rho(g) = \mathbf{0}, \tag{6.14}$$

where $\mathbf{0}$ is the null operator.

Proof. This is a corollary of Schur's lemma [82]. Since $\mathbb{E}_g\, \rho(g)$ commutes with any element $\rho(h)$, we have $\mathbb{E}_g\, \rho(g) = \alpha \mathbb{I}$. Proving the lemma is equivalent to showing that α is zero. By *reductio ad absurdum*, suppose α is not zero. Then for any $h \in G$, $(\mathbb{E}_g\, \rho(g))\rho(h) = \alpha \mathbb{I} \rho(h) = \rho(h) = \alpha \mathbb{I}$. Thus, for any $g \in G$, $\rho(g) = \alpha \mathbb{I}$, which implies ρ is reducible. This is a contradiction, as the statement assumes ρ is an irrep. The contradiction arises from assuming $\alpha \neq 0$. Thus, $\alpha = 0$, and therefore $\mathbb{E}_g\, \rho(g) = \mathbf{0}$. □

Lemma 6.2. *Let G be a group and σ a representation, possibly reducible, of G. Then the*

operator

$$\mathop{\mathbb{E}}_{g\in G}\sigma(g) \tag{6.15}$$

is a projector onto the vector space spanned by invariant vectors under the action of σ*.*

Proof. Summing over the representatives of an irrep is equal to the null operator unless the irrep is the trivial irrep (see Lemma 6.1). Therefore, assume the trivial irrep $\sigma_{\mathbb{I}}$ appears $k > 0$ times in σ. It is always possible to block-diagonalise (here achieved by the matrix U) a reducible representation [82]:

$$U\left(\mathop{\mathbb{E}}_{g}\sigma(g)\right)U^\dagger = \mathrm{diag}[1,\ldots,1]\oplus\sigma^\perp, \tag{6.16}$$

where $\sigma^\perp$ is the orthogonal complement of the direct sum of trivial irreps appearing in σ. Then we have

$$(U\mathop{\mathbb{E}}_{g}\sigma(g)U^\dagger)\left|i\right\rangle = \left|i\right\rangle, \tag{6.17}$$

in the computational basis ($i \in [k]$). Defining

$$\left|\psi_i\right\rangle \coloneqq U^\dagger\left|i\right\rangle, \tag{6.18}$$

the form of the operator is

$$U(\mathop{\mathbb{E}}_{g}\sigma(g))U^\dagger = \sum_i \left|\psi_i\right\rangle\!\left\langle\psi_i\right|, \tag{6.19}$$

which is a projector onto the vector space spanned by invariant vectors under the action of σ. □

The following lemma is used in Eqs. (6.22) and (6.25).

Lemma 6.3. *Let* G *be a group and* σ *and* ρ *be two* <u>real</u> *irreps of* G*. Then one of the following results holds:*

1. *If* σ *and* ρ *are not isomorphic, then* $\mathbb{E}_g\,\sigma(g)\otimes\rho(g) = \mathbf{0}$*.*

2. *If σ and ρ are isomorphic, then the trivial irrep appears once in the decomposition of $\sigma \otimes \rho$.*

Proof. Computing the trace of $\mathbb{E}_g \sigma(g) \otimes \rho(g)$ is equivalent to taking the inner product between the characters of σ and ρ, which is the number of times ρ appears in σ. If σ is not isomorphic to ρ, then the sum of traces (divided by the order of the group) equals 0. However, if the irreps are equivalent, the average of the trace of the representation equals 1, corresponding to the number of times the trivial irrep appears in $\sigma \otimes \rho$. □

Now that the tools for the method are mostly stated, I describe the big picture of the method. There are two main differences compared to other randomised benchmarking schemes that I introduced in this scheme. The first difference is the amount of *a posteriori* data analysis required. Another important difference is the absence of an inversion gate.

Grosso modo, the scheme requires randomly drawing gates and applying them to some state, finishing with a measurement. This part is similar to randomised benchmarking but does not require the computation and implementation of an inversion gate. The randomly drawn gates are stored, and an almost identical ideal sequence is then computationally evaluated. The adverb "almost" is justified because an additional matrix A pre-multiplies the representatives used in the simulation. Then, for each sequence of gates, the experimental estimate of the sequence fidelity and the simulation, using A, are multiplied, and the results are averaged over each sequence. The matrix A is chosen to compute part of the trace of the Kronecker product of A and the Pauli-Liouville representation of the noise.

Now, I discuss the construction of the auxiliary representation to compute the convolution sequence, which I denote by C. I show—in 5.3—that the Pauli-Liouville representation Γ of ${}_d$rHDG (acting on $\mathfrak{h}$) has three invariant subspaces: $\mathfrak{h}_{\mathbb{I}}$, $\mathfrak{h}_0$, and $\mathfrak{h}_+$. These irreps decompose the endomorphisms of the Hilbert space $\mathfrak{h}$ as:

$$\mathrm{end}(\mathfrak{h}) = \mathfrak{h}_{\mathbb{I}} \oplus \mathfrak{h}_0 \oplus \mathfrak{h}_+. \tag{6.20}$$

Therefore, to construct a convolution of sequence fidelity that allows the estimation of the eigenvalue η_ϖ, I can use σ_ϖ as the auxiliary representation.

I compute the expression for the sequence fidelity to estimate the parameter η_0. For a mapping f defined on domain X with $Y \subset X$, I denote the mapping with domain Y as $f \restriction_Y$, defined as $f(y)$ if $y \in Y$ and $f(y) = \mathbf{0}$ if $y \in X \setminus Y$. Using this notation, consider the irrep:

$$\sigma_\varpi \coloneqq \Gamma \restriction_{\mathfrak{h}_\varpi} . \tag{6.21}$$

I use σ_ϖ to estimate the new sequence fidelity.

The computation of the sequence fidelity starts with the computation of the following mapping:

$$\begin{aligned} \mathbf{\Pi}_0 &\coloneqq \mathop{\mathbb{E}}_g \sigma_0(g) \otimes \Gamma(g) && (6.22)\\ &= \mathop{\mathbb{E}}_g(\sigma_0 \otimes \sigma_{\mathbb{I}}) \oplus \mathop{\mathbb{E}}_g(\sigma_0 \otimes \sigma_0) \oplus \mathop{\mathbb{E}}_g(\sigma_0 \otimes \sigma_+) && (6.23)\\ &= \mathbf{0}_{d-1} \oplus \Pi_0 \oplus \mathbf{0}_{(d-1)(d^2-d)}, && (6.24) \end{aligned}$$

where I used Lemma 6.3 and $\mathbf{0}_i$ is the projector onto the null subspace with dimension i. Repeating the analysis with the irrep σ_+, I get

$$\mathbf{\Pi}_+ \coloneqq \mathop{\mathbb{E}}_g \sigma_+(g) \otimes \Gamma(g) = \mathbf{0}_{d-1} \oplus \mathbf{0}_{(d-1)(d^2-d)} \oplus \Pi_+. \tag{6.25}$$

Using mappings $\mathbf{\Pi}_0$ and $\mathbf{\Pi}_+$ I obtain the expression for the convolution sequence required to obtain the expression for the sequence fidelity.

Now, I compute the convolution sequence fidelity for shadow estimate. In this case, as mentioned at the beginning of this subsection, the expression is of the form

$$P_\varpi(m) = \langle\langle 0, \varpi | (\mathbb{I} \otimes \mathcal{E})(\mathbf{\Pi}_\varpi A \otimes \mathcal{E}\mathbf{\Pi}_\varpi)^{m-1} | 0, \varpi \rangle\rangle. \tag{6.26}$$

Therefore, the form of the convoluted sequence fidelity is

$$P_{\varpi}(m;\mathcal{E},A,|\varpi\rangle\,,\langle\varpi|)=a_A(\mathcal{E})\left(\frac{\eta_{\varpi}}{\dim\varpi}\right)^{m-1}, \tag{6.27}$$

where $\varpi\in\{0,+\}$ and $a_A(\mathcal{E})$ is a parameter that depends on A and $\mathcal{E}$—$a_A(\mathcal{E})$ is irrelevant in the characterisation scheme; η_{ϖ} is the same eigenvalue as in Chapter 5: Eq. (5.30).

For definitiveness, I show how to use this shadow estimate scheme to approximate the average gate fidelity of the ${}_d$rHDG gates. For the qutrit case I mention the matrices required to estimate the average gate fidelity. The matrices required to access the eigenvalues η_{ϖ}, written in Eq. (6.26), are labelled by the subspace

$$A_0\coloneqq 18\mathbb{I},\quad A_+\coloneqq 12\mathbb{I}. \tag{6.28}$$

Then the sequence fidelity is

$$P_0(m)=6^{-1}(6\eta_0)^m,\quad P_+(m)=(18)^{-1}((18)\eta_+)^m. \tag{6.29}$$

Now, I illustrate the procedure to estimate η_{ϖ}.

Let $\mathbf{M}$ denote a list of positive integers corresponding to the circuit's depths to use. I describe the steps for the state $|0\rangle$ and the irrep $\mathfrak{h}_0$ and those steps are the same for $|+\rangle$ and the irrep $\mathfrak{h}_+$. Let $m\in\mathbf{M}$ be a circuit depth. Let $\boldsymbol{\sigma\alpha}_m$ denote an ordered multiset (a sequence) of ${}_d$rHDG elements. Assume that each $(\sigma,\boldsymbol{\alpha})\in\boldsymbol{\sigma\alpha}_m$ has been randomly drawn, where each pair $(\sigma,\boldsymbol{\alpha})$ denotes a ${}_d$rHDG element. Then two quantities are to be obtained: one is numerical, computed and depends on A; and the other is experimental; first, I discuss the experimental and then the numerical.

Now, I discuss the quantity that should be experimentally estimated. Let $\widetilde{C}$ denote the

circuit corresponding to the sequence of gates to be applied to the initial state:

$$\widetilde{C} = \bigcirc_{\sigma,\boldsymbol{\alpha}\in\boldsymbol{\sigma\alpha}_m} \tilde{\gamma}(\sigma,\boldsymbol{\alpha}), \tag{6.30}$$

where $\bigcirc_{\sigma,\boldsymbol{\alpha}\in\boldsymbol{\sigma\alpha}_m}$ denotes the composition indexed by $\boldsymbol{\sigma\alpha}_m$. The state $|0\rangle$ is prepared, the noisy circuit $\widetilde{C}$ is applied to $|0\rangle$, and then the outcome $\widetilde{C}(|0\rangle)$ is measured with respect to $|0\rangle$. Repeating the experiment results in a quantity that approximates

$$\tilde{p}(\boldsymbol{\sigma\alpha}_m) \coloneqq \mathrm{tr}\Big[|0\rangle\langle 0|\,\widetilde{C}(|0\rangle\langle 0|)\Big]. \tag{6.31}$$

After estimating $\tilde{p}(\boldsymbol{\sigma\alpha}_m)$ the next step is computing the quantity that is required for the convolution.

The analytical quantity is similarly computed with a few changes. Let $\hat{C}$ be the following sequence of gates

$$\hat{C} = \bigcirc_{\sigma,\boldsymbol{\alpha}\in\boldsymbol{\sigma\alpha}_m} A\Gamma(\hat{\gamma}\restriction_{\mathfrak{h}_0})(\sigma,\alpha). \tag{6.32}$$

Then the following quantity is computed:

$$\hat{p}(\boldsymbol{\sigma\alpha}_m) \coloneqq \langle\!\langle 0|\hat{C}|0\rangle\!\rangle. \tag{6.33}$$

The quantities $\tilde{p}(\boldsymbol{\sigma\alpha}_m)$ and $\hat{p}(\boldsymbol{\sigma\alpha}_m)$ are estimated and computed for each $m \in \mathbf{M}$. From these two numerical values, the following quantity is defined:

$$\bar{p}(\boldsymbol{\sigma\alpha}_m) \coloneqq \mathop{\mathbb{E}}_{\sigma,\boldsymbol{\alpha}} \tilde{p}(\boldsymbol{\sigma\alpha}_m)\hat{p}(\boldsymbol{\sigma\alpha}_m). \tag{6.34}$$

$\bar{p}(\boldsymbol{\sigma\alpha}_m)$ is used to estimate the fidelity.

By fitting the expression Eq. (6.27) to the graph $(m, \bar{p}(\boldsymbol{\sigma\alpha}_m))$ the eigenvalue η_0 can be estimated. Similarly, by repeating the procedure using as initial state $|+\rangle$ the average gate fidelity of ${}_d$rHDG can be determined with Eq. (5.29). This concludes my exposition of the

method called shadow estimate."

6.4 Universal cycle benchmarking

In this section, I discuss my extension of cycle benchmarking to qudits. My extension is valid for both Clifford gates and non-Clifford gates. The cycle benchmarking scheme, in its original formulation, allows the individual characterisation of a Clifford gate by using only Pauli gates and the target Clifford gate to characterise. In this section, I concentrate on my extension to characterise a T gate. My scheme uses the same circuit-design as the original cycle benchmarking scheme, and in terms of difficulty, it substitutes the Clifford gate with the qudit X gate.

Whereas in the standard formulation of cycle benchmarking, the basic gate set is the Pauli gate set, in my scheme, I use a slight variation of it. The basic gate set is generated by X^T and Z :

$$\mathcal{R}_d \coloneqq \langle X^T, Z \rangle, \tag{6.35}$$

where I am denoting conjugation

$$g^T \coloneqq T^\dagger g T, \tag{6.36}$$

do not confuse with the transposition operation for matrices. Now, I list the properties of this gate set and their relevance to the scheme. I drop the subscript $_d$ when the result applies for any dimension or if the dimension is given in the context.

The gate set $\mathcal{R}$ is—under conjugation—T-invariant. Given $g \in \mathcal{R}$, the condition refers to $g^T \in \mathcal{R}$. This condition is employed during the computation of the sequence fidelity. The second property is that a basis for the set of $d \times d$ matrices is a subset of $\mathcal{R}$.

In my case, the subset is formed by

$$\mathcal{B} \coloneqq \{(X^T)^i Z^j : i, j \in [d]\}. \tag{6.37}$$

I use $\mathcal{B}$ to construct the Pauli-Liouville representation of the channels. The reason is that using the set of Pauli matrices as a basis for the Pauli-Liouville representation results in a non-diagonal matrix. This makes it a bit more tedious to obtain and explain the resulting expressions.

The last property of $\mathcal{R}$ is related to representation theory. I require that the number of irreps appearing in the decomposition of the Pauli-Liouville representation of $\mathcal{R}$ is the same as for $\mathcal{P}$, the Heisenberg-Weyl matrices. Verifying this for $\mathcal{R}$ is a trivial computation.

6.5 Conclusion

In this chapter, I showed three applications of the ${}_d$rHDG in randomised benchmarking-like schemes. These applications serve to illustrate the importance of ${}_d$rHDG beyond my original goal. Some applications are key to the relevance of the scheme as they address more practical considerations. In this section, I briefly discuss several of the results and their implications.

Consider first interleaved benchmarking, where the assumption in randomised benchmarking schemes—that all gates have equal noise—is unrealistic. Interleaved benchmarking allows for a more realistic noise model by considering different noise levels for different gates. It makes sense to assume Clifford gates have the same noise, different from that of a T gate, in an error-correcting scheme. In this case, the Clifford gates are encoded, but the T gate is implemented using magic state distillation.

Next, shadow estimation can be used to partially reconstruct a channel, which is an essential aspect of quantum process tomography. I also learned that the inversion gate is useful for reducing the number of samples required in such reconstructions. Avoiding the inversion gate in shadow estimation results in an almost four-fold increase in the number of samples needed compared to randomised benchmarking. However, this comparison is not entirely fair, as shadow estimation should be compared to tomographic techniques.

Lastly, in cycle benchmarking, I exploit several properties of groups to generalise the tech-

nique to characterise non-Clifford gates. This method does not use the ${}_d$rHDG explicitly but draws on similar group structures. The resulting gate set has the same structure as ${}_d$rHDG, although the irrep structure (the number of distinct irreps appearing in the Pauli-Liouville representation) is different. This demonstrates the flexibility of the ${}_d$rHDG framework in adapting to various benchmarking schemes.

Chapter 7

Summary and open problems

In this final chapter of the thesis, I summarise the results presented in the following chapters: §4, §5, and §6. The main result of §4 is the gate set HDG. By using a set of gates labelled by HDG elements, I developed a characterization scheme for a qutrit T gate similar to dihedral benchmarking for qubits in terms of experimental implementation. In §5, I not only addressed an issue in the qutrit case but also generalized dihedral benchmarking for any level system and any number of qudits. This was achieved using the group I introduced: ${}_d$rHDG. This is the main chapter of this thesis. Finally, in §6, I discuss several applications of ${}_d$rHDG. This aims to both increase the applicability of my results and extend certain protocols to a universal gate set. Now, I will discuss the summary for each chapter in more detail.

In Chapter 2, I reviewed some experimental implementations of randomised benchmarking. There, I showed the circuit depths that can be achieved and how the graph of circuit depth *vs* sequence fidelity justifies trace-preserving and Markovian noise conditions. These two facts, at least for these platforms, show that the assumptions of Clifford randomised benchmarking are valid, and thus my schemes could be used to characterise those platforms. With randomised benchmarking mitigating further issues such as drift and leakage [128, 131], even under highly adversarial noise conditions, the average gate fidelity can be estimated by

adding randomised compiling to my schemes.

7.1 Universal qutrit randomised benchmarking

Before my work on qutrits, there was no method to characterise a non-Clifford qutrit gate. With multiple qutrit implementations now a reality, it makes sense to have a generalisation for qutrit systems. In this project, I faced the challenge of formalising dihedral benchmarking, as it was initially unclear what generalising dihedral benchmarking entailed. Moreover, in contemporary literature, the scheme was introduced with limited explanations and justifications, making my work necessary to address those gaps.

In Chapter 4, I demonstrated that using the same construction for the gate set as in the qubit case leads to a group that has a Pauli-Liouville representation with five parameters, two of which are complex conjugate pairs. Using this group with a high-fidelity gate set, the characterisation reduces to that of the qubit case. By applying the gate set construction, I recover either the dihedral group or a gate set that leads to the same irrep decomposition of the Pauli-Liouville representation. I also provide expressions for both sequence fidelity and gate fidelity, showing how these elements tie into existing frameworks. Using GAP, I demonstrated that the HDG is the smallest group (i.e. lowest order) that achieves such decomposition of the Pauli-Liouville representation.

In addition, I introduced criteria in Chapter 4 to certify the properties that a given abstract group and its representations should satisfy. These criteria played a crucial role in the generalisation for qudits of the HDG representations. This approach was essential in ensuring that the qutrit benchmarking scheme could be extended effectively to higher-dimensional systems. Without such criteria, the applicability of the scheme would have been much more limited.

Three primary limitations to my HDG work exist. The most tangible issue is the phase that appears in the sequence fidelity expression. This phase complicates the fitting pro-

cess because it is difficult to estimate, contradicting the simplicity sought with randomised benchmarking schemes. The difficulty in estimating this phase introduces a complexity that challenges the ease of use of the overall method.

The second limitation concerns entangling gates. My scheme is currently limited to characterising single qutrit gates, which leaves out multi-qubit systems. However, with current implementations of entangling gates in qutrit systems, this limitation is increasingly significant and limits broader applicability. Addressing this limitation would be an important step for further development.

The last limitation is specific to qutrits. The construction $\langle X, T \rangle$ leads to more than five parameters for ququart systems and continues to increase for larger systems. This means that, beyond qutrits, the HDG method becomes increasingly complex and unwieldy. Therefore, it would be preferable to find another gate set for qudits with $d > 3$ that simplifies the parameter count while maintaining fidelity to the benchmarking objectives.

7.2 Universal qudit randomised benchmarking

Before my work, only for qubits and qutrits was there a scheme to characterise non-Clifford gates. Currently, there is increasing interest in qudit systems across multiple platforms, even those with entangling gates [147]. Therefore, a generalisation of dihedral benchmarking is urgently needed. Moreover, the proof technique that dihedral benchmarking used to characterise multi-qubit gates does not work for qudits, as it relies on the fact that each Pauli gate is its own inverse, a property that does not hold for qudits.

In Chapter 5, I presented my generalisation of dihedral benchmarking for any dimension and any number of qudits. The outcomes of this scheme include the gate set necessary to characterise a non-Clifford gate and the corresponding expressions for sequence fidelity and average gate fidelity. This is a crucial development, as no such generalisation existed before for higher-dimensional systems. My approach allows for efficient characterisation across

multiple qudit platforms, addressing a significant gap in the field.

The gate set in this scheme is a representation of a group. This group, the ${}_d$rHDG, forms a semidirect product of the symmetric group and a product of cyclic groups. Under certain circumstances, depending on the T gate used, the group is the generalised symmetric group. This group structure simplifies the complexity of the benchmarking scheme while ensuring it is adaptable to various multi-qudit systems.

The ${}_d$rHDG scheme requires estimating only two parameters. Estimating each parameter requires a different initial state: the first state is $|0\rangle$, and the second requires applying the Fourier operator to $|0\rangle$. This result is applicable for any dimension and across any multi-qudit platform, making it a versatile tool for characterisation. These parameters streamline the process of evaluating gate fidelity, significantly reducing the overhead in experimental setups.

The scheme is grounded in the decomposition into irreps of the Pauli-Liouville representation of a unirrep of the abstract group ${}_d$rHDG. This decomposition makes ${}_d$rHDG unique: using an arbitrary group could require up to d^2 parameters to fit. For an n-qudit platform, the parameter count could increase to d^{2n}, making the approach impractical. My results on decomposition are mathematically formal, ensuring robustness in the scheme.

One key result is the computation of the character of the irrep γ of the ${}_d$rHDG. With that knowledge, I employed a celebrated identity relating the number of partitions of a set and the powers of fixed points of permutations. These results allowed me to prove that the Pauli-Liouville representation of γ decomposes into three inequivalent irreps. Each non-trivial irrep corresponds to one parameter, and one parameter is equal to 1 if the noise is trace-preserving.

The scheme has two main limitations: one inherited from the assumptions of randomised benchmarking and the other related to current knowledge of qudit gates. In the standard formulation of randomised benchmarking, it is assumed that gates experience the same Markovian noise. Consequently, the estimated quality is only an approximation of the average gate

fidelity of the target gate, which is typically a non-Clifford gate.

The second limitation relates to the dimension of the system. For non-POP systems, a finite universal gate set is currently unknown, and even an appropriate Pauli gate set (with a normalizer that is also a unitary 2-design) has not been established. Therefore, while my scheme can characterise any diagonal gate with order greater than two in any dimension, characterising universal gates is still limited to POP systems. This limitation highlights the need for further research in developing gate sets for higher-dimensional systems.

Finally, the justification for the applicability of my work to current hardware is two-fold. First, empirical evidence (see Sec. 3.5) demonstrates that sequence fidelity decays as a single exponential, indicating that non-CPTP or non-Markovian behaviour is negligible. Second, my scheme is compatible with randomised compiling [58, 108, 113], ensuring that even in the presence of non-CPTP or non-Markovian behaviour, my scheme remains valid as a characterisation tool in terms of average gate fidelity estimation.

7.3 Applications of the real HyperDihedral group

In Chapter 6, I discuss the applications of the ${}_d$rHDG. I present four generalisations of randomised benchmarking, which are useful because they employ the same methods as randomised benchmarking but for more general error configurations. This means that I relax the assumption that the noise is the same for each gate. These generalisations extend the applicability of the scheme to a wider range of practical scenarios.

First, the most important generalisation is that of interleaved benchmarking for two reasons. One is that interleaved benchmarking offers the most practical application for the characterisation of universal gates. This is crucial in the context of error correction schemes, where Clifford gates can be encoded, but the T-gate cannot. In this case, I can clearly distinguish between the types of noise affecting Clifford and non-Clifford gates.

The second generalisation is shadow estimation. This technique is significant because it

allows for noise recovery or reconstruction, making it similar to tomography. However, unlike tomography, shadow estimation avoids the issues associated with SPAM errors, making it more reliable in practice.

Finally, I discuss cycle benchmarking. Cycle benchmarking plays an essential role in interleaved benchmarking, particularly for multi-qubit gates. Although I do not directly use ${}_d$rHDG in this case, I employ a similar construction, leading to results equivalent to those in the Clifford gate scenario. This shows the versatility of the method in various contexts.

7.4 Open problems

The open problems in my research are as follows. The most practical and immediately relevant problem is creating a user-friendly and effective estimate of the number of samples required to attain a certain confidence level in the fitting procedure. Currently, the variance is unknown, which complicates this estimate. Additionally, the number of samples needed to obtain an estimate from non-linear or even linear fitting, when the variance is not constant, prevents obtaining an optimal estimate for these quantities.

Another issue, more mathematical in nature, is identifying the smallest group that has a decomposition (in terms of irreps) identical to the decomposition obtained by using the ${}_d$rHDG. For the qutrit case, I demonstrated that the HDG is the smallest because I was able to list all groups of that order. However, for higher-level systems, the task of enumeration becomes extremely complicated and does not scale well, presenting a significant challenge.

The final open problem concerns a universal gate set for non-prime power systems. Currently, the only known proofs for non-trivial universal gate sets are restricted to systems with prime or prime power dimensions. This leaves a gap for systems such as quhex (six-level systems), where no generating set has yet been established. Solving this problem would

expand the applicability of universal gate sets to a broader range of quantum systems.

Bibliography

[1] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Theoretical framework for quantum networks. *Phys. Rev. A*, 80:022339, 2009.

[2] Arnaud Carignan-Dugas, Joel J Wallman, and Joseph Emerson. Characterizing universal gate sets via dihedral benchmarking. *Phys. Rev. A*, 92:060302, 2015.

[3] Michael Tinkham. *Group Theory and Quantum Mechanics*. Dover, 1992.

[4] Michael A Nielsen and Isaac L Chuang. *Quantum Computation and Quantum Information*. Cambridge UP, Cambridge, UK, 10th anniversary ed. edition, 2010.

[5] P W Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, 1994.

[6] Dorit Aharonov and Michael Ben-Or. Fault-tolerant quantum computation with constant error rate. *SIAM J. Comput.*, 38(4):1207–1282, 2008.

[7] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Phys. Rev. A*, 70:052328, 2004.

[8] Daniel Gottesman. *Stabilizer Codes and Quantum Error Correction*. PhD thesis, California Institute of Technology, California, 1997.

[9] Robin Blume-Kohout and Kevin C Young. A volumetric framework for quantum computer benchmarks. *Quantum*, 4:362, 2020.

[10] David Amaro-Alcalá, Barry C Sanders, and Hubert de Guise. Benchmarking of universal qutrit gates. *Phys. Rev. A*, 109:012621, 2024.

[11] Ranee K Brylinski and Goong Chen, editors. *Mathematics of Quantum Computation*. Computational Mathematics. Chapman, Philadelphia, PA, 2002.

[12] Joseph Lindon, Arina Tashchilina, Logan W Cooke, et al. Complete unitary qutrit control in ultracold atoms. *Phys. Rev. Appl.*, 19:034089, 2023.

[13] A Morvan, V V Ramasesh, M S Blok, et al. Qutrit randomized benchmarking. *Phys. Rev. Lett.*, 126(21):210504, 2021.

[14] Martin Ringbauer, Michael Meth, Lukas Postler, et al. A universal qudit quantum processor with trapped ions. *Nat. Phys.*, 18(9):1053–1057, 2022.

[15] M Veldhorst, J C C Hwang, C H Yang, et al. An addressable quantum dot qubit with fault-tolerant control fidelity. *Nat. Nanotechnol.*, 9(12):981–985, 2014.

[16] James D Louck. *Applications of Unitary Symmetry and Combinatorics*. World Sci., Singapore, Singapore, 2011.

[17] Yulin Chi, Jieshan Huang, Zhanchuan Zhang, et al. A programmable qudit-based quantum processor. *Nat. Commun.*, 13(1):1166, 2022.

[18] Daniele Cozzolino, Beatrice Da Lio, Davide Bacco, et al. High-dimensional quantum communication: benefits, progress, and future challenges. *Adv. Quantum Technol.*, 2(12):1900038, 2019.

[19] Yi-Han Luo, Han-Sen Zhong, Manuel Erhard, et al. Quantum teleportation in high dimensions. *Phys. Rev. Lett.*, 123(7):070505, 2019.

[20] Xiao-Min Hu, Chao Zhang, Bi-Heng Liu, et al. Experimental high-dimensional quantum teleportation. *Phys. Rev. Lett.*, 125(23):230501, 2020.

[21] E A Vashukevich, E N Bashmakova, T Yu Golubeva, et al. High-fidelity quantum gates for OAM single qudits on quantum memory. *Laser Phys. Lett.*, 19(2):025202, 2022.

[22] Matthew Otten, Keshav Kapoor, A. Barış Özgüler, et al. Impacts of noise and structure on quantum information encoded in a quantum memory. *Phys. Rev. A*, 104(1):012605, 2021.

[23] Hao Zhang, Chao Zhang, Xiao-Min Hu, et al. Arbitrary two-particle high-dimensional Bell-state measurement by auxiliary entanglement. *Phys. Rev. A*, 99(5):052301, 2019.

[24] C Senko, P Richerme, J Smith, et al. Realization of a quantum integer-spin chain with controllable interactions. *Phys. Rev. X*, 5(2):021026, 2015.

[25] Poolad Imany, Jose A Jaramillo-Villegas, Mohammed S. Alshaykh, et al. High-dimensional optical quantum logic in large operational spaces. *npj Quantum Inf.*, 5(1):59, 2019.

[26] B P Lanyon, T J Weinhold, N K Langford, et al. Manipulating biphotonic qutrits. *Phys. Rev. Lett.*, 100(6):060504, 2008.

[27] Earl T Campbell, Hussain Anwar, and Dan E Browne. Magic-state distillation in all prime dimensions using quantum Reed-Muller codes. *Phys. Rev. X*, 2:041021, 2012.

[28] Earl T Campbell. Enhanced fault-tolerant quantum computing in d-level systems. *Phys. Rev. Lett.*, 113:230501, 2014.

[29] Eliot Kapit. Hardware-efficient and fully autonomous quantum error correction in superconducting circuits. *Phys. Rev. Lett.*, 116:150501, 2016.

[30] Ritajit Majumdar, Saikat Basu, Shibashis Ghosh, et al. Quantum error-correcting code for ternary logic. *Phys. Rev. A*, 97:052302, 2018.

[31] Benedikt Fauseweh and Jian-Xin Zhu. Quantum computing Floquet energy spectra. *Quantum*, 7:1063, 2023.

[32] Frédéric Bouchard, Alicia Sit, Khabat Heshami, et al. Round-robin differential-phase-shift quantum key distribution with twisted photons. *Phys. Rev. A*, 98:010301, 2018.

[33] Isaac Nape, Eileen Otte, Adam Vallés, et al. Self-healing high-dimensional quantum key distribution using hybrid spin-orbit Bessel states. *Opt. Express*, 26(21):26946–26960, 2018.

[34] Mikka Stasiuk, Felix Hufnagel, Xiaoqin Gao, et al. High-dimensional encoding in the round-robin differential-phase-shift protocol. *Quantum*, 7:1207, 2023.

[35] Eric T Holland, Kyle A Wendt, Konstantinos Kravvaris, et al. Optimal control for the quantum simulation of nuclear dynamics. *Phys. Rev. A*, 101(6), 2020.

[36] Daniel González-Cuadra, Torsten V Zache, Jose Carrasco, et al. Hardware efficient quantum simulation of non-Abelian gauge theories with qudits on Rydberg platforms. *Phys. Rev. Lett.*, 129(16), 2022.

[37] Julian Bender, Erez Zohar, Alessandro Farace, et al. Digital quantum simulation of lattice gauge theories in three spatial dimensions. *New J. Phys.*, 20(9):093001, 2018.

[38] Erik J Gustafson. Prospects for simulating a qudit-based model of (1+1)D scalar QED. *Phys. Rev. D*, 103:114505, 2021.

[39] Erez Zohar, J Ignacio Cirac, and Benni Reznik. Quantum simulations of gauge theories with ultracold atoms: Local gauge invariance from angular-momentum conservation. *Phys. Rev. A*, 88:023617, 2013.

[40] F Turro, A Roggero, V Amitrano, et al. Imaginary-time propagation on a quantum chip. *Phys. Rev. A*, 105:022440, 2022.

[41] Doga Murat Kurkcuoglu, M Sohaib Alam, Joshua Adam Job, et al. Quantum simulation of ϕ^4 theories in qudit systems. *arXiv:2108.13357*, 2021.

[42] Christian W Bauer, Zohreh Davoudi, A Baha Balantekin, et al. Quantum simulation for high-energy physics. *PRX Quantum*, 4(2), 2023.

[43] Christian W Bauer, Zohreh Davoudi, Natalie Klco, et al. Quantum simulation of fundamental particles and forces. *Nat. Rev. Phys.*, 5(7):420–432, 2023.

[44] Anastasiia S Nikolaeva, Evgeniy O Kiktenko, and Aleksey K Fedorov. Efficient realization of quantum algorithms with qudits. *EPJ Quantum Technol.*, 11(1), 2024.

[45] Pawel Horodecki, Lukasz Rudnicki, and Karol Życzkowski. Five open problems in theory of quantum information. *arXiv:2002.03233*, 2020.

[46] Fernando Pastawski, Alastair Kay, Norbert Schuch, et al. How long can a quantum memory withstand depolarizing noise? *Phys. Rev. Lett.*, 103:080501, 2009.

[47] John Watrous. *The Theory of Quantum Information*. Cambridge UP, Cambridge, 2018.

[48] Isaac L Chuang and M A Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *J. Mod. Opt.*, 44(11-12):2455–2467, 1997.

[49] E Knill, D Leibfried, R Reichle, et al. Randomized benchmarking of quantum gates. *Phys. Rev. A*, 77:012307, 2008.

[50] Joseph Emerson, Robert Alicki, and Karol Życzkowski. Scalable noise estimation with random unitary operators. *J. Opt. B: Quantum Semiclassical Opt.*, 7(10):S347–S352, 2005.

[51] Easwar Magesan, Jay M Gambetta, and Joseph Emerson. Characterizing quantum gates via randomized benchmarking. *Phys. Rev. A*, 85:042311, 2012.

[52] Jeffrey M Epstein, Andrew W Cross, Easwar Magesan, et al. Investigating the limits of randomized benchmarking protocols. *Phys. Rev. A*, 89(6):062321, 2014.

[53] Michael A Nielsen. A simple formula for the average gate fidelity of a quantum dynamical operation. *Phys. Lett. A*, 303(4):249–252, 2002.

[54] Mahnaz Jafarzadeh, Ya-Dong Wu, Yuval R Sanders, et al. Randomized benchmarking for qudit Clifford gates. *New J. Phys.*, 22(6):063014, 2020.

[55] Timothy Proctor, Kenneth Rudinger, Kevin Young, et al. What randomized benchmarking actually measures. *Phys. Rev. Lett.*, 119(13):130502, 2017.

[56] J Helsen, I Roth, E Onorati, et al. General framework for randomized benchmarking. *PRX Quantum*, 3:020357, 2022.

[57] Bas Dirkse, Jonas Helsen, and Stephanie Wehner. Efficient unitarity randomized benchmarking of few-qubit Clifford gates. *Phys. Rev. A*, 99(1), 2019.

[58] Joel J Wallman and Joseph Emerson. Noise tailoring for scalable quantum computation via randomized compiling. *Phys. Rev. A*, 94:052325, 2016.

[59] Yuval R Sanders, Joel J Wallman, and Barry C Sanders. Bounding quantum gate error rate based on reported average fidelity. *New J. Phys.*, 18(1):012002, 2015.

[60] Easwar Magesan, Jay M Gambetta, B R Johnson, et al. Efficient measurement of quantum gate error by interleaved randomized benchmarking. *Phys. Rev. Lett.*, 109(8):080505–080505, 2012.

[61] Linghang Kong. A framework for randomized benchmarking over compact groups. *arXiv:2111.10357*, 2021.

[62] Jianxin Chen, Dawei Ding, and Cupjin Huang. Randomized Benchmarking beyond Groups. *PRX Quantum*, 3:030320, 2022.

[63] Jahan Claes, Eleanor Rieffel, and Zhihui Wang. Character randomized benchmarking for non-multiplicity-free groups with applications to subspace, leakage, and matchgate randomized benchmarking. *PRX Quantum*, 2:010351, 2021.

[64] Christopher J Wood and Jay M Gambetta. Quantification and characterization of leakage errors. *Phys. Rev. A*, 97(3):032306, 2018.

[65] Bujiao Wu, Xiaoyang Wang, Xiao Yuan, et al. Leakage benchmarking for universal gate sets. *Entropy*, 26(11):71, 2024.

[66] Andrew W Cross, Easwar Magesan, Lev S Bishop, et al. Scalable randomised benchmarking of non-Clifford gates. *npj Quantum Inf.*, 2(1):1–5, 2016.

[67] Alexander Erhard, Joel J Wallman, Lukas Postler, et al. Characterizing large-scale quantum computers via cycle benchmarking. *Nat. Commun.*, 10(1), 2019.

[68] Teiko Heinosaari and Mario Ziman. *The Mathematical Language of Quantum Theory*. Cambridge UP, 2012.

[69] Karl Kraus, A Böhm, J D Dollard, and W H Wootters, editors. *States, Effects, and Operations Fundamental Notions of Quantum Theory*. Springer, 1983.

[70] Junan Lin, Brandon Buonacorsi, Raymond Laflamme, et al. On the freedom in representing quantum operations. *New J. Phys.*, 21(2):023006, 2019.

[71] Simon L Altmann. *Induced Representations in Crystals and Molecules: Point, Space, and Nonrigid Molecule Groups*. Acad. Press, London, 1978.

[72] Louis Comtet. *Advanced Combinatorics*. Springer, Dordrecht, 1974.

[73] Gian-Carlo Rota. The number of partitions of a set. *Am. Math. Mon.*, 71(5):498–, 1964.

[74] V V Fedorov. *Theory of Optimal Experiments*. Probability and mathematical statistics. Acad. Press, San Diego, CA, 1972.

[75] John A Howell. Spans in the module Zms. *Linear Multilinear Algebra*, 19(1):67–77, 1986.

[76] Stephen H Friedberg, Arnold J Insel, and Lawrence E Spence. *Linear Algebra*. Pearson, Upper Saddle River, NJ, 4 edition, November 2002.

[77] H F Baker. *The Collected Mathematical Papers of James Joseph Sylvester*, volume 3. Cambridge UP, Cambridge, 1909.

[78] Julian Schwinger. Unitary operator bases. *Proc. Natl. Acad. Sci. U.S.A.*, 46(4):570–579, 1960.

[79] J Patera and H Zassenhaus. The Pauli matrices in n dimensions and finest gradings of simple Lie algebras of type An-1. *J. Math. Phys.*, 29(3):665–673, 1988.

[80] Reinhold A Bertlmann and Philipp Krammer. Bloch vectors for qudits. *J. Phys. A: Math. Theor.*, 41(23):235303, 2008.

[81] Fereshte Shahbeigi, David Amaro-Alcalá, Zbigniew Puchała, and Karol Życzkowski. Log-convex set of Lindblad semigroups acting on N-level system. *J. Math. Phys.*, 62(7), 2021.

[82] Jean-Pierre Serre. *Linear Representations of Finite Groups*. Graduate Texts in Mathematics. Springer, New York, 1977.

[83] Daniel Gottesman and Isaac L Chuang. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature*, 402(6760):390–393, November 1999.

[84] Charles H Bennett, David P DiVincenzo, John A Smolin, et al. Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, 54:3824–3851, 1996.

[85] Gabriele Carcassi, Lorenzo Maccone, and Christine A. Aidala. Four postulates of quantum mechanics are three. *Phys. Rev. Lett.*, 126:110402, Mar 2021.

[86] Valter Moretti. *Spectral Theory and Quantum Mechanics*. La Matematica per il 3+2, 110. Springer, Cham, 2nd ed. 2017. edition, 2017.

[87] Stephen D Bartlett, Terry Rudolph, and Robert W Spekkens. Reference frames, superselection rules, and quantum information. *Rev. Mod. Phys.*, 79:555–609, 2007.

[88] Seth T Merkel, Emily J Pritchett, and Bryan H Fong. Randomized benchmarking as convolution: Fourier analysis of gate dependent errors. *Quantum*, 5:581, 2021.

[89] Joel J Wallman. Randomized benchmarking with gate-dependent noise. *Quantum*, 2:47, 2018.

[90] Robin Harper and Steven T Flammia. Estimating the fidelity of T gates using standard interleaved randomized benchmarking. *Quantum Sci. Technol.*, 2(1):015008, 2017.

[91] A Barut and R Rączka. *Theory of Group Representations and Applications*. World Sci., 1986.

[92] John F Cornwell. *Group Theory in Physics: An Introduction*. Acad. Press, 1997.

[93] Shelby Kimmel, Marcus P Da Silva, Colm A Ryan, et al. Robust extraction of tomographic information via randomized benchmarking. *Phys. Rev. X*, 4(1):011050, 2014.

[94] J F Poyatos, J I Cirac, and P Zoller. Complete characterization of a quantum process: The two-bit quantum gate. *Phys. Rev. Lett.*, 78:390–393, 1997.

[95] K R W Jones. Fundamental limits upon the measurement of state vectors. *Phys. Rev. A*, 50:3682–3699, 1994.

[96] Christoph Dankert, Richard Cleve, Joseph Emerson, et al. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Phys. Rev. A*, 80(1):012304, 2009.

[97] Eiichi Bannai and Etsuko Bannai. A survey on spherical designs and algebraic combinatorics on spheres. *Eur. J. Comb.*, 30(6):1392–1425, 2009.

[98] David P DiVincenzo, Patrick Hayden, and Barbara M Terhal. Hiding quantum data. *Found. Phys.*, 33(11):1629–1647, 2003.

[99] Easwar Magesan. Gaining information about a quantum channel via twirling. Master's thesis, University of Waterloo, 2008.

[100] Andrew W Cross, Lev S Bishop, Sarah Sheldon, Paul D Nation, and Jay M Gambetta. Validating quantum computers using randomized model circuits. *Phys. Rev. A*, 100:032328, 2019.

[101] Elijah Pelofske, Andreas Bärtschi, and Stephan Eidenbenz. Quantum volume in practice: what users can expect from NISQ devices. *IEEE Trans. Quantum Eng.*, 3:1–19, 2022.

[102] A. Shabani, R. L. Kosut, M. Mohseni, H. Rabitz, M. A. Broome, M. P. Almeida, A. Fedrizzi, and A. G. White. Efficient measurement of quantum dynamics via compressive sensing. *Phys. Rev. Lett.*, 106:100401, Mar 2011.

[103] Robin Blume-Kohout, John King Gamble, Erik Nielsen, Kenneth Rudinger, Jonathan Mizrahi, Kevin Fortier, and Peter Maunz. Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography. *Nat. Comm.*, 8(1), 2017.

[104] Erik Nielsen, Kenneth Rudinger, Timothy Proctor, Antonio Russo, Kevin Young, and Robin Blume-Kohout. Probing quantum processor performance with pygsti. *Quantum Sci. Technol.*, 5(4):044002, 2020.

[105] M Mohseni, A T Rezakhani, and D A Lidar. Quantum-process tomography: Resource analysis of different strategies. *Phys. Rev. A*, 77:032322, 2008.

[106] Easwar Magesan, J M Gambetta, and Joseph Emerson. Scalable and robust randomized benchmarking of quantum processes. *Phys. Rev. Lett.*, 106:180504, 2011.

[107] Ingemar Bengtsson and Karol Życzkowski. *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge UP, Cambridge, second edition, 2017.

[108] Akel Hashim, Ravi K Naik, Alexis Morvan, et al. Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor. *Phys. Rev. X*, 11(4), 2021.

[109] Arnaud Carignan-Dugas, Kristine Boone, Joel J Wallman, et al. From randomized benchmarking experiments to gate-set circuit fidelity: how to interpret randomized benchmarking decay parameters. *New J. Phys.*, 20(9):092001, 2018.

[110] P Figueroa-Romero, M Papič, A Auer, M-H Hsieh, K Modi, and I de Vega. Operational Markovianization in randomized benchmarking. *Quantum Sci. Technol.*, 9(3):035020, 2024.

[111] Pedro Figueroa-Romero, Kavan Modi, Robert J. Harris, Thomas M. Stace, and Min-Hsiu Hsieh. Randomized benchmarking for non-markovian noise. *PRX Quantum*, 2:040351, 2021.

[112] Pedro Figueroa-Romero, Kavan Modi, and Min-Hsiu Hsieh. Towards a general framework of randomized benchmarking incorporating non-markovian noise. *Quantum*, 6:868, 2022.

[113] Adam Winick, Joel J Wallman, Dar Dahlen, Ian Hincks, Egor Ospadov, and Joseph Emerson. Concepts and conditions for error suppression through randomized compiling. arXiv, 2022.

[114] J. I. Cirac and P. Zoller. Quantum computations with cold trapped ions. *Phys. Rev. Lett.*, 74:4091–4094, 1995.

[115] Wolfgang Paul. Electromagnetic traps for charged and neutral particles. *Rev. Mod. Phys.*, 62:531–540, 1990.

[116] Samuel Earnshaw. On the nature of the molecular forces which regulate the constitution of the luminiferous ether. *Trans. Camb. Philos. Soc.*, 7:97, 1848.

[117] Liudmila Zhukas. *Novel Methods in Trapped-Ion Quantum Computing: Single-Photon-Sensitive Time-Resolving Camera, Sympathetic Cooling, and Qutrit*. PhD thesis, University of Washington, 2021.

[118] Seyed Shakib Vedaie. *Framework for Learning and Control in the Classical and Quantum Domains*. PhD thesis, University of Calgary, 2024.

[119] D. Leibfried, R. Blatt, C. Monroe, and D. Wineland. Quantum dynamics of single trapped ions. *Rev. Mod. Phys.*, 75:281–324, 2003.

[120] H Haffner, C Roos, and R Blatt. Quantum computing with trapped ions. *Phys. Rep.*, 469(4):155–203, 2008.

[121] I. Pogorelov, T. Feldker, Ch. D. Marciniak, L. Postler, G. Jacob, O. Krieglsteiner, V. Podlesnic, M. Meth, V. Negnevitsky, M. Stadler, B. Höfer, C. Wächter, K. Lakhmanskiy, R. Blatt, P. Schindler, and T. Monz. Compact ion-trap quantum computing demonstrator. *PRX Quantum*, 2:020343, 2021.

[122] Anders Sørensen and Klaus Mølmer. Entanglement and quantum computation with ions in thermal motion. *Phys. Rev. A*, 62:022311, 2000.

[123] Christian F Roos. Ion trap quantum gates with amplitude-modulated laser beams. *New J. Phys.*, 10(1):013002, 2008.

[124] Gabriele Nebe, E M Rains, and N J A Sloane. The invariants of the Clifford groups. *Des. Codes Cryptogr.*, 24(1):99–122, 2001.

[125] J Tolar. On Clifford groups in quantum computing. *J. Phys.: Conf. Ser.*, 1071:012022, 2018. arXiv: 1810.10259.

[126] The GAP Group. *GAP – Groups, Algorithms, and Programming, Version 4.11.0*, 2020.

[127] Dudley Littlewood. *The Theory of Group Characters and Matrix Representations of Groups*. AMS Chelsea Publishing. American Mathematical Society, Providence, RI, March 2006.

[128] M A Fogarty, M Veldhorst, R Harper, et al. Nonexponential fidelity decay in randomized benchmarking with low-frequency noise. *Phys. Rev. A*, 92(2):022326, 2015.

[129] Jonas Helsen, Xiao Xue, Lieven M K Vandersypen, et al. A new class of efficient randomized benchmarking protocols. *npj Quantum Inf.*, 5(1):1–9, 2019.

[130] Joel J Wallman, Marie Barnhill, and Joseph Emerson. Robust characterization of leakage errors. *New J. Phys.*, 18(4):043021, 2016.

[131] Mohan Sarovar, Timothy Proctor, Kenneth Rudinger, Kevin Young, Erik Nielsen, and Robin Blume-Kohout. Detecting crosstalk errors in quantum information processors. *Quantum*, 4:321, 2020.

[132] Jay M Gambetta, A D Córcoles, S T Merkel, et al. Characterization of Addressability by Simultaneous Randomized Benchmarking. *Phys. Rev. Lett.*, 109(24):240504, 2012.

[133] Reinier W Heeres, Philip Reinhold, Nissim Ofek, et al. Implementing a universal gate set on a logical qubit encoded in an oscillator. *Nat. Commun.*, 8(1), 2017.

[134] Shelby Kimmel, Marcus P da Silva, Colm A Ryan, et al. Robust extraction of tomographic information via randomized benchmarking. *Phys. Rev. X*, 4(1):011050, 2014.

[135] E Onorati, A H Werner, and J Eisert. Randomized benchmarking for individual quantum gates. *Phys. Rev. Lett.*, 123:060501, 2019.

[136] Juan José García Ripoll. *Quantum Information and Quantum Optics with Superconducting Circuits*. Cambridge UP, 2022.

[137] Pei Liu, Ruixia Wang, Jing-Ning Zhang, Yingshan Zhang, Xiaoxia Cai, Huikai Xu, Zhiyuan Li, Jiaxiu Han, Xuegang Li, Guangming Xue, Weiyang Liu, Li You, Yirong Jin, and Haifeng Yu. Performing SU(d) operations and rudimentary algorithms in a superconducting transmon qudit for $d = 3$ and $d = 4$. *Phys. Rev. X*, 13:021028, 2023.

[138] M Kononenko, M A Yurtalan, S Ren, et al. Characterization of control in a superconducting qutrit using randomized benchmarking. *Phys. Rev. Res.*, 3(4):L042007, 2021.

[139] Lennart Maximilian Seifert, Ziqian Li, Tanay Roy, David I Schuster, Frederic T Chong, and Jonathan M Baker. Exploring ququart computation on a transmon using optimal control. *Phys. Rev. A*, 108:062609, 2023.

[140] Vivek V Shende, Igor L Markov, and Stephen S Bullock. Minimal universal two-qubit controlled-NOT-based circuits. *Phys. Rev. A*, 69:062321, 2004.

[141] Frank Celler, Charles R. Leedham-Green, Scott H. Murray, Alice C. Niemeyer, and E.A. O'brien. Generating random elements of a finite group. *Commun. Algebra*, 23(13):4931–4948, 1995.

[142] I. Pak. The product replacement algorithm is polynomial. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, page 476, Los Alamitos, CA, USA, nov 2000. IEEE Computer Society.

[143] Jonas Helsen, Joel J Wallman, Steven T Flammia, et al. Multi-qubit randomized benchmarking using few samples. *Phys. Rev. A*, 100(3):032304, 2019.

[144] Mark Howard, Joel Wallman, Victor Veitch, et al. Contextuality supplies the 'magic' for quantum computation. *Nature*, 510(7505):351–355, 2014.

[145] Shelly Garion, Naoki Kanazawa, Haggai Landa, et al. Experimental implementation of non-Clifford interleaved randomized benchmarking with a controlled-S gate. *Phys. Rev. Res.*, 3:013204, 2021.

[146] J Helsen, M Ioannou, J Kitzinger, et al. Shadow estimation of gate-set properties from random sequences. *Nat. Commun.*, 14(1), 2023.

[147] Noah Goss, Alexis Morvan, Brian Marinelli, et al. High-fidelity qutrit entangling gates for superconducting circuits. *Nat. Commun.*, 13(1):7481, 2022.

[148] Steven M Roman. *The Umbral Calculus*. Acad. Press, Orlando, 1984.

[149] Jacques Touchard. Sur les cycles des substitutions. *Acta Math.*, 70(none):243 – 297, 1939.

[150] N G Van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier, San Diego, 2007.

Appendix A

Proof of an identity for Bell numbers

In this appendix, I prove the mathematical result on which our results are grounded. I follow standard notation [72].

Consider the set $\{1, \ldots, n\}$ with n elements. Given a permutation σ of n elements with k cycles we define a vector. Let σ be a permutation and $c_i \in \{0, \ldots, n\}$ be be the number of i-cycles in σ. I define the following list of $\boldsymbol{c} = (c_1, c_2, \ldots, c_n)$. For an integer partition λ, the number of times the integer i appears in λ is $c(\lambda, i)$. Notice, in the context of the same permutation σ in a conjugacy class labelled by λ, $c_i = c(\lambda, i)$. I use j for the number of times a number appear in an integer partition, I use i for c_i.

Definition A.1. The number of permutations with cycle decomposition $\boldsymbol{c} = (c_1, c_2, \ldots)$ with k different cycles is denoted by

$$q(n, k, \boldsymbol{c}) \coloneqq \frac{n!}{\prod_i i^{c_i} c_i!}. \tag{A.1}$$

As an example, consider the partition $\lambda = (2, 1)$ of $n = 3$; it contains one 1-cycle and one 2-cycle and no 3-cycle so $c_1 = 1, c_2 = 1, c_3 = 0$. Therefore, $q(3, 2, (1, 1, 0)) = 3$ and we know there are three permutations with cycle decomposition $\boldsymbol{c}$: $(1, 2)$, $(1, 3)$, and $(2, 3)$.

Definition A.2. I define

$$p(n,k,s;\boldsymbol{c}) \coloneqq c_1^s q(n,k,\boldsymbol{c}) = \frac{c_1^s n!}{\prod_i i^{c_i} c_i!}. \tag{A.2}$$

I assign an integer partition λ to each $\boldsymbol{c}$ as follows: the entry c_i denotes the number of times i appears in the λ.

Lemma A.3. *Let Φ be the generating function of p. Then*

$$\Phi(t,1;x_i=1) = \sum_{k<s}\left(\sum_{k_1}^{k} S(s,k_1)\right) t^k + \sum_{k\geq s} B_s t^k, \tag{A.3}$$

where B_s is the s-th Bell number and $S(s,k_1)$ is a Stirling second-kind number.

Proof. We compute the generating function of p with an infinite number of variables

$$\Phi\left(t,u;\boldsymbol{x}=(x_1,\ldots)\right) \coloneqq \sum_{n,k,\boldsymbol{c}} p(n,k,s;\boldsymbol{c}) u^k \frac{t^n}{n!} x_1^{c_1} x_2^{c_2} \cdots, \tag{A.4}$$

where $\boldsymbol{x}$ is a vector of an infinite countable dummy variables. By definition q of Eq. (A.2) satisfies $q(n,k;\boldsymbol{c}) \neq 0$ only if

$$\sum_i c_i = k \text{ and } \sum_i i c_i = n. \tag{A.5}$$

The justification is that, given a $\boldsymbol{c}$, the values of k and n are fixed. Consider for instance $\boldsymbol{c} = (1,1,0)$. Is invalid to have either $n \neq 3$ or $k \neq 2$.

I substitute the values of n and k (as given in Eq. (A.5)) in terms of $\boldsymbol{c}$ in Eq. (A.19). I get:

$$\Phi(t,u;\boldsymbol{x}) = \sum_{\boldsymbol{c}} \frac{c_1^s u^{\sum_j c_j}}{\prod_i c_i! i^{c_i}} t^{\sum_i i c_i} \left(\prod_{l\geq 1} x_l^{c_l}\right). \tag{A.6}$$

For clarity, I expand Eq. (A.6)

$$\Phi(t,u;\boldsymbol{x}) = \sum_{c_1,c_2,\ldots\geq 0} \frac{c_1^s n!}{c_1!c_2!\cdots 1^{c_1}2^{c_2}\cdots} u^{c_1+c_2+\cdots} \frac{t^{c_1+2c_2+\cdots}}{n!}. \tag{A.7}$$

Grouping terms in Eq. (A.7) according to c_i:

$$\Phi(t,u;\boldsymbol{x}) = \sum_{c_1,c_2,\ldots\geq 0} \frac{c_1^s}{c_1!}(tux_1)^{c_1}\frac{1}{c_2!}(tux_2)^{c_2}\frac{1}{c_3!}(tux_3)^{c_3}\cdots \tag{A.8}$$

Is useful to shorten the expression Eq. (A.8) in the following way:

$$\Phi(t,u;\boldsymbol{x}) = \prod_{i\geq 1}\sum_{c_i\geq 0} c_1^s \frac{[(t_i/i)ux_i]^{c_i}}{c_i!}. \tag{A.9}$$

Now, I make two observations. My first observation applies to $i > 1$:

$$\sum_{c_i\geq 0}\frac{[(t_i/i)ux_i]^{c_i}}{c_i!} = \exp\left[\frac{t^i}{i}ux_i\right]. \tag{A.10}$$

Applying Eq. (A.10) in Eq. (A.9)

$$\Phi(t,u;\boldsymbol{x}) = \sum_{c_1\geq 0}\frac{c_1^s tux_1}{c_1!}\prod_{i>1}\exp\left[\frac{t^i}{i}ux_i\right] \tag{A.11}$$

The second observation uses Dobiński's formula [148]:

$$\sum_{c_1\geq 0} c_1^s\frac{tux_1}{c_1!} = \exp[tux_1]B_s(tux_1). \tag{A.12}$$

Applying Eq. (A.12) on Eq. (A.11) I get

$$\Phi(t,u;\boldsymbol{x}) = B_s(tux_1)\prod_{i\geq 1}\exp\left[\frac{t^i}{i}ux_i\right]. \tag{A.13}$$

Setting $x_i = 1$ for all $i \geq 1$ in Eq. (A.13), I get

$$\Phi(t, u; x_i = 1) = B_s(tu)(1-t)^{-u}. \tag{A.14}$$

Similarly setting $u = 1$ in Eq. (A.14)

$$\Phi(t, 1; x_i = 1) = \frac{B_s(t)}{1-t}. \tag{A.15}$$

I recall that $B_s(1) = B_s$ is the s-th Bell number. Also,

$$\frac{1}{1-t} = \sum_n t^n$$

is the exponential generating function of $n!$. Bell polynomials (also named Touchard polynomials [149]) are given in terms of Stirling second kind numbers—$S(n, k)$ the number of equivalence relations with k classes on $[s]$ [72]—as:

$$B_s(t) \coloneqq \sum_{k=0}^{s} S(s, k) t^k; \tag{A.16a}$$

Bell numbers are thus computed as

$$B_s \coloneqq B_s(0) = \sum_{k=0}^{s} S(s, k). \tag{A.16b}$$

Thus, substituting Eqs. (A.16) in Eq. (A.15), I obtain:

$$\Phi(t,1;x_i=1) = \sum_{k_0\geq 0} B_s(t)t^{k_0} = \sum_{k_0\geq 0}\left(\sum_{k_1}^{s} S(s,k_1)t^{k_1}\right)t^{k_0}, \tag{A.17a}$$

$$= \sum_{k_0\geq 0}\left(\sum_{k_1}^{s} S(s,k_1)\right)t^{k_0+k_1}, \tag{A.17b}$$

$$= \sum_{k_0\geq 0}\left(\sum_{k_1}^{\min(s,k_0)} S(s,k_1)\right)t^{k_0}, \tag{A.17c}$$

$$= \sum_{k<s}\left(\sum_{k_1}^{k} S(s,k_1)\right)t^{k} + \sum_{k\geq s} B_s t^k. \tag{A.17d}$$

□

Theorem A.4. *Let n be a positive integer, B_s the s-th Bell number with $s \geq n$. For a given integer partition $\lambda \vdash n$, let $c(\lambda, j)$ denote the number of times j appears in λ. Then*

$$B_s = \sum_{\lambda\vdash n} \frac{c(\lambda,1)^s}{\prod_j j^{c(\lambda,j)}c(\lambda,j)!}. \tag{A.18}$$

Proof. I compute the generating function of p

$$\Phi\left(t,u;\boldsymbol{x}=(x_1,\dots)\right) \coloneqq \sum_{n,k,\boldsymbol{c}} p(n,k,s;\boldsymbol{c})u^k\frac{t^n}{n!}x_1^{c_1}x_2^{c_2}\cdots, \tag{A.19}$$

where $\boldsymbol{x}$ is a vector containing a countable number dummy variables.

Setting, for all i, $x_i = 1$ in Eq. (A.19), I obtain

$$\Phi(t,u;1,1,\dots) = \sum_{n,k,\boldsymbol{c}} p(n,k;\boldsymbol{c})u^k\frac{t^n}{n!}. \tag{A.20}$$

Given n and k, I define

$$p'(n,k) \coloneqq \sum_{\boldsymbol{c}} p(n,k;\boldsymbol{c}), \tag{A.21}$$

where the value of n and k constrain the values of $\boldsymbol{c}$, as in Eq. (A.5). There is no contradiction

between Eqs. (A.21) and (A.6), as in Eq. (A.6) I am also (implicitly) summing over n and k in Eq. (A.20):

$$\Phi(t,u;x_i=1)=\sum_{n,k}p(n,k)'\frac{t^n}{n!}u^k=\sum_n\left(\sum_k p(n,k)'u^k\right)\frac{t^n}{n!}, \tag{A.22}$$

setting $u=1$

$$\Phi(t,1;x_i=1)=\sum_n\left(\sum_k p'(n,k)\right)\frac{t^n}{n!}, \tag{A.23}$$

Therefore, matching coefficients of Eq. (A.17d) in Lemma A.3 and Eq. (A.23), I obtain—for $n\geq s$—

$$\sum_k p'(n,k)=\sum_{\boldsymbol{c}}\frac{c_1^s n!}{\prod_i i^{c_i}c_i!}=\sum_\lambda\frac{c(\lambda,1)^s n!}{\prod_j j^{c(\lambda,j)}c(\lambda,j)}=B_s n!. \tag{A.24}$$

□

Appendix B

Miscellaneous computations

B.1 Motivation for the Clifford hierarchy

In this section, I justify the unconventional definition of the Clifford hierarchy mentioned for first time in my thesis on chapter 1. By unconventional I refer to the fact it would be more natural to define the hierarchy in terms of mapping from one level to a previous one, instead of a most natural transformations mapping Pauli to the previous level of the hierarchy.

I consider the teleportation circuit as a mapping that takes as input a state $|\alpha\rangle$ and a gate U. The result is $UP|\alpha\rangle$, where P is an element of the Pauli group. I denote this stochastic mapping $TP(|\alpha\rangle, U) = UP|\alpha\rangle$.

If U is a member of the second level of the Clifford hierarchy $UP = P'U$ where P' is another member of the Pauli group. If T is a member of the third level of the Clifford hierarchy $TP = CT$ where C is a member of the second level of the Clifford hierarchy. We apply the mapping once again but the input is now $CT|\alpha\rangle$ and C^{-1}. The resulting state is $C^{-1}PCT|\alpha\rangle$. Using C is a Clifford gate $C^{-1}PCT|\alpha\rangle = C^{-1}CP'T|\alpha\rangle = P'T|\alpha\rangle$. As the higher levels of the Clifford hierarchy do not form a group, C must be used instead of C^{-1} in the commutation with P'.

Therefore, each level of the Clifford hierarchy to which some gate is a member is equal

to the number of times the teleportation protocol circuit needs to be applied to recover the action on the initial state $|\alpha\rangle$.

In this brief comment, I highlight that a qubit operation that is Clifford may be no longer Clifford in a higher-level system. This discussion is important since it shows that the basis that substitutes the Pauli matrices satisfies that its normaliser forms a unitary 2-design.

I show that $X_{(12)}$ (the qubit X gate embedded in a ququart system) (5.1) is not a Clifford operation. To this end, we introduce a recursive map

$$B(g,h,i) \coloneqq B(g,g^h,i-1),\ B(g,h,1) \coloneqq g^h = \bar{h}gh. \tag{B.1}$$

I say a gate g belongs to the k-level of the Clifford hierarchy if $B(g,h,k)$—h is a Heisenberg-Weyl matrix—is a member of the qudit Pauli group. Using B with $h = X$, a trivial computation reveals $X_{(12)}$ is not a member of any level of the Clifford hierarchy.

B.2 Proof and estimate of the total variation distance

In Ref. [59] there is a statement linking the total variation distance and the diamond distance. I explicitly shows this link as follows. The original statement is

$$\text{As } N \to \infty, r_{\min} \to \|p_{\text{id}} - p_{\text{ac}}\|_{\text{TV}}, \tag{B.2}$$

where $\|.\|_{\text{TV}}$ is the total variation distance between two distributions. I prove this claim using elementary calculus notions.

Let

$$E_N \coloneqq \left\| \frac{n(x)}{N} - p_{\text{ac}}(x) \right\|. \tag{B.3}$$

Consider r such that

$$E_N = \left\| \frac{n(x)}{N} + r - p_{\text{id}}(x) \right\|, \tag{B.4}$$

which should be thought as the fraction of manipulated outcomes needed to transform the distribution p_{ac} into p_{id}. Without loss of generality, I assume

$$\frac{n(x)}{N} < p_{\text{id}}(x). \tag{B.5}$$

Then we can solve for r to obtain two solutions:

$$r = p_{\text{id}}(x) - \frac{n(x)}{N} \pm E_N. \tag{B.6}$$

Therefore, we can write the equivalent equation to Eq. (B.2)

$$\lim_{N\to\infty} r = \|p_{\text{id}} - p_{\text{ac}}\|_{\text{TV}}. \tag{B.7}$$

Proof. The proof relies on two steps. First,

$$\lim_{N\to\infty} \frac{n(x)}{N} = p_{\text{ac}}(x). \tag{B.8}$$

Lastly,

$$\lim_{N\to\infty} E_N = \lim_{N\to\infty} \left\| \frac{n}{N} - p_{\text{ac}} \right\| = 0, \tag{B.9}$$

where the last equality follows from the definition of the quotient and the continuity of the norm. □

We notice that for a pair of Bernoulli random variables p and q

$$\|p - q\|_{\text{TV}} = \|p(x) - q(x)\|, \tag{B.10}$$

for any x in the shared domain of p and q.

B.3 GAP code for qutrit HDG

In this appendix, I write the code used to compute the decomposition into irreps and the characters. First we introduce the generators of the group into `GAP` [126]:

```
  xgate := [[0,0,1],[1,0,0],[0,1,0]];
  tgate := DiagonalMat([1,E(9), E(9)^8]);
group := Group(xgate, tgate);
```

The next step is to compute the character of this representation. From this character, I compute the character of the PL representation. This character is the product of the character and its complex conjugate.

```
  character := NaturalCharacter(group);;
  plrep := character*ComplexConjugate(character);;
```

Then I compute the decomposition using the multiset of characters in `plrep`:

```
  decomposition := ConstituentsOfCharacters(plrep);;
```

Finally, I compute the projectors in `Mathematica` by exporting the character table.

B.4 Proof number of independent parameters in chi-representation

In this section I prove that the number of independent parameters of a chi-representation for a quantum channel $\mathcal{E}$ is $d^4 - d^2$, with d the dimension of the Hilbert space. The proof is based on two claims: the chi-representation is an Hermitian matrix, and that there are d^2 constraints due the trace preserving condition on the Kraus operators $\{A_i\}$ of $\mathcal{E}$.

Definition B.1. Let $\mathcal{E}$ be a channel with Kraus operators $\{A_i\}$. Consider the basis $\{|i\rangle\langle j|\}$ of operators. Expressing each

$$A_i = \sum_{n,m} e_{i,nm} \, |n\rangle\langle m| \, . \tag{B.11}$$

Then the chi-representation of $\mathcal{E}$ is a matrix with entries defined as

$$\chi(\mathcal{E})_{nm,uv} \coloneqq \sum_i e_{i,nm} \bar{e}_{i,uv}. \tag{B.12}$$

Proposition B.2. *Let $\mathcal{E}$ be a channel with Kraus operators $\{A_i \colon i \in I\}$, with I an index-set. Then the matrix χ is Hermitian.*

Proof. I simply compute the complex conjugate:

$$\bar{\chi}_{mn,uv} = \overline{\sum_i e_{i,mn} \bar{e}_{i,uv}}, \tag{B.13}$$

$$= \sum_i \bar{e}_{i,mn} e_{i,uv}, \tag{B.14}$$

$$= \sum_i e_{i,uv} \bar{e}_{i,mn}, \tag{B.15}$$

$$= \chi_{uv,mn}. \tag{B.16}$$

□

Proposition B.3. *The trace-preserving condition on the Kraus operators of a channel $\mathcal{E}$ fixes d^2 parameters of χ.*

Proof. Here I use a explicitly the basis on which χ is defined. Consider the Kraus operators expressed on a bi-label basis

$$A_i = \sum_i e_{i,kl} \, |k\rangle\langle l| \, , \tag{B.17}$$

by this, then I use four indices to write the χ representation

$$\chi_{kl,uv} = \sum_i e_{i,kl} \bar{e}_{i,uv} = \sum_i \, \langle k|A_i|l\rangle \, \langle v|A_i^\dagger|u\rangle \, , \tag{B.18}$$

Then, from the CPTP condition ($\sum_i A_i^\dagger A_i = \mathbb{I}$) I obtain

$$\delta_{k,l} = \langle k| \sum_i A_i^\dagger A_i |l\rangle \,, \tag{B.19}$$

$$= \sum_{i,j} \langle k|A_i^\dagger|j\rangle \, \langle j|A_i|l\rangle \,, \tag{B.20}$$

$$= \sum_j \chi_{jl,jk} = \delta_{k,l}. \tag{B.21}$$

This finishes the proof since l and k take d parameters. □

B.5 Discussion numerical confidence intervals

In this appendix, I describe a numerical analysis that justifies the feasibility of my scheme. A numerical analysis of the feasibility must consider three experimental values: the number of different circuit depths used, the number of randomly sampled circuits per circuit depth, and the number of shots. The number of shots refers to the number of times a state is prepared, the circuit applied, and then measured.

In what follows, I present the numerical analysis of these variables based on the average gate fidelity of the gates, their confidence intervals, and the error of the estimate.

Shots	Circuits	Error given a probability			Fidelity
s	r	0.95	0.999	1	
100	10	0.013	0.020	0.024	0.89
10	100	0.013	0.021	0.021	
100	100	0.003	0.006	0.008	

Table B.1: Table reporting the confidence values for the estimate of the parameter η_0. The noise corresponds to a randomly sampled channel with average gate fidelity 0.89.

Shots	Circuits	Error given a probability			Fidelity
s	r	0.95	0.999	1	
100	10	0.008	0.013	0.014	
10	100	0.008	0.014	0.015	0.931339
100	100	0.002	0.004	0.004	
20	100	0.005	0.008	0.009	
20	20	0.013	0.022	0.022	

Table B.2: Table reporting the confidence values for the estimate of the parameter η_0. The noise model is a composition of a totally depolarising and an amplitude damping channel, which is a comprehensive Markovian and completely positive noise model. The value of fidelity corresponds to the parameters 0.01 for the depolarising and amplitude damping channel.

Shots	Circuits	Error given a probability			Fidelity
s	r	0.95	0.999	1	
100	10	0.005	0.007	0.008	
10	100	0.005	0.008	0.009	0.958284
100	100	0.001	0.002	0.002	
20	100	0.003	0.004	0.005	
20	20	0.008	0.011	0.012	

Table B.3: Table reporting the confidence values for the estimate of the parameter η_0. The noise model is a composition of a totally depolarising and an amplitude damping channel, which is a comprehensive Markovian and completely positive noise model. The value of fidelity corresponds to the parameters 0.03 for the depolarising and amplitude damping channel.

Shots	Circuits	Error given a probability			Fidelity
s	r	0.95	0.999	1	
100	10	0.002	0.002	0.003	
10	100	0.002	0.003	0.004	0.985921
100	100	0.000	0.000	0.001	
20	100	0.001	0.002	0.002	
20	20	0.003	0.006	0.006	

Table B.4: Table reporting the confidence values for the estimate of the parameter η_0. The noise model is a composition of a totally depolarising and an amplitude damping channel, which is a comprehensive Markovian and completely positive noise model. The value of fidelity corresponds to the parameters 0.05 for the depolarising and amplitude damping channel.

Each table should be read as follows. If the noise is known or it is suspected to have an average gate fidelity value in one of the values on Tables B.1-B.4 then, according to the confidence expected to obtain (we use 0.95, 0.999, and 1), the number of shots and circuits required to obtain such confidence intervals are displayed. For concreteness, we explain the reading of Tab. B.1: fixing the number of shots and circuits to 100 and for a noise with fidelity 0.89, the values resulting from the fitting of Eq. (5.27) have an error of 0.003 with a frequency of 0.95, error 0.006 with frequency 0.999, and error 0.008 in any case. This could also be written as $\text{Prob}(|\eta_0 - \check{\eta}_0| > 0.003) = 1 - 0.95$, $\text{Prob}(|\eta_0 - \check{\eta}_0| > 0.006) = 1 - 0.999$, and $\text{Prob}(|\eta_0 - \check{\eta}_0| > 0.008) = 0$, respectively.

For the ququart case, we get similar results, as presented in Tables B.5-B.7:

Shots	Circuits	Error given a probability			Fidelity
s	r	0.95	0.999	1	
100	10	0.008	0.014	0.015	0.90
10	100	0.009	0.014	0.018	
10	10	0.027	0.050	0.052	
20	20	0.013	0.021	0.023	

Table B.5: Table reporting the confidence values for the estimate of the parameter. In contradistinction with the qutrit case, the noise model corresponded to a randomly sampled channel with an average gate fidelity 0.90.

Shots	Circuits	Error given a probability			Fidelity
s	r	0.95	0.999	1	
100	10	0.001	0.002	0.002	0.99
10	100	0.001	0.002	0.002	
10	10	0.004	0.006	0.008	
20	20	0.002	0.003	0.004	

Table B.6: Table reporting the confidence values for the estimate of the parameter. In contradistinction with the qutrit case, the noise model corresponded to a randomly sampled channel with an average gate fidelity 0.99.

Shots	Circuits	Error given a probability			Fidelity
s	r	0.95	0.999	1	
100	10	0.004	0.007	0.007	0.95
10	100	0.004	0.007	0.007	
10	10	0.013	0.026	0.034	
20	20	0.006	0.012	0.013	

Table B.7: Table reporting the confidence values for the estimate of the parameter. In contradistinction with the qutrit case, the noise model corresponded to a randomly sampled channel with an average gate fidelity 0.95.

From the numerical evidence gathered above we see that the number of shots and random circuits required is around 20. Our numerical results suggest that doubling the shots and circuits exponentially decreases the error.

Strategy	Error given a probability		
circuit depth	0.95	0.999	1
i	0.002	0.003	0.004
ii	0.002	0.003	0.004
iii	0.003	0.004	0.006
iv	0.001	0.001	0.002

Table B.8: Comparison of the strategies. The result is based on the confidence intervals. For the noise we used a depolarising channel composed with a phase-damping; the depolarising and dephasing parameters are set to 0.1.

Appendix C

Classical randomised benchmarking

In this appendix, I introduce a scheme that does not rely on quantum mechanics but that illustrates the main components of an RB schemes. I call it classical randomised benchmarking. Without going into full generality, I present the simplest case. I will finish with a table that for reference related the classical quantities with the quantities used in randomised benchmarking schemes.

I consider a faulty switch (FS). A switch is understood as a device that if the status is ON and you press the switch you get OFF and if the status is OFF then the status after pressing the switch is ON. Assume that, with probability p, the switch realises the mapping ON to OFF and with probability q the mapping OFF to ON. This is modeled as the transfer matrix [150]

$$M_{\mathrm{FS}} \coloneqq \begin{array}{c} \begin{array}{cc} \mathrm{ON} & \mathrm{OFF} \end{array} \\ \begin{bmatrix} 1-p & q \\ p & 1-q \end{bmatrix} \begin{array}{c} \mathrm{ON} \\ \mathrm{OFF} \end{array} \end{array} . \tag{C.1}$$

The probability vector is a vector with entries denoting the probability that the state is ON and OFF:

$$v_{\mathrm{FS}} \coloneqq \begin{bmatrix} x \\ y \end{bmatrix} \begin{array}{c} \mathrm{ON} \\ \mathrm{OFF} \end{array} . \tag{C.2}$$

For this example, characterisation means obtaining p, since this parameter quantifies the quality of the switch.

The final ingredient is the Bhattacharyya coefficient (BC), also known as classical fidelity [107]. Consider two probability distribution q_i and p_i. Then, the BC between p and q is

$$\mathrm{BC}(p,q) = \sum_i \sqrt{p_i q_i}. \tag{C.3}$$

The square of the BC is known as the classical fidelity [107]. Now, I describe the algorithm to estimate q; that is, I describe the algorithm to estimate the performance of the faulty switch.

- Pick m_{c} an even positive integer.
- Pick N_{c} a positive integer.

For each $m_i \in [m]^{\mathrm{even}}$, press the switch m_i times, record 1 if the output of the final sequence of operations is ON and 0 if the outcome is OFF. Let $N^{(m_i)}$ be the sum of ones divided by N_{c}.

Then, the curve m_i *vs* $N^{(m_i)}$ is an exponential function. This can be seen by computing the following inner product

$$v_{\mathrm{FS}}^{\top} M_{\mathrm{FS}}^{m} v_{\mathrm{FS}} = A + Bq^m. \tag{C.4}$$

By the experiment mentioned above, the parameter q can be estimated. Thus, by the scheme described above, the quality of the FS is estimated.

[. . .] y así, por toda dulzura
nunca yo me perderé [. . .]